



Continuous meteorological surface and soil records (2004–2024) at the Met Office surface site of Cardington, UK.

Simon R. Osborne¹, Jennifer K. Brooke¹, Bernard M. Claxton¹, Tony Jones¹, Amanda M. Kerr-Munslow¹,
5 James R. McGregor¹, Emily G. Norton², Nicola Phillips¹, Martyn A. Pickering¹, Jeremy D. Price¹, Jenna
Thornton¹, Graham P. Weedon³

¹ Met Office, FitzRoy Road, Exeter, EX1 3PB, UK

10 ² National Centre for Atmospheric Science (NCAS), University of Manchester, Oxford Road, Manchester, M13 9PL, UK

³ 2 Millbrook Dale, Axminster, Devon, EX13 5EF, UK

Correspondence to: Simon R. Osborne (simon.osborne@metoffice.gov.uk)

Abstract. A continuous meteorological and hydrological observational record is described of the Met Office semi-rural field site of Cardington in southern England between 2004 and 2024. The site was designed to carry out boundary layer, fog and
15 air-surface exchange research to improve the representation of process-based physics within the Met Office Unified Model. The site lay in a flat river basin and was laid mainly to cropped grass and was surrounded by arable fields intermixed with small trees and shrubs through most wind sectors. Observations utilised flux masts at various heights, visibility, radiosondes, very near-surface and subsoil in situ sensors in addition to more specialist remote sensing instruments to retrieve atmospheric properties. In addition to boundary layer and surface data, soil properties such as temperature, moisture and water table depth
20 were obtained. All components of the surface energy balance could be determined. Availability of data based on 30 minute time steps over 20 yr, for the combined components of the energy balance not flagged as either bad or missing, amounts to 77 %. The momentum roughness length as determined at the 10 m height for the prevailing wind sector increased from 3 cm to 8 cm over the period predominately due to 52 ha of woodland growth within 1 km of the site. An overview of the site, instrumentation, data availability, quality control, data storage at the UK CEDA repository, and potential uses of the dataset
25 are described. A set meteorological forcing files have also been compiled suitable for driving standalone land surface models configured for a single point.

Copyright statement. The works published in this journal are distributed under the Creative Commons Attribution 4.0 License. This licence does not affect the Crown copyright work, which is re-usable under the Open Government Licence (OGL). The
30 Creative Commons Attribution 4.0 License and the OGL are interoperable and do not conflict with, reduce or limit each other.

© Crown Copyright 2025.



1 Introduction

It is well known that geographically dispersed ground-based weather data from both professional and private (‘citizen science’) automatic weather station sources are crucial to the initialisation and therefore operation of numerical weather prediction (NWP) models (Rawlins et al., 2007; Bell et al., 2015). On the other hand, specialist surface sites at locations such as Cabauw in the Netherlands (Bosveld et al., 2020), the SIRTa observatory 20 km southwest of Paris, France (Chiriaco et al., 2018), the three fixed ARM sites of the Atmospheric Measurement Facility (Miller et al., 2016), and Cardington in southern England provide datasets for in-depth research in order to improve the explicit and parametrised physics within NWP. This can be achieved either via a statistical approach using long-term data, or special intensive observation periods (IOPs) lasting between hours and a few days, or indeed a combination of both methods. Although a fixed surface site cannot be readily applied to large spatial scales, except via remote sensing such as with radars and lidars, it allows analysis over a wide range of time scales from minutes (e.g. fog development) to the months or years (e.g. changes in deep soil moisture content). This therefore allows for model forecast evaluation across a range of time scales as well as the development of parametrizations, whether these be “full parametrizations” or partially resolving such as within the turbulent grey zone (Wyngaard, 2004). Although the Cardington site has proven useful for both evaluation (Price et al., 2018) and parametrization development (Haywood et al., 2008; Boutle et al., 2014), there remains much untapped data that the wider community is now welcome to access. Table 1 shows a list of research projects and campaigns from the past 20 yrs that used Cardington data. The datasets can be download from the UK-based Centre for Environmental Data Analysis (CEDA) repository for atmospheric and earth sciences observation data.

Years used	Short title	Campaign or otherwise	Data type / instruments	Reference
2003	LES simulation of evening transition case studies	IOP <i>ad hoc</i>	In situ; turbulent fluxes	Beare et al (2006)
2006–07	NWP forecasting of visibility	VISURB	In situ; visiometer, aerosols, nephelometer	Haywood et al (2008)
2006–07	NWP forecasts based on seasonal diurnal cycles	Multi-year	In situ, radiation; surface energy balance, screen and skin temperature biases	Edwards et al (2011)
2006–07	Comparison of two closely located observation sites	SIREX	In situ, radiation; energy balance	Horlacher et al (2012)
2008	Stability within radiation fog	IOP <i>ad hoc</i>	In situ, radiation. Turbulent fluxes, soil temperature	Price (2011)



2011	Evolution of stratocumulus over land	COALESC	In situ, radiation, remote sensing. Microwave radiometer, Doppler lidar, radiosondes	Osborne et al (2014)
2012	Persistent fog cases	IOP ad hoc	In situ, radiation. Fluxes, radiation	Price et al (2015)
2012–14	Dew meter description	IOP ad hoc	In situ. Dew and frost deposition/evaporation	Price and Clark (2015)
2008–14	Effect of gravity-wave drag on surface winds	Multi-year	In situ. Fluxes, state parameters	Lapworth et al (2015)
2010–15	Evidence for gravity-wave drag in a NWP model	Multi-year	In situ. Fluxes	Lapworth and Osborne (2016)
2014	NWP and LES modelling of radiation fog case studies	LANFEX	In situ, radiation. Flux masts	Boutle et al (2018)
2016	Comparing observations and NWP of an undular bore	Opportunistic event	In situ, radiation, remote sensing. Doppler lidar, wind profiles	Osborne and Lapworth (2017)
2014–16	Observations of radiation fog	LANFEX	In situ, radiation, remote sensing. Dew meters, infrared cameras, flux masts	Price et al (2018)
2018–19	Evapotranspiration during a meteorological drought	Seasonal	In situ, radiation. Near-surface, surface and subsoil	Osborne et al (2020)
2020	Prototype deployment of a differential absorption lidar	Seasonal	In situ, remote sensing. Radiosondes, water vapour profiles	Gaffard et al (2021)
2020	Comparison of observation and model during a dry spring	Seasonal	In situ, remote sensing, radiation. Near-surface variables Doppler lidar, radiosondes	Brooke and Osborne (2021)



2016–22	Dew, frost, fog and the lifted temperature minimum	Multi-year	In situ, radiation. Thermodynamic profiles, fog detection, visibility	Weedon et al (2024)
---------	--	------------	---	---------------------

Table 1: A modern history of research at Cardington— significant research projects leading to publications that use Cardington site data from the period 2004–2024. IOP = Intensive Observation Period, VISURB = Visibility in Urban areas, LES = Large Eddy Simulation. Other abbreviations described in the caption of Fig. 5.

Meteorological research under the leadership of Met Office scientist Maurice Giblett started in January 1925 at Cardington, where the Short brothers had established airship construction in 1916, in order to investigate effects of wind flow on airship flight using a network of masts (Giblett 1932). Two large hangars were constructed to house the R100 (Hangar No. 1) and R101 (Hangar No. 2) airships. Research into turbulence and meteorology in general therefore has a history of 100 years at Cardington. Barrage balloon development for military purposes started in 1936 and continued with operations during World War II. This led in 1943 to the use of such balloons under the RAF meteorological research unit and later the Balloon Development Establishment, established in the mid-1950s (Jones and Butler, 1958; Smith and Hay, 1961). Newly developed anemometers (Jones, 1965) were mounted on the steel tether securing the barrage balloon to the ground. This meant research into turbulent kinetic energy within the lower 1 km could be made (Readings and Rayment, 1969). Fog research was a focus in the 1970s when surface and mast instrumentation expanded at the site (Roach et al., 1976), with fog research continuing intermittently ever since, e.g. Price (2011). Despite continued use of the site for research in the following two decades (e.g. Caughey et al., 1982; Turton and Brown, 1987), continuous near surface monitoring that could fully capture the energy balance was not established until September 1996. Although turbulence research using large (200 m³) and subsequently small (50 m³) tethered helium balloons continued to be made at Cardington until 2021 (e.g. Price, 2000; Price et al, 2018; Smith et al., 2020), the associated data do not form any part of this paper.

The Cardington site was overhauled in terms of instrumentation, logging hardware and software in the spring of 2004 and so this is deemed the "modern era" for data collection that is described here. In particular, the flux masts were fixed at 10, 25 and 50 m heights, amongst other designations such as the latent heat flux at 10 m and the introduction of ground heat flux measurement, until the site closed at the end of 2024. The subsoil pits were first dug and temperature and moisture sensors installed in 1999, although again, we only include such data from 2004. Increased urbanisation immediately surrounding the site (mainly within the northwest and northeast sector) in recent years with future large scale housing developments 1–5 km planned upwind to the west (i.e. the prevailing wind direction) meant that surface-based meteorological research was compromised and therefore it was decided that the site be decommissioned at the end of 2024.

A key role of a land surface model (LSM) is to partition the surface energy balance via the fluxes of heat, moisture and momentum. Therefore in principle, field sites that observe all the components of the surface energy should be used to test and



improve such schemes. All components of energy exchange in an LSM are parametrised and assume a two-dimensional exchange of energy that is perpendicular to the surface. Observations that attempt to close the surface energy balance always encounter problems because of: (i) sensor error and drift, (ii) energy components having different footprint sizes, principally because energy is partitioned between radiative, conductive and turbulent components, i.e. there is a scale problem, and (iii) the effects of atmospheric advection. LSMs obey the conservation of energy at each time step, whereas the observations do not: the observed energy balance should be treated as unclosed (Mauder et al., 2020a).

The most accurately measured component of the observed energy budget is the net radiation. If we consider that the ground heat energy for the temperate grass site that is Cardington varies between 5 % and 10 % (depending on canopy health and soil water content) of the net radiation, then the main problem of trying to observe the remaining available energy reduces to estimating the sensible and latent heat fluxes. Depending on conditions of course, there are minor complications such as from dew fall and heat storage within the canopy. Although the radiative energy used by photosynthesis is less than 1 %, it is the transpiration that occurs whilst the plant is photosynthesising that is of significance. The importance of soil moisture on evapotranspiration and ground heat storage was the reason for installing subsoil sensors at Cardington. Historical observations of land surface evaporation are relatively scarce (Blyth et al., 2010) in contrast to precipitation and runoff data. In brief, the soil water content strongly modulates how the surface responds to atmospheric forcings. This should be born in mind when carrying out long-term simulations using LSMs: the relatively slowly changing nature of soil water content compared to atmospheric time scales, means that an anonymously wet period high soil water values can persist from weeks to seasonal time scales (e.g. Niu et al., 2011).

100

The turbulent heat fluxes are often deemed to be systematically underestimated by up to 20 % because they miss some of the energy involved (Wilson et al., 2002; Blyth et al., 2010). Studies have attempted to correct for the missing energy, for example, by the Bowen ratio method (Maayar et al., 2008) or the residual energy method (Twine et al., 2000). No attempt has been made to correct the turbulent fluxes in such a manner to the Cardington dataset, although this should be considered by the user. Confidence in the site observation of the energy terms is nonetheless reasonably good because a long-term (~ 1 yr) calculation of the energy budget remainder results in a value of 10 W m⁻² or less (Horlacher et al., 2012).

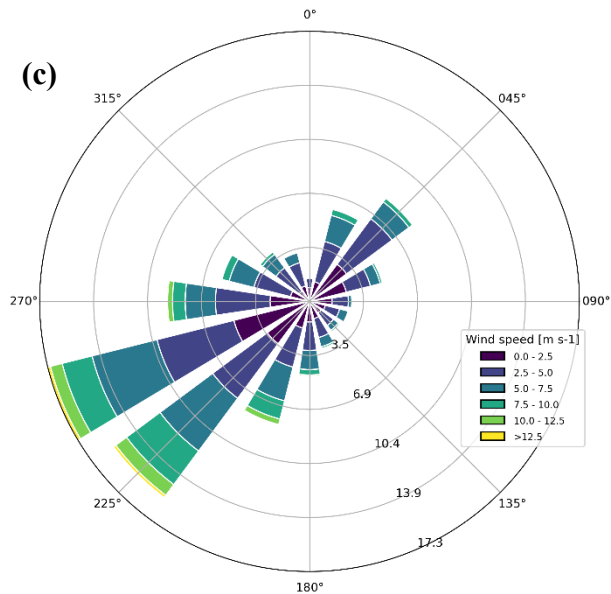
105

This paper is laid out as follows: Section 2 describes the site, Section 3 gives a breakdown of the core instrumentation, followed by how a LSM forcing dataset was derived and test from the principal dataset in Section 4. In Section 5 the large specialist radiometers are described, followed by the radiosonde archive (Section 6), an example use of the turbulence data (Section 7), and finally a description of the file formats and DOIs used in the archived products (Section 8). The supplementary sections S1–S10 tabulate the variables and date/time structure in each of the archived NetCDF files that will be referred to in relevant sections below (and summarised in Section 8).

110



Figure 1: Satellite imagery with inlaid magnetic compass of the site location taken in (a) January 2003, and (b) August 2023. The airship hangars (red outline), Cardington site (orange) and Shocott Spring woodland (yellow) are annotated. (c) A wind rose calculated over the years 2005–2009 inclusive. Satellite Images courtesy of Google Earth: © 2025 The GeoInformation Group and © 2025 Maxar Technologies.



115 2 Site description

The Cardington 18 ha site in Bedfordshire southern England ($52^{\circ} 06' 17.9''$ N $0^{\circ} 25' 26.8''$ W) was the location of the 10 m flux tower in the centre of the site) has an elevation of $29 \text{ m} \pm 1 \text{ m}$ above mean sea level and was laid mainly to manicured grass maintained at 5–10 cm height throughout the year. This area of the UK receives amongst the lowest rainfall of the country with around 550 mm per annum, alongside 1320 h of annual sunshine and a peak monthly occurrence of radiation fog occurrence of 160 hr based on the 20-yr October average. The site sits within a broad (10 km), shallow valley that is a tributary



of the River Great Ouse. The down-valley slope in general is about 0.15° . Although the immediate surroundings of the site are fairly flat, there is a ridge to the southeast with elevations of up to 100 m above mean sea level; this ridge runs along a southwest-northeast orientation and passes within 5 km of the instrumented site at its closest approach. The screen and 10 m sensors were situated in the middle of the site to be as isolated as possible. An investigation by Grant (1994) showed that the terrain surrounding Cardington can influence the wind field by channelling surface flow along the ridge in a south-west to north-east direction for a stably stratified boundary layer.

The sector 030° through 280° is in general open fetch with arable fields with changing crop types (alternating wheat, rapeseed or left fallow). The sector $280\text{--}350^\circ$ is housing (>1 km away) and sector $350\text{--}025^\circ$ is dominated by the two large airship sheds (each 247 m long by 84 m wide by 55 m high) 400–600 m away. Consequently, wind and turbulence data should be analysed with caution when the wind direction comes from the hangars ($350\text{--}025^\circ$)—though the majority of wind vectors lie outside this range. Figure 1 shows two satellite images of an area 2.5 km by 4.5 km with the site in the centre. The left image is from 2003, and the right image is from 2023. There are dramatic differences in the colour of many of the fields in the image: due to changing crop types and the time of year the photograph was taken. The increase in urbanisation immediately to the north and preparation of an area for further housing immediately to the northeast are clearly seen. The hangars and an area called Shocott Spring are annotated. Shocott Spring is an area of growing woodland situated in the sector $170\text{--}240^\circ$. The change in turbulence due to the growth of these trees is discussed in section 8. Figure 1 includes a wind rose covering the years 2005–2009, indicating that the Shocott Spring woodland lies partially within the prevailing wind direction.

The soils at the site are described geologically as loamy solifluction deposits over river valley gravels. Impervious Oxford Clay Formation underlies the whole area at an unknown depth. Soil sample analysis (Burton, 1999) shows that the topsoil (0–20 cm) is clay loam with 3–4 % intimate humus (organic matter), depths between 20 and 66 cm is medium clay loam (roughly equal fractions of clay, sand and soil), whilst deeper soil down to 170 cm is sandy gravel with 70–80 % sand content though locally there is chalky diamictite (boulder clay). The soil composition partially controls water infiltration, percolation, soil moisture content and evaporation. The vegetation canopy also affects infiltration and the plant water uptake—itsself dependent on the moisture content within the rooting zone—controls transpiration. The exchange of water vapour between the soil and atmosphere is often a poorly constrained mechanism of LSMs and therefore is a weak link in the simulated hydrological cycle. Soil hydraulic properties can be derived from the observed soil composition and such soil properties can be used to initialise LSMs; this is discussed in more detail in Section 4. A small stream runs through the research site which was situated to the north of all the instrumentation. Water table depth data (from two locations on the site labelled as ‘south’ and ‘west’) shows that the hydraulic gradient, and hence the flow of subsurface water, is towards the stream despite the surface of the site being essentially flat.



155 3 Site set-up and data logging

3.1 Site set-up and data logging

The core hydrometeorology instrumentation for logging purposes was divided into six groups:

- 160 (i) 50 m ultrasonic anemometer, temperature, humidity.
- (ii) 25 m ultrasonic anemometer, temperature, humidity.
- (iii) 10 m ultrasonic anemometer, temperature, fast hygrometer.
- (iv) Broadband radiative fluxes
- (v) ‘Screen-level’ temperature, humidity, aerosol, visibility, pressure, rainfall and other miscellaneous; ultrasonic
 165 anemometer from 2011.
- (vi) Subsoil profiles of moisture and temperature.

Data were logged almost continuously—allowing of course for sensor failure, calibration and power outages—between May 2004 and the end of 2024 creating the ‘core dataset’. Although data logging started in May 2004, the number of variables was initially limited as instrumental spin-up occurred over a period of a few months. Data prior to this period was stored in an
 170 archaic format deemed too costly to recover. The so-called ‘screen level’ was set at a height of 1.2 m for pressure, temperature and humidity throughout the period. Although for logging purposes they were included in the “screen” data, the aerosol, visibility and present weather sensors (Section 3.2.6) were at 2 m, the raingauge was at the surface, and the sonic anemometer fitted in 2011 was also at 2 m. Therefore because of the large number of atmospheric variables at or below 2 m, the data in
 175 Fig. 2 for example has been split into ‘screen’ and ‘2 m’, where ‘screen’ refers to sensors at 1.2 m. The sonic anemometer sensor heights for masts nominally stated at heights of 25 m and 50 m were more accurately at 26.2 m and 51.2 m above ground level. The sonics were logged using in-house software on Linux-based MOXA UC740 embedded computers while the remainder was logged using commercially available DT85-series data loggers manufactured by dataTaker. The DT85 could monitor a wide variety of analogue inputs (voltages, currents, resistances) at varying rates. One minute averaged data was
 180 logged based on a raw sampling rate of 0.5 Hz.

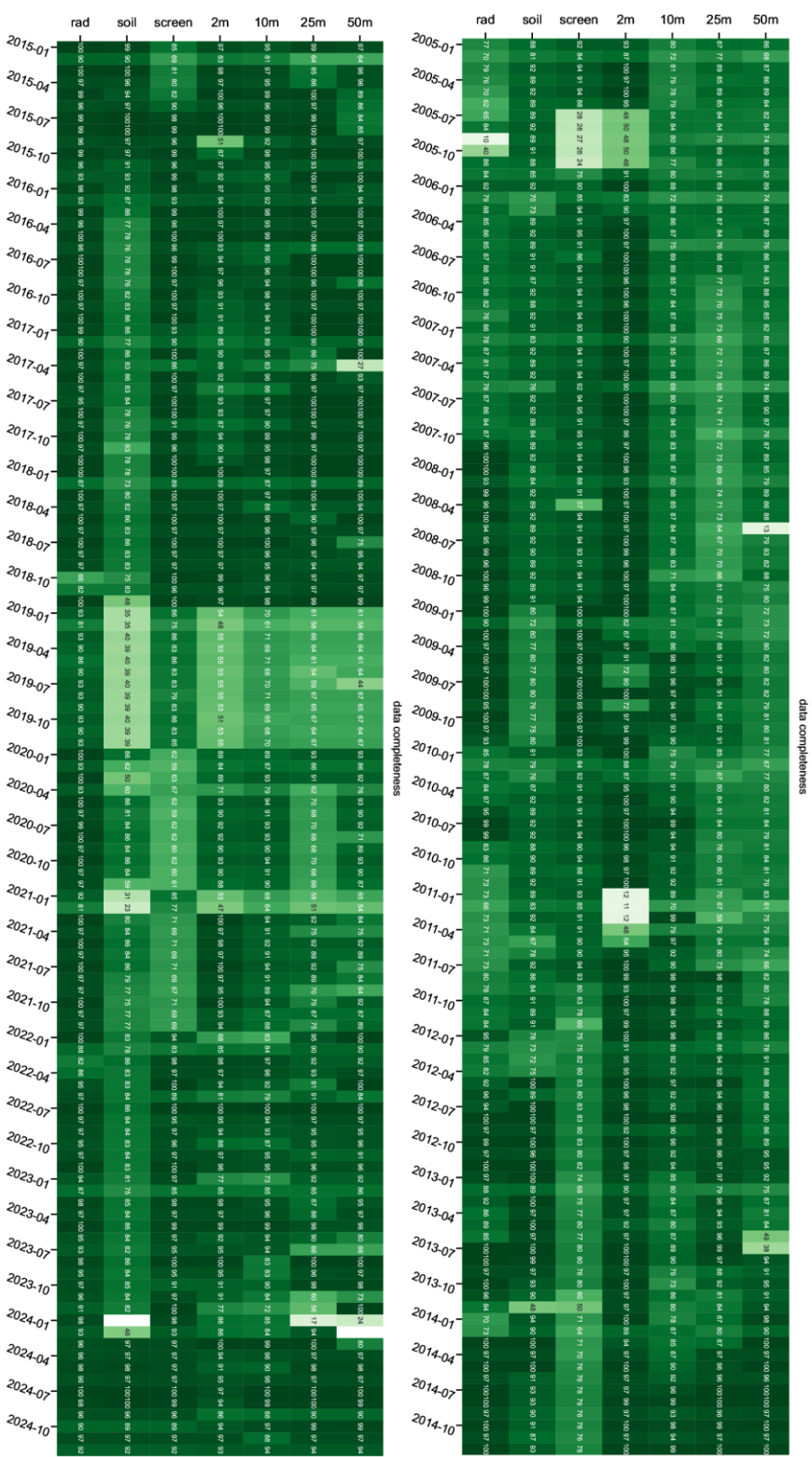


Figure 2: Data availability divided into two 10-year periods for the core surface site instrumentation as split into ‘50m’, ‘25m’, ‘10m’, ‘2m’, ‘screen’ ($\leq 1.2\text{m}$) levels, and also ‘soil’ (all buried subsoil sensors) and ‘rad’ (radiative fluxes) categories. Each bar contains the data availability as a percentage of data not flagged as bad or missing (see Table 6), graded on a green scale from white (0%) to dark green (100%).



All data were regularly transferred from the loggers to a central data storage computer. The central processing PC clock was routinely synchronised to an external Network Time Protocol (NTP) server, with the individual logger clocks in turn adjusted to the PC time. The data processing routines created four files per day i.e. for data averaged over periods of 1, 5, 10 and 30 minutes. These four timestep intervals have been preserved when creating the archived NetCDF files: therefore, apart from major data losses due to power cuts for example, there exist four NetCDF files per day for the core site variables throughout the 20-yr period. Data from certain slow response sensors, like the soil sensors for example, are not included in the 1 minute files; and we have only calculated variances and covariances from the mast data at 10 and 30 minutes. During relatively developed turbulence, 10 minutes is often sufficient to capture the majority of the turbulent energy and therefore derive reliable (co)variances (Raabe et al., 2002); otherwise in more benign conditions 30 minute intervals are recommended that will capture most large scale, low frequency contributions (El-Madany et al., 2013). Linear detrending of the sonic data over either the 10 or 30 minute interval was carried out before the variance calculations. A comprehensive list of all variables of the core dataset is shown in supplementary material S1.

In addition to the sensors contributing to the core dataset, we had various microwave radiometers, lidars, ceilometers and disdrometer running at various times. These more specialist instruments were logged separately on their own PCs. Sometimes these additional instruments would be used elsewhere on detachment, so in general their use was not intended to be as continuous as the core dataset instrumentation. They are described in more detail in Section 5. With the exception of a special subset of the core data for driving LSMs described in Section 4 there has been no gap-filling applied to either the core data or the additional data in Section 5. A system of quality flags has been used for the core data.

3.2 Core dataset instrumentation

This section will provide details of core dataset instrumentation listed in Table 2 (meteorology), Table 3 (aerosol and visibility), Table 4 (radiation) and Table 5 (subsoil) that require additional description to enhance their application. The descriptions below are not designed to be exhaustive: the metadata in the NetCDF archive files is often all that is needed. Table 6 lists the flagging system used in the archived files for the core dataset instrumentation. Figure 2 shows data availability, i.e. percentage of data not flagged as missing or bad, organised by month and by meteorological level—all variables in Fig. 2 are divided into 50 m, 25 m, 10 m, 2 m, screen, subsoil and radiation. This allows a basic grasp of how reliable the site was as function of time. We have split the data into two panels covering a decade each to improve legibility. To put Fig. 2 in some additional perspective, the data availability (i.e. not flagged as either bad ‘X’ or missing ‘m’) of the components of the surface energy balance—latent heat flux at 10 m, sensible heat flux at 10 m, ground heat flux, net shortwave and longwave radiation and surface (grass) temperature—are 93.8, 96.0, 95.7, 95.5 and 92.8 %, respectively, based on the full 20-yr dataset. The combined data availability of these components, i.e. the percentage of 30 minute time steps where all the components are not flagged as



220 bad or missing at the same time, is 81.9 %. When we in addition ignore data that is flagged as suspect ('?') then this combined value becomes 77.5 %.

Measurement	Manufacturer	Model	Variable	Derived properties	Height	Logging interval	Notes
Tri-axis sonic anemometer	Gill Solent	HS-50	3D winds; virtual temperature	mean wind; (co)variances of u, v, w & T	50, 25, 10, 2, 0.4 m	10Hz	Time period: 2004–2024 2m height from 2011; 0.4 m height from 2022. See Section 3.2.1.
High-speed infrared hygrometer	Licor	LI-7500	Specific humidity, carbon dioxide	Specific humidity, w'q' covariance, CO ₂ mixing ratio	10 m	10Hz	Time period: 2004–2024 Affected by precipitation and dew on optics; not an absolute instrument. See Section 3.2.3.
Platinum resistance thermometer	Vector Instruments	T302	Temperature	n/a	50, 25, 10, 1.2 m	60 s	0.1°C accuracy, screened, aspirated. Section 3.2.2
Platinum resistance thermometer	Rotronics	Hygroclip2, PT100 type	Temperature	n/a	0.4, 0.15, 0.08 m	60 s	Time period: 2016–2024 0.1°C accuracy, screened, aspirated 0.08m is the air at “grass tips”. Section 3.2.2
Relative humidity (RH)	Vaisala	HMP155 capacitive R2 Humicap	RH	n/a	50, 25, 10, 1.2 m	60 s	1% error for RH< 90 %; 1.7% error for RH 90–100 %, screened, aspirated. Section 3.2.2
RH	Rotronics	Hygroclip2, HT-1 capacitive	RH	n/a	0.4, 0.15, 0.08 m	60 s	1 % error, screened, aspirated, heated humicap sensor. Section 3.2.2
Barometric pressure	Setra	Model 270 transducer	pressure	n/a	1.5 m	60 s	0.1 hPa accuracy
Rainfall	Met Office	Mk V	Rainfall	n/a	Surface	60 s	0.2 mm resolution
Chilled mirror hygrometer	Michell		Dew and frost point temperature	RH	1.2m	60 s	0.1°C accuracy. Section 3.2.2

Table 2: Core meteorological instrumentation. Note the logging interval is not equal to the archived time step (=30 minutes). Supplementary section S1 shows an exhaustive list of all core variables with start and end times.



3.2.1 Sonic anemometers

225

230

235

240

245

250

255

The Gill tri-axis ultrasonic anemometers have an asymmetrical design that ensures minimal flow distortion except for a small angle centred at the mount point. The manufacturer quotes accuracy of the u , v and w components of better than 1% (root mean square) for wind speeds between 0 and 45 m s⁻¹. When the wind direction was coming from the mast and mount of the sonic, it was flagged as such in the data (see Table 6). Although the sonics are capable of a 50 Hz processing rate, 10 Hz is sufficient to capture all the energy within the inertial subrange as it starts to blend into the dissipation frequencies (Mauder et al., 2020b). The orientation of the anemometer is logged in the header of the data files. Directions were measured with a compass and the bearing in degrees magnetic is noted. The magnetic variation can also be entered, and this is applied when calculating the true wind direction for the data display. The magnetic variation is not logged, so subsequent data processing routines are needed to take it into account in order to obtain the most accurate wind directions.

The additional sensors that were co-located with the sonics were connected to the analogue inputs of the sonic anemometers. These additional sensors were the PRT temperature sensors (at all heights), the humidity Humicaps (at 2 m, 25 m and 50 m), and a Licor hygrometer specific humidity. The Licor was positioned at 10 m for the duration of the dataset, although an additional device was deployed at 2 m (mounted alongside the sonic anemometer) for seven months during 2024. It was hoped the latter would help elucidate the hard-to-measure moisture fluxes in stable conditions when the flux often hovers around zero, despite appreciable dew fall accumulations being observed at the surface (Osborne and Weedon, 2021). The 2 m moisture flux data remain unutilised. The PRTs and humicaps were calibrated in-house to an accuracy of 0.1°C and 2 %, respectively. The coincident logging allowed the sensor outputs to be incorporated into the same data stream as the sonic data. This ensures accurate time synchronisation when calculating, for example, humidity covariances. For latent heat flux calculations, 10-m covariance using the Licor hygrometer should therefore be used as standard. This eddy covariance technique is the most common method globally for measuring evapotranspiration (Pastorello et al., 2020). Although a PRT sensor was used as an absolute temperature measurement at each mast height, it is the fast-response sonic temperature that was used to calculate variances and covariances. This sonic temperature is related, but not identical, to virtual temperature and is derived from the measurement of the speed of sound (the principle of operation of the sonic anemometer). Sonic temperature is suitable for monitoring variations because of a very short time constant but it is unsuitable as a measure of absolute air temperature. Offsets in sonic temperature from virtual and true temperature can amount to 0.5°C and 2.4°C, respectively, because of effects of humidity (Kaimal and Gaynor, 1991). When calculating turbulent fluxes, however, the effect of humid air on the sonic temperature variance falls to the order of 0.01°C or less than 2 % of the flux, well within experimental error (Horlacher et al., 2012). All the appropriate sensors were fast-response and small enough to cause no flow disturbances, meeting the requirements for the eddy-covariance technique (Lorrai et al., 2010). Finally, to re-iterate what was stated in the introduction: no attempt was made to correct the turbulent fluxes for potentially missing energy which would lead to a modest underestimation in the fluxes. This is left to the user to consider.



Measurement	Manufacturer	Model	Variable	Derived properties	Height	Logging interval	Note
Disdrometer	Thies	Laser precipitation probe	droplet size and movement	Rain/drizzle rate, droplet size distribution, fall speed	2 m	60 s	Time period: 2019–2024 (intermittently). 0.786 μm wavelength; 160 μm minimum diameter; 0.001 mm h^{-1} sensitivity; includes snow/hail diagnostics; 15% error in rain, 30% in snow. Section 3.2.5
Aerosol scattering	MRI	1550B	475 nm total scattering coeff		3 m	60 s	Time period: 2004–2010 heated to 30C (RH<60%)
Aerosol scattering	Optec	NGN-3a	550 nm total scattering coeff		4 m	60s	Time period: 2011–2020 Heated to 38°C, dehydrated (RH<40%) aerosols < 2.5 μm using a spiral impactor
Visible range	Belfort	6230A	visibility		2 m	60 s	Time period: 2004–2024. 2 Hz native frequency. Active forward scatter from xenon lamp
Present weather sensor	Biral	HSS VPF-730	visibility	Hydrometeor weather code	2 m	60 s	Time periods: 2011–2014 and 2017–2021. 0.88 μm active sensing
Present weather sensor	Campbell	CS125	visibility	Hydrometeor weather code	2 m	60 s	Time period: 2021–2024. 5 m to 100 km range; 0.05 mm h^{-1} sensitivity to rain rate, includes snow diagnostics; 0.05 mm h^{-1} precip sensitivity; 8% accuracy for vis<600m; 10% accuracy for vis<10km

Table 3: Aerosol and visibility instruments (logged as part of the core data). Note the logging interval is not equal to the archived time step (=30 minutes).



3.2.2 Relative humidity

265

270

275

280

285

290

295

The screened and aspirated HMP155s relative humidity (RH) sensors were mounted as standard at 50 m, 25 m, 10 m and 1.2 m. For the near surface sensors at 40 cm, 15 cm and 8 cm, the Hygroclip2 sensor (for both temperature and RH) was used. The 8 cm sensor effectively records the canopy temperature and RH, i.e. the air that is in contact with the grass blade tips. The 8 cm data collection began in 2014 whilst the 40 cm and 15 cm sensors were not deployed until 2021. The reason for deploying the Hygroclip2 sensors below the traditional screen height of 1.2 m was to investigate the thermodynamic conditions that lead to dew fall, fog (Price, 2019) and the lifted temperature minimum (Weedon et al., 2024). Two HMP155s were co-located at 25 m and 50 m from March 2021 in case of sensor failure and data from these sensors is included in the archive (labelled as sensors A and B in supplementary S1).

When the HMP155 and Hygroclip2 RH sensors were calibrated in the laboratory, the upper RH limit of the calibration chamber was commonly 95%. Therefore, determining RH values with accuracy as saturation approaches is problematic as extrapolation is assumed which makes assumptions about both the sensor under calibration and the calibration machine. As with many aspirated sensors, in particular temperature, sensor wetting via condensation or otherwise leads to ‘wet-bulbing’ and hence miss-reading due to evaporative cooling of the sensing element.

Until 2013 laboratory calibrations of all temperature and RH sensors were carried out in an in-house designed and built environmental chamber. This device had two-chambers with an inner test volume (23.4L) of circulating air. This large volume allowed several sensors to be tested at the same time. A Michell S3020 chilled mirror hygrometer was used as the reference humidity within the test chamber. From August 2013 onwards a commercially available Rotronic HygroGen2-HG2-S was used (sensing volume of 1.6L), and from 2018 a larger HygroGen2-HG2-XL version was used. The Rotronic devices could be programmed for easier and quicker calibrations. Calibrations of RH sensors typically range from 10 to 95 % in 10 % stages for two temperatures i.e. 5°C and 25°C.

In brief, as the error in the measurement of RH should be treated as 2 % at RHs above 95 %, readings in general above 100 % can be accounted for via this sensor error. RH above 100 % is flagged as ‘query’ (Table 6). Do not assume that 100 % is the point of saturation. If a flat, stable RH is achieved in the data at or around 100% then saturated air can be assumed and therefore the time of saturation can be estimated (e.g. pertaining to fog formation studies). There is always danger in reliance on one sensor; therefore, humidity studies using the Cardington dataset should scrutinise all available sensors, i.e. between canopy and 50 m, and compare the evolution of RH as a function of both height and time.



The Michell chilled mirror dew point hygrometer was deployed between 2004 and 2014. Although this is a slower response instrument compared to many humicap-type devices, it can be considered as the reference dew point. Because deriving the specific or RH from this dew point requires temperature and pressure, respectively, the error will increase accordingly.

300 3.2.3 Licor high speed hygrometer

The Licor-7500 hygrometer is an active near-infrared open-path instrument that detects changes in carbon dioxide and water vapour. Data was logged at 10 Hz, although the unit can log up to 20 Hz. The fundamental gas sampling rate is 150 Hz. The Li-7500 is not an absolute device and should be used primarily for variances of CO₂ and specific humidity and therefore
 305 covariances when collocated with a sonic anemometer, such as $w'q'$ from which the latent heat flux can be estimated. Yearly zeroing of the Li-7500 is achieved in the laboratory by using soda lime and magnesium perchlorate to scrub a controlled flow of CO₂ and water vapour, respectively. The Li-7500 contains similar internal scrubbers to void the sample optics of detection gases that were changed every year. The Li-7500 was mounted at a 15° angle so that rain or fog deposition water on the sapphire optics readily flows off; nonetheless, data during periods of precipitation or mist/fog should be treated as suspect. As
 310 with the sonic anemometer data, although the raw 10 Hz data for the Li-7500 are stored at the Met Office, the CEDA archive only contains the processed (co)variances (at 10 or 30 minute time intervals) and the nominal mean specific humidity and CO₂ mixing ratio at all time intervals. Accuracy of the specific humidity of the Licor, based on the calibrations, is estimated as 0.2–0.3 g kg⁻¹, although this error could drift to larger unknown values between calibrations. The Licor CO₂ data in general remains to be exploited.

315

3.2.4 Aerosol measurements

Visible total scattering coefficients were measured with integrating nephelometers (two types depending on the date as shown in Table 3), both using heated sample air in an attempt to reduce the RH to below 40% to minimise deliquescent/hysteresis
 320 effect on aerosol particle growth. Periodic calibration of the nephelometers was carried out with clean air (low span gas) and CO₂ (high span gas). The later Optec instrument included temperature and RH sensors in the heater-controlled scattering chamber. The earlier MRI instrument did not do this, but the sample air was heated just prior to entry to the unit. Measuring a dry aerosol sample allows subsequent theoretical estimation of the aerosol growth factor as a function of RH, e.g. to the 85% RH standard. Assumptions must be made about the aerosol chemistry in order to do this. This is preferred to measuring a
 325 highly variable, ambient sample RH. This standardised aerosol growth estimation, however, is not provided in the processed files; it is left to the user to obtain this.

A Cimel CE318 sun-photometer was installed at the end of 2020 as part of the Centre for Connected and Autonomous Vehicles (CCAV) project (Jones, 2022). The aerosol retrieval data from the CE318, which also includes cloud optical depth retrievals



330 from diffuse zenith views in overcast conditions, does not form part of the CEDA archive but is nonetheless freely available
 from the Aerosol Robotic NETwork (AERONET) website. The instrument has been calibrated at the University of Lille after
 the decommissioning of the Cardington site (January 2025) and the AERONET team are currently processing the data with a
 plan to publish it on their website.

335 3.2.5 Precipitation

The only measure of distinguishing sleet/snow/hail from liquid precipitation was using the Belfort and Campbell present
 weather sensors using their sophisticated scattering detectors (Table 3). These report four types of precipitation (rain, freezing
 rain, ice pellets, snow), fog, mist, haze (smoke) or dry. A combination of the traditional tipping-bucket raingauge and the
 340 present weather sensor can help with analysis; as can using the much more recent Thies laser disdrometer (also included in
 Table 3) that can in addition detect fine drizzle, drop fall speed and drop size distribution. The disdrometer is the most
 sophisticated precipitation instrument deployed at Cardington and comes recommended for all research stations. Note that its
 deployment time in the field was intermittent from 2019 until the end of 2024 and so there will be limited research use from
 the dataset archive. The disdrometer data is nonetheless archived as separate netCDF files, with the variables listed in
 345 supplementary section S7.

Other than falling snow detection mentioned above, no device measured snowfall depth lying on the ground. This is because
 snowfall at Cardington was relatively unusual, with lying snow being particularly scarce.

350 3.2.6 Present weather sensors

The Biral HSS VPF-730 instrument is used for both the measurement of the visual range through air, and for the determination
 of present weather (Table 3). This is given in terms of both precipitation type and rate. This instrument consists of an optical
 transmitter and two receivers. The light source is a flashlamp in the infrared band with a central wavelength of 0.85 μm . One
 355 of the receivers measures the forward scatter (45°) of the light caused by atmospheric particles, which gives the atmospheric
 extinction coefficient. From this the horizontal visual range (visibility) is calculated between 10 m and 75 km. The second
 receiver measures the backscatter off precipitation (or indeed fog and aerosol) particles. The amplitude and duration of the
 light pulses created by each precipitation particle as they pass through the sample volume are measured, and from this, the
 particle size and velocities are determined. An algorithm is used to determine the precipitation rate and type. A further method
 360 is also used for deducing precipitation type, by measuring the ratio of the backscatter extinction coefficient to the forward
 scatter, with a ratio above a certain value indicating ice particles. 15 present weather codes (WMO Table 4680: see
https://artefacts.ceda.ac.uk/badc_datadocs/surface/code.html for all meteorological codes) were generated by the device (a



subset of the disdrometer codes as outlined in supplementary S7) that cover haze, and various rates of drizzle, rain, snow and hail.

365

Calibration of the instrument was carried out annually. The calibration procedure involves attaching a scatter plate to the instrument to simulate a known scattering coefficient. A zero calibration is also performed by completely obscuring the receiver heads. Routine maintenance involved the inspection and cleaning of the receiver and transmitter windows. This was typically carried out on a weekly basis. Dirt on the windows, and cobwebs inside the window hoods, can degrade the performance of the instrument and cause spurious data. The sensor was orientated to avoid exposure of the receiver heads to light from the setting/rising sun.

370

The Campbell CS125 also operates at a wavelength of 0.85 μm and provides similar derived variables to the Biral, i.e. a visual range (from 75 km down to 5 m)— that is calculated using Koschmeider’s Law from an extinction coefficient— and also a present weather code. It does not use a backscatter detector but instead derives everything from the 42° forward scatter signal. The CS125 uses fall speed, particle size and air temperature to identify the type of particle and 56 SYNOP codes are available from WMO Table 4680.

375

Variable	Manufacturer	Model	Height	Logging interval	Note
Shortwave irradiance	Kipp&Zonen	CM ² 2 pyranometer	4 m	60 s	Downwelling hemispherical, downwelling diffuse, upwelling
Shortwave irradiance	Kipp&Zonen	CM ² 1 pyranometer	4 m	60 s	downwelling hemispherical diffuse; upwelling (reflected) hemispherical
Longwave irradiance	Kipp&Zonen	CG4 pyrgeometer	2 m upwelling 4 m dnwelling	60 s	4.5–42 μm ; downwelling and upwelling hemispherical
Surface radiometric temperature	Heitronics	KT15D pyrometer	2m	60 s	1 m ² of grass scene; 8–14 μm window region; concrete scene between 2004–2005 only

Table 4: Core radiation instruments (logged within core data) between 2004 and 2024 (except as noted). Note the logging interval is not equal to the archived time step (=30 minutes). Supplementary section S1 shows an exhaustive list of all core variables (which the radiation instruments form a part of) with start and end times.

380



3.2.7 Radiation

385

390

395

Conditioned measurements of the radiative energy flux density ($\text{J s}^{-1} \text{m}^{-2}$), or irradiance, through a horizontal plane over a grass surface were made over the full 2004–2024 period for the shortwave (solar) and longwave (thermal) spectral bands. Conditioning means the units were aspirated and heated to minimise rain, dew, frost and fog water on the glass domes. Table 4 summarises the various units deployed. Such hemispherical irradiances are often called total or global irradiances, with diffuse downwelling shortwave in addition being made using an automated Kipp&Zonen Solys2 solar tracker that blocks out any contribution from direct from the solar disc. Thus the direct solar beam contribution, i.e. that arc subtended by a cone having a linear angle of 2.5° (a solid angle of $2.0 \times 10^{-3} \pi$ steradian), can be calculated to an accuracy of 0.02° with the solar tracker. This angle is larger than the solar disc itself (0.5°) to allow for circumsolar diffuse irradiance as defined by the WMO. The pyrgeometer measuring downwelling longwave irradiance was also mounted on the solar tracker (it could accommodate up to three instruments) to minimise the effect of window heating from direct sun (although the effect of this heating was ordinarily within 4 W m^{-2}).

400

All glass domes were cleaned on a weekly basis. Calibration, and therefore instrumental offset, of the pyranometers was determined periodically before and after major campaigns using an outdoors comparison in clear sky conditions with a secondary standard instrument to ISO-9847 standards. This secondary standard was ordinarily stored on site but in turn was calibrated to a primary standard again using outdoor real-world data against the World Radiometric Reference (WRR) at the World Radiation Centre (WRC) in Davos, Switzerland and issued with a calibration certificate. No such calibration exists for the pyrgeometers, although they required a sensitivity test (1 bit per W m^{-2}) and an internal desiccant check every year.

405

410

The long-term standard way of measuring grass canopy, or skin, temperature at Cardington was radiometrically with the Heitronics KT15 pyrometer. The KT15 was housed in a waterproof shield and mounted on a mast at a height of 2.5 m above the ground. It is tilted at an angle of approximately 20° to the vertical and the surface below is short grass, which is deemed representative of the site. The detector is a standard 26 mm pyroelectric type A and a germanium M6 close-focus lens is used as the front-end optics on the unit. This setup this gives an effective target area on the ground having a diameter of about 1 m. For practicality, the surface emissivity was set to 1.0 across all data collected. An adjustment can be made for the reflected sky component by making assumptions about the grass emissivity, which can be set to 0.965 in typical conditions at Cardington according to Edwards et al. (2011). See Weedon et al (2024) for more details on correcting the KT15 data.

415



Variable	Manufacturer	Model	Depth	Logging interval	Note
Temperature	Delta-T	PRT	1, 4, 7, 10, 17, 35, 65, 100 cm	60 s	Time period: 2004–March 2012
Temperature	Delta-T	ST2-396 thermistor	1, 4, 7, 10, 17, 35, 65, 100 cm	60 s	Time period: From March 2012–2024
Soil water content	Delta-T	ThetaProbe ML2/ML3	2.5, 10, 22, 57, 160 cm	60 s	Time period 2004–2024 except for ‘2.5 cm’ sensor (from January 2020) positioned vertically into the soil and is a nominal depth
Soil water content	Delta-T	PR2	10, 20, 30, 40, 60, 100cm	60 s	Time period: from 2016–2024. Column probe with six sensing depths; South site only
Ground heat flux	Hukseflux	HFP01SC flux plate	2 cm	60 s	Time period: 2012–2024; self-calibration every 13 h
Water table depth	Druck	1830 pressure transducer		60 s	Time period 2004–2024. Pressure transducer at two locations labelled as ‘south’ and ‘west’

Table 5: Subsoil sensors (logged as part of the core data). Note the logging interval is not equal to the archived time step

(=30 minutes). Supplementary section S1 shows an exhaustive list of all core variables with start and end times, of which the soil sensors form a part of.

3.2.8 Subsoil sensors

Two soil pits, called the West and South pits, were originally dug and fitted out with an identical suite of sensors in the late 1990s. Table 5 summarises the subsoil sensors. Soil temperature was recorded at depths of 1, 4, 7, 10, 17, 35, 65 and 100 cm using thermistor-based sensors. Volumetric soil moisture content was recorded at depths of 10, 22, 57 and 160 cm using ThetaProbes that utilise the change of refractive index with soil water. Logging continued essentially unchanged until 2022 when the West pit was decommissioned. New sensors were installed at the South pit from early in 2023, albeit limited to a depth of 1.0 m compared to 1.6 m within the bulk of the dataset. Water table depth was observed continuously using a pressure transducer buried at a depth of 2 m at both soil pits. All soil sensors were located under manicured grass.



The Delta-T ThetaProbe was the most accurate ($\approx 1\%$) method of measuring soil moisture content at Cardington. The PR2 stainless steel column with 5 detectors position along it was acquired for testing alongside the ThetaProbes and was intended for detachment use away from Cardington because of its ease of installation. Although several of these PR2 columns were used in campaigns, the one at Cardington remained in position from 2016 onwards and provides another measure of soil water. There is evidence that the signal drifts with time (seasonal to yearly timescales), perhaps due to varying degrees of contact with the surrounding soil, so the data should be used with caution as an absolute device. An additional ThetaProbe at the South pit for the final few years (from 2021) was installed vertically from the surface and so represents the top few cm of soil (labelled as 2 cm depth in the archived files). This surface sensor is responsive to light to moderate accumulations of rain that do not penetrate to the 10 cm depth.

flag_meanings (<i>nc_char</i>)	flag_values (<i>nc_byte</i>)	data value
not_used	0b	n/a
good_data	1b	instrument output
suspect_data	2b	instrument output
suspect_data_calibration	3b	instrument output
suspect_data_object_upwind	5b	instrument output
suspect_data_sonic_anemometer _orientation	6b	instrument output
missing_data	7b	fill value (1×10^{11})
bad_data_do_not_use	8b	fill value (1×10^{11})

Table 6: Data flags used in the core hydrometeorology NetCDF files.

From mid-2017 until the end of 2024 there was a Cosmic-ray Soil Moisture Observing System (COSMOS) installed at Cardington that was part of a nationwide network of soil moisture monitoring sites operated by the UK Centre of Ecology and Hydrology. COSMOS harnesses naturally-produced neutrons from cosmic-ray interaction with the atmosphere to sense soil water content over a ‘field scale’ area with radius up to 100–200 m. The COSMOS soil water data does not form part of the archived dataset but is available upon request from <https://cosmos.ceh.ac.uk/data/data-request> (see Cooper et al., 2021 for more details on the technique).

Two Hukseflux ground heat flux plates were deployed at a nominal 2 cm depth from 2012. This allows an alternative method of determining soil heat flux to the change in the soil temperature profile with time based on one-dimensional heat conduction. After the West pit was decommissioned, two flux plates were installed side-by-side at the same depth at the South pit. An active heating self-calibration of the heat plates lasting typically 20 minutes was carried out every 13 h. During this calibration, no data are available. The self-calibrations are designed to account for changes in the soil conductivity, mainly because of



changes in the soil water content. The plates were mounted horizontally, but like all subsoil sensors are prone to movement by soil heave, water flow etc. The fitting coefficients were chosen that most closely align with the soil type at Cardington (as detailed in Burton, 1999). Although the grass canopy was meant to be kept to a nominal 5 cm height, there will be unreported times when this was not strictly maintained— this will suppress the diurnal range in ground heat flux compared to a short canopy. Canopy heat storage is not something that can be observed directly. Correction of the heat flux plate data for heat storage in the overlying top 2 cm of soil is possible based on the change in the co-located vertical temperature gradient (from the 1 cm temperature sensor downwards) with time. This correction is not applied to the processed files but should nonetheless be considered by the user. All these aspects of measuring ground heat flux imply the technique carries significant error. The timing of the change of sign of the ground heat flux during diurnal cycles, as it is for the surface turbulent heat fluxes too, is an important as part of energy balance studies.

3.2.9 $j(\text{NO}_2)$ radiometer

The photodissociation of absorbing trace gas molecules into reactive species, such as the dissociation of NO_2 into NO and $\text{O}(^3P)$, is a crucial part of atmospheric chemistry cycles. The reaction of $\text{O}(^3P)$ and molecular oxygen to form ozone is the next stage and therefore the NO_2 photolysis frequency, designated as $j(\text{NO}_2)$, controls the primary production of the tropospheric ozone pollutant. An ultraviolet/visible spectroradiometer manufactured by Meteorologie Consult GmbH was deployed at Cardington for a limited period to retrieve the atmospheric photolysis frequency of NO_2 molecules. The data has been used to validate the prediction of $j(\text{NO}_2)$ using the online NAME (Jones et al., 2007) and offline AQUM (Savage et al., 2013) air quality schemes developed by the Met Office. The solar actinic flux is the radiation available for initiation of molecular photodissociation. The measurement of the 2π steradian radiative flux as a function of wavelength allows the calculation of photolysis rate when combined with molecular parameters such as the molecular absorption cross section for NO_2 . The instrument consisted of a hemispherical flux entrance optic, a single monochromator, a 512-pixel diode array detection system. The diode array measured wavelengths from 285 to 450 nm in consecutive 0.5-, 1-, 3- and 5-s integration times with a spectral band pass of 2.2 nm (Shetter et al., 2003). The photolysis rate $j(\text{NO}_2)$ is included in the core dataset for the time of deployment i.e. May 2015 until January 2021.

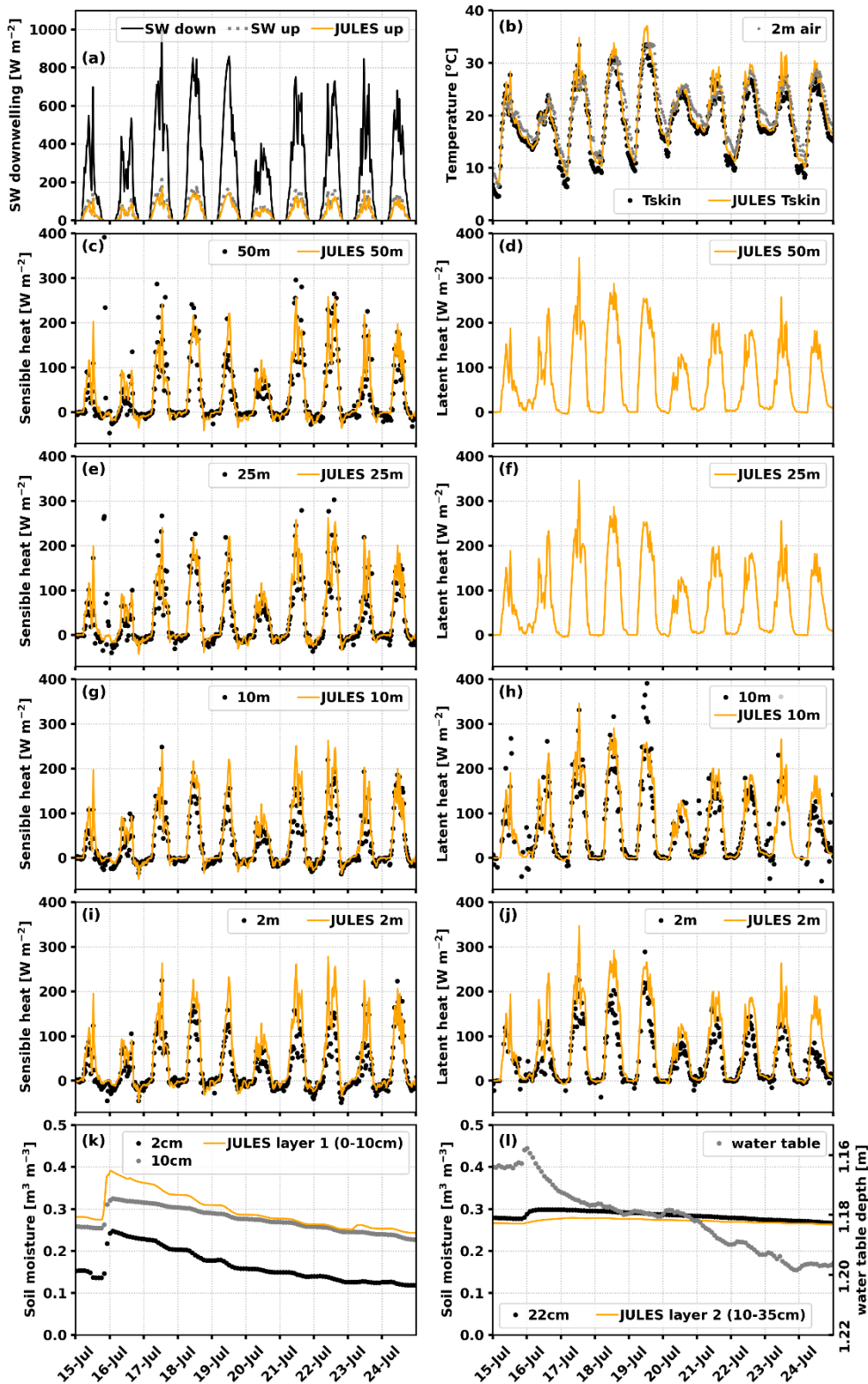


Figure 3: Comparison of JULES model output (yellow) and Cardington observations (black and grey) over ten days in July 2024 where JULES has been forced with site data at 30 minute time steps. (a) Observed SW downwelling top-left shows periods of cloudy and predominantly cloud-free conditions. (b) Skin (grass) and air temperatures, (c, e, g, i) sensible and (d, f, h, j) latent heat fluxes at the available heights, and soil moisture content at (k) level 1 and (l) level 2 are shown. The observed water depth is also included in (l) but this does not have a simulated equivalent.



4 Land surface model forcing data

JULES (Joint UK Land Environment Simulator) is a community LSM that is used in the Met Office Unified Model (MetUM) from short-range weather forecasts through to climate predictions (Best et al., 2011; Blyth et al., 2010). This could be clearer
 495 e.g., JULES can be run offline for a gridded domain using meteorological forcing datasets (e.g., WFDEI, Weedon et al., 2014) or for a single point location using observed meteorological forcing data. JULES requires the following seven atmospheric input variables at every time step for it to be able to run using prescribed meteorology from field observations: downwelling shortwave irradiance, downwelling longwave irradiance, rainfall, air temperature, mean horizontal wind, surface barometric pressure, and specific humidity. We have compiled a separate Cardington forcing meteorological dataset with a 30 minute
 500 time step to drive JULES standalone, so it that covers the same period as the core archived files. The JULES drive dataset (see also Supplementary section S2) comprises a NetCDF file for each of the four drive heights (2, 10, 25 and 50 m), such that temperature, wind and humidity drive variables are taken from the different mast heights, and the pressure, radiation and rainfall remaining unchanged as they were only available from fixed levels (i.e. pressure at 1.2 m, downwelling radiation at 4 m, upwelling radiation at 2 m, and rainfall at the surface). Due to the instrumentation deployed, the 2 m level drive data is only
 505 available for the whole years 2012–2024. Although the NetCDF forcing dataset has been configured to run with JULES, it should be straightforward to apply the data within other LSMs that can be run offline and forced by prescribed meteorology for a single point (Grimmond et al., 2011; Yang et al., 2011).

The forcing dataset specifies the downwelling radiation at every time step. The partitioning of the remaining energy into
 510 reflected radiation (partly dependent on skin temperature), turbulent fluxes (partly dependent on evapotranspiration), ground heat and canopy storage components from the LSM diagnostics can be tested by comparison to the full core dataset. It is also possible to prescribe the surface albedo within JULES for every time step using observations, such that subsequent analysis of the energy partitioning becomes more constrained. Yet since vegetation photosynthesises and transpires during the daytime, the latent heat flux is controlled by plant physiology as well as bare soil evaporation. The JULES forcing dataset is gap-filled
 515 where data are either missing (Flag 7b, Table 6) or deemed unreliable (Flag 8b, Table 6) to ensure that every time step is populated. Short gaps (≤ 3 h) were filled via linear interpolation; longer gaps were filled with the long-term (20-yr) mean values calculated from available measurements at each time step. The latter method of gap-filling ensures the preservation of daily and annual cycles. Each driving data variable has a simple flag to indicate whether gap filling has been applied, or not, at each time step. The driving dataset could potentially be used to apply an optional spin-up to JULES, for example by repeatedly
 520 driving the LSM with the first two years of data so that the soil temperature and soil moisture reach stability.

Two approaches to configuring JULES for offline runs can be adopted: either use the soil and vegetation parameters as they are prescribed operationally in the MetUM (be that either in a regional configuration such the UKV or a global configuration) or tune these parameters where practicable to local site properties. For example, soil composition and thereby derived



525 hydrology properties, and canopy information are usually available at research sites such as Cardington. Table 7 shows a range
 of soil and canopy parameters as derived from the observations at the Cardington site. The soil properties are deemed
 appropriate for the top 1 m of soil. These parameters are commonly used in LSMs to configure the subsoil and plant parameters
 to initialise the simulations. Alternatively, estimated soil properties can be taken from auxiliary global datasets (e.g. FAO
 & IIASA 2023) when running LSMs as part of NWP in a coupled model. Key assumptions are often made in LSMs, such as
 530 assuming the soil properties are constant with both depth and time because of a lack of real-world characterisation, apart from
 allowing some properties such as the thermal conductivity to vary with soil water content as a function of time. More guidance
 on how Cardington site data can be used to initialise and force JULES is found in Osborne and Weedon (2021).

Soil property	derived value	units
Soil dry heat capacity	1.235×10^6	$\text{J m}^{-3} \text{K}^{-1}$
Soil dry thermal conductivity	0.234	$\text{W m}^{-1} \text{K}^{-1}$
Soil hydraulic conductivity at saturation	0.00312 ± 0.0255	$\text{Kg m}^{-2} \text{s}^{-1}$
Soil matric suction at saturation	0.26714 ± 0.0255	m
Soil moisture at saturation	0.4454 ± 0.0556	$\text{m}^3 \text{m}^{-3}$
Soil moisture at critical point	0.3801	$\text{m}^3 \text{m}^{-3}$
Soil moisture at wilting point	0.1942	$\text{m}^3 \text{m}^{-3}$
Canopy height	0.05–0.10	m
Leaf area index	2.36–1.62	unitless
Rooting depth	0.2	m

Table 7: Soil and C3 grass canopy parameters as derived from local site soil properties at Cardington. Soil values are
 535 appropriate for the top 1 m of soil. Leaf area index shows the range from typical healthy grass through to senescence in drought
 conditions. Rooting depth is an *e*-folding depth derived from Osborne and Weedon (2021).

Figure 3 shows an example of observations and JULES output diagnostics over ten days in the summer of 2024 when the 2-m
 latent flux is available in addition to that at the standard 10m height. The configuration of JULES here is the MetUM-JULES
 540 Regional Atmosphere and Land configuration as described in Bush et al (2025). Figure 3 is illustrative and is not a scientific
 scrutiny of the JULES surface scheme. A brief description of the figure, highlighting a ten-day period in July 2024, is provided.
 The period is dry apart from 18 mm of rainfall over a period of 12 h between 15–16 July. The effects of the rain can be seen
 in the increases in soil water content observed at 2, 10 and 22 cm and likewise in the simulated soil water. Although soil water
 observed (and modelled) at 57 cm depth (not shown) did not register any response to the rain, it is interesting that the water
 545 table shows a small rise (presumably responding to rain in the local area and therefore demonstrating lateral soil water flow)
 before decreasing over the remainder of the period. There is an increase in the daytime latent heat flux as observed at 2 and 10



m in the largely cloud-free days immediately after the rain as evapotranspiration strengthens. Although some of the highest observed latent heat values are not captured by the model (19 July at 10 m), some other flux data are well matched (17 July for both latent and sensible heat at 10 m). That being said, the JULES latent heat flux is in general well simulated at 10 m, although it is too large around midday at the 2 m height. This suggests that the near-surface gradients in the heat flux vertical profiles are not large enough in the model. So, although the peak values of the JULES 25 m and 50 m sensible heat fluxes are suppressed, the sensible heat in JULES are overall close to the observations at 10 m. This is understandable if we look at the simulated skin temperatures that tend to be too warm in the middle of day (and too warm at night). Details aside, the change in Bowen ratio from the days immediately after rain (17–18 July) that have a Bowen ratio < 1 , to the last day shown (24 July) when the ratio is > 1 , is captured by the model.

Type	Manufacturer	Model	Serial No.	Derived properties	Time interval	Note
Doppler lidar	Halo Photonics	Streamline II lidar	01	Wind profiles, backscatter, radial turbulence	8-9 s (vertical stare)	Time period: 2009–2021 (non-continuous). 1.55 μm laser; DBS wind scans
			30	Wind profiles, backscatter, radial turbulence, depolarisation	1-2 s (vertical stare)	Time period: 2011–2022 intermittent. 1.55 μm laser; DBS wind scans
			35	Wind profiles, backscatter, radial turbulence, depolarisation	1-2 s (vertical stare)	Time period: 2012–2024 intermittent. 1.55 μm laser; DBS and VAD wind scans
Ceilometer	Vaisala	CT25k lidar	n/a	Cloud base height; backscatter	30 s	Time period: 2013–2024. Two devices labelled ‘A’ and ‘B’. Up to 3 cloud bases; 15 m – 7.5 km; 15m resolution; 0.905 μm laser
	Impulsphysik	LD40 lidar			60 s	Time period: 2010–2014. Up to 3 cloud bases; 7.5 m - 13 km; 7.5 m resolution; 0.855 μm laser
		LD25 lidar			60 s	Time periods: 2004–2010 and 2012–2014. Up to 3 cloud bases; 15 m – 7.5 km; 15 m resolution; 0.855 μm laser



Passive microwave radiometer	Radiometrics	WVR-1100		LWP & IWV time series, basic surface meteorology	10 s	Time period. 2004–2021; principal frequencies 23.8 and 31.4 GHz
		TP/WVP-3000	3008		30 s	Time period: 2012–2016 sporadically.
	RPG	Humpro	002-0003	LWP & IWV time series, basic surface meteorology	1 s	Time period: 2016–2024 intermittent.
				humidity profiles	60 s	

Table 8: Non-core remote sensing instruments (logged individually). The time interval represents both the logging rate and the archived time step.

560 **5 Non-core remote sensing instruments**

Figure 4 shows the data availability of the Halo Doppler lidars, ceilometers and microwave radiometers in a similar manner to the core data in Fig. 2. Table 8 lists the various large devices by manufacturer that were operational at the site. Figure 4 shows fairly comprehensive data coverage for the ceilometers and the microwave radiometers (thanks to the reliability in particular of the WVR-1100). Doppler lidar data coverage was also substantial once they were installed in 2007. These more specialist instruments were not logged and processed centrally like the core dataset. Therefore, the core flagging method was not used for the non-core data in this section, and neither was it used for the radiosondes described later in Section 6. The various ceilometers were standard unmodified lidars ordinarily used on the Met Office operational network to detect cloud base height, but they also provide attenuated backscatter coefficient profiles from aerosol, precipitation and thin cloud. The other non-core dataset instruments are described in more detail below.

575

580

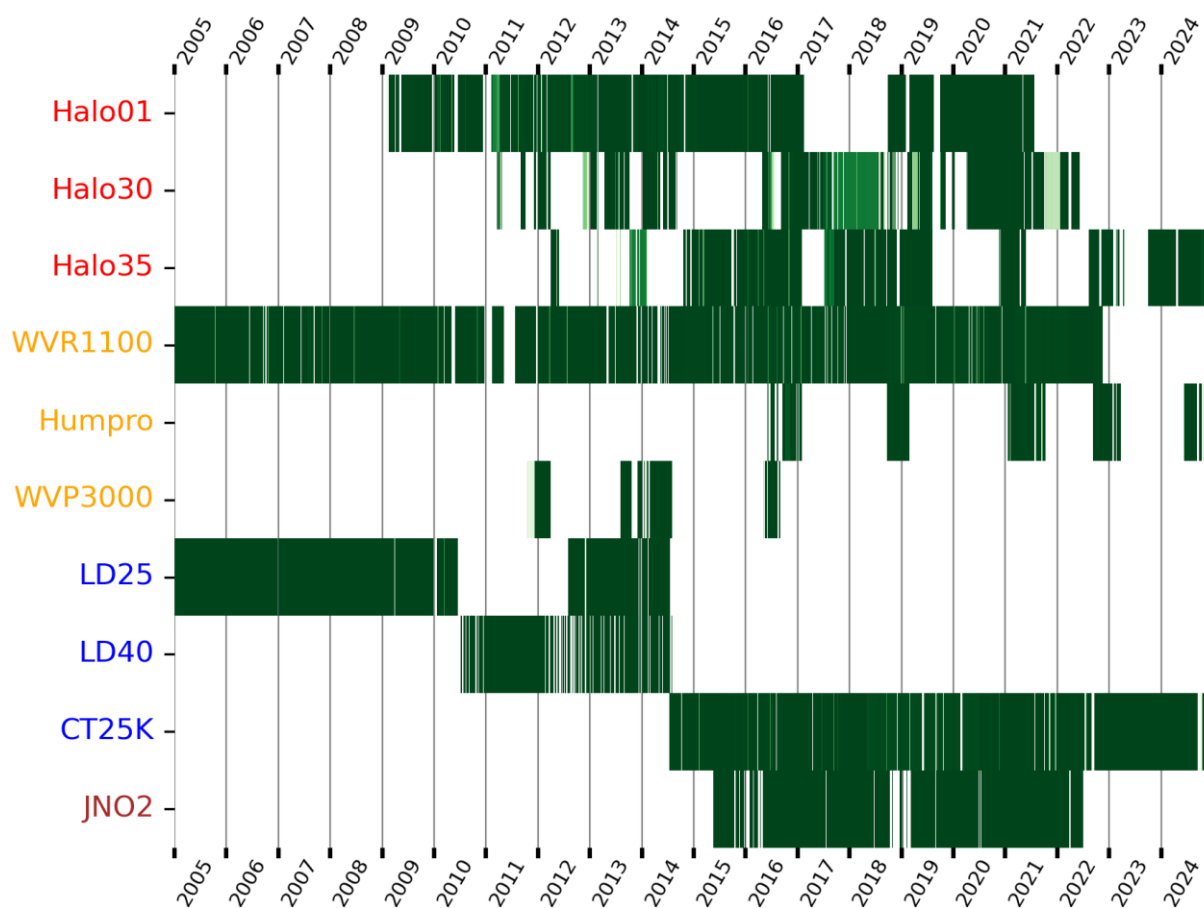


Figure 4: Data availability for individual, non-core instrumentation, graded on a green scale from white (0%) to dark green (100%). The three Halo Doppler lidars are in red, the three ceilometers are in blue, and the three microwave radiometers are in orange. Data from the $j(\text{NO}_2)$ instrument in brown is contained within the core dataset files, not separately like the other instruments included here.

5.1 Halo Doppler lidars

Three Halo Photonics Streamline doppler lidars (Pearson et al., 2009) have been deployed at various times at the site. Table 9 summarises the Halo specifications for the three models operated at Cardington. Daily netCDF files have been archived for each unit; supplementary section S6 lists all the variables. All three are based on a 1565 nm laser emitting linearly polarized pulsed light through an 8 cm diameter lens with a heterodyne detector. Laser beam returns from the atmosphere are range-



gated velocity and back scattered power. The laser beam divergence from the lens was 33 μ rad. Most of the beam returns are
 595 a result of aerosol particles acting as targets where the scattered light intensity and frequency shift are used to determine the
 attenuated backscattering coefficient and radial air velocity. Multiple pulses are averaged over a time interval called a ray. The
 Halo is capable of full hemispheric scanning of the backscatter coefficient and radial velocity as a function of beam range. The
 Halo laser interacts with relatively large aerosol particles compared to typical ceilometers (≈ 900 nm) and aerosol lidars
 (typically 355 or 532 nm) having shorter wavelength lasers. This often restricts the Halo instruments to boundary layer
 600 measurements because the free troposphere is typically very low in coarse mode aerosol concentrations, apart from sporadic
 elevated plumes such as those containing volcanic ash and mineral dust.

The radial velocity data during vertical stares can be used statistically over sufficiently long averaging intervals (10–60 minutes
 depending on the SNR) to compute variance, skewness and kurtosis throughout the boundary layer and some distance into
 605 cloud (2–4 gates) before attenuation becomes significant. Therefore, quantities derived from the vertical velocity and
 backscatter coefficient—diagnosing updraughts and downdraughts, times of crossover and onset (Brooke et al., 2023) up to
 and including cloud base, and diagnosing boundary-layer type (Harvey et al., 2013)—can be determined.

Serial No.	#01	#30	#35
No. gates, typically	200	200	200
Gate length (m)	30	30	36
Velocity resolution (m s^{-1})	0.023	0.038	0.0318
Divergence (rad)	33	33	33
Max range (m)	9,600	15,000	15,000
Pulse repetition (kHz)	20	15	10
Sampling frequency (MHz)	30	50	50
Depolarisation capability	no	no	yes

Table 9: Halo Photonics Doppler lidar specifications

610 The usual operation at Cardington was vertical stares (zenith angle=0°) with periodic wind scans that invoke various options
 of off-axis views. Wind profiles performed every 30 minutes was the default operation for wind scans, although this was not
 strictly always the case. Most profiles of horizontal wind within the historical dataset are based on doppler beam swinging
 (DBS) scans which use a tri-axis azimuthally orthogonal technique using the single lidar beam to retrieve horizontal mean
 615 wind components. This scan was chosen for the bulk of the time because it only takes about 21 s to complete, which leaves
 98% of the available time to vertical stares if one wind scan is completed every 30 minute. More recent scans have however
 used multi-axis velocity azimuth display (VAD) scans, which are effectively a more involved version of the DBS scans and
 use 6 or 12 point off-zenith views. Whatever the method employed, there is the assumption of a horizontally homogeneous



wind flow and constant vertical velocity over the sampling volume, i.e. the volume defined by the conical ‘chunky slice’ defined by the geometry of the lidar beam divergence and gate length. The two scan methods used to estimate vertical profiles of the horizontal wind speed and directions are described in more detail below. A third, rarely used, type of scan was the range height indicator (RHI) where the elevation angle is stepped for a fixed azimuth angle. The vertical stares, DBS, VAD and RHI data are stored in separate archived NetCDF file names, as listed in Section 8. The scan files contain the same variables as the vertical stare files i.e. range, radial Doppler velocity, backscatter, signal-to-noise ratio for each of the scan positions. Derived profiles of horizontal wind speed and direction are stored in separate files as described in Section 8 and the complete set of variables in the netCDF files are shown in S6 in the Supplementary section.

(i) DBS scanning

For a given measured vertical velocity component (r_w , from a vertical stare) and two orthogonal off-zenith radial velocities (r_n , r_e) at an elevation angle α (commonly 75°), then the horizontal components are calculated from the two beam axes by:

$$v_n = \frac{r_n}{\cos \alpha} - r_w \tan \alpha \quad (1)$$

$$v_e = \frac{r_e}{\cos \alpha} - r_w \tan \alpha \quad (2)$$

which then means we then calculate the two horizontal wind components thus:

$$u = \frac{v_e \cos q_n - v_n \cos q_e}{\sin (q_e - q_n)} \quad (3)$$

$$v = \frac{v_n \sin q_e - v_e \sin q_n}{\sin (q_e - q_n)} \quad (4)$$

which allows the mean vector wind speed ($\sqrt{u^2 + v^2}$) and direction ($\tan^{-1} \frac{v}{u}$) to be calculated.

(ii) VAD scanning

As already stated, the Halo retrieves the radial component of the ambient wind field based on the movement of suspended aerosol particles. The orthogonal wind components (u , v , w) that we desire are related to this radial, or line of sight, velocity (v_r) thus:

$$v_r = \sin \varphi (u \sin \theta + v \cos \theta) + w \cos \varphi$$



where ϕ is the zenith angle (vertical= 0°) and θ is the azimuth angle (north= 0°). u is defined as the east-west component, v as
 650 the north-south component, and w is the vertical component. The coefficient matrix containing the relationship between the
 radial velocities and the wind vector components is compiled of $\mathbf{A} = [\sin\phi \sin\theta, \sin\phi \cos\theta, \cos\phi]$ for the specified number of
 azimuth and zenith angles. This is usually over an azimuthal scan of either 6 or 12 points for a fixed zenith angle (15° from
 vertical). A least squares approximation is then sought as the solution to the linear vector matrix equation $\mathbf{A}\mathbf{v}=\mathbf{V}_r$ that has no
 absolute solution, where $\mathbf{V}_r$ is the vector of the measured radial velocities for the separate beams and $\mathbf{v}$ is the 3-dimensional
 655 wind vector containing the u , v and w components we require, such that $\mathbf{v}=(\mathbf{A}^T\mathbf{A})^{-1} \mathbf{A}^T\mathbf{V}_r$, with T indicating the transposed
 matrix. The mean wind speed and direction can then be calculated

The principal use of the Halo lidars has been to visualise boundary layer meteorology as function of time using plots of the
 vertical component of the retrieved vertical velocity turbulence (often designated as the velocity variance σ_w^2) and the
 660 backscatter coefficient ($\text{m}^{-1} \text{sr}^{-1}$). Thereby turbulent mixing, growth of diurnal boundary layers, development of the profiles of
 morning and evening transitions, nocturnal low-level jets, sea breeze fronts and other phenomena can be observed.

The linear depolarization ratio (Vakkari et al., 2021) was also possible with #30 and #35, although this was not switched on
 by default. The co- and cross-components of the returned laser pulses from non-spherical aerosol particles or ice crystals was
 665 achieved with a fibre-optic switching polarizer. Depolarization ratio as a function of zenith angle of orientated ice crystals in
 cirrus clouds (such as used in Westbrook et al., 2010a) has been studied to some degree using #35 data, although this technique
 was not fully developed because the scanning had to be done manually and was not able at the time to be automated using the
 available control software. The cross-component data gathered at Cardington is nonetheless included as separate archive files
 (see Section 8).

670 Another unpublished project was using dual lidar differential absorption (using the simultaneous Halo and CT25k ceilometer
 data) to retrieve precipitation rates, such as fine drizzle falling from stratocumulus cloud. Drizzle from warm stratocumulus at
 an inland site such as Cardington is unusual compared to marine stratocumulus, so when it occurs, for example due to
 particularly thick cloud layer, it becomes of interest because of (i) NWP forecast errors, (ii) traditional methods to measure
 675 rainfall are likely to be insensitive to the drizzle at the surface. Calibration of the attenuated backscatter coefficient is required
 for this differential absorption technique, which is achievable using the integrated backscatter in optically thick stratocumulus
 based on the predictions of Mie scattering theory. There is potential within the datasets to continue this analysis in a similar
 manner to Westbrook et al (2010b).

680



5.2 Microwave radiometers

Passive microwave radiometers allow continuous monitoring of integrated water vapour (IWV), liquid water path (LWP), and in addition sometimes profiles of humidity and temperature. Although humidity data is available (depending on the model), temperature is not included in any of the Cardington data. Various types of internal and external calibrations are required for these radiometers due to the huge gains required to do the required retrievals. Three models of radiometer that were deployed at Cardington are described below, all of which used zenith views for the retrievals with occasional off-zenith views for calibration purposes.

The Radiometrics WVR-1100 passive radiometer was the longest serving such device and measured the atmospheric emissions at two frequencies (23.8 and 31.4 GHz) which provide brightness temperature at these channels and thereby information of the column water vapour and liquid water. The WVR-1100 used a bi-linear regression method based on local radiosonde launches to retrieve column integrations of liquid water and water vapour (Price, 2003). A large number of past radiosonde launches were required that had been carried out from the site at which the radiometer was located; concurrent launches are not required in general in order to operate microwave radiometers. The WVR-1100 in addition performed ‘tipping curve’ observations using off-zenith slant scans where the optical depth for each frequency varies in a known way with atmospheric geometrical thickness. Tipping curves assume the atmosphere is horizontally homogeneous. The overall error in liquid water path is estimated to be 0.015 kg m^{-2} . Water vapour and liquid water column amounts were logged typically every 9–10 s. As with all microwave radiometers, absolute calibrations for the absorbing channels were done occasionally (such as when the radiometer was moved) using an external black body cooled with liquid nitrogen. See Supplementary S3 for full list of variables for the WVR-1100.

The Radiometrics TP/WVP-3000 microwave radiometer was mostly used on detached duty and therefore relatively little at Cardington; we nonetheless still include the available data in the archive. It used a neural network to retrieve profiles of water vapour and temperature. The neural network was trained with a radiative transfer model using multiple years of radiosonde data. The TP/WVP-3000 was set up to take readings in the vertical approximately every 8 s. Regular tipping curve scans were done over a range of zenith angles (30, 45, 90, 135, 150°) to compare the atmospheric radiances to that of known values at relatively opaque water vapour frequencies (with the opacity being a linear function of the slant path), in addition using frequent views of an internal temperature-controlled black body. See Supplementary S4 for full list of variables for the TP/WVP-3000.

The RPG Humpro profiling radiometer retrieved humidity profiles in addition to the usual liquid water and integrated water vapour paths (see Supplementary S5 for full list of variables) using brightness temperatures measured at seven microwave frequencies between 22.24 and 31.4 GHz (this band in general being sensitive to water vapour and cloud). The liquid and



vapour water path retrievals used a supplied neural network algorithm (which is trained with radiosonde data using a radiative transfer scheme). Two archive files were produced, based around the time series (water vapour and liquid water) and profile (humidity) data; see file name listings in Section 8.

720 5.3 Ceilometers

The three models of near-infrared diode laser ceilometers installed at Cardington (called the LD25, LD40 and CT25K as shown in Table 8) are able to retrieve not only cloud base height (at up to three levels if penetration power is sufficient), but also cloud penetration depth per cloud layer, the vertical visibility, and a measure of the vertical profiles of backscattered intensity in a similar manner to the Halo Doppler lidars. There were two CT25K ceilometers installed, with the second unit deployed from October 2015 and is called CT25K_B in the archived netCDF files. The CT25K_B was tilted 4° from the zenith to avoid specular backscatter from cirrus clouds. The other ceilometers all pointed in the true zenith. For the cloud-base height retrievals from the CT25K_B, the height above ground level was corrected for the instrument tilt. Supplementary section S8 lists the variables in the ceilometer netCDF files.

730

5.4 Radar Wind Profiler

The National Centre for Atmospheric Science (NCAS) mobile Degreane Horizon PCL1300 Radar Wind Profiler (RWP) owned by the University of Manchester was originally purchased by Aberystwyth University in 2002. The RWP was deployed at Cardington for non-continuous periods between 2002 and 2016 as part of collaborative work with the Met Office. The advantage of the RWP over a lidar is that it can measure in and above cloud. Technically a L-band radar operating at 1290 MHz, these RWPs are commonly called UHF Doppler radars in the literature. At this frequency radars detect clear air echoes from variations in refractive index on a scale of 23 cm. In the lower atmosphere these irregularities are mainly due to humidity fluctuations. In the presence of hydrometeors stronger Rayleigh scattering dominates the signal.

The RWP consists of three static arrays of dipole antennae panels that both emit and receive three separate beams. The vertical panel measures the vertical component of the wind, and the other panels at elevations of 73° and orthogonal azimuths provide a direct measurement of the mean radial velocity along the radar beam. The RWP cycles between the antenna directions and data is combined to calculate full wind vectors. The RWP measures wind speed (direction) to an intrinsic accuracy of $< 1 \text{ m s}^{-1}$ ($< 10^\circ$) in all weather conditions. In principle, the minimum altitude was 75 m depending on ground clutter signals and atmospheric conditions with a minimum vertical gate spacing of 75 m. The radar typically returned wind profiles from around 75 to 4500 m depending on atmospheric conditions.



Data is archived as daily NetCDF files (see Supplementary section S10) using 15 minute averages and can be found at the CEDA repository, albeit not as part of the Cardington archive otherwise described in this paper. Applications of RWP data can be found in Norton et al (2008), Morcrette et al (2006), Parton et al (2009) and Osborne and Lapworth (2017).

750

6 Radiosonde archive

We include the historical archive of radiosonde launches at Cardington going back to 1996. Although this goes back seven years more than the surface site core dataset archive, this was relatively straightforward to do due the consistency in data format and also due to the large number of routine daily or twice-daily launches during 1997 and especially 1998 that have the potential to be used statistically by future data users. Examples from the past of using large numbers of radiosonde soundings for instrument validation include microwave radiometer retrievals (Price, 2003; Gaffard and Hewison, 2003) and lidar profiling of water vapour (Gaffard et al., 2021).

Table 10 summarises the major Vaisala sonde package types and associated generation of the ground receiving station. Figure 5 displays in the top panel the number count of launches as a function of date, coloured by campaign year. The same colours are used in the bottom two panels showing histograms of the hourly time of launch and sounding termination height. The termination height was determined by setting a timer during the sonde initialisation, or terminating the sonde manually during flight, or when the balloon bursts naturally at altitude. Additionally, sometimes an increase in pressure, and therefore a decrease in sonde altitude, would cause the software to terminate the sounding.

Figure 5 shows the largely sporadic nature of radiosonde launches at Cardington, which shows how launches were focussed on campaigns with distinct research goals (as annotated by research type or campaign name) that would vary the time and frequency of launches. A notable exception was in 1997 and 1998 where launches were carried out daily (sometimes early morning but mostly at midday and midnight) for an extended period. The lack of sonde launches between 2021 and 2023 was because of detached campaigns which meant the focus was away from Cardington. Launches not annotated in Fig. 5 were opportunistic launches based on interesting weather events (e.g. thick fog) or at the request of operational forecasters in the Met Office who wanted to understand the structure of the atmosphere better during periods of troublesome weather conditions.

775



years	sonde	mass	location/winds	ground station	sensors	battery, notes
1996–2002	RS80	220g	Loran-C -GPS from 1999	MW11/M W15	ThermoCap, HumiCap, BaroCap	2 nd gen, Wet battery, lead-acid 19V
2006–2014	RS92- SGPD	280g	GPS	MW31	ThermoCap, HumiCap, capacitive pressure	3 rd gen, Alkaline 9V, lifetime 135 minutes
2014–2024	RS41- SG(P)	109g	GPS	MW41	PRT, SGP=silicon capacitive pressure, HumiCap	4 th gen, Lithium 3V, lifetime >240 minutes

Table 10: Summary of radiosonde and receiving station versions.

The majority of sonde launches at Cardington were performed with small 100 g latex balloons that were filled with enough helium to provide a mean ascent rate of 2.5 m s^{-1} . This is about half the rate of the operational Met Office launches, for example, that use 700 g balloons. The slower ascent rate was used to improve the vertical sampling resolution in the atmospheric boundary layer whilst maintaining a sufficient ventilation rate over the sensors. Because almost all of research at Cardington concerned the boundary layer, many sonde launches were terminated at low altitude so that rapid-fire launches could be carried out (e.g. 30–45 minute intervals between launches during fog research).

During the ground preparation of each sonde unit the humicap and PRT sensors were heated to 150°C to remove contaminants, a procedure called conditioning, then zero humidity and in-built temperature checks were performed. A ground truth for the pressure sensor is also entered and once a stable pressure reading is secured, and the sensor boom has cooled down, then the sonde is ready for launch. Unless a fault with one or more of the sonde variables was noted during flight, with the possibility of terminating the sonde and launching another to replace it, no routine quality control was carried out on the data in the sonde archive. Each launch sounding has its own netCDF file and the Supplementary section 9 lists the variables in full.

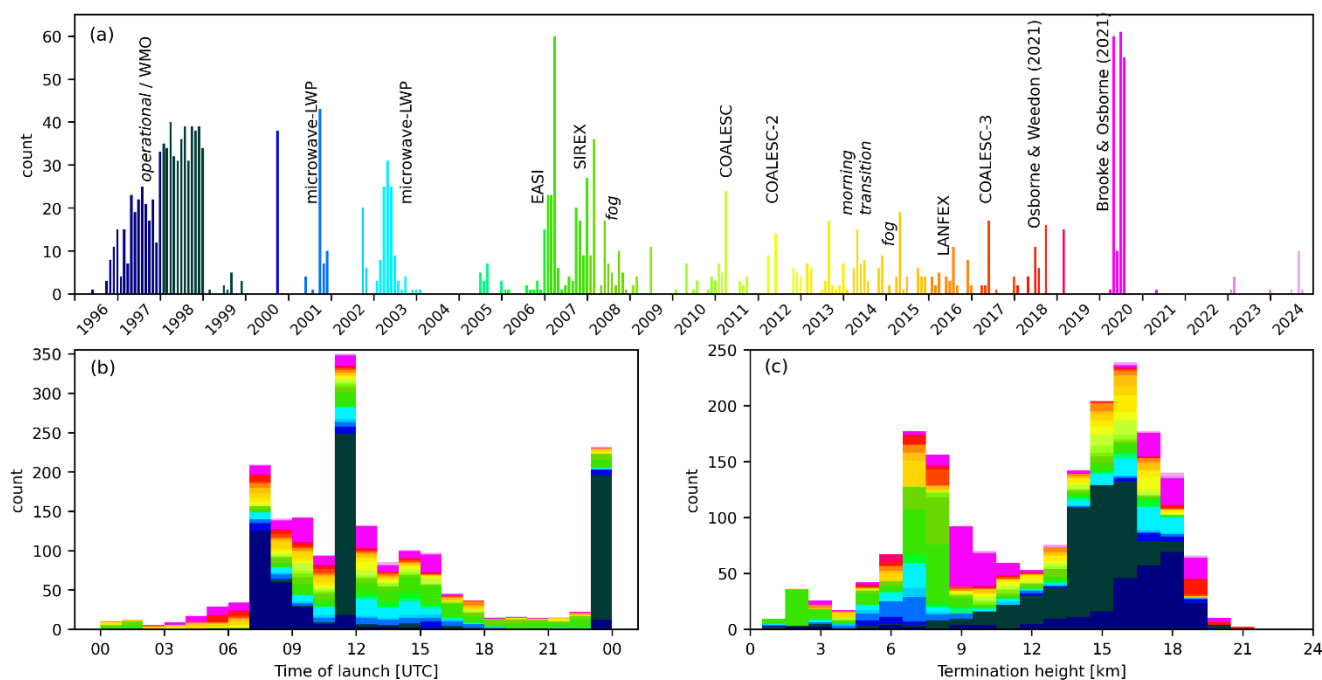


Figure 5: Summary of radiosonde launches at Cardington: (a) Number of launches per month with annotations of significant campaigns—colours are used to distinguish years in the manner as the next two panels, (b) histogram of time of launch with 1-hourly bins, (c) histogram of sonde termination height with 1-km bins. COALESC=Combined Observations of the Atmospheric boundary-Layer to study Evolution of StratoCumulus; EASI= East Anglia Stratocumulus Investigation; LANFEX=Local and Non-local Fog Experiment; SIREX=Surface Inhomogeneity Research Experiment; WMO=World Meteorological Organization database.

7 Example of turbulence data—roughness length

The multi-height wind and turbulence dataset begins in earnest at the start of 2004 when sonic anemometer data was standardised at 10 m, 25 m and 50 m and ends in December 2024 when the site closed. Mast data at 2 m is also available from 2011 onwards. Since 2012 urbanisation in the form of housing has expanded to the north of the site within 0.7–1.5 km. The sector from 350° to 025° was excluded to remove the effects of the hangars, and thereby also much of the urbanised area to the north. Such filtering is common when analysing wind and turbulence data at Cardington. We have been yet more exclusive for the analysis shown here and only included mast data with wind directions from 155 to 280°. The 155° to 280° sector contains the prevailing winds and amounts to 58 % of the total turbulence dataset (Fig. 1c). This removes potential disturbances from housing to the northwest and also highly localised effects of the site accommodation buildings within 150 m of the masts from the east, notwithstanding the otherwise largely undeveloped land in general out to the east. Figure 6 shows probability



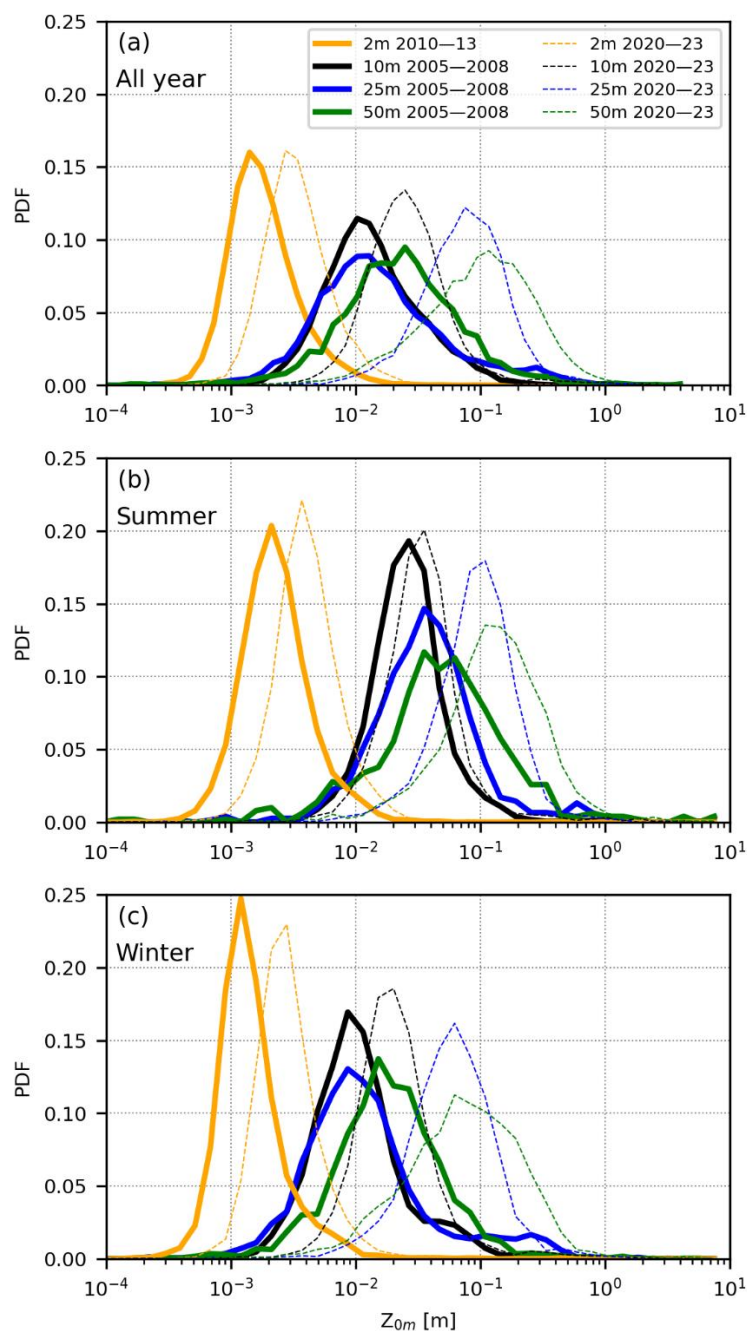
distribution functions (PDF) of the roughness length for momentum ($z_{0,m}$) calculated in neutral conditions ($|z/L| < 0.1$, where L = Ubokhov length and $z = 10$ m) at the four mast heights for (a) all year round, (b) summer months and (c) winter months.
 815 The effect of the onsite single story buildings to the immediate east of the sensors affected the turbulence at the 2 m and 10 m mast heights in particular, i.e. a hump at around $+1\sigma$ of the modal value.

Of interest for the turbulence data within the prevailing conditions of open fetch is an area of woodland called Shocott Spring that was planted in stages from 2005 to 2011. This area now amounts to 52 Ha, is 0.5–1.3 km away from Cardington and lies
 820 within the south to southwest sector (see the change in land use in the two photographs in Fig. 1). It is this gradual growth of Shocott Spring up to and including 2024 that has impacted on the turbulence measured at Cardington within the prevailing wind direction.

Turbulence data in Fig. 6 has been averaged over two 4-year periods of 2005–2009 (or 2010–2013 at the 2 m height) and
 825 2020–2023. We might describe summer and winter as ‘leafy’ and ‘non-leafy’ seasons in that the change in foliage explains the increase in $z_{0,m}$ in the summer relative to the winter at all heights from both 4-year periods. There are also modest increases in summer compared to all-year, although at 50 m for the 2020–2023 period the modal value is about the same— but note the broader distributions in general for the all-year data. The log-normal distributions in general are close to symmetrical, but with some negative skew at the 50 m height.

830 Figure 6 shows a distinct increase in $z_{0,m}$ between the two time periods at all four heights, with the increase being especially large at 25 and 50 m i.e. nearly an order of magnitude at 25 m. The increase in $z_{0,m}$ with measurement height within each period is to be expected. Growing vegetation within Shocott Spring and other minor land use changes can explain the overall increase in $z_{0,m}$ over time. The 2 m derived roughness length will be representative of turbulence generated mainly within the site itself,
 835 so its increase over the period may have resulted from growth of the hedgerow that is the site boundary.

There is a dramatic increase in the modal value of $z_{0,m}$ from 2 m to 10 m, but a smaller increase from 10 m up to 50 m. This is because turbulence at 2 m is not fully developed and so is too low to be representative of the general fetch. The 10 and 25 m modes of $z_{0,m}$ are about equal for all-year and winter data in the 2005–2009 period. During the 2020–2023 period, however,
 840 substantial differences appear with much larger values at 25 m compared to 10 m, i.e. probably from growing vegetation (trees) further upwind that affect the turbulence more at 25 m than at 10 m.



845

Figure 6: Probability distributions functions (PDFs) of momentum roughness length (z_{0m}) at four different heights and split into two time periods of 4 yrs each. (a) The ‘All year’ data has been sub-divided into (b) ‘Summer’ (May, Jun, Jul, Aug), and (c) ‘Winter’ (Nov, Dec, Jan, Feb), where we assume deciduous vegetation is ‘leafy’ and ‘non-leafy’, respectively.



850 The increase in $z_{0,m}$ between the two periods is seen at 50 m and 25 m during the winter as well as the summer. The winter increase possibly shows the turbulent effects of deciduous trees even when bare. At 10 m the increase is only significant in winter, but the winter modes are universally smaller than their summer counterparts. For the 2005–2009 period the winter mode values at 10 m and 25 m are about equal but they separate out for 2020–2023 because of the large increase at 25 m already mentioned.

855

In summary, Fig. 6 shows a relatively simple illustration of the Cardington turbulence based on the sonic anemometer data and how the roughness length for momentum varies with chosen sensor height and time. Such variations should be considered when modelling results that have a simplistic treatment of roughness length are compared to observations. There is much further potential in the Cardington data, for example for retrieving roughness lengths for scalar properties such as heat and
 860 water vapour and also for retrieving aerodynamic surface resistances for moisture, heat and momentum (e.g. Liu et al., 2006).

8 Data availability

All data described in this paper is stored on the UK-based CEDA repository as NetCDFs, all of which comply to the CF-1.8
 865 ACDD-1.3 conventions. The CEDA archive can be accessed at: <https://archive.ceda.ac.uk> with the Dataset Collection for Met Office Meteorological Research Unit, Cardington at <http://catalogue.ceda.ac.uk/uuid/5487380511084413a502c4b229273bc6> (Met Office, 2025).

Use of these data is covered by the following licence: <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>.
 870 [licence/version/3/](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/). When using these data, you must cite them correctly using the citation given on the CEDA Data Catalogue record. Separate DOIs are provided for each of the datasets outlined above, i.e. at the four different time steps of the core dataset, with the large radiometers, radiosondes and derived LSM forcing files supplied with separate links. All data are divided into daily files, except for the LSM forcing files (necessarily continuous at 30 minute time intervals from 2004 to 2024) and the radiosondes (one file per launch).

875

8.1 File formatting

The NetCDF file naming conventions are listed below, where $YYYY$ = year (e.g. 2005), MM = month (01–12), DD = day of month (01–31), hh = hour (01–23), mm = minutes (01–59), ss = seconds (01–59), ht = height in m (drive level for JULES):

880

- (i) Core surface site NetCDF files (Supplementary section S1) under the description “Dataset Collection Record: Continuous hydrometeorological record (2004–2024) at the Met Office surface site of Cardington, UK Dataset Collection”:



metoffice-obr-fluxtower_cardington_YYYYMMDD_mm.nc

885 The data flags shown in Table 6 summarise the codes used in the NetCDF files for all variables in the core surface site NetCDF files. *mm* here refers to the four averaging periods, i.e. ‘01’, ‘05’, ‘10’ or ‘30’ minutes. Section S2 is an exhaustive list of variables within these files. Data files are available using these DOI links:

1 minutes: <https://dx.doi.org/10.5285/fe75afd7723140c19edfdeb75fed1e48>

5 minutes: <https://dx.doi.org/10.5285/e75de035395f48dbbb43f1a190406632>

890 10 minutes: <https://dx.doi.org/10.5285/27dfc610944446a6a7862c97f93325a6>

30 minutes: <https://dx.doi.org/10.5285/1f92b91149704c4bb5048300615a1945>

(ii) JULES standalone forcing file derived from core surface site NetCDF files (Supplementary section S2):

metoffice-obr-Forcing-ht-W-T-Q-gap-filled-Cardington-2005-2024_v1.nc

895 Data variables have an associated flag at every 30 minute time step: ‘0’ means actual value, ‘1’ means gap-filled. Data available at: <https://dx.doi.org/10.5285/19c5dc39bb8c4c40a5643678c31168e7>

(iii) WVR-1100 microwave radiometer NetCDF files (Supplementary section S3):

metoffice-obr-microwave-radiometer-wvr1100_cardington_YYYYMMDD.nc

900 Data available at: <https://dx.doi.org/10.5285/21c423889e6a4035ac7f4761e467de2b>

(iv) Humpro microwave radiometer NetCDF files (Supplementary section S4):

metoffice-obr-microwave-radiometer-humpro-timeseries_cardington_YYYYMMDD.nc

metoffice-obr-microwave-radiometer-humpro-humidity-profile_cardington_YYYYMMDD.nc

905 Data available at: <https://dx.doi.org/10.5285/9bf50847dd4d49a281d5663d512e1646>

(v) TP/WVP-3000 microwave radiometer NetCDF files (Supplementary section S5):

metoffice-obr-microwave-radiometer-tp-wvp-3000_cardington_timeseries-profiles_YYYYMMDD.nc

Data available at: <https://dx.doi.org/10.5285/d44b15c8f183404ca47291bc677f93e0>

910

(vi) Halo Doppler lidar NetCDF files (Supplementary section S6):

metoffice-obr-halo-lidar-serial_cardington_scantype_YYYYMMDD.nc

where *serial* can be “01”, “30” or “35” and *scantype* can be “stare”, “dbs-scans”, “vad-scans”, “rhi-scans”, “cross-stare” or “windprofiles”

915 Data available for Halo #01 at: <https://dx.doi.org/10.5285/ba87087355ed4e748d1650d012adc4ef>,

for Halo #30 at: <https://dx.doi.org/10.5285/6ebd987dac6f4d1692d878258bf7112c>,

for Halo #35 at: <https://dx.doi.org/10.5285/77bef4103ec2426281a5e74ccc0ba5c7>



(vii) Thies laser precipitation monitor (disdrometer) NetCDF files (Supplementary section S7):

920 **metoffice-obr-disdrometer_cardington_precipitation_YYYYMMDD.nc**

Data available at: <https://dx.doi.org/10.5285/5d8997e0cd974835999a8d8ba677b26f>

(viii) Ceilometer data files (Supplementary section S8):

metoffice-obr-ceilometer-model_cardington_cloudbase-backscatter_YYYYMMDD.nc

925 where *model* can be “ld24”, “ld40” or “ct25k” i.e. the three ceilometer devices have their own NetCDF files, but with similar attributes and variable names as show in S9.

LD25 data available at: <https://dx.doi.org/10.5285/06a71fd559884416ad798e452aa21bef>

LD40 data available at: <https://dx.doi.org/10.5285/51d139c161e746e0a9d91e1156958a88>

CT25K_A data available at: <https://dx.doi.org/10.5285/a71afe47c63a4e4a9c9d5d18625cd8f9>

930 CT25K_B data available at: <https://dx.doi.org/10.5285/691e8c36bb9446efaa46cc67d37aadae>

(ix) Radiosonde sounding data files (Supplementary section S9):

metoffice-obr-radiosonde_cardington_sounding_YYYYMMDD_hhmmss.nc

Data available at: <https://dx.doi.org/10.5285/5934d2a5706c4a3c9caa15188d9ed24b>

935

(x) NCAS (non-Met Office) radar wind profiler data files (Supplementary section S10):

ncas-radar-wind-profiler-1_cardington_YYYYMMDD_snr-winds_high-range-mode-15min_v8.0.nc

Data available using these links for various time periods:

20060407–20060927: <https://dx.doi.org/10.5285/eb352545ce1b4476b2580a3e5885c00d/>

940 20070821–20071029: <https://dx.doi.org/10.5285/c4846909fcad4480903857e7ef486743/>

20080924–20090130: <https://dx.doi.org/10.5285/38083202924c4785bc61b6e511ad3389/>

20100416–20110324: <https://dx.doi.org/10.5285/36d4e72b2ea8477aaba3eb6d0f052fad/>

20110325–20110525: <https://dx.doi.org/10.5285/3d38805185b34efda9d608830608eccc/>

20130509–20130619: <https://dx.doi.org/10.5285/4fd27b94fd794197aad1556a75abef27/>

945 20131106–20140831: <https://dx.doi.org/10.5285/9885fb709fbc4caba054bac772cefdd5/>

20150101–20160118: <https://dx.doi.org/10.5285/ca9b5288ad62491f8fb226eff22a0486/>



950 9 Summary

Well instrumented meteorological sites are important for the development and testing of land surface models. This paper describes the 20-yr (2004–2024) Met Office ground-based surface and near-surface meteorological and subsoil hydrological dataset from Cardington in central England. The variables that encompass atmospheric turbulence fluxes over a range of heights and the components of the near-surface energy balance, together with subsoil temperature and moisture sensors that include water table depth, make the dataset relatively unusual and potentially useful for both model evaluation and development of LSMs. Meteorological records of evapotranspiration and CO₂ fluxes using the eddy covariance method having a continuous duration of 20 yrs or more, like we have shown here in the Cardington dataset, are scarce (Li et al., 2025).

960 The Cardington core dataset has been quality controlled and is archived with open-access at the UK CEDA repository. This core data consist of four daily netCDF files based on four averaging periods (1, 5, 10 and 30 minutes). Use of the open-access data is subject to referencing this paper and acknowledgement of the CEDA repository. A subset of the core dataset has been used to generate four forcing files suitable for driving LSMs at a 30 minute time step. These files represent the whole 20-yr period for the three main drive heights, i.e. 10 m, 25 m and 50 m, plus the 2-m level from 2012 i.e. only whole years are included in the forcing files. The soil properties of the site were summarised in Table 7, which allows users to configure their simulations to local conditions. Specialist Doppler lidars, ceilometers, precipitation disdrometer and microwave radiometers are in addition archived into daily netCDF files based on the averaging interval as set for each instrument. Finally, there are the sporadic radiosonde soundings that are also archived, with one netCDF file per radiosonde launch. These amount to around 1800 site launches since 1996.

970 Use of the Cardington data in the past has tended to focus on shorter time periods of between the order of a day up to a few months. Decadal analysis of soil conditions and turbulence characteristics, for example, remain untapped. We showed an illustration of the potential use of the momentum roughness length trends in Fig. 6. The roughness length was shown to vary with measurement height and with time, averaged over both winter and summer seasons, as land use changed. Further analysis could break the data down into further wind sectors as we only used data from the prevailing 155–280° sector. We hope this, alongside the snapshot of turbulent heat flux and soil data in Fig. 3, serves as an advertisement for others to use the available data.

Acknowledgements

980 The data is archived at the Centre for Environmental Data Analysis (CEDA), which is part of the Natural Environment Research Council (NERC) Environmental Data Service (EDS).



Competing interests

985 The corresponding author has declared that there are no competing interests for all authors.

Author contributions

SRO principal writing; GW, SRO, JT, JB concept and construction; JMcG, JP, NP, MAP, TJ data collection; AKM JMcG, JB quality control, archiving; SRO, GW JULES forcing file generation and testing. All authors contributed to reviewing the paper.

990

References

- Beare, R. J., Edwards, J. M., and Lapworth, A. J.: Simulation of the observed evening transition and nocturnal boundary layers: Large-eddy simulation. *Quart J Royal Meteorol. Soc.*, 132, 81–99, <https://doi.org/10.1256/qj.05.64>, 2006.
- 995 Bell, S., Cornford, D., and Bastin, L.: How good are citizen weather stations? Addressing a biased opinion. *Weather*, 70(3), 75–84., which has been published in final form at <http://dx.doi.org/10.1002/wea.2316>, 2015.
- Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, *Geosci. Model Dev.*, 4, 677–699, <https://doi.org/10.5194/gmd-4-677-2011>, 2011.
- 1000 Blyth, E., Gash, J., Lloyd, A., Pryor, M., Weedon, G. P. and Shuttleworth, J.: Evaluating the JULES Land Surface Model Energy Fluxes Using FLUXNET Data. *J. Hydrometeorol.*, 11, 509–519. <https://doi.org/10.1175/2009JHM1183.1>, 2010.
- 1005 Bosveld, F. C., Baas, P., Beljaars, A. C. M. et al. Fifty Years of Atmospheric Boundary-Layer Research at Cabauw Serving Weather, Air Quality and Climate. *Boundary-Layer Meteorol* 177, 583–612, <https://doi.org/10.1007/s10546-020-00541-w>, 2020.
- 1010 Boutle, I. A., Eyre, J. E. J. and Lock, A. P.: Seamless Stratocumulus Simulation across the Turbulent Gray Zone. *Monthly Weather Review*, 142, 1655–1668. <https://doi.org/10.1175/MWR-D-13-00229.1>, 2014.
- Boutle, I. Price, J., Kudsztsa, I., Kokkola, H. and Romakkaniemi, S.: Aerosol–fog interaction and the transition to well-mixed radiation fog. *Atmos. Chem., Phys.*, 18, 7827–7840. <https://doi.org/10.5194/acp-18-7827-2018>, 2018.
- 1015 Brooke, J. K. and Osborne, S. R.: Observations and numerical weather forecasts of land-surface and boundary-layer evolution during an unusually dry spring at a site in central England. *Quart J Royal Meteorol. Soc.*, 148, 15–34, <https://doi.org/10.1002/qj.4190>, 2021.
- Brooke, J. K., Best, M. J., Lock, A. P., Osborne, S. R., Price, J., Cuxart, Boone, A., Canut-Rocafor, G., Hartogensis, O. K., Royet, A.: Irrigation contrasts through the morning transition. *Quart. J. Royal Meteorol. Soc.*, 150(758), 170–194. <https://doi.org/10.1002/qj.4590>, 2023.
- 1020 Burton, R. G. O.: Assessment of soil characteristics at the Meteorological Office Research Unit experimental site. Cardington Tech. Note 49, 25 pp., https://digital.nmla.metoffice.gov.uk/IO_5dcd1143-a38f-4936-91ea-53d3c5bb1655/, 1999.
- 1025 Bush, M., Flack, D. L. A., Lewis, H. W., Bohnenstengel, S. I., Short, C. J., Franklin, C., Lock, A. P., Best, M., Field, P., McCabe, A., Van Weverberg, K., Berthou, S., Boutle, I., Brooke, J. K., Cole, S., Cooper, S., Dow, G., Edwards, J., Finnenkoetter, A., Furtado, K., Halladay,



- 1030 K., Hanley, K., Hendry, M. A., Hill, A., Jayakumar, A., Jones, R. W., Lean, H., Lee, J. C. K., Malcolm, A., Mittermaier, M., Mohandas, S.,
 Moore, S., Morcrette, C., North, R., Porson, A., Rennie, S., Roberts, N., Roux, B., Sanchez, C., Su, C.-H., Tucker, S., Vosper, S., Walters,
 D., Warner, J., Webster, S., Weeks, M., Wilkinson, J., Whittall, M., Williams, K. D., and Zhang, H.: The third Met Office Unified Model–
 JULES Regional Atmosphere and Land Configuration, RAL3, *Geosci. Model Dev.*, 18, 3819–3855, [https://doi.org/10.5194/gmd-18-3819-](https://doi.org/10.5194/gmd-18-3819-2025)
 2025, 2025.
- 1035 Caughey, S. J., Crease, B. A. and Roach, W. T.: A field study of nocturnal stratocumulus II. Turbulence structure and entrainment. *Quart J*
Royal Meteorol. Soc., 108, 125–144. <https://doi.org/10.1002/qj.49710845508>, 1982.
- 1040 Chiriaco, M., Dupont, J.-C., Bastin, S., Badosa, J., Lopez, J., Haeffelin, M., Chepfer, H., and Guzman, R.: ReOBS: a new approach to
 synthesize long-term multi-variable dataset and application to the SIRTa supersite, *Earth Syst. Sci. Data*, 10, 919–940,
<https://doi.org/10.5194/essd-10-919-2018>, 2018.
- Cooper, H., Bennett, E., Blake, J., and coauthors: COSMOS-UK: national soil moisture and hydrometeorology data for environmental
 science research. *Earth System Science Data*, 13, <https://doi.org/10.5194/essd-13-1737-2021>, 2021.
- 1045 Edwards, J. M., McGregor, J. R., Bush, M. R., Bornemann, F. J.: Assessment of numerical weather forecasts against observations from
 Cardington: seasonal diurnal cycles of screen-level and surface temperatures and surface fluxes. *Quart J Royal Meteorol. Soc.*, 137, 656–
 672. <https://doi.org/10.1002/qj.742>, 2011.
- El-Madany, T. S., Griessbaum, F., Fratini, G., Juang, J. Y., Chang, S. C., and Klemm, O.: Comparison of sonic anemometer performance
 under foggy conditions. *Agr. For. Meteorol.*, 173, 63–73, 2013.
- 1050 FAO & IIASA: Harmonized World Soil Database Version 2.0. Rome and Laxenburg. <https://doi.org/10.4060/cc3823en>, 2023.
- Gaffard, C. and Hewison, T. J.: “Radiometrics MP3000 Microwave Radiometer Trial Report”, Met Office (OD) Technical Report TR26,
 Available from National Meteorological Library, UK , 2003.
- 1055 Gaffard, C., Li, Z., Harrison, D. L., Lehtinen R. and Roininen, R.: Evaluation of a Prototype Broadband Water-Vapour Profiling Differential
 Absorption Lidar at Cardington, UK, *Atmosphere*, 2021, 12, 1521, <https://api.semanticscholar.org/CorpusID:244451612>, 2021.
- Giblett, M. A.: The Structure of Wind over Level Country. *Nature* 130, 265–266, <https://doi.org/10.1038/130265a0>, 1932.
- 1060 Grant A. L. M.: Wind profiles in the stable boundary layer, and the effect of low relief. *Q J Roy Meteorol Soc* 120:27–46, 1994.
- Grimmond, C. S. B., and coauthors: Initial results from phase 2 of the International Urban Energy Balance Model Comparison. *Int. J.*
Climatol., 31, 244–272, DOI: <http://dx.doi.org/10.1002/joc.2227>, 2011.
- 1065 Harvey, N. J., Hogan, R.J. and Dacre, H. F.: A method to diagnose boundary-layer type using Doppler lidar. *Quart J Royal Meteorol. Soc.*,
 139, 1681–1693. <https://doi.org/10.1002/qj.2068>, 2013.
- Haywood, J., Bush, M., Abel, S., Claxton, B., Coe, H., Crosier, J., Harrison, M., Macpherson, M. B., Naylor, M. and Osborne, S.: Prediction
 of visibility and aerosol within the operational Met Office Unified Model. II: Validation of model performance using observational data.
Quart J Royal Meteorol. Soc., 134, 1817–1832. <https://doi.org/10.1002/qj.275>, 2008.
- 1070 Horlacher, V., Osborne, S. and Price, J.D.: Comparison of Two Closely Located Meteorological Measurement Sites and Consequences for
 Their Areal Representativity. *Boundary-Layer Meteorol.* 142, 469–493, <https://doi.org/10.1007/s10546-011-9684-3>, 2012.
- 1075



- Jones, J. I. P. and Butler, H. E.: The measurement of gustiness in the first few thousand feet of the atmosphere. *Quart. J. Royal Meteorol. Soc.*, 84(359), 17–24., 10.1088/0950-7671/42/6/311, 1958.
- 1080 Jones, J. I. P.: A portable sensitive anemometer with proportional d.c. output and a matching wind velocity-component resolver. *J. Sci. Instr.*, 42, 414–417, 10.1088/0950-7671/42/6/311, 1965.
- Jones, A., Thomson, D., Hort, M. and Devenish, B.: The U.K. Met Office's Next-Generation Atmospheric Dispersion Model, NAME III. In *Air pollution modeling and its application XVII*, pp 580–589. Boston, MA: Springer US., 2007.
- 1085 Jones, D.: CCAV document [Assessing-the-ODD-of-AV-perception-sensors-with-confidence-Sensor-Assurance-Framework-project-Met-Office-National-Physical-Laboratory-2024.pdf.aspx](#), 2022.
- Kaimal, J. C. and Gaynor, J. E.: Another look at sonic anemometry. *Boundary-Layer Meteorol.*, 56, 401–410., <https://doi.org/10.1007/BF00119215>, 1991.
- 1090 Lapworth, A., Claxton, B. M., and McGregor, J. R. The Effect of Gravity Wave Drag on Near-Surface Winds and Wind Profiles in the Nocturnal Boundary Layer over Land. *Boundary-Layer Meteorol* 156, 325–335, <https://doi.org/10.1007/s10546-015-0026-8>, 2015.
- Lapworth, A., and Osborne, S. R.: Evidence for gravity wave drag in the boundary layer of a numerical forecast model: a comparison with observations. *Quart J Royal Meteorol. Soc* <https://doi.org/10.1002/qj.2909>, 2016.
- 1095 Li, W., Yao, Z., Qu, Y., Yang, H., Song, Y., Song, L., Wu, L., and Cui, Y.: A benchmark dataset for global evapotranspiration estimation based on FLUXNET2015 from 2000 to 2022, *Earth Syst. Sci. Data*, 17, 3835–3855, <https://doi.org/10.5194/essd-17-3835-2025>, 2025.
- 1100 Liu, S., Mao, D., Lu, L.: Measurement and estimation of the aerodynamic resistance. *Hydrology and Earth System Sciences Discussions*, 3(3), 681–705. <https://hal.science/hal-00298684v1>, 2006.
- Lorria, C., McGinnis, D. F., Berg, P., Brand, A. and Wuest, A.: Application of Oxygen Eddy Correlation in Aquatic Systems. *J. Atmos. Oceanic Technol.*, 27, 1533–1546. <https://doi.org/10.1175/2010JTECHO723.1>, 2010.
- 1105 Lorria, C., McGinnis, D. F., Berg, P., Brand, A. and Wuest, A.: Application of Oxygen Eddy Correlation in Aquatic Systems. *J. Atmos. Oceanic Technol.*, 27, 1533–1546. <https://doi.org/10.1175/2010JTECHO723.1>, 2010.
- Maayar, M. E., Chen, J. M., and Price, D. T.: [On the use of field measurements of energy fluxes to evaluate land surface models](#). *Ecological Modelling*, 214, 293–304. <https://doi.org/10.1016/j.ecolmodel.2008.02.008>, 2008.
- 1110 Mauder, M., Foken, T., and Cuxart, J.: Surface-Energy_Balance closure over land: a review. *Boundary-Layer Meteorol.*, 177, 395–426. <https://doi.org/10.1007/s10546-020-00529-6>, 2020a.
- 1115 Mauder, M., Eggert, M., Gutsch, C., Oertel, S., Wilhelm, P., Voelksch, I., Wanner, L., Tambke, J., and Bogoev, I.: Comparison of turbulence measurements by a CSAT3B sonic anemometer and a high-resolution bistatic Doppler lidar, *Atmos. Meas. Tech.*, 13, 969–983, <https://doi.org/10.5194/amt-13-969-2020>, 2020b.
- Met Office; McGregor, J. Kerr-Munslow, A. Price, J. Osborne, S. Brooke, J.: Continuous hydrometeorological record (2004–2024) at the Met Office surface site of Cardington, UK Dataset Collection. NERC EDS Centre for Environmental Data Analysis, <http://catalogue.ceda.ac.uk/uuid/5487380511084413a502c4b229273bc6>, 2025.
- 1120



- 1125 Miller, M. A., Nitschke, K., Ackerman, T. P., Ferrell, W. R., Hickmon, N. and Ivey, M.: The ARM mobile facilities. Meteorological Monographs, 57, pp.9.1–9.15, <https://doi.org/10.1175/AMSMONOGRAPHS-D-15-0051.1>, 2016.
- Morcrette, C. J., Browning, K. A., Blyth, A. M., Bozier, K. E., Clark, P. A., Ladd, D., Norton, E. G., and Pavelin, E.: Secondary initiation of multiple bands of cumulonimbus over southern Britain. I: An observational case-study. Quart. J. Royal Meteorol. Soc.: A journal of the atmospheric sciences, applied meteorology and physical oceanography. 132(617):1021–51. <https://doi.org/10.1256/qj.05.151>, 2006.
- 1130 Niu, G-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y., 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. J. Geophys. Res., 116, <https://doi.org/10.1029/2010JD015139>.
- 1135 Norton, E. G., Vaughan, G., Methven, J., Coe, H., Brooks, B., Gallagher, M., and Longley, I.: Boundary layer structure and decoupling from synoptic scale flow during NAMBLEX. Atmos. Chem. Phys., 6, 433–445, <https://doi.org/10.5194/acp-6-433-2006>, 2006.
- Osborne, S. R., Abel, S. J., Boutle, I. A., and Marengo, F.: Evolution of Stratocumulus Over Land: Comparison of Ground and Aircraft Observations with Numerical Weather Prediction Simulations. Boundary-Layer Meteorol 153, 165–193, <https://doi.org/10.1007/s10546-014-9944-0>, 2014.
- 1140 Osborne, S. R. and Lapworth, A.: Initiation and Propagation of an Atmospheric Bore in a Numerical Forecast Model: A Comparison with Observations. J. Appl. Meteorol. and Clim., 56, 2999–3016 DOI: <https://doi.org/10.1175/JAMC-D-17-0045.1>, 2017.
- Osborne, S. R. and Weedon, G. P.: Observations and Modeling of Evapotranspiration and Dewfall during the 2018 Meteorological Drought in Southern England. J. Hydrometeorol., 22, 279–295. <https://doi.org/10.1175/JHM-D-20-0148.1>, 2021.
- 1145 Parton, G. A., Vaughan, G., Norton, E. G., Browning, K. A. and Clark, P. A.: Wind profiler observations of a sting jet. Quart. J. Royal Meteorol. Soc.. 135(640), pp.663–680, <https://doi.org/10.1002/qj.398>, 2009.
- 1150 Pastorello, G., Trotta, C., Canfora, E., and coauthors 2020: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, Sci. Data, 7, 225, <https://doi.org/10.1038/s41597-020-0534-3>, 2020.
- Pearson, G., Davies, F., and Collier, C.: 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. <https://doi.org/10.1175/2008JTECHA1128.1>, 2009.
- 1155 Price, J. D.: A case-study of a partially cloudy boundary layer. Quart J Royal Meteorol. Soc., 126, 393–413. <https://doi.org/10.1002/qj.49712656303>, 2000.
- Price, J.: Operation of a Radiometrics WVR-1100 passive microwave radiometer at the Met Office Research Unit, Cardington. Cardington Technical Note No. 61. Available from National Meteorological Library, UK, 2003.
- 1160 Price, J.: Radiation Fog. Part I: Observations of Stability and Drop Size Distributions. Boundary-Layer Meteorol 139, 167–191, <https://doi.org/10.1007/s10546-010-9580-2>, 2011.
- Price, J. D., and Clark, R.: On the measurement of dewfall and fog-droplet deposition. Bound.-Layer Meteor., 152, 367–393. <https://doi.org/10.1007/s10546-014-9930-6>, 2014.
- 1165 Price, J., Porson, A. and Lock, A.: An Observational Case Study of Persistent Fog and Comparison with an Ensemble Forecast Model. Boundary-Layer Meteorol., 155, 301–327, <https://doi.org/10.1007/s10546-014-9995-2>, 2015.



- 1170 Price, J. D., Lane, S., Boutle, I. A., Smith, D. K. E., Bergot, T., Lac, C., Duconge, L., McGregor, J., Kerr-Munslow, A., Pickering, M. and Clark, R.: LANFEX: A Field and Modeling Study to Improve Our Understanding and Forecasting of Radiation Fog. <https://doi.org/10.1175/BAMS-D-16-0299.1>, pp 2061–2077, 2018.
- Price, J. D.: On the Formation and Development of Radiation Fog: An Observational Study. *Boundary-Layer Meteorol* 172, 167–197.
- 1175 <https://doi.org/10.1007/s10546-019-00444-5>, 2019.
- Raabe, A., Arnold, K and Ziemann, A.: Horizontal Turbulent Fluxes of Sensible Heat and Horizontal Homogeneity in Micrometeorological Experiments. *J. Atmos Oceanic Technol.*, 19, 1225–1230. [https://doi.org/10.1175/1520-0426\(2002\)019<1225:HTFOSH>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<1225:HTFOSH>2.0.CO;2), 2002.
- 1180 Rawlins, F., Ballard, S. P., Bovis, K. J., Clayton, A. M., Li, D., Inverarity, G. W., Lorenc, A. C. and Payne, T. J.: The Met Office global four-dimensional variational data assimilation scheme. *Quart. J. Royal Meteorol. Soc.*, 133, 347–362. <https://doi.org/10.1002/qj.32>, 2007.
- Readings, C. J. and Rayment, D. R.: The high-frequency fluctuation of the wind in the first kilometer of the atmosphere. *Radio Science*, 4, 1127–1131. <https://doi.org/10.1029/RS004i012p01127>, 1969.
- 1185 Roach, W. T., Brown, R., Caughey, S. J., Garland, J. A., and Readings, C. J.: The physics of radiation fog: I—A field study. *Quart. J. Roy. Meteor. Soc.*, 102, 313–333, <https://doi.org/10.1002/qj.49710243204>, 1976.
- Savage, N. H., Agnew, P., Davis, L. S., Ordóñez, C., Thorpe, R., Johnson, C.E., O'Connor, F. M., and Dalvi M.: Air quality modelling using the Met Office Unified Model (AQUM OS24-26): model description and initial evaluation. *Geoscientific Model Development*, 6, 353–372. <https://doi.org/10.5194/gmd-6-353-2013>, 2013.
- 1190 Shetter, R.E., Junkermann, W., Swartz, W. H., and coauthors: Photolysis frequency of NO₂: Measurement and modeling during the International Photolysis Frequency Measurement and Modeling Intercomparison (IPMMI). *J. Geophys. Res.*, 108, <https://doi.org/10.1029/2002JD002932>, 2003.
- 1195 Smith, F. B. and Hay, J. S.: The expansion of clusters of particles in the atmosphere. *Quart. J. Royal Meteorol. Soc.*, 87(371), 82–101. <https://doi.org/10.1002/qj.49708737110>, 1961.
- 1200 Smith, D. K. E., Renfrew, I. A., Dorling, S. R., Price, J. D. and Boutle, I. A.: Sub-km scale numerical weather prediction model simulations of radiation fog. *Quart J Royal Meteorol. Soc.*, 147, 746–763, <https://doi.org/10.1002/qj.3943>, 2020.
- Turton, J. D. and Brown, R.: A comparison of a numerical model of radiation fog with detailed observations. *Quart. J. Roy. Meteor. Soc.*, 113, 37–54. <https://doi.org/10.1002/qj.49711347504>, 1987.
- 1205 Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P. R., Meyers, T. P., Prueger, J. H., Starks, P. J., Wesely, M. L.: Correcting eddy-covariance flux underestimates over a grassland. *Agric. For. Meteorol.* 103, 279–300, [https://doi.org/10.1016/S0168-1923\(00\)00123-4](https://doi.org/10.1016/S0168-1923(00)00123-4), 2000.
- 1210 Vakkari, V., Baars, H., Bohlmann, S., Bühl, J., Komppula, M., Mamouri, R.E. and O'Connor, E.J.: Aerosol particle depolarization ratio at 1565 nm measured with a Halo Doppler lidar. *Atmospheric Chemistry and Physics*, 21(8), pp.5807–5820, 2021.
- Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J. and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, 50(9), pp.7505–7514, 2014.
- 1215 Weedon, G. P., Osborne, S. R., and Best, M. J.: Dew, frost, fog and lifted temperature minima: Observations in southern England and implications for modelling. *Quart J Royal Meteorol. Soc.*, 150, 2168–2184. <https://doi.org/10.1002/qj.4702>, 2024.



- 1220 Westbrook, C. D., Illingworth, A. J., O'Connor, E. J. and Hogan, R. J.: Doppler lidar measurements of oriented planar ice crystals falling from supercooled and glaciated layer clouds Quart J Royal Meteorol. Soc., 136, 260–276, <https://doi.org/10.1002/qj.528>, 2010a.
- Westbrook, C. D., Hogan, R. J., O'Connor, E. J., and Illingworth, A. J.: Estimating drizzle drop size and precipitation rate using two-colour lidar measurements, Atmos. Meas. Tech., 3, 671–681, <https://doi.org/10.5194/amt-3-671-2010>, 2010b.
- 1225 Wilson, K. B., Baldocchi, D. D., Aubinet, M., Berbigier, P., Bernhofer, C., Dolman, H., Falge, E., Field, C., Goldstein, H., Granier, A., Grelle, A., Halldor, T., Hollinger, D., Katul, G., Law, B.E., Lindroth, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., Verma, S., Vesala, T., Wofsy, S.: Energy partitioning between latent and sensible heat flux during the warm season at FLUXNET sites. Water Resour. Res. 38, 1294, <https://doi.org/10.1029/2001WR000989>, 2002.
- 1230 Wyngaard, J. C.: Towards numerical modelling in the “Terra Incognita.” J. Atmos. Sci., 61, 1816–1826, [https://doi.org/10.1175/1520-0469\(2004\)061<1816:TNMITT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2), 2004.
- Yang., Z-L, Niu, G-Y., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Longueverge, L., Manning, K., Niyogi, D., Tewari, M. and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. J. Geophys. Res., 116, <https://doi.org/10.1029/2010JD015140>, 2011.
- 1235