University of Reading

A forward model for lidar aerosol attenuated backscatter: application in a megacity

Submitted for the degree of Doctor of Philosophy

Department of Meteorology

Elliott Warren

October 2019

Abstract

Aerosols are extremely important to monitor and predict given their impact on health, radiation and visibility. Chemical transport and numerical weather prediction (NWP) models are becoming increasingly sophisticated, providing detailed forecasts of aerosol dispersion and characteristics. Consequently, the need for model evaluation and data assimilation using appropriately detailed observations is high. Automatic lidar and ceilometers (ALC) observe attenuated backscatter which contains information on aerosols in the absence of hydrometeors. However, a forward operator is needed to directly relate ALC observations to forecast aerosol characteristics.

In this thesis, a flexible and computationally cheap ALC aerosol forward operator (aerFO) is developed to estimate attenuated backscatter (β_m) in clear-sky conditions. Vertical profiles of dry aerosol mass mixing ratio (m) and relative humidity are inputs with default, tuneable assumptions about aerosol constituents. Parameterisations initially estimate physical characteristics, then further estimate optical properties. An extinction enhancement factor ($f(RH)$) is used to represent the change in optical properties due to aerosol swelling. Large components of work involve evaluating the Met Office UK variable resolution regional NWP model (UKV) (1.5 km resolution) in London.

Sensitivity of aerosol characteristics to estimates of aerosol optical properties is explored. Aerosol speciation, total number concentration and dry mean radius are critical and optical properties can vary greatly with wavelength. Evaluation of UKV output over London suggests insufficient mixing in the model can lead to large errors in near-surface β_m and the NWP surface scheme is important in urban areas. Additionally, β_m evaluation is highly dependent on ALC data quality.

Given uncertainties in optimal ALC network design, spatial patterns in β_m variability over Greater London were explored in the UKV and the Met Office research London Model (333 m resolution), then clustered to produce informative maps relevant to instrument placement. Spatial patterns were strongly related to orography, wind and relative emission locations.

Declaration of original authorship

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged:

Elliott Warren

Acknowledgements

Firstly, I would like to thank Sue Grimmond, for the opportunity to carry out this PhD and her unequivocally strong support and dedication to help me improve the standard of my work - willing to help me out no matter the time or place (including the weekend). This PhD has immensely benefited from her guidance for which I will always be grateful. I would like to thank Cristina Charlton-Perez for letting me have a go at improving her forward operator, for her endless help and patience including answering my many, many questions, along with Lee Hawkness-Smith, no matter how simple or complex. I would still be stuck on paper one without her. I am also thankful to Humphrey Lean who urged me to think deeply about the weather models used and how to interpret analysis; our drop-in chats are fond memories of mine. I must thank Simone Kotthaus for teaching me about ceilometers, helping me use the data effectively and still doing her best to be involved as a supervisor despite being a new mom and working in France. I would also like to thank Sue Ballard who sadly passed away before she was able to see the outcome of this work. She was always willing to discuss my early ideas and make me think critically about them, which helped me produce a higher quality of work I hope she could be been proud of.

I want to thank my co-authors Ben Johnson and Franco Marenco from the Met Office, Claire Ryder at University of Reading and David Green at Environmental Research Group in London for their help and support with my work. Their positivity, willingness, eagerness to answer my questions greatly helped me understand aerosols and lidars. I will always be thankful to them for that.

I also want to thank the vast Micromet team and all our scientific visitors, past and present, who have looked after the London Urban Meteorological Observatory, including but not limited to Will Morrison, Kjell Zum Berge, Christoph Kent, Beth Saunders, Meg Stretton, Andy Gabey, Ben Crawford, David Meyer, Natalie Theeuwus, Mathew Lipson, Izzy Capel-timms, Ting Sun and Jessica Brown. You have all been an absolute blast to work alongside.

I must also thank my amazing office mates in 2U08 for the entertaining conversations, banter and food!

I want to thank Luke Warren for his mathematical support and for checking through the thesis even while finishing his own.

Finally, I would like to thank my girlfriend Emma Locking for her continual support throughout my PhD journey, through both the good and the bad. I truly could not have finished the PhD without her.

Contents

List of Figures

List of Tables

Abbreviations

aerFO	Aerosol forward operator
ALC	Automatic lidars and ceilometers
AOD	Aerosol optical depth
APS	Aerodynamic particle sizer
AQUM	Air Quality Unified Model
asl	Above sea level
agl	Above ground level
BCFL	Barbican Frobisher Cresent Lower
BCT	Barbican Cromwell Tower
BGH	Barbican Gilbert House
CA	Cluster analysis
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CBH	Cloud base height
Ch	Chilbolton
CIR	Complex index of refraction
CPC	Condensation particle counter
CTM	Chemistry transport model
DA	Data assimilation
DEFRA	Department for environment, food and rural affairs
DTED	Digital Terrain Elevation Data
DMPS	Differential mobility particle sizer
del.	Deliquescence relative humidity limit
EC	Elemental Carbon
ECMWF	European Centre for Medium-Range Weather Forecasts
EOF	Empirical orthogonal function
ERG	King's College London Environmental Research Group
EU	European Union
eff.	Efflorescence relative humidity limit
FO	Forward operator
FWHM	Full width half maximum
Ha	Harwell
ID	Experiment identifier
IFS	Integrated Forecasting System
IML	Islington Michael Cliffe, Lower Level
IMU	Islington Michael Cliffe, Upper Level
IQR	Inter-quartile range
JULES	Joint UK Land Environment Simulator
KC1	North Kensington (LAQN identifier)
KC2	Kensington and Chelsea – Cromwell Road (LAQN identifer)
KCL	King's College London
KSK	Kings Building, Kings College London (Strand Campus)
KSNW	North Wing Building, Kings College London (Strand Campus)
KSSW	West Roof Strand Building, Kings College London (Strand Campus)
LAQN	London Air Quality Network
LBLRTM	Line-by-line radiative transfer model
LM	London model
LUT	Look up table
LUMO	London Urban Meteorological Observatory
medBE	Median bias error
MO	Met Office

MORUSES	Met Office – Reading Urban Surface Exchange Scheme
MR	Marylebone Road
MY1	Marylebone Road (LAQN identifier)
NAEI	National Atmospheric Emissions Inventory
NK	North Kensington
NPL	National Physics Laboratory
nSD	Normalised standard deviation
NWP	Numerical weather prediction
OC	Organic carbon
PC	Principal component
PCA	Principal component analysis
RGS	Royal Geographical Society
SMPS	Scanning mobility particle sizer
SOCRATES	Suite of community radiative transfer codes
SNR	Signal-to-noise ratio
SR	Sunrise
SS	Sunset
UK	United Kingdom
UKV	UK Variable Resolution model
uEOF	Unrotated eigenvector/ unrotated empirical orthogonal function
WHO	World Health Organisation
WS	Weather station

Notation

B	Activation parameter
Cc	Cunningham slip correction factor
C_i	ith Cluster group
C_{LM}	Cluster group derived using LM data
C_{UKV}	Cluster group derived using UKV data
$C_{ext,aer}$	Particle extinction cross section
$C_{back,aer}$	Particle backscatter cross section
D	Particle diameter
$D_{dry,aer}$	Aerosol species dry particle diameter
D_a	Particle aerodynamic equivalent diameter
D_m	Particle mobility equivalent diameter
D_v	Particle volume equivalent diameter
$D_{wet,aer}$	Aerosol species particle diameter at ambient *RH*
d_g	Geometric standard deviation
e_i	Eigenvector
$f_{ext,rh}$	Extinction enhancement factor
g_{aer}	Aerosol species' particle physical growth factor with respect to water
g_{OC}	Organic carbon physical growth factor with respect to water
j	Number of specified clusters
$M_{abs,aer}$	Particle mass absorption
$M_{abs,wv}$	Water vapour mass absorption
$M_{ext,aer}$	Particle mass extinction
$M_{scat,aer}$	Particle mass scattering
m	Aerosol mass mixing ratio
m_0	'Standard' mass mixing ratio
m_{MURK}	MURK aerosol mass mixing ratio
m_{aer}	Aerosol species mass
N	Total particle number concentration for aerosol mode
N_0	Total particle number concentration
N_{aer}	Aerosol species' dry number distribution
N_{obs}	Non-speciated number distribution
N_{tot}	Total number of particles across all size bins
n	Number of particles in bin
n_{aer}	Aerosol species refractive index
n_i	Partial refractive index of the bulk aerosol
n_{MURK}	Refractive index for the bulk aerosol
n_{water}	Refractive index of water
$n_{wet,aer}$	Aerosol species mixed partial refractive index with water
PM_{10}	Total mass of particles with diameter less than 10 μm
p	Scaling power
p_{air}	Air pressure
q	Specific humidity
Q_{ext}	Particle extinction efficiency
$Q_{ext,dry}$	Dry particle extinction efficiency
Q_H	Sensible heat flux
$Q_{H,m}$	Modelled sensible heat flux

$Q_{H,o}$	Observed sensible heat flux
R_s	Spearman correlation coefficient
R_P	Pearson correlation coefficient
RH	Relative Humidity
r_0	'Standard' mass particle radius
r_g	Geometric mass mean radius
r_{md}	Dry mean particle radius
r_v	Water vapour mixing ratio
S	Lidar ratio
SD	Standard deviation
s	Parameterised wet/dry hygroscopic growth state
T	Two-way transmission
T_{air}	Air temperature
V	Volume
V_{aer}	Aerosol species volume
V_{tot}	Total aerosol volume
$V_{weight,aer}$	Aerosol species volume weighting
v_i	Volume fractional contribution
X_{aer}	Aerosol species' size parameter
x	Original dataset used to derive eigenvalues and eigenvectors
Z	forecast start validity time
z	Height
$\propto$	Proportional to
β	Backscatter
$\beta_{m,unatt}$	Forward modelled unattenuated backscatter
β_m	Forward modelled attenuated backscatter
β_o	Observed attenuated backscatter
ρ_{air}	Air density
ρ_{aer}	Aerosol species density
η	Proportionality constant
λ	Wavelength
λ_i	Eigenvalue
ρ_{air}	Air density
σ_{back}	Particle backscatter coefficient
$\sigma_{back,aer}$	Aerosol species' particle backscatter coefficient
σ_{ext}	Extinction coefficient
$\sigma_{ext,aer}$	Particle extinction coefficient
$\sigma_{ext,wv}$	Water vapour extinction coefficient
τ	Optical depth
χ	Dynamic shape factor

Chapter 1 Introduction

1.1 Motivation

Aerosols impact public health (AQEG, 2005; Atkinson et al., 2014; W. Chen et al., 2015; Dockery and Pope III, 1994; Elliot et al., 2016; Koenig, 2000; Mauderly and Chow, 2008; Shiraiwa et al., 2017; Yan et al., 2019; section 2.1; section 3.1) and meteorological variables such as radiation (Kuniyal and Guleria, 2019; Myhre et al., 2013; Roessler and Faxvog, 1981; Seinfeld and Pandis, 2016; Stier et al., 2007; Thornhill et al., 2018; section 2.1; section 3.1) and visibility (Haywood et al., 2008; Sabbagh-Kupelwieser et al., 2010; Stull, 1988). Therefore, it is critical to observe, monitor and predict them accurately. Aerosol emission sources and sinks, and their impacts can be highly spatially and temporally variable at local to global scales. A wide range of natural and anthropogenic (human) related emission sources exist (section 2.1; section 3.1). Natural sources include mineral dust (Liu et al., 2017; Papayannis et al., 2005; Ryder et al., 2018), ocean wave breaking (Clarke et al., 2006; Gantt et al., 2015) and vegetation (Jim Haywood et al., 2003; Marenco et al., 2016; Seinfeld and Pandis, 2016; Thornhill et al., 2018). Secondary aerosols may be produced through physical and chemical interactions between combinations of particles and gases (Knipping and Dabdub, 2003; Tang et al., 2017). Most aerosol particles are within the planetary boundary layer, the lowest layer of the atmosphere that is directly influenced by the Earth's surface (Wallace and Hobbs, 2006).

Globally, 55 % of the population live in urban areas, and this is projected to be 68 % by 2050 (United Nations, 2018). Urbanisation changes local land-use and vegetation, modifying both local and regional meteorological processes (Oke, 1978; Oke et al., 2017), which can impact the spatio-temporal distribution of aerosol emissions. Typically, as urban areas grow and develop, their aerosol emissions increase to be greater than their surrounding areas (Pandis et al., 2016) with additional anthropogenic sources including vehicles (Beevers et al., 2009; Brito et al., 2018), industry (Oke et al., 2017; D. Zhao et al., 2018) and cooking (Abdullahi et al., 2013; Cheng et al., 2016). Although many aerosol sources and sinks can be local (section 2.5.1), aerosols emitted elsewhere can advect into urban areas (Redington et al., 2016). Given the large exposure to urban citizens, it is critical to understand how aerosols vary in cities.

Numerical weather prediction (NWP) and chemistry transport models (CTM) that forecast aerosols require evaluation against observations (section 2.1; section 3.1). Although CTM have more complete atmospheric chemistry than NWP models, they require NWP forcing (e.g. via coupling) and are computationally more expensive to run (Clark et al., 2008). Whereas in NWP models, aerosols (if included) are typically treated more simply, but can still be sufficient for air quality forecasting (e.g. Met Office (MO) Air Quality Unified Model, (AQUM), Savage et al., 2013). As the critical aerosol characteristics may differ with application (e.g. health vs radiation), it is important to

understand parameterisation uncertainties due to aerosol characteristic inter-dependencies (e.g. optical characteristics dependence on physical aerosol properties).

Many NWP and CTM aerosol forecast evaluations use in-situ observations that are ground-based (Angelini et al., 2009; Burton et al., 2012; Hood et al., 2018; Noh, 2014; Chapter 4.1), or aircraft mounted (Alvarado et al., 2016; J Haywood et al., 2003; Li et al., 2015; Thornhill et al., 2018). Additionally, evaluations can use satellite remote sensing observations (Bender et al., 2019; Chudnovsky et al., 2013; Jose et al., 2016; Palacios-Penã et al., 2019; You et al., 2016; section 2.1, section 3.1, section 4.1) . However, as most sensors have spatial and temporal coverage and/or cost limitations, obtaining additional observations of aerosol characteristics are always of interest.

Multi-wavelength Raman lidars can provide multiple aerosol optical properties. These include the particle extinction coefficient (σ_{ext}, ability of a particle to absorb and scatter light [m^{-1}]), attenuated backscatter (β_o, light reflected directly back from particles to the lidar at 180° angle following some extinction of the emitted lidar pulse from particles below [m^{-1} sr^{-1}]), and particle extinction-to-backscatter ratio (lidar ratio, S) (Dionisi et al., 2018; Müller et al., 2007; Papayannis et al., 2005; Povey et al., 2014; Wang et al., 2016; section 2.1; section 3.1). Other useful Raman lidar measurements include profiles of air temperature (Arshinov et al., 1983; He et al., 2019) and water vapour mixing ratio (Goldsmith et al., 1994; Melfi and Whiteman, 1985). However, Raman lidars are expensive and therefore sparsely distributed. Automatic lidars and ceilometers (ALC), originally designed to measure cloud base heights (Emeis et al., 2012; Van Tricht et al., 2014), have been used to explore aerosol variations (section 2.1, section 3.1, section 4.1). Recently the ALC, with various central wavelengths (Campbell Scientific, 2013; Lufft, 2016; Vaisala, 2017, 2006) and improved ability to measure β_o (Heese et al., 2010), have become both comparable to more sophisticated lidars (Madonna et al., 2018, 2014) and more widely used. Applications include: to explore boundary layer height using aerosol as a tracer (Kotthaus and Grimmond, 2018a; Tang et al., 2017; Wagner and Schäfer, 2015), relations between β_o and particle mass (Münkel et al., 2004; You et al., 2016) and sea breezes (Lemonsu et al., 2006; Uzan et al., 2016; Zéphoris et al., 2005).

Ceilometers are single wavelength elastic-backscatter lidars, which can only emit and measure reflected light at the same fixed wavelength. Therefore, to obtain additional aerosol optical properties such as σ_{ext}, transmission or the aerosol optical depth, S needs to be known as a priori. However, S can vary considerably in space and time as it is a complex function of multiple variables (e.g. particle size distribution, relative humidity and complex index of refraction of the particles and water) (section 3.2.2), that themselves depend on aerosol sources, sinks and advection. Given this complexity, S is often derived more generally for regional types (e.g. urban, continental or maritime) rather than for specific locations (Müller et al., 2007; section 3.1). S is often derived from Raman lidars (Müller et al., 2007; Papayannis et al., 2005; Wang et al., 2016), high spectral resolution lidars (Burton et al., 2012; Hair et al., 2008; Reid et al., 2017), or sunphotometers with additional aerosol

measurements or modelling (Cattrall et al., 2005; H. Zhao et al., 2018). Although, the number of optical instruments are limited, especially in urban areas, because of cost and siting constraints (e.g. appropriate communication, power and security, (Muller et al., 2013; section 3.1), in-situ aerosol observations are increasing (section 3.1). As S is a function of aerosol physical and optical properties, the additional aerosol observations could be used to obtain S.

Strictly, ALC β_o is an optical property of the atmospheric column above the instrument, measured at the wavelength of the sensor, and not a measurement of aerosol physical characteristics (e.g. number concentration). To relate β_o to aerosol physical characteristics a model is needed (section 2.1; section 3.1; section 4.1). Lidar aerosol forward operators (FO) estimate attenuated backscatter and require variables (e.g. mass of different aerosol species) from NWP models or CTM as inputs. The results can be evaluated using β_o or assimilated into models. Several lidar aerosol FO exist for specific numerical models, ALC sensors or aerosol types (Benedetti et al., 2009; Chan et al., 2018; Geisinger et al., 2017; section 2.1; section 3.1; section 4.1). Given the large number of different ALC sensor types within and between networks, a FO that is usable across any wavelength for multiple aerosol characteristics would provide network consistency for both instruments and models.

As the number of ALC and networks have increased (Flentje et al., 2010; Guerrero-Rascado et al., 2016; Illingworth et al., 2015, 2007; Kotthaus and Grimmond, 2018b; Nishizawa et al., 2016; Osborne et al., 2018; Pappalardo et al., 2014; Welton et al., 2000; section 2.5.1; section 3.2.1), there is a large amount of data being gathered at different scales that could provide insight into aerosol variations. In urban areas, there are few observations of the upper 90 % of the boundary layer (Barlow, 2014), partially due to logistical constraints (e.g. siting). ALC networks could help fill this gap, but network design depends on the scales of meteorological processes to be observed. With better knowledge of the spatio-temporal scales and structure of meteorological features that occur in an area, we can inform instrument deployment.

1.2 Objectives

In this research a computationally cheap and flexible aerosol forward model is developed for ALC attenuated backscatter, to evaluate model output. Although the developed FO is usable in any area (e.g. rural) with any appropriate numerical model (e.g. CTM), the work focuses on evaluating NWP output in an urban area (London).

The objectives of this thesis are to (addressed in the indicated chapters):

I. Design an aerosol forward operator (aerFO) to estimate lidar attenuated backscatter from aerosol content and relative humidity, with a focus on minimising computational expense and maximising forward model input flexibility (Chapter 2, Chapter 3).

II. Quantify the relative importance of different aerosol characteristics in accurately estimating bulk aerosol physical and optical properties (Chapter 2, Chapter 3).

III. Evaluate aerFO modelled attenuated backscatter from NWP in London using ALC attenuated backscatter observations (Chapter 2, Chapter 3).

IV. Analyse the spatial variability in aerFO modelled attenuated backscatter across Greater London informed by two NWP models of different resolutions (Chapter 4).

V. Develop an ALC network design method that considers the spatio-temporal variability of aerFO modelled backscatter from NWP (Chapter 4).

1.3 Thesis structure

Background literature to this thesis is presented in the individual chapters as indicated in Table 1.1.

A detailed description of the aerFO developed (Chapter 2[1]) includes a sensitivity analysis of the impact of different aerosol characteristics to estimating optical properties. The aerFO modelled attenuated backscatter (β_m) uses the MO UKV (1.5 km) NWP model and is evaluated in London. This includes evaluating the impact of the NWP urban surface schemes.

A method to estimate the particle extinction-to-backscatter ratio S from aerosol observations is provided (Chapter 3[2]). From S, other optical properties can be derived. Additionally, the relative importance of different aerosol characteristics to accurately forward model β_m in an urban area are also explored, to inform MO NWP aerosol parameterisation development (Chapter 3[2]). An updated version of the aerFO is used (section 3.2.3).

The spatio-temporal variability of β_m from two NWP models at different resolutions (1.5 km and 333 m) are analysed using principal component analysis (PCA) and cluster analysis (CA) (Chapter 4[3]).

---

[1] Chapter 2: **Warren, E.**, Charlton-Perez, C., Kotthaus, S., Lean, H., Ballard, S., Hopkin, E., Grimmond, S., 2018. Evaluation of forward-modelled attenuated backscatter using an urban ceilometer network in London under clear-sky conditions. Atmos. Environ. 191, 532–547. https://doi.org/10.1016/j.atmosenv.2018.04.045

[2] Chapter 3: **Warren, E.**, Charlton-Perez, C., Kotthaus, S., Marenco, F., Ryder, C., Johnson, B., Lean, H., Ballard, S., Grimmond, S., 2019. Observed aerosol characteristics to improve forward-modelled attenuated backscatter in urban areas. Atmos. Environ. 117177. https://doi.org/10.1016/j.atmosenv.2019.117177

[3] Chapter 4: **Warren, E.**, Charlton-Perez, C., Kotthaus, S., Lean, H., Grimmond, S., n.d. Spatial variability of forward modelled attenuated backscatter in clear-sky conditions over a megacity: Implications for observation network design. *In preparation*

The PCA-CA provides a method to consider ALC sensor network design. The analysis is undertaken for Greater London. The aerFO used is the same as in Chapter 3 (section 3.2.3)

The contributions, conclusions, limitations and recommendation for future work are presented in Chapter 5.

Table 1.1: Background literature by topic is presented in individual chapters in the sections indicated.

Topic	Chapter		
	2	**3**	**4**
	Chapter sub-section		
Aerosols and their characteristics			
Aerosol impacts	1	1	
Aerosol sources	1	1	
Aerosol physical properties	2		
Aerosol hygroscopic growth and hysteresis		2.2	
Shape factor		2.1	
Aerosol optical properties			
Extinction enhancement factor	2		
Dependence of aerosol optical properties on wavelength	2		
Dependence of aerosol optical properties on aerosol species	2		
Aerosol extinction-to-backscatter ratio (lidar ratio)	2	1	
Deriving or estimating the lidar ratio		1, 2.2	
Observing aerosols			
In-situ measurements of aerosols	1	1	
Diameter conversions		2.1	
Remote sensing aerosols	1		
Automatic lidar and ceilometers			
Observing meteorological features	1	1	1
Remote sensing of aerosols	1	1	1
Automatic lidar and ceilometer networks	1		1
Data assimilation of automatic lidar and ceilometer data	1		
Numerical weather prediction and chemistry transport models			
Representation of aerosol	1	1	
Met Office MURK visibility scheme	1, 3.3, Appendix A	1	
SOCRATES radiation suite	2		
Met Office CLASSIC aerosol scheme		2.1	
JULES urban land surface schemes	3.3		
Aerosol forward operators for automatic lidar and ceilometers			
Existing aerosol forward operators	1	1	1
The aerosol forward operator (aerFO)	2	2.3	2.2
Data and observations			
London Urban Meteorological Observatory (LUMO)	3.1	2.1	
Ceilometer network data and processing	3.1	2.1	
Ceilometer firmware and hardware	3.1		
Automatic weather station data	3.2	2.1	
Aerosol measurements in London			
London Air Quality Network (LAQN)	3.2		
National Physics Laboratory (NPL)		2.1	

Department for Environment, Food & Rural Affairs (DEFRA)		2.1	
Met Office UKV (1.5 km) data	3.3		2.1
Met Office London Model (333 m) data			2.1
Observation networks			
Optimising observation networks			1
Methods			
Principal component analysis			1, 2.3
VARIMAX rotation			2.3
Cluster analysis			2.4

1.4 Data used in this thesis

Central to this thesis are data from Met Office UK regional forecast NWP model (UKV, Tang et al., 2013), a configuration of the Met Office Unified Model (Davies et al., 2005) (section 2.5.1; section 4.2.1). This model has a 1.5 km inner domain covering the UK and a coarser variable resolution outer domain (1.5 or 4 km; see Figure 1 and 2 of Tang et al., 2013). It has 70 vertical model levels and extends to 40 km (Met Office, 2019). The UKV is also deterministic, providing a singular forecast of the atmospheric state without further information on forecast uncertainty. Additionally, the UKV is 'convection-permitting' and can partially represent convection through model physics, though this means the UKV still relies on parameterisations which are more simplified mathematical representations of reality. Data used from the UKV are from lowest level to ~3 km for the London (Chapter 2; Chapter 4).

Additionally, data from the Met Office London Model (LM) (Boutle et al., 2016) with a higher horizontal resolution (333 m) are used (Chapter 5). This research model has a domain of 100 × 66 km^2 focussed on Greater London (Boutle et al., 2016) (section 4.2.1). The LM is one-way nested in the UKV via 11 prognostic variables, so the LM boundary conditions at the edge of the domain are updated by the UKV every 15 mins. Like the UKV, the LM has 70 vertical levels extending to 40 km.

Both the UKV and LM have the Joint UK Land Environment Simulator (JULES) coupled as the land surface model (Best et al., 2011; Walters et al., 2017). Prior to 16 March 2016, the UKV-JULES urban energy surface balance scheme was the 1-tile Best scheme (Best, 1998; Best et al., 2011, 2006; section 2.3.3; section 4.2.1). Subsequently, the newer MORUSES urban energy surface scheme is used with London parameters, which can take into account both roofs and street canyons (Bohnenstengel et al., 2014, 2011; section 2.3.3; section 4.2.1). Hence, the archived UKV data used varies by date, but impacts are considered briefly (section 2.7.1). The LM case studies analysed all use the Best scheme, as the LM-MORUSES archive of forecast data is still minimal.

The aerosol emission ancillary in the UKV and LM is derived from the 1 km resolution National Atmospheric Emissions Inventory (NAEI; Neal, 2019, 2016). It is coarsened for use in the UKV and

interpolated for the LM and was updated on 3 February 2015 to use the most recent data. The emissions are given as total dry mass of aerosol, and this aerosol is primarily used by the MURK visibility scheme as a proxy (Clark et al., 2008; Claxton, 2013; section 2.1; Appendix 2.A).

The other major source of data used in the thesis are from observations. Vertical profiles of modelled attenuated backscatter (β_m) derived from the aerFO (section 2.4) are evaluated (section 2.7, section 3.3.2) against observed attenuated backscatter (β_o) measured by Vaisala CL31 ceilometers from the London Urban Meteorological Network (LUMO; Kotthaus and Grimmond, 2014). Two generations of hardware are used in the LUMO CL31 instruments with two different firmware versions (Kotthaus et al., 2016; section 2.5.1). β_o is processed and corrected for the near-range and instrument related background before a centred moving average is applied with a time window of 25 mins (101 time steps) and range window of 110 m (11 range gates), in order to improve the signal to noise (Kotthaus et al., 2016; section 2.5.1). β_o is then calibrated using daily calibration coefficients derived using the 'cloud' method (Hopkin et al., 2019; O'Connor et al., 2004; section 2.5.1). In addition, mixing layer heights are derived from β_o using the CABLAB algorithm (Kotthaus and Grimmond, 2018b) to help identify the boundary layer (section 4.2.1) and support evaluation of β_m (section 2.7.3).

Aerosol observations support aerFO β_m estimation and evaluation. This includes King's College London Environmental Research Group (ERG) London Air Quality Network (LAQN; Mittal et al., 2016) observations of particulate matter (total mass of dry aerosol below 10 μm, PM_{10}) used as a proxy for the total dry mass of aerosol (section 2.7.2). ERG observations of number distribution and mass of aerosol species are used to derive urban climatological values and monthly aerosol speciation for aerFO (section 3.2). This required the number distribution from a size mobility particle sizer (SMPS, ~0.016 – 0.6 μm) and an aerodynamic particle sizer (APS, ~0.5 – 20.0 μm) (Table 2.1) to be combined into one distribution with common size units (section 3.2.1; Table 2.1).

Additional meteorological observations from LUMO also supported β_m evaluation. *RH* observations from automatic weather stations support near-surface β_m evaluation (section 2.5.2). Further, *RH* measurements are used with aerosol observations to parameterise aerosol hygroscopic growth (physical change in size due to the absorption of water vapor) to estimate *S* and β_m (section 3.2.1; section 3.2.2). Sensible heat fluxes derived using eddy covariance techniques were used to aid evaluation of the NWP urban land surface scheme (Kotthaus and Grimmond, 2014; section 2.5.2).

Chapter 2 Evaluation of forward-modelled attenuated backscatter using an urban ceilometer network in London under clear-sky conditions

Abstract

Numerical weather prediction (NWP) of urban aerosols is increasingly sophisticated and accurate. In the absence of large particles (e.g. rain, cloud droplets), information on atmospheric aerosols can be obtained from single wavelength automatic lidars and ceilometers (ALC) that measure vertical profiles of attenuated backscatter (β_o). To assess the suitability of ALC profile observations for forecast evaluation and data assimilation, a forward operator is required to convert model variables into the measured quantity. Here, an aerosol forward operator (aerFO) is developed and tested with Met Office NWP data (UKV 1.5 km) to obtain synthetic attenuated backscatter profiles (β_m). To compute β_m, aerFO requires input of the bulk aerosol mass mixing ratio and relative humidity, together with air temperature and pressure to calculate the effect of water vapour absorption. Bulk aerosol characteristics (e.g. mean radius and number concentration) are used to estimate optical properties. ALC profile observations in London are used to assess the accuracy of β_m. A wavelength-dependent extinction enhancement factor accounts for the change in optical properties due to aerosol swelling. Sensitivity studies show the aerFO unattenuated backscatter is very sensitive to the aerosol mass and relative humidity above ~60-80 %. The extinction efficiency is sensitive to the choice of aerosol constituents and to ALC wavelength. Given the aerosol is a tracer for boundary layer dynamics, application of the aerFO has proven very useful to evaluate the performance of urban surface parameterisation schemes and their ability to drive growth of the mixing layer. The implications of changing the urban surface scheme within the UKV are explored using two spring cases. For the original scheme, morning β_m is too high, likely a result of delayed vertical mixing. The new scheme reduced this persistence of high morning β_m, demonstrating the importance of surface heating processes. Analysis of profiles at five sites on 12 clear-sky days shows a positive, statistically significant relation between the differences of modelled and measured near-surface attenuated backscatter [β_m - β_o] and near-surface aerosol mass. This suggests that errors in near-surface attenuated backscatter can be attributed to errors in the amount of aerosol estimated by the NWP scheme. Furthermore, this correlation increases when cases of high relative humidity in the NWP model are excluded. Given the impact on aerosol optical properties demonstrated, results suggest that the use of a fixed, bulk aerosol for urban areas in the UKV should be revisited and the lidar ratio should be constrained. As the quality of the observed attenuated backscatter is shown to be critical for performing model evaluation, careful sensor operation and data processing is vital to avoid drawing false conclusions about model performance.

2.1 Introduction

Urban aerosols impact health (Atkinson et al., 2014; C. H. Chen et al., 2015; Elliot et al., 2016; Xing et al., 2016; Yin et al., 2017), visibility (Sabbagh-Kupelwieser et al., 2010; Stull, 1988) and local radiation budgets (Boucher et al., 2013; Haywood and Boucher, 2000; Marley and Gaffney, 2006), and with a high density of people in urban areas, many may be impacted. Most countries, such as EU member states (EEA, 2015), have legal maxima set for several air quality constituents. Identifying when behavioural changes are needed (e.g. reduction of car usage) in advance of critical exceedances requires improved understanding, and accurate prediction of, aerosol concentrations in urban areas.

Aerosol concentration and transport can be predicted with atmospheric chemistry models that are *one-* or *two-way* coupled to a numerical weather prediction (NWP) model (Benedetti et al., 2009; Grell et al., 2004; Appendix 2.A). As including an atmospheric chemistry model is significantly more computationally expensive than standalone NWP (Clark et al., 2008), this reduces their operational utility. However, NWP resolution has increased and aerosols have been included more explicitly (Benedetti et al., 2009; Lu et al., 2016). Some contain information on various aerosol species, for instance the AQUM configuration of the Met Office (MO) Unified Model (Savage et al., 2013) and the Integrated Forecasting System (IFS) at the European Centre for Medium-Range Weather Forecasts (ECMWF). These include mass mixing ratios for different aerosol species such as salt, dust, organic carbon, black carbon, sulphates and nitrates for use in radiation parameterisations (Benedetti et al., 2009; Morcrette et al., 2008), which makes them more computationally expensive. Another example is the MO operational variable resolution (UKV) NWP model (1.5 km inner, 4 km outer domain) (Tang et al. 2013) that has an aerosol proxy (Clark et al., 2008) with a single mass mixing ratio (m_{MURK}, kg kg^{-1}) derived from emission inventories (Neal, 2019) (Appendix 2.A). This aerosol scheme, called MURK, is designed to be computationally inexpensive for use in a visibility parameterisation and data assimilation scheme (Claxton, 2013).

Vertical profiles of atmospheric aerosols are monitored with remote sensing instruments on satellites (Chudnovsky et al., 2013; Goto et al., 2011; Jose et al., 2016; You et al., 2016), aircraft (Li et al., 2015), and the ground (Wiegner et al., 2014). While satellite observations provide extensive spatial coverage, their temporal resolution is limited. When the near surface atmosphere is of interest, ground-based profilers are useful as the signal is less likely to be saturated by clouds. Automatic lidars and ceilometers (ALC) are single-wavelength lidars, designed to provide cloud base height (CBH; Emeis, 2011) using the strong scattering properties of cloud droplets at ALC wavelengths (Martucci et al., 2010; Sharma et al., 2016; Willén et al., 2005). With increased sensitivity, it is possible to derive information about smaller aerosols (Wiegner et al., 2014). The capability of measuring attenuated backscatter (β_o) profiles is being exploited, for example, to track volcanic ash clouds (Flentje et al., 2010a), or Saharan dust (Jin et al., 2015), in the expanding ALC networks (Flentje et al., 2010b; Pappalardo et al., 2014; Sugimoto et al., 2008). Urban aerosol characteristics

such as $PM_{2.5}$ and PM_{10} concentrations (Münkel et al., 2007; You et al., 2016), aerosol layers (McKendry et al., 2009; Zéphoris et al., 2005) and mixed-layer height (Poltera et al., 2017; Schafer et al., 2011; Schäfer et al., 2012; Wagner and Schäfer, 2015) have previously been explored using ALC attenuated backscatter measurements.

Data assimilation (DA) in NWP has used CBH data (Francis, 2012; Janisková et al., 2002; Lopez, 2002), and more recently attenuated backscatter profiles (Benedetti and Dabas, 2016; Charlton-Perez et al., 2016; Janisková et al., 2010; Janisková and Stiller, 2010). As specific aerosol properties such as radius and number concentration distributions are not observed by ALC, a forward operator (FO) is required to enable quantitative comparison to modelled values. An FO aims to estimate what an instrument would measure, given a set of inputs representing an atmospheric state. This can be done using physical or empirical relationships between variables, or a combination of both. Several FOs for atmospheric-chemistry and NWP model evaluation exist. For example, the Geisinger et al. (2017) FO estimates attenuated backscatter (β_m) from volcanic ash clouds in the free troposphere using COSMO-ART over Germany and ALC data, but omits the effect of backscatter from boundary layer aerosols. Another FO estimates backscatter from dust and sea salt with the ECMWF IFS using CALIOP data from the CALIPSO satellite (Table 1 in Morcrette et al., 2009). The Chan et al. (2016) FO estimates attenuated backscatter to evaluate against Lufft CHM15K ALC measurements, and has a prognostic mass mixing ratio of different aerosol species from NWP. This FO also utilises either Mie or T-matrix calculations with fixed size distributions and growth factors for a set number of species. The Charlton-Perez et al. (2015) FO, used to estimate vertical profiles of β_m, accounts for the effects of aerosols (using MURK; Clark et al. 2008), liquid cloud and rain. Initial testing at rural sites suggests this FO produced realistic β_m compared to β_o observed with Vaisala CL31 (Charlton-Perez et al., 2015, 2016). However, this FO is effectively wavelength-independent, with a fixed scattering efficiency of 2, a value typically not suitable for aerosols. Although the FO only requires the total bulk mass mixing ratio as aerosol input, it does not speciate the aerosol, so that physical growth and scattering properties are assumed to be invariant between particles. The latter is a critical assumption, since scattering properties vary significantly between different aerosol species (Seinfeld and Pandis, 2016).

This chapter has two main parts. The first part focuses on the initial analysis of ALC observations and the aerosol component of the Charlton-Perez et al. (2015) FO (hereafter aerFO). The aims are to (i) demonstrate that observed ALC attenuated backscatter adequately reflects changes in atmospheric properties using two case studies, and (ii) demonstrate the aerFO can be used to evaluate the aerosol mass mixing ratio in NWP using ALC measurements, and explore meteorological phenomena in urban areas. Aim (ii) will use a sea breeze case study, examining the UKV and a 100 m experimental research model, in a city (London, UK). Part one is in sections 2.2 and 2.3.

Following the initial assessment, the aerFO was extended (section 2.4) to include aerosol speciation, wavelength dependency and water vapour absorption. The second part of this chapter then aims to (iii) further demonstrate the ability to evaluate the aerosol mass mixing ratio in NWP using ALC measurements, and (iv) comment on the potential application of ALC-measured attenuated backscatter profiles in DA. In the more thorough assessment, a sensitivity study is undertaken to assess the impact of several assumptions in the aerFO (section 2.5). Subsequently, aerFO is used to evaluate m_{MURK} and relative humidity (RH) from the UKV in London, based on a comparison of two spring day cases (section 4) and a composite of 11 cloud-free days (section 2.7).

It is concluded (section 2.8) that the detailed modelling approach provides a valuable tool for future studies in real city settings.

2.2 Suitability of Vaisala CL31 observations for verification

2.2.1 Measurement error and inter-comparison

CL31 ceilometers are used in Chapter 2 and 3 for verification. Although a fair quantitative comparison cannot be made between the CL31 ceilometers and lidar backscatter due to the differences in wavelength of the instruments (CL31: 905 nm; lidars including Raman: 355, 532, 1064 nm), CL31 β_o is found to be qualitatively comparable to lidar measurements when examining aerosol layers and boundary layer structures (Madonna et al., 2014; McKendry et al., 2009; Schoenfelder et al., 2009; Wiegner et al., 2019). Nonetheless, there are several sources of error in ceilometer measurements that need correction before justifying ceilometer use in quantitatively verifying β_m.

To reduce instrument error, each CL31 is factory calibrated before shipment due to its sensitivity to the sensor optics and electronics. Incomplete optical overlap (range gates below 70 m) is also corrected for using internal software. Window transmission (cleanliness [%]) is reported by the instrument to inform when cleaning is necessary and to minimise unnecessary beam attenuation. CL31 ceilometers also operate at wavelengths which can be affected by water vapour absorption (905 nm), with a potential error of ~ 20 % in mid latitudes (Wiegner and Gasteiger, 2015). However, this can be compensated through observation processing (Wiegner et al., 2019; Wiegner and Gasteiger, 2015), or incorporated into forward modelled estimates (section 2.4).

The CL31 observations are also processed to further reduce the error. Four LUMO CL31 ceilometers used in this thesis underwent an inter-comparison (Kotthaus et al., 2016), following an international effort to standardise ceilometer processing and improve data quality (TOPROF, 2020). The inter-comparison included recommendations to CL31 ceilometer processing including a background correction, range correction, optical overlap assessment, near-range correction, absolute (factory) calibration assessment, and a signal-to-noise ratio and an (SNR) assessment which included recommended averaging windows in height and time and SNR thresholds, (complete details in

Kotthaus et al., 2016). All recommendations were made with sensitivity to different software versions and hardware component combinations. The recommendations were since used in the European Aerosol Research Lidar Network (Pappalardo et al., 2014) and implemented in processing of all LUMO CL31 ceilometers. Each set of corrections are specifically tailored to each ceilometer, to minimise their respective error sources.

Prior to analysis in this thesis, another inter-comparison was done to determine whether the processed measurements were reasonable. Each instrument was paired ~5 m away from CL31-A for at least 2 weeks. Qualitative comparisons with meteorological measurements, as well as mean differences and correlations of the vertical profiles of β_o were made to determine that the same meteorological features were being measured appropriately. Diagnostic information was also assessed to ensure the measurements were not grossly erroneous and the instrument was functional, including checks of window cleanliness, laser diode temperature, blower condition and tilt angle.

Following the absolute calibration, measurement processing and field inter-comparison checks, the instruments are self-calibrated using the 'cloud-method' (see section 2.5.1). The variability in estimated calibration factors using the cloud-method is typically ~ ± 5 % for CL31 instruments (over 2015), and has had a reported uncertainty between 5 – 10 %, depending on the amount of multiple scattering, for a CT75K ceilometer operating at 905 nm (Hopkin et al., 2019; O'Connor et al., 2004).

2.2.2 Qualitative case study assessment

To further determine whether the Vaisala CL31 ceilometer attenuated backscatter measurements were suitable for verification of forward modelled backscatter and whether they vary with varying meteorological conditions, several case studies of uncalibrated observed ceilometer attenuated backscatter from Vaisala CL31 ALC instruments (further description in section 2.5.1) were examined. Uncalibrated attenuated backscatter was used as only qualitative assessment was being done to identify meteorological features. Of the cases examined, two are shown below.

Figure 2-1a shows uncalibrated observed attenuated backscatter for the 9 July 2017 from the Reading University Observatory (URAO), and clearly shows several meteorological features. Between 00:00 and 09:00 the high backscatter near the surface indicates fog formation. This coincides with high levels of *RH* (Figure 2-1c) and results from the ambient aerosol swelling with water vapour, increasing in size and increasing the attenuated backscatter. Additionally, the high attenuated backscatter (~ 10^{-5} m^{-1} sr^{-1}) between 1500 and 2000 m is cloud. Notably during midday, the cloud's effect on the radiation is clear with large variations in observed incoming shortwave radiation (Figure 2-1b). Finally, the attenuated backscatter also reflects the higher amounts of aerosol level (and water vapour) within the boundary layer (~5×10^{-7} m^{-1} sr^{-1}) compared to above it (<1×10^{-7} m^{-1} sr^{-1}), again showing the ability of the attenuated backscatter to distinguish between different meteorological features.

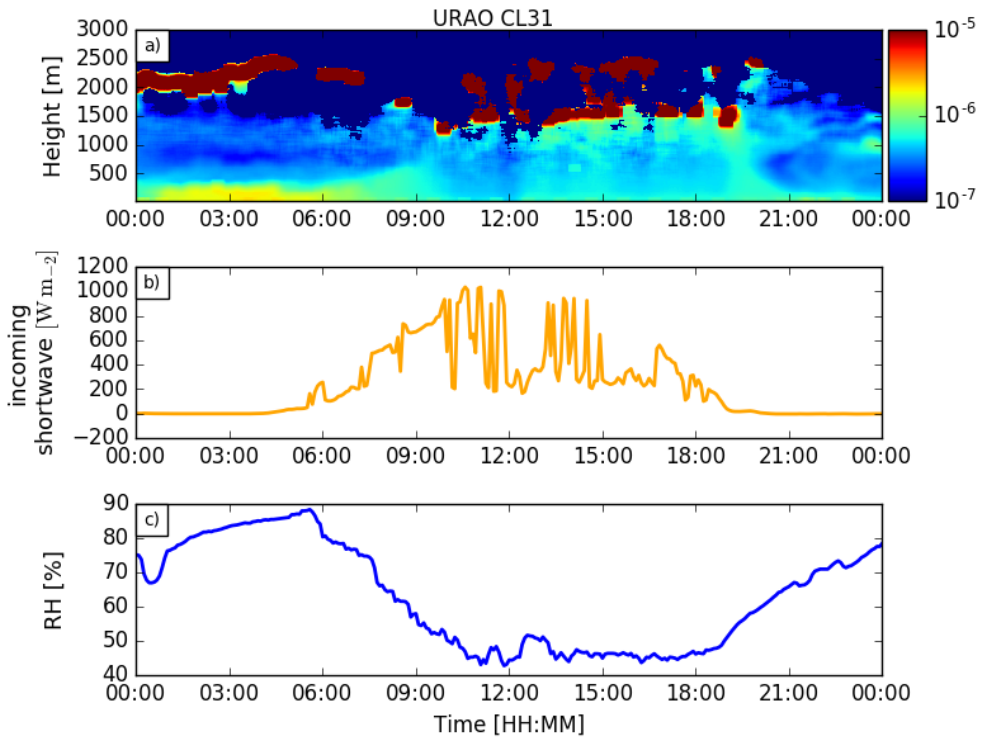

Figure 2-1: a) Vertical profiles of Vaisala CL31 observed, uncalibrated attenuated backscatter [m^{-1} sr^{-1}], b) in-situ incoming shortwave radiation (5 min average) and c) in-situ relative humidity (RH, 5 min average) for 9 July 2017 at University of Reading Observatory (URAO). Height is above ground level [m] (agl). In-situ measurements taken adjacent to CL31 instrument.

Figure 2-2 shows a second case study from Islington Michael Cliffe, Upper Level (IMU) in London, from 13 September 2016, with PM_{10} measurements from the Holloway Road London Air Quality Network (LAQN) site ~4.0 km north of IMU. This case also shows the change in backscatter with respect to varying meteorological phenomena, as well as varying aerosol, with little cloud and no rain. Relatively higher attenuated backscatter at the lowest heights above the instrument (< 50 m) are due to the relatively high aerosol and *RH* in the surface layer compared to the residual layer above. In the surface layer, aerosol increased from ~ 18 µg kg^{-1} at 00:00 to ~ 30 µg kg^{-1} by 12:00 (Figure 2-2b), while *RH* increased up to 90 %. Between 06:00 and 12:00 *RH* reduced from 90 to 50 % but as aerosol was still increasing, near-surface backscatter remained high. Between 12:00 and 18:00 near-surface aerosol decreased from ~ 35 – 40 µg kg^{-1} to ~ 25 µg kg^{-1}. This is due to convection vertically mixing near-surface air with air in the residual layer, corresponding with a reduction in near surface backscatter and increase in backscatter above the surface. After ~ 18:00, aerosol then increased again to ~ 45 µg kg^{-1}, likely due to increased road traffic, and *RH* began to increase, which corresponded with an increase in surface layer backscatter. Finally, above ~1500 m is the free atmosphere where *RH* and aerosol amounts are low, and consequently backscatter was low. However, where the signal to noise ratio is low, the data can be easily identified and removed further analysis.

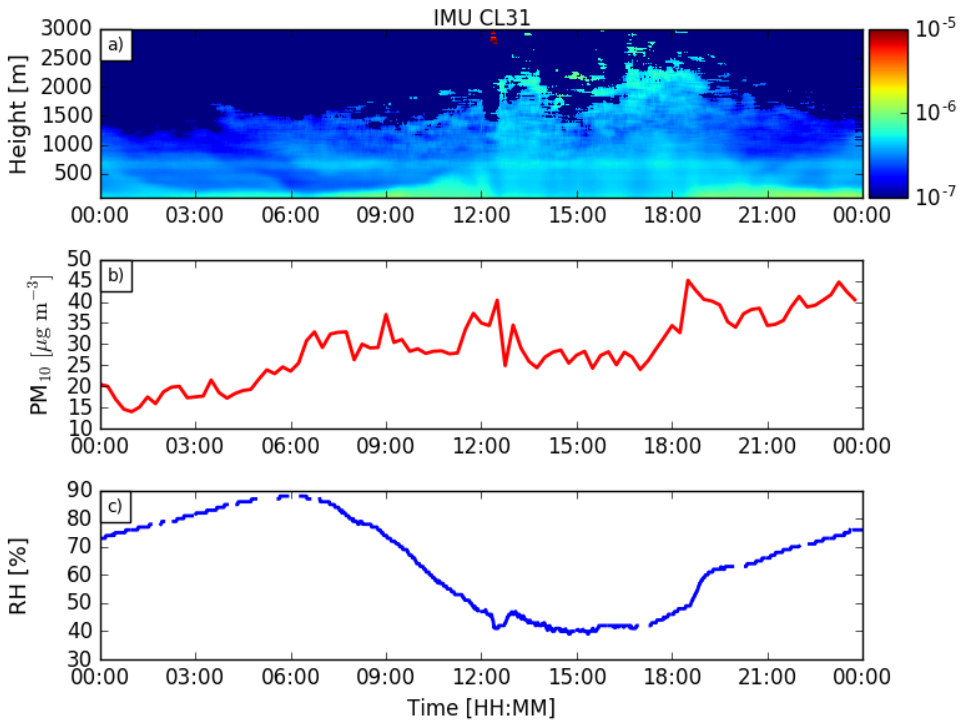

Figure 2-2: a) Vertical profiles of Vaisala CL31 observed, uncalibrated attenuated backscatter [m^{-1} sr^{-1}], b) PM_{10} [$\mu g\ m^{-3}$] (15 min average) and c) relative humidity (RH, 1 min average) for 13 September 2016 at Islington Michael Cliffe, Upper Level (IMU). Height is above ground level [m] (agl). In-situ RH measurements taken at 91 m agl (adjacent to CL31 instrument). Note PM_{10} measurements from Holloway Road London Air Quality Network site, ~4.0 km north of IMU, 3 m agl.

Following the assessment of Vaisala CL31 backscatter to appropriately represent changes in meteorological phenomena and aerosol concentrations, the instruments were deemed suitable for use in evaluating β_m.

2.3 Assessment of the original aerFO

The early version of the aerFO (version 0.1), was a direct copy of the aerosol part of the forward operator presented in Charlton-Perez et al. (2015). To determine whether the aerFO could realistically represent urban meteorological phenomena, it was used to

1) Estimate β_m during a sea breeze event with both the 1.5 km UKV (β_{UKV}) (Davies et al., 2005) and 100 m experimental grid length model (β_{100}) (Boutle et al., 2016). Both estimates were then compared to ceilometer observations of β_o (sections 2.3.2 - 2.3.3).
2) Estimate β_m for a range of clear-sky case studies with two cases presented (section 2.3.4).

The differences between the early aerFO and the more expansive aerFO outlined in section 2.4 (version 1.0), are presented below.

2.3.1 Original aerosol forward operator (aerFO, version 0.1)

aerFO version 0.1 has three differences from version 1.0 (section 2.4):

(1) An analytical solution to the Köhler curve is used to swell the particles (Clark et al., 2008):

$$r_m = \begin{cases} r_{md}, & if\ RH < 38\ \% \\ r_{md}\left(1 - \dfrac{B}{\ln(RH)}\right)^{1/3}, & if\ RH \geq 38\ \% \end{cases} \quad (2.1)$$

where B is the activation parameter and set to 0.5 assuming all aerosol is ammonium sulphate. Particles above a critical RH threshold of 38 % are swollen, consistent with ammonium sulphate aerosol, in accordance with laboratory experiments (Tang and Munkelwitz 1994). If RH is less than this, the final swollen radius (r_m) is equal to the dry radius (r_{md}). r_m is used to estimate σ_{ext}.

$$\sigma_{ext}(z) = \int \pi r_m^2 Q\ N\ dr \quad (2.2)$$

(2) The particle extinction efficiency (Q) is set to 2 and is independent of particle size (Clark et al., 2008).

(3) The role of water vapour absorption in the atmosphere is not considered.

2.3.2 Sea breeze case data

The two configurations of the MO Unified Model use are:

(1) UKV the current MO operational forecast model.
(2) 100 m research model with approximately twice the UKV model levels (Lean et al., 2019).

β_m is compared to β_o from four uncalibrated Vaisala CL31 ceilometers, deployed within the London Urban Meteorological Observatory (LUMO, Table 2.1). Comparison is arbitrarily limited to the lowest 2000 m, as the boundary layer is of primary interest and was determined to be located below this height, given the low quantity of aerosol above this level. To aid comparison, PM_{10} observations from LAQN (King's College London, 2016) and surface RH observations from LUMO are also used.

The two generations of ALC sensors have different transmitters (Table 2.3), with higher noise levels and 'layering/ringing' artefacts in backscatter in 311 sensors but not in 321 sensors (Kotthaus et al., 2016).

For the test case, β_m is calculated for North Kensington (NK) for 24 h (from 10:00 am 25 July 2012), for both models, from a single forecast run (Figure 2-8). During this period, a sea breeze passed

across the observation network with aerosol concentration and *RH* changes. The sea breeze front arrived at NK at 18:00 25 July 2012.

New emission ancillaries for MURK aerosol (3rd February 2015) use more recent emissions data (internal communication with MO, Neal 2016). The aerosol emissions are of the total dry mass mixing ratio of aerosol (m). A MO comparison between the old and newer methods suggest MURK emissions of m calculated using the original method was too high. Therefore, as m was calculated using the old method during the sea breeze case, m is scaled at all heights using 100 % (m_{100}), 50 % (m_{50}) and 20 % (m_{20}) to retain the relative shape of the m profile, and to test whether m was also too high in the forecasts during the case study.

2.3.3 Sea breeze case results

Example profiles (runs 1-3) with m_{100}, m_{50} and m_{20} for MO 100 m (Figure 2-3) show variations of both β_m and β_o backscatter profiles as the sea breeze passes. More aerosol rich, lower humidity air is replaced by less aerosol rich and more humid air from the sea. Rescaling m down (20 or 50% of its original value) reduces the difference between model results and observations, both in terms of absolute difference and shape of the profile. The difference between $\beta_m(m_{100})$ and β_o is approximately 1-2 orders of magnitude higher than β_o, whereas the differences between $\beta_m(m_{20})$ and β_o is approximately 0-1 order of magnitude. The better accuracy of $\beta_m(m_{50})$ and $\beta_m(m_{20})$ supports the decision to rescale the MURK ancillaries down on 3rd February 2015.

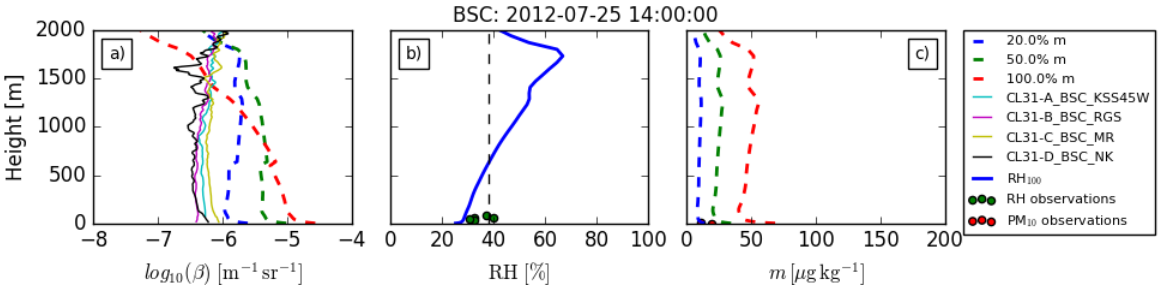

Figure 2-3: Examples of profiles at 14 UTC of (a) attenuated backscatter, (b) relative humidity, and (c) mass mixing ratio. In (a) dashed: model data β_{100} (runs 1-3) using different amounts of m, solid = β_o from ceilometers; (b) line: model and points: observations; (c) dashed: relative m used to model β_{100} and points: PM_{10} surface observations.

Figure 2-4 shows two sets of time-series of Spearman correlation coefficients of vertical profiles of β_o from four CL31 in LUMO (Table 2.3) with vertical profiles of β_{100} (Figure 2-4a) and β_{UKV} (Figure 2-4b) using m_{50}. Before the sea breeze arrival, the highest correlations are with CL31-D at NK (β_m calculated for) and CL31-C at MR, whereas the other two sites had lower and even negative correlations. The lower correlations could be from being more geographically distant or because CL31-A and CL31-B are older generation CL31 (Table 2.3) with a 'layering' artefact, reducing the

correlations. The spike and dip in correlations match the arrival of the sea breeze front and are likely explained by the different times it passes each instrument.

After the sea breeze arrival, correlations between β_{100} and the different ceilometers are often within 0.1 and relatively more positive (0.2 – 0.7). This reflects the more spatially homogeneous maritime air being advected horizontally, replacing the more heterogeneous urban air. The reduction between 22:00 and 07:00 UTC may reflect the surface air mass mixture in the model becoming more dissimilar to reality, as m increased in the surface layer at a greater rate in the model than PM_{10} did in reality (not shown).

The correlation trend between β_{UKV} and β_o is similar to β_{100} and β_o before the sea breeze arrival, but different afterwards. After the sea breeze front passes the β_{UKV} correlations are initially around 0, ~0.4 lower than between β_{100} and β_o, although they steadily rise to ~0.6 between 20:00 and 10:00 UTC. This suggests β_{UKV} did not capture the different layers of 'dirty' and 'cleaner' air after the front as well as β_{100}, alluding to a possible benefit of the higher vertical resolution of the 100 m model; cf. UKV. The steady rise in correlation afterward between β_{UKV} and β_o might also be because m_{50} increases at a slower rate in the UKV compared to the 100 m model, and more similarly to PM_{10} observations (not shown), leading to a more similar surface layer mixture of aerosols. However, the high surface emission may be because the aerosol is dispersed too shallowly in the 100 m model due to the relatively higher vertical resolution.

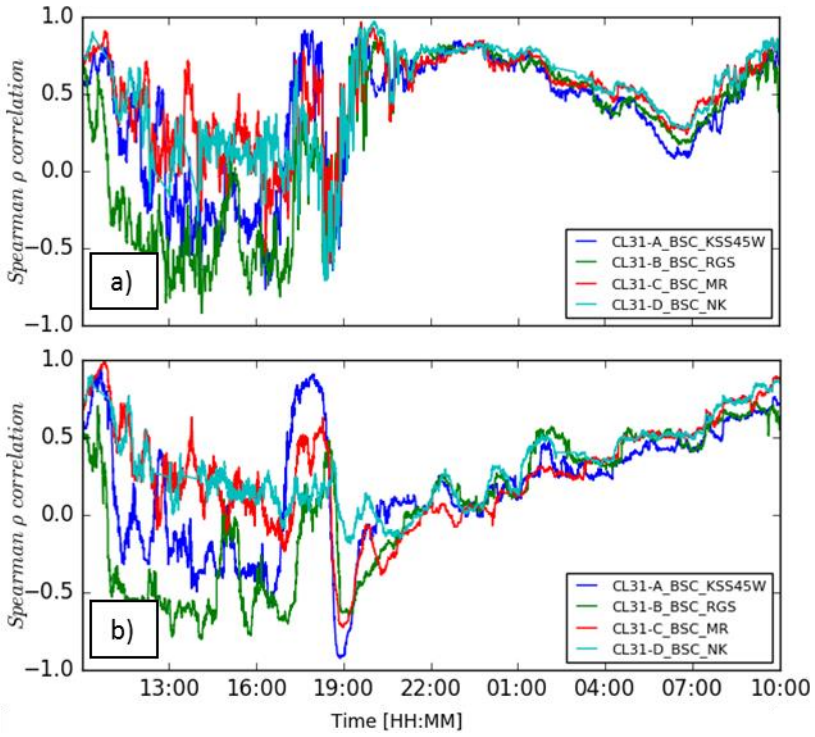

Figure 2-4: Time-series of Spearman correlation coefficient between modelled (β_m) at NK and observed attenuated backscatter using m_{50} for a) MO 100 m and b) UKV. Sea breeze arrival defined as 18:00.

Mean bias error (MBE, defined as $\log_{10}(\beta_m)$ - $\log_{10}(\beta_o)$) is calculated for four equal layers of 500 m thickness (not shown). The m_{20} runs have the lowest MBE before the sea breeze arrival, with a typical maximum difference of 1 [$\log_{10}(\beta)$ m^{-1} sr^{-1}] between 0 and 500 m, and 0.5 between 500 and 200 m. After the sea breeze arrival, m_{50} runs have the lowest MBE. The typical difference below 500 m is 0.5, however the difference reduces with height from around 0.3 between 500 and 1000 m, to around 0 between 1500 and 2000 m.

The evaluation of β_m estimated from the UKV and 100 m model showed the sea breeze event passing the site in both models and additionally that the MURK aerosol was too high during the event. This utility during the sea breeze event provided support for further development and use of the aerFO in a larger range of case studies.

2.3.4 Qualitative case study assessment

Further qualitative case study comparisons were made between β_m derived from UKV output, and Vaisala CL31 β_o across four LUMO sites to determine whether β_m could demonstrate realistic changes in more meteorological phenomena in addition to sea breezes (section 2.3.2). The results from two cases for NK are shown here.

Figure 2-5 shows (a) β_o and (b) β_m for 20 April 2015. The overall shape and height of the boundary layer in β_m appears consistent with β_o. The surface layer is also visible between 00:00 and 09:00 below 200 m in β_o and below ~500 m in β_m although with a greater magnitude. This difference in magnitude between β_o and β_m could partially be due to β_o being uncalibrated but also the near-surface *RH*, or aerosol, from the UKV might be higher than reality, leading to a higher estimate of β_m. Further, a small aerosol layer is present in β_m at ~ 1000 m between ~ 03:00 and 08:00, which is partially reflected in β_o, and could indicate aerosol advection. From 09:00 to 15:00, the growth of the boundary layer is clear in β_m as near surface aerosol and moisture is mixed vertically through convection. This vertical mixing also produces a large difference between the growing boundary layer and residual layer above. Finally, from 15:00 to 00:00 the decay of the boundary layer is visible as the vertical extent of the mixing reduces, and the near-surface β_m appears to increase as fresh surface emissions begin to build in the surface layer.

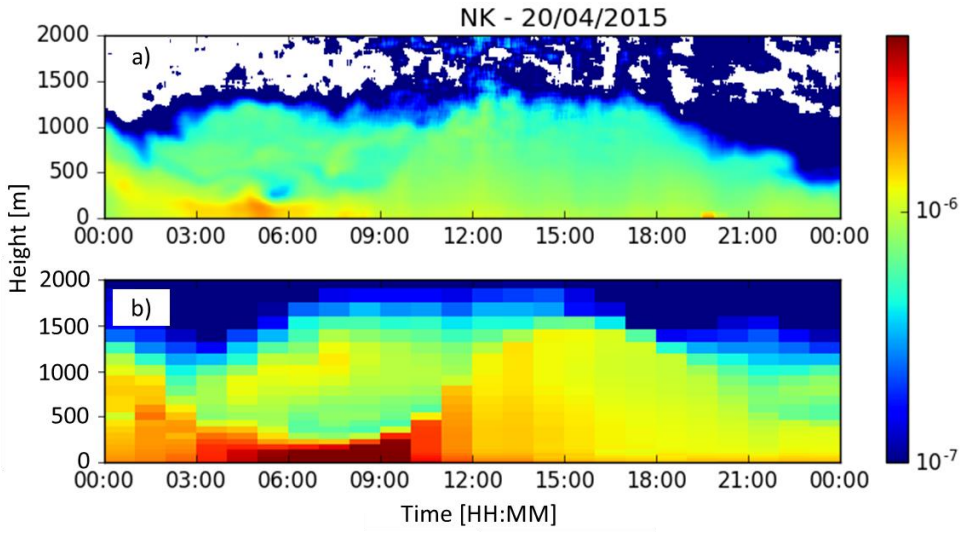

Figure 2-5: Vertical profiles of attenuated backscatter [m^{-1} sr^{-1}] that are (a) uncalibrated observed (β_o) and (b) forward modelled (β_m) using the original aerFO (section 2.3.1), using UKV output, for 20 April 2015, at North Kensington, London. Note the backscatter are plotted on a log scale.

Figure 2-6 shows a second case from 21 April 2015 at NK. Similar to Figure 2-5, the surface layer is present below ~ 300 m in the morning, with a growing boundary layer to late afternoon, and boundary layer decay in the evening. Increases in vertical backscatter are present at 13:00 and 18:00 in β_m and likely represent the stronger vertical convection that occurred at those times. In addition to the boundary layer features, another aerosol layer is present between 01:00 and 04:00 at ~ 700 m in β_o. In β_m the same aerosol layer appears deeper and more prevalent between 500 and 1500 m, though differences would exist due to differences between UKV output and reality.

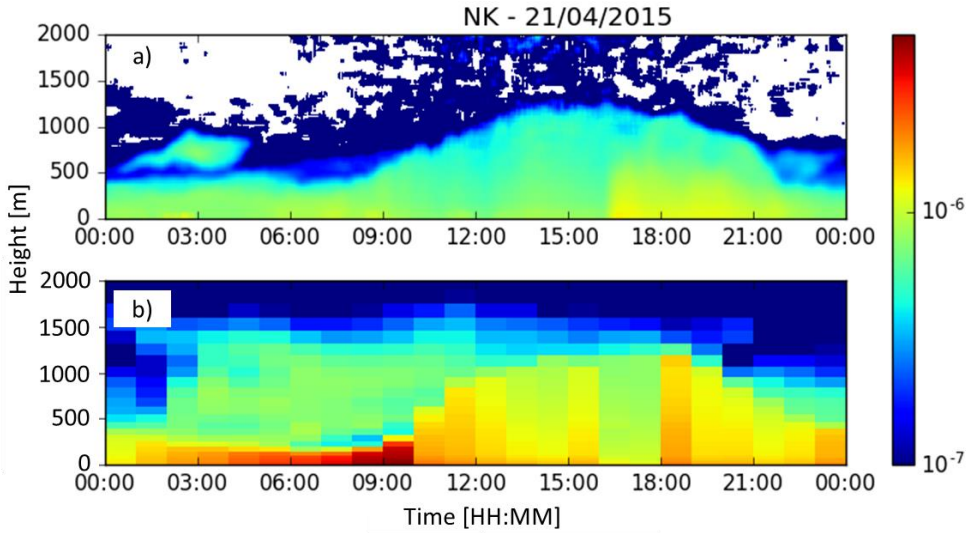

Figure 2-6: Same as Figure 2-5 but for 21 April 2015.

The cases analysed suggest that the aerFO can reasonably represent meteorological phenomena, including boundary layer development and the advection of aerosol layers. This supports its further development and testing in a greater range of case studies.

2.4 The cloud-free aerosol Forward Operator (aerFO, version 1.0)

Following the adequate performance of the early aerFO in section 2.3, it was determined that the aerFO should be improved for future usage and tested more thoroughly. This section provides a full overview of the expanded aerFO used in the remainder of this chapter, whereas sections 2.5 - 2.7 provide the methods and results of the more thorough analysis.

The aerosol forward operator (aerFO) accounts for aerosols only. It follows that, it is suitable for use in sub-saturated, cloud-free conditions, in the absence of hydrometeors. To compute β_m (Figure 2-7) the main inputs are vertical profiles of a bulk aerosol mass mixing ratio (such as m_{MURK}) and *RH.* Secondary inputs of specific water vapour (q), air temperature (T_{air}) and air pressure (p_{air}) are used to correct for water vapour absorption. The aerFO is extracted and built from the aerosol part of the FO produced in Charlton-Perez et al. (2015). New additions to the forward operator presented in this chapter are highlighted in blue, in Figure 2-7.

In aerFO the bulk aerosol characteristics (i.e. mean dry particle radius (r_{md}) and number concentration (N)) are determined first (part 1 in Figure 2-7). The mean dry particle radius is a function of the standard mass mixing ratio [m_o, kg kg^{-1}] (Claxton, 2013), the standard aerosol particle radius (r_o) and a scaling factor (p) (Figure 2-7). A similar scaling is applied to estimate N as a function of the background number concentration (N_0). When p = 1/6, both N and r_{md} increase as m increases (Clark et al., 2008; Haywood et al., 2008). The aerFO aims to represent the accumulation mode of the aerosol distribution using N_0 and r_0, as this contains a large proportion of the mass due to the relatively low loss of particles in this size range (Seinfeld and Pandis, 2016). Here, the accumulation range is defined as 0.04-0.7 μm as interpreted from aerosol distributions taken during the winter intensive observation period of the ClearfLo (Bohnenstengel et al., 2015) and REPARTEE II (Harrison et al., 2012) campaigns. Default values for aerFO parameters are shown in Table 2.1.

Using r_{md}, *RH* and the assumed aerosol constituents, the scattering efficiency (Q_{ext}) is calculated (part 2 in Figure 2-7) as a function of ALC wavelength (section 2.5.1). This includes using a *RH*-dependent extinction enhancement factor ($f_{ext,rh}$) coefficient, which represents the change in optical properties due to hygroscopic growth of the aerosol. This includes the changing of physical particle size and the change in chemistry as the particle solution contains relatively more water compared to the original solute. The particle extinction coefficient ($\sigma_{ext,aer}$) is then calculated from Q_{ext} (part 3 in Figure 2-7). Since water vapour strongly attenuates radiation at certain wavelengths, the water vapour

extinction coefficient is calculated explicitly ($\sigma_{ext,wv}$) and used with $\sigma_{ext,aer}$ in the determination of the total extinction coefficient (σ_{ext}) (part 3 in Figure 2-7). When integrated over the atmospheric column, from ground (z = 0 m) to a model level height (z_i), σ_{ext} provides the aerosol optical depth (τ) up to height z_i. Aerosol optical depth is then used to obtain the two-way transmission, T (part 4 in Figure 2-7), to represent the effectiveness of light to penetrate through the optically thick air. The backscatter ($\beta_{m,unatt}$) is derived using $\sigma_{ext,aer}$ and an assumed extinction-to-backscatter ratio (lidar ratio, S). Currently, the extinction and backscatter of air molecules is not considered in the aerFO.

Assuming aerosol is the only significant source of backscatter, a lidar ratio S for continental aerosol is applied (Doherty et al., 1999). However, S varies considerably in space and time as it is strongly dependent on particle size, shape and refractive index (Müller et al., 2007); from 20 sr (marine) to over 100 sr (volcanic aerosols) (Ansmann et al., 2010; Boyouk et al., 2011; Dawson et al., 2015; Geisinger et al., 2017; Povey et al., 2014; Wang et al., 2016). Using the fixed lidar ratio (Table 1.1) will introduce increased uncertainty as continental aerosol becomes less likely (e.g. during sea breezes), but without known values, a fixed S is used in this study. Finally, attenuated backscatter (β_m) is the product of T and $\beta_{m,unatt}$ (part 4 in Figure 2-7).

Table 2.1: Parameters required by the aerosol forward operator (aerFO) and their default values. Calculated from ClearfLo Winter intensive observation campaign data at North Kensington, London (10 Jan 2012 – 8 Feb 2012) using the DMPS and TSI APS* (Bohnenstengel et al., 2015).

Symbol	Parameter	Default value	Value source
S	Lidar ratio	60 sr	Doherty et al. (1999)
η	Proportionality constant	0.75	Clark et al. (2008)
r_0	Standard mass particle radius	1.1×10^{-7} m	Haywood et al. (2008)
N_0	Total number concentration	4.461×10^{9} m^{-3}	ClearfLo (Winter) *
m_0	Standard mass mixing ratio	1.8958×10^{-5} kg kg^{-1}	Met Office NWP setting
p	Scaling power	1/6	Clark et al. (2008)

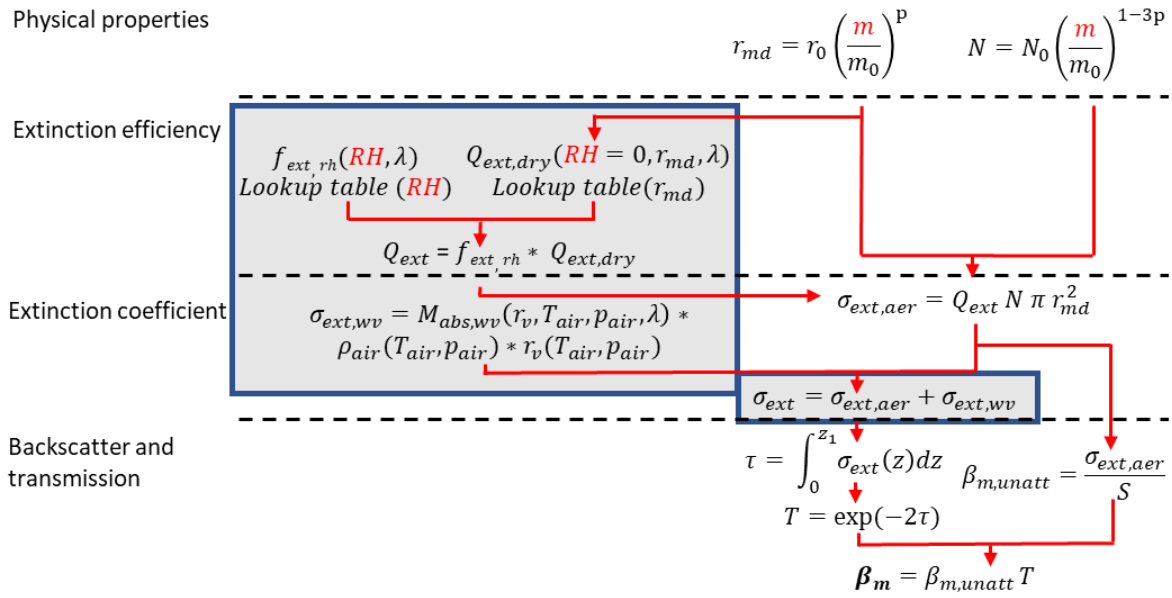

Figure 2-7: The aerosol forward operator (aerFO) to estimate attenuated backscatter from the NWP inputs of aerosol mass mixing ratio (m) and relative humidity (RH). Inputs to aerFO (red symbols). Improvements made to the FO of Charlton-Perez et al. (2015) are highlighted (blue box). See text for details and symbol definitions.

As ALC observed scattering is a function of particle size, laser wavelength and atmospheric composition, these effects are incorporated in aerFO through Q_{ext}. Q_{ext} is a function of laser wavelength, aerosol radius and complex index of refraction (Jacobson, 2005). In humid air, hydrophilic aerosols typically swell through hygroscopic growth, which increases particle size and can change the complex index of refraction. Both effects typically increase Q_{ext} for individual particles and, consequently, for the bulk particle mixture. For this reason, in the aerFO, a *dry* extinction efficiency ($Q_{ext,dry}$) is calculated and then multiplied by an extinction enhancement factor, $f_{ext,rh}$ (Figure 2-7, often called *f(RH)* in the literature). $f_{ext,rh}$ is a function of *RH* and is calculated for each aerosol species separately. The assumed proportions of each aerosol are used to create a weighted $f_{ext,rh}$ for the bulk aerosol.

The actual aerosol composition (i.e. constituents and their proportions) is important (Seinfeld and Pandis, 2016) to both $Q_{ext,dry}$ and $f_{ext,\mathrm{rh}}$. $Q_{ext,dry}$ is a function of the complex index of refraction, which differs between aerosol types. In sufficiently moist conditions $f_{ext,\mathrm{rh}}$ depends on both the particular aerosol's hygroscopicity and the critical *RH* required for that aerosol constituent to change phase into a saturated solution, at which point the aerosol's optical characteristics change (Fitzgerald, 1975).

To incorporate a mixed aerosol composition in the calculation of $Q_{ext,dry}$, aerFO allows the user to specify the aerosol types and proportions, including the complex index of refraction for each aerosol species. The relative mass of each species is used to calculate a bulk complex index of refraction (n) for the bulk aerosol mixture using the volume mixing method (Liu and Daum, 2008):

$$n = \sum_i v_i \, n_i \tag{2.3}$$

where n_i is the partial refractive index and v_i is the volume fractional contribution of the ith constituent. Here each aerosol constituent's complex index of refraction by wavelength (Table 2.2) is obtained from the MO NWP radiation suite SOCRATES (Manners et al., 2015). The aerosol species and their proportions implemented in the current MURK scheme are derived from airborne observations (Haywood et al., 2008, their Table 1). $Q_{ext,dry}$ is calculated from n based on Mie theory at discrete wavelengths (0.2 nm resolution), assuming spherical aerosols (Jacobson, 2005). This assumption is often not met in reality (Aptowicz et al., 2006), so it introduces additional uncertainty to the aerFO estimates of β_m. However, at higher RH values this uncertainty decreases, as aerosols become swollen and hence more spherical. Comparison of the Mie code used to calculate $Q_{ext,dry}$ to Figure 13 of Haywood et al., (2008) found qualitatively minimal difference. Look up tables (LUT) of $Q_{ext,dry}$ are produced for dry radii ranging between 5×10^{-10} m and 2×10^{-6} m, with a resolution of 5×10^{-10} m.

The MO SOCRATES radiation code is also used to create $f_{ext,rh}$. Based on the Edwards and Slingo, (1996) code, Mie theory calculations can be carried out with SOCRATES to derive mass scattering and absorption coefficients for different aerosol types, when ambient RH is between 0 and 100 %. The effect of hygroscopic swelling is parameterised for NH_4NO_3 and $(NH_4)_2SO_4$ following Fitzgerald (1975), and organic carbon following Varutbangkul et al. (2006). Both parameterisations are empirically derived from laboratory experiments for estimating hygroscopic growth below saturation, specifically for those aerosol types. The approach to use traditional Köhler theory was not taken, despite its popularity in visibility parameterisation schemes. This is because Köhler theory has underlying assumptions based on droplets having a relatively low solute concentration which would not be the case below saturation (Lewis, 2008), and would lead to errors in estimates of hygroscopic growth at sub-saturation humidities where the aerFO is intended to operate. Subsequent changes in refractive index and density due to hygroscopic swelling are used to calculate the optical properties using volume weighting. The SOCRATES suite is routinely updated and hence produces accurate $f_{ext,rh}$ curves based upon current knowledge. To create $f_{ext,rh}$, the complex indices of refraction by wavelength, geometric mass mean radius (r_g), geometric standard deviation (d_g) and a mean number concentration of the aerosol are required for each of the contributing species (Table 2.2). Sensitivity tests showed no difference in $f_{ext,rh}$ using different number concentrations, therefore an arbitrary value of 8000 cm^{-3} is used. For all the aerosols, r_g is set to 0.11 μm (Haywood et al., 2008), i.e. the same value as r_0 (Table 2.1), and d_g to 1.6. The latter is estimated from the first mode of the observed mass distribution in Regent's Park, London (Harrison et al., 2012). As with the aerosol type and proportions, r_g and d_g values can be set by the user.

*Table 2.2: Sources of complex index of refraction from the SOCRATES code by aerosol species used for different ceilometer central wavelengths († CL31: 895-915 nm; Lufft CHM15K: 1064 nm). *Absorption set to ~1/3 of the value for aged biomass burning organic matter based on MO scientific judgement. As volatile organic compounds were unavailable in SOCRATES, aged fossil fuel organic carbon is used instead (MO internal discussion, 04/2017). In this study, the bulk aerosol is assumed to contain ammonium sulphate, ammonium nitrate and organic compounds based on airborne observations from Haywood et al. (2008) (section 2.5.3).*

Aerosol Type	Chemical Symbol	Refractive Indices †		Part: Source
		895-915 nm	1064 nm	
Ammonium Nitrate	NH_4NO_3	1.61, 1.70×10^{-6} i	1.61, 4.59×10^{-6} i	Real: Weasted (1977); Imaginary: Gosse et al. (1997)
Ammonium Sulphate	$(NH_4)_2SO_4$	1.53, 2.31×10^{-7} i	1.51, 2.22×10^{-6} i	Real and Imaginary: Toon et al. (1976)
Aged Fossil Fuel Organic Carbon	N/A	1.54, 6.0×10^{-3} i	1.54, 6.0×10^{-3} i	Real and Imaginary: Haywood et al. (2003)*

SOCRATES produces particle mass-specific scattering $M_{scat,aer}(RH)$ and absorption $M_{abs,aer}(RH)$ coefficients for RH between 0 and 100 %. These are used to calculate mass extinction coefficients $M_{ext,aer}(RH)$:

$$M_{ext,aer}(RH) = M_{scat,aer}(RH) + M_{abs,aer}(RH) \quad (2.4)$$

The extinction enhancement factor $f_{ext,rh}$ is the ratio of $M_{ext,aer}$ at a given RH level, to $M_{ext,aer}$ in the complete absence of humidity:

$$f_{ext,rh} = \frac{M_{ext,aer}(RH)}{M_{ext,aer}(RH = 0)} \quad (2.5)$$

$f_{ext,rh}$ is stored in a LUT for use in the aerFO.

As the amount of water vapour attenuation can be significant for some ALC wavelengths (e.g. Vaisala CL31, Wiegner and Gasteiger, 2015), the water vapour extinction ($\sigma_{ext,wv}$) needs to be incorporated into the aerFO:

$$\sigma_{ext,wv} = M_{abs,wv}(r_v, T_{air}, p_{air}) * \rho_{air}(T_{air}, p_{air}) * r_v(T_{air}, p_{air}) \quad (2.6)$$

where $M_{abs,wv}$ is the mass water vapour absorption coefficient, r_v is the water vapour mixing ratio, T_{air} the air temperature and ρ_{air} the air density. As the scattering of light by water vapour is negligible at ALC wavelengths, $M_{abs,wv} = M_{ext,wv}$. ρ_{air} and r_v can be calculated using the NWP model outputs of T_{air}, p_{air} and specific humidity (q). $M_{abs,wv}$ values were obtained using a line-by-line radiative transfer model LBLRTM 12.2 (Clough et al., 2005, 1992) and were stored in a LUT after evaluation against atmospheric radiance spectra observed by an interferometer sounder.

Finally, the ALC wavelength range must be considered when estimating Q_{ext} and $\sigma_{ext,wv}$ as all the lasers do not strictly operate on one singular wavelength. For example, the central wavelength of Vaisala CL31 ceilometers is reported to be 905 ± 10 nm with a spectral width of 4 nm at full width half maximum (FWHM; section 3.1; Kotthaus et al., 2016). To incorporate the effect of spectral width, Gaussian weighting with a FWHM of 4 nm is used when calculating $Q_{ext,dry}$, $f_{ext,rh}$ and $\sigma_{ext,wv}$. The weight is applied to the LUT of $Q_{ext,dry}$, to $f_{ext,rh}$ when calculating $M_{scat,aer}(RH)$ *and* $M_{abs,aer}(RH)$ in SOCRATES, and to the LUT of $C_{abs,wv}$ before calculating $\sigma_{ext,wv}$.

There are several key improvements in the aerFO from the original FO by Charlton-Perez et al. (2015). The main improvement is the calculation of Q_{ext}. Originally in the Charlton-Perez FO, Q_{ext} was fixed to 2 which is only suitable for geometric scattering of large particles. However in the aerFO, Q_{ext} is calculated using $Q_{ext,dry}$ to represent the extinction efficiency of different aerosol types when the particles are dry, and $f_{ext,rh}$ to represent the change in extinction efficiency due to the absorption of water vapour. The $f_{ext,rh}$ also includes the effect of particle growth, which in the Charlton-Perez FO was parameterised using the analytical solution to the Köhler curve. However, the original Köhler parameterisation was fixed to only calculate the solution for ammonium sulphate. In the aerFO, Q_{ext} can now be calculated for different aerosol species and wavelengths. Further new aerFO additions included parameterising a wavelength dependent $\sigma_{ext,wv}$, and updating N_0 so the aerFO was more suited for estimating β_m in urban environments.

2.5 Methods

2.5.1 ALC data

Vertical profiles of aerFO-derived β_m (section 2.4) are compared to β_o from five London Urban Meteorological Observatory (LUMO) (Kotthaus and Grimmond, 2014; Table 2.3) Vaisala CL31 ALC, deployed at four sites in central London (Figure 2-8). Both hardware and firmware are critical for interpretation of the attenuated backscatter profiles observed, with older hardware (generation 311) generally having lower SNR (Kotthaus et al., 2016). In addition, ripple effects occur in some transmitters (generation CLT321) (section 3.1 in Kotthaus et al., 2016) that may deteriorate β_o quality. Two generations of hardware are operated in LUMO during the study period analysed (Feb 2015 – Dec 2016; Table 2.3).

Attenuated backscatter profiles, recorded at 15 s and 10 m resolution, are corrected for instrument-related background and near-range artefacts (Kotthaus et al., 2016). A centred moving average is applied with windows of 25 min (101 time steps) and 110 m (11 range gates) to increase the SNR (section 4.2 in Kotthaus et al., 2016). The vertical range is converted to height above ground level (agl). Comparison is undertaken for the lowest 2000 m agl as the atmospheric boundary layer did not

exceed this limit during the case study days analysed (Kotthaus and Grimmond, 2018b) . The FWHM of the Vaisala CL31 laser wavelength is 4 nm, where FWHM is defined as the width across the wavelength spectrum where the dependent variable is half of the peak value. The uncertainty in the Vaisala CL31 central wavelength is up to ±10 nm (Kotthaus et al., 2016), so here the average central wavelength of 905 nm is taken, as specified by the manufacturer. This is used as the centre in the Gaussian weighting function, along with the FWHM of 4 nm, to calculate $Q_{ext,dry}$, $f_{ext,rh}$ and $\sigma_{ext,wv}$. For the comparison, the nearest β_o vertical profile in time was used for each β_m profile. In section 2.7.2, the near-surface model values are compared to the second ceilometer range gate (20 m above sensor) as the first gate is prone to augmented noise (Kotthaus et al., 2016). The CL31 reaches complete optical overlap at 70 m and observations below this height are corrected. Further, near-range artefacts (Kotthaus et al., 2016) are accounted for so that the profile data can be used from the first range gate. For the comparison of vertical profiles (section 2.7.3) the value of β_o at the nearest vertical height to the β_m level is selected.

The smoothed ALC signal profiles of attenuated backscatter are calibrated using a daily calibration coefficient. The calibration coefficients are first calculated using the 'cloud method' (Hopkin et al., 2017; O'Connor et al., 2004). Here, an apparent lidar ratio specifically for stratocumulus cloud is estimated from the observed attenuated backscatter over the day, and then scaled to the actual known lidar ratio of stratocumulus cloud (18.8 sr). This scaling is then used as the daily calibration coefficient for the backscatter profiles and has been shown to produce stable estimates (±5 % over 2015, Hopkin et al., 2017). In this study, the calibration coefficients are calculated using the 'cloud method', on the attenuated backscatter profiles corrected for instrument-related background (Kotthaus et al., 2016). The automatic procedure of Hopkin et al. (2017) requires stratocumulus clouds to be present and some knowledge of vertical profiles of water vapour. Therefore, UKV water vapour vertical profiles for the nearest grid-point to KSSW (Figure 2-8) were used to account for water vapour absorption. As we focus on clear and stratocumulus-free days, the 'cloud method' was expanded. Calibration coefficients calculated for days with stratocumulus cloud were interpolated to estimate the calibration coefficients for clear days. To do this, time periods were defined for each instrument based on when key software and hardware changes had been made. Then to estimate the calibration coefficient for each time period, either linear regression was used against the window transmission (i.e. cleanliness [%]), or for periods when the window transmission remained high throughout, a time period average was used. Finally, the relevant regression equation or mean was used for the daily calibration coefficient for the clear-sky cases.

Table 2.3: Ceilometers at each site (Figure 2) from LUMO network (Kotthaus and Grimmond, 2014). ALC data analysed for the spring comparison (*) and 12 cloud-free days between 5/2/2015 – 31/ 12/2016 [2015: 14/4, 15/4, 21/4, 11/6, 2016: 19/1; 4/5; 23/8, 11/9, 25/11, 29/11, 30/11, 4/12]. All heights are metres above ground level (m agl). † Older (311) and newer (321) generation.

Site	Ceilometer ID	hardware generation †	firmware versions	Height [m agl]	Analysis
KSS45W	CL31-A	311	1.71, 1.72	32.9	4 days
IMU	CL31-A	311	1.72, 1.74	91.0	7 days
RGS	CL31-B	311	1.71, 1.72, 1.74	8.7	8 days
MR	CL31-C	321	2.02, 2.03, 2.04	4.5	12 days *
NK	CL31-D	321	2.02, 2.03, 2.04	3.8	5 days
NK	CL31-E	321	2.03, 2.04	3.8	7 days

2.5.2 Ancillary data

To aid interpretation of β_m and β_o differences, *in-situ* surface observations (Table 2.4, Figure 2-8) are used, including hourly PM_{10} from three London Air Quality Network sites (LAQN, King's College London, 2016) near LUMO ceilometer sites (Figure 2-8). As the MURK bulk aerosol is the total mass of aerosol of all sizes and PM_{10} is the total mass of aerosol with diameters below 10 μm, the two are not directly comparable. *RH* sampled using the Vaisala WXT520 sensors at 5 s and averaged to 60 s (Kotthaus and Grimmond, 2014) is used to assess modelled near-surface *RH*. Turbulent sensible heat fluxes derived from CSAT3 3D sonic anemometer observations (flux processing details in Kotthaus and Grimmond, 2014) are compared with turbulent surface fluxes from the model.

Table 2.4: Ancillary instrumentation at each site from LUMO (Kotthaus and Grimmond, 2014) and LAQN (Mittal et al., 2016) networks. WS: weather station model. All heights are meters above ground level (m agl)

Site	Height [m agl]	Site	Height [m agl]	LAQN site	Closest LUMO site	Height [m agl]
WS: Davis Vintage Pro		WS: Vaisala WXT520		Filter Dynamics Measurement Systems		
BCT	145.0	KSK	38.8	MY1	MR	4.5
BFCL	21.0	KSNW	47.1	KC1	NK	2.8
BGH	49.0	KSSW	41.3	KC2	RGS	3.6
IML	25.5	Sonic anemometer: Campbell Sci. CSAT3				
IMU	91.0	KSSW	46.2			

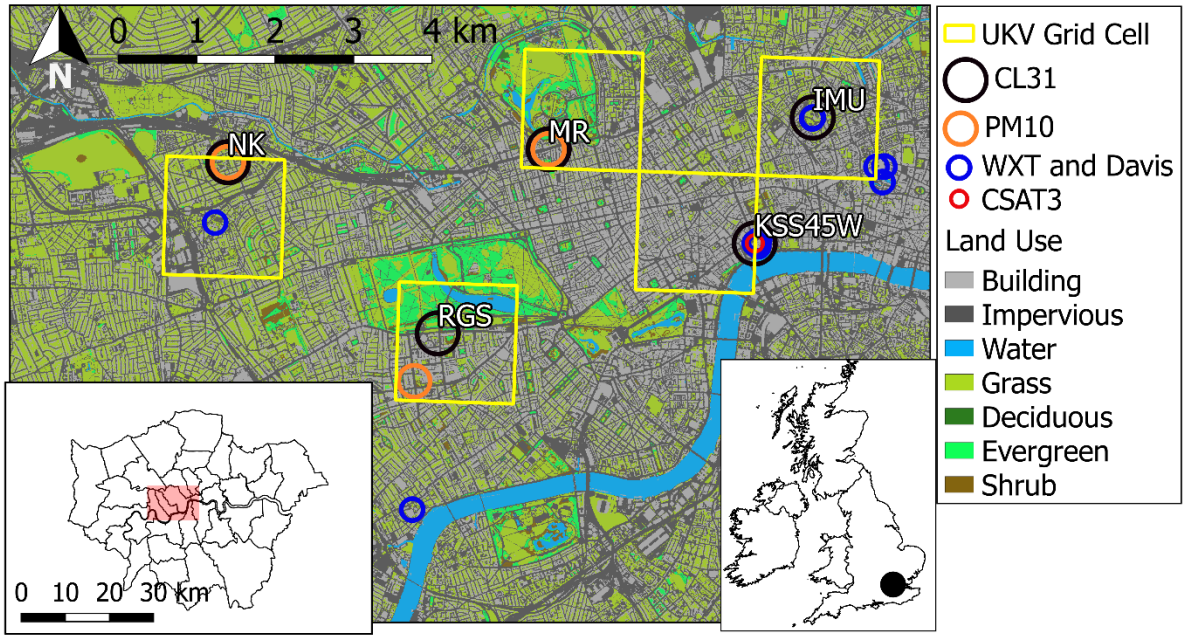

Figure 2-8: LUMO and LAQN measurement sites (ceilometer instruments: Table 3; ancillary instruments: Table 2.4) in central London with relevant UKV 1.5 x 1.5 km grid cells (yellow), and surface cover type (data source described in Lindberg and Grimmond, 2011); within Greater London and the British Isles (insets).

2.5.3 Data from the NWP model

The NWP inputs (*RH*, m_{MURK}) used to drive the aerFO are from the 3 hourly, 36 h MO operational UKV forecast (1.5 km). The aerFO calculations of β_m are performed hourly, from 3 h after the forecast begins, using archived UKV data for 5 grid cells (Figure 2-8).

The MO UKV MURK aerosol is a single mass mixing ratio (m_{MURK}, kg kg^{-1}) and a passive tracer. It represents a bulk aerosol composed of ammonium sulphate, ammonium nitrate and organic carbon (Table 2.2) based on the aerosol species measured by Haywood et al. (2008) during four flights around the UK's coast.

The data analysis is undertaken after the aerosol model had changed (5 February 2015) and the MURK ancillaries were updated to improve the visibility forecast (Appendix 2.A). For comparison against PM_{10} concentrations (section 2.7.2), m_{MURK} units are converted to kg m^{-3} by multiplying it with the UVK air density, derived from UKV air temperature and pressure fields.

In the UKV, JULES (Best et al., 2011) accounts for the land surface effects. JULES represents the surface using 9 different tile types, one of which is for urban areas. On 16 March 2016 the urban land surface option in JULES was upgraded from the 1-tile scheme (Best, 1998) (hereafter the *Best* scheme) to MORUSES (Bohnenstengel et al., 2011; Porson et al., 2010). The latter aims to improve the parameterisation of surface-atmosphere interactions in cities through a better representation of the heat and moisture fluxes from urban surfaces. This is done by estimating the radiative properties

of the street canyon (walls, road) and the roof respectively, and then coupling them together. The combined radiative effect is then used for representing urban tile effect, for the grid cell.

2.5.4 Study periods

To evaluate the NWP output of m_{MURK} and RH, and explore the possible implications of changing the urban surface scheme, analysis is undertaken for two spring cases (14 April 2015 and 04 May 2016) from before and after the change to MORUSES respectively (section 2.7.1). A more general evaluation of β_m is performed based on 11 cloud-free days between 5 February 2015 and 31 December 2016 (Table 2.3). These cases have data for at least 2 ALC but without clouds at any height for > 99 % of the day (section 2.7).

2.6 Sensitivity Analyses

Several simplifying assumptions are made within the aerFO and about its inputs (section 2.4). To understand the implications and uncertainties associated with these, sensitivity analyses are conducted, addressing potential issues with the aerosol species (section 2.6.2), laser wavelength (section 2.6.3) and NWP model variables required as input by the aerFO (section 2.6.4 and 2.6.5).

2.6.1 Sensitivity of total number concentration to defined accumulation radii range

To calculate N_0, the accumulation range of the aerosol distribution needs to be defined, so that N_0 is representative of the same mode as r_0. However, aerosol distributions change in time and space (Rodríguez et al., 2007), such that the range of radii for the accumulation mode varies. Consequently, it is important to test the implications of this assumption in calculating N_0.

A comparison of different radii ranges reveals that N_0 is highly sensitive to the range of radius used to define the accumulation range, in particular to the lowest limit used (Table 2.5). Reducing the lower limit of the range can extend into the finer mode of the aerosol distribution, which dominates the number distribution of total aerosol (e.g. Figure 13 in Harrison et al., 2012). Increasing the upper limit has a smaller effect due to the fewer number of aerosols at larger sizes. The choice of radii range could consequently have a large impact on β_m as $N_0 \propto N \propto \sigma_{ext}$. For example, $\beta_{m,unatt}$ is 15.3 % smaller if the radii range is 0.05 – 0.7 μm instead of 0.04 – 0.7 μm, and 53.4 % larger if a radii range of 0.02 – 0.7 μm is used, as a result of the variation in N_0 (Table 2.5). Given the high sensitivity of N_0 to radii range, future analysis of additional aerosol distribution observations could inform the definition of a more suitable radius range and reduce the uncertainty in estimating β_m.

Table 2.5: Total number concentration (N_0) calculated using different ranges of radius to define the accumulation range of the total aerosol distribution. Example extinction (σ_{ext}) and unattenuated backscatter ($\beta_{m,unatt}$) coefficients calculated for m = 18.0 μg kg^{-1} and RH = 60 % . † Currently used in aerFO.

Total number concentration (N_0, cm^{-3})	Radii range (μm)	Extinction coefficient (σ_{ext}, m^{-1})	Unattenuated backscatter coefficient ($\beta_{m,unatt}$, m^{-1} sr^{-1})
3769	0.05 – 0.7	1.49×10^{-5}	2.48×10^{-7}
4461	0.04 – 0.7 †	1.76×10^{-5}	2.94×10^{-7}
5426	0.03 – 0.7	2.15×10^{-5}	3.58×10^{-7}
6824	0.02 – 0.7	2.70×10^{-5}	4.50×10^{-7}
4471	0.04 – 1.0	1.77×10^{-5}	2.95×10^{-7}

2.6.2 Sensitivity of extinction efficiency to aerosol species

To calculate Q_{ext}, aerosol types and their relative volumes are required from the NWP model (section 2.4). Currently, for consistency with m_{MURK} assumptions (Appendix 2.A), aerFO assumes that the bulk aerosol is externally mixed and consists of ammonium sulphate ($(NH_4)_2SO_4$), ammonium nitrate (NH_4NO_3) and organic carbon (OC), with fixed relative volumes in space and time. However, as other aerosol types may be present in addition (e.g. salt and black carbon, Seinfeld and Pandis, 2016) and the relative fractions are likely to vary (e.g. at monthly, daily and hourly time scales; Harrison et al., 2012; Young et al., 2015), implications of these assumptions are assessed.

For this sensitivity study, $f_{ext,rh}$ and $Q_{ext,dry}$ are varied separately for seven commonly observed aerosol types (Seinfeld and Pandis, 2016). SOCRATES is used to calculate $f_{ext,rh}$ (section 2.4). Each aerosol is assumed to have an idealised distribution with a fixed geometric mean radius (r_g) of 0.11 μm (Haywood et al., 2008) and fixed geometric standard deviation (d_g) of 1.6 except for salt. Salt (generic NaCl), can be relatively coarse compared to $(NH_4)_2SO_4$, (NH_4NO_3) and OC, so here r_g is set to 8.0 μm. Given black carbon's highly hygrophobic nature, $f_{ext,rh}$ is set to 1 (Forster et al., 2007) for this species. Calculations are undertaken for the Vaisala CL31 ceilometer laser wavelengths (λ = 905 nm with FWHM of 4 nm, section 2.5.1).

Large differences in $f_{ext,rh}$ occur between the different aerosol species, with the highest discrepancies between salt and black carbon (Figure 2-9). The variation in $f_{ext,rh}$ reflects differences in the hygroscopic nature of the aerosol types and hygroscopic growth with water condensation (Haeffelin et al., 2016). These large differences in $f_{ext,rh}$ between commonly observed aerosol types suggest that using only $(NH_4)_2 SO_4$, NH_4NO_3 and aged fossil-fuel (for OC) could lead to a poor estimation of the "true" $f_{ext,rh}$.

The uncertainty in $f_{ext,rh}$ will increase as the unaccounted aerosol proportion increases, especially if the aerosol $f_{ext,rh}$ is significantly different from the applied proxies. For instance, the mean mass fraction of black carbon can be large in some cities (up to 17% during winter in London, Liu et al.,

2014), and therefore omitting it (as e.g. in the MURK-mixture; Appendix 2.A) could lead to an overestimation of hygroscopic growth. Similarly, the sea salt fraction can increase from episodic events (e.g. sea breezes and gales) or be naturally high in cities exposed to maritime flows such as Glasgow (AQEG, 2005). However, some combinations of aerosols may have compensatory effects on $f_{ext,rh}$; for example, salt (high $f_{ext,rh}$) and black carbon (low $f_{ext,rh}$) in approximately equal amounts.

A further increase in the uncertainty of $f_{ext,rh}$ occurs as *RH* increases (Figure 2-9); this is due to the differing aerosol species growth rates above the critical *RH* value of approximately 40%. Black carbon is notably different to all other species tested as it is insensitive to changes in *RH*. The uncertainty is highest near saturated conditions, where the particles would continue to grow into cloud and rain droplets. This indicates that β_m needs to be interpreted carefully in high humidity conditions, although, the large uncertainty at high *RH* is not considered critical here as the aerFO focuses on cloud-free conditions only.

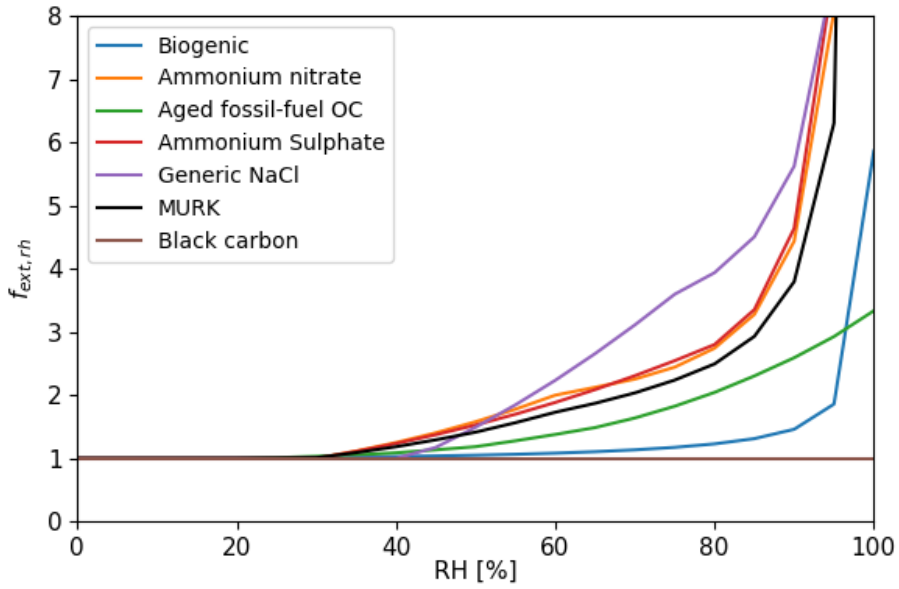

Figure 2-9: Extinction enhancement factors ($f_{ext,rh}$) as a function of relative humidity (RH) for several aerosol types commonly observed in urban areas for a mean wavelength of 905 nm, with a full width half maximum of 4 nm. MURK (black) is a bulk aerosol type including ammonium sulphate ($(NH4)_2SO_4$), ammonium nitrate (NH_4NO_3) and aged fossil-fuel organic compounds (to represent organic carbon (OC)) used to represent the aerosol field in the UKV. The dry mean geometric radius (r_g) of all aerosol types is assumed to be 0.11 μm except for generic NaCl (r_g = 8.0 μm). Geometric standard deviation (d_g) is set to 1.6 μm for all aerosols.

The variation of urban aerosol types (Harrison et al., 2012; Seinfeld and Pandis, 2016; Young et al., 2015) also impacts the dry scattering and absorption properties of the bulk particle (defined by their complex index of refraction, Jacobson, 2005) and causes differences in extinction. A clear impact of particle type on $Q_{ext,dry}$ is evident (Figure 2-10). Notably, the extinction efficiency of black carbon shows a different pattern compared to the other species because of its relatively high absorptivity and low scattering properties. In cities where black carbon is a critical proportion of the mass, for example

in London (Liu et al., 2014), Ostrava (Kucbel et al., 2017) or Beijing (Ji et al., 2017), bulk aerosol characteristics should account for black carbon with its distinct characteristics.

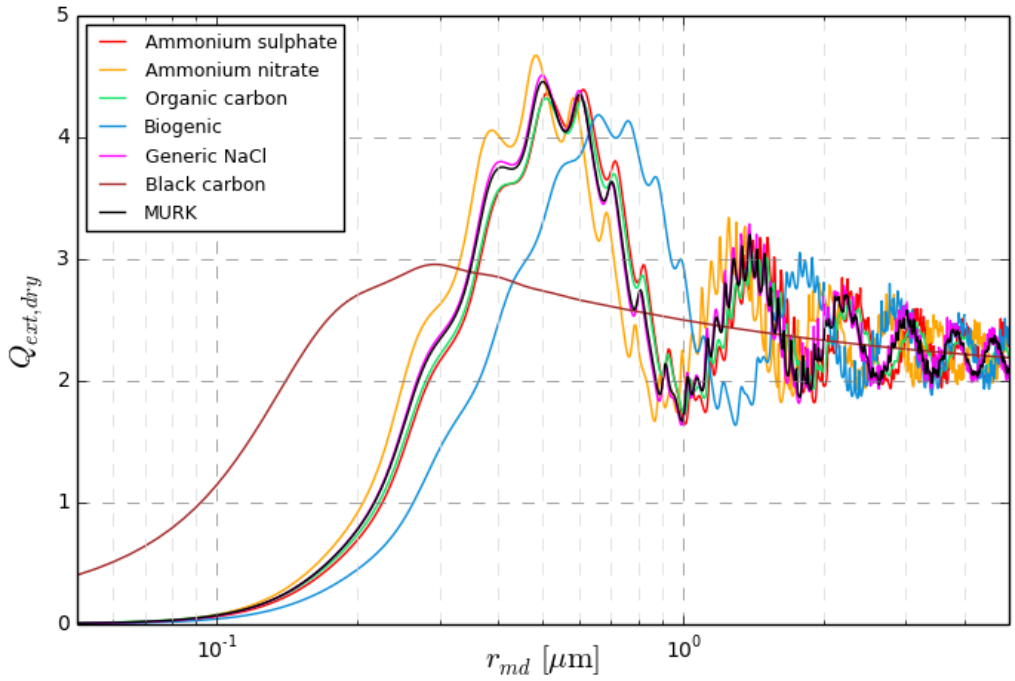

Figure 2-10: Dry extinction efficiency ($Q_{ext,dry}$) calculated for several aerosol types (no particle swelling) observed in urban areas and the bulk aerosol composition of MURK (section 2.4), for a range of radius r_{md}. Calculations are based on Mie theory and a laser wavelength of 905 nm.

2.6.3 Sensitivity of extinction efficiency to lidar central wavelength

The central wavelength of ALC lasers differs between models (e.g. Vaisala CT25K, CL31, CL51 at 905 nm; Lufft CHM15K at 1064 nm (Flynn, 2004; Lufft, 2016)). Furthermore, some uncertainty is associated with the nominal centre wavelength of individual sensors of the same model (Kotthaus et al., 2016). For this reason, the implications of wavelength variation on $f_{ext,rh}$ and $Q_{ext,dry}$ uncertainty are explored.

The variation with respect to wavelength from the nominal 905 nm (e.g. Vaisala) was assessed across the range 895 to 915 nm (1 nm resolution) because the centre wavelength uncertainty may vary by up to ±10 nm (Kotthaus et al., 2016). $f_{ext,rh}$ across this wavelength range was found to differ from the factor at the nominal centre wavelength by at most ~2%, so this uncertainty can be neglected (Appendix 2.B). However, the variation in $Q_{ext,dry}$ across the wavelength range was larger, and greatest at a mean radius of ~0.8 – 2 µm (Figure 2-11). A detected wavelike pattern with respect to radius can be explained by $Q_{ext,dry}$ becoming increasingly out of phase for each wavelength, with respect to values calculated at 905 nm. For typical r_{md} estimates (0.08 – 0.15 µm) and mass mixing

ratios (1 µg kg^{-1} < m_{MURK} < 100 µg kg^{-1}), the effect of varying wavelength on $Q_{ext,dry}$ reaches up to 6% for MURK and up to 8% for ammonium sulphate and ammonium nitrate. This relatively high sensitivity of $Q_{ext,dry}$ to small variations in λ suggests results of the aerFO could be improved if the exact centre wavelength of the laser was known. However, where this information is not available from the manufacturer, this uncertainty remains.

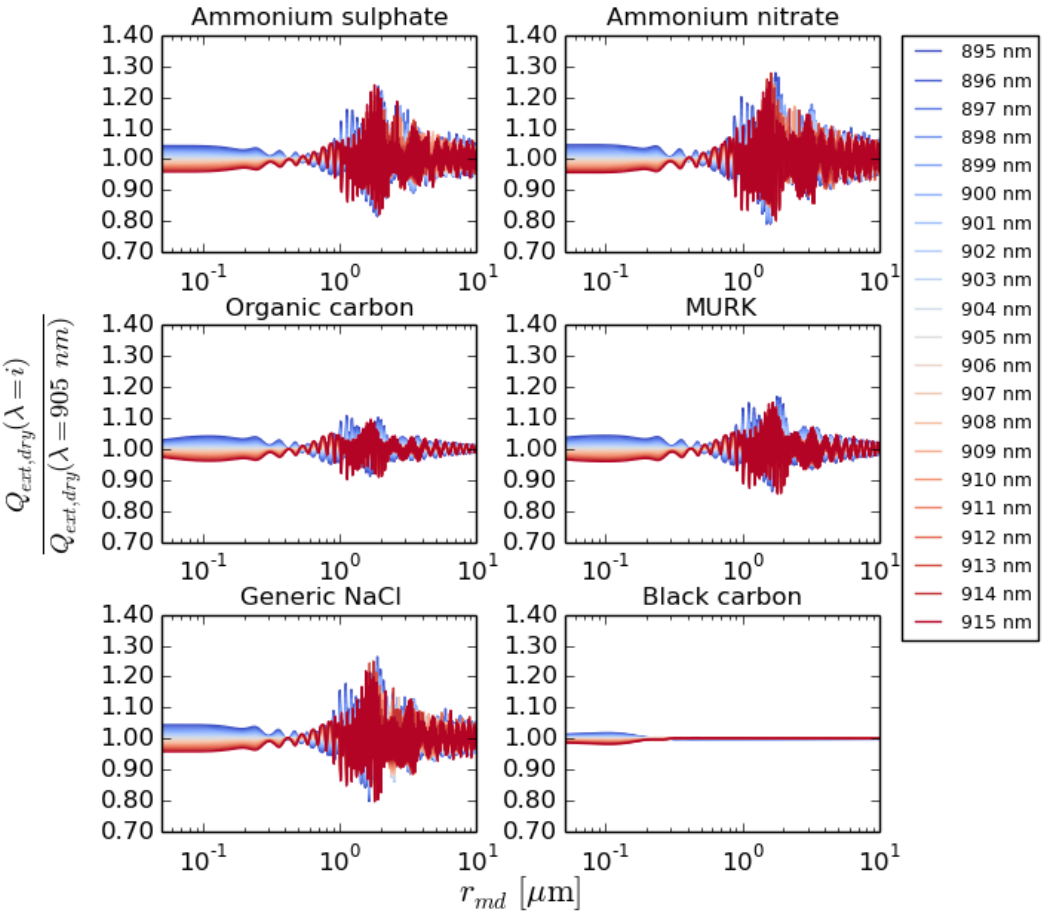

Figure 2-11: Dry extinction coefficient ($Q_{ext,dry}$) variation with mean aerosol radius r_{md} [µm] for different monochromatic wavelengths (colour) normalized by results at 905 nm for MURK; MURK's three constituents: ammonium sulphate, ammonium nitrate and organic carbon; and two additional common aerosol species (generic sea salt and black carbon). Particle swelling is not accounted for in this test.

Since multiple ALC networks can contain different instruments, with largely different wavelengths, the importance of this inter-instrument wavelength variation on estimating β_m was briefly explored. To do this, the impact of laser wavelength on $f_{ext,rh}$ and $Q_{ext,dry}$ for the two most common ALC wavelength (1064 nm, FWHM of 0.1 nm e.g. Lufft, and 905 nm, FWHM of 4 nm, e.g. Vaisala) was compared. The ratio of the enhancement factor for the two wavelengths is calculated for MURK and the three MURK aerosols constituents (Figure 2-12). As $f_{ext,rh}$ differs by up to ~30 % between these wavelengths, the aerFO allows the user to specify the centre laser wavelength and FWHM to enable appropriate comparison of modelled and observed attenuated backscatter.

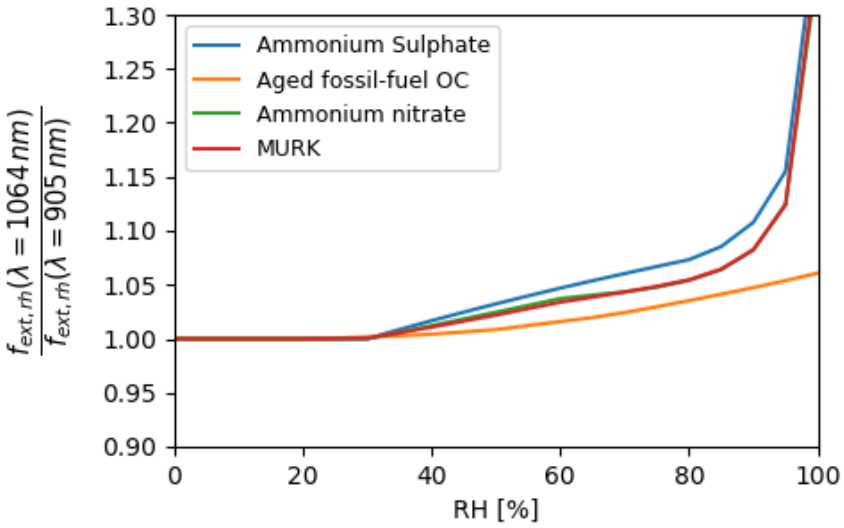

Figure 2-12: Extinction enhancement coefficient ($f_{ext,rh}$) ratio between 1064 nm with FWHM = 0.1 nm and 905 nm with FWHM = 4 nm, against RH, for MURK and the different aerosol species within MURK. Note: results for ammonium nitrate (green) are almost identical to those for MURK (red).

Similar to $f_{ext,rh}$, $Q_{ext,dry}$ results strongly differ when calculated for the two most common ALC centre wavelengths (i.e. 905 and 1064 nm; Figure 2-13) for common aerosol radii tested in the aerFO (0.08 – 0.15 µm). At 1064 nm, $Q_{ext,dry}$ of MURK can be up to 50% lower compared to the results at 905 nm (Figure 2-13). This reaffirms the conclusion that the aerFO must account for ALC laser wavelength in order to accurately estimate $Q_{ext,dry}$. As a result, sensor networks with a mix of ALC types must account for these differences by calculating $Q_{ext,dry}$ separately for the operated laser wavelengths.

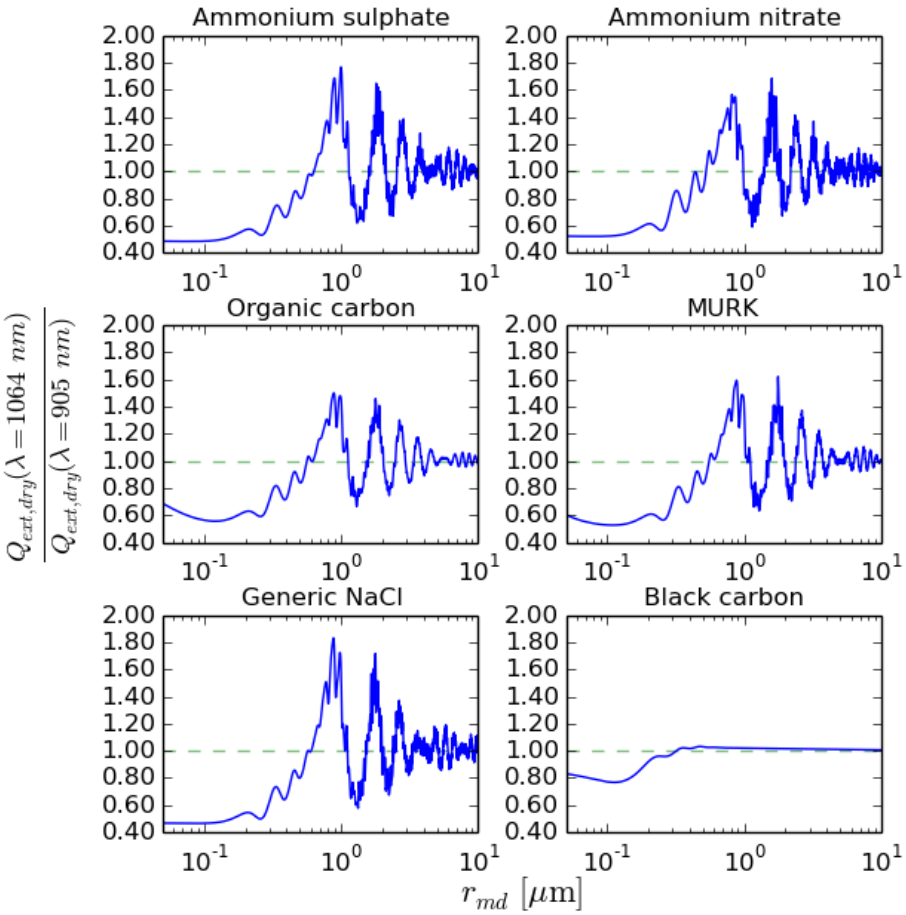

Figure 2-13: As Figure 2-11, but for different central wavelengths commonly used by ALC, i.e. 905 nm (e.g. Vaisala) and 1064 nm (e.g. Lufft). Dashed line represents a constant ratio of 1.

2.6.4 Sensitivity of forward-modelled unattenuated backscatter to relative humidity and aerosol mass mixing ratio

The sensitivity of aerFO-calculated unattenuated backscatter ($\beta_{m,unatt}$) to aerosol mass mixing ratio m and relative humidity RH was tested to inform analysis of β_m and to attribute inaccuracies correctly to either m or RH. It was shown that $\beta_{m,unatt}$ increases with increasing m and RH (Figure 2-14). At higher m values there is greater sensitivity, mainly because $\beta_{m,\ unatt} \propto \sigma_{ext}$, while $\sigma_{ext} \propto N$ and $\sigma_{ext} \propto r_{md}^2$ (Figure 2-7). Additionally, the relation of σ_{ext} to $Q_{ext,dry}$ also increases exponentially for common aerosol size regimes ($r_{md} \cong 0.07 - 0.15$ µm, with $1 > m > 100$ µg kg^{-1}).

The increase in $\beta_{m,\ unatt}$ with RH reveals a stronger exponential relation. The sensitivity of $\beta_{m,\ unatt}$ to RH is low when RH < 60 %, but greatly increases if RH > 60 %. This reflects both the physical growth of particle size and the change in the complex index of refraction caused by hygroscopic swelling, which is parameterised through $f_{ext,rh}$ (Figure 2-14). This relation suggests small

inaccuracies in *RH* at high *RH* levels could have a large impact on $\beta_{m,\ unatt}$ and, therefore, β_m. However, variations in *RH* at drier conditions (<~60 %) have little effect on the $\beta_{m,\ unatt}$ estimated.

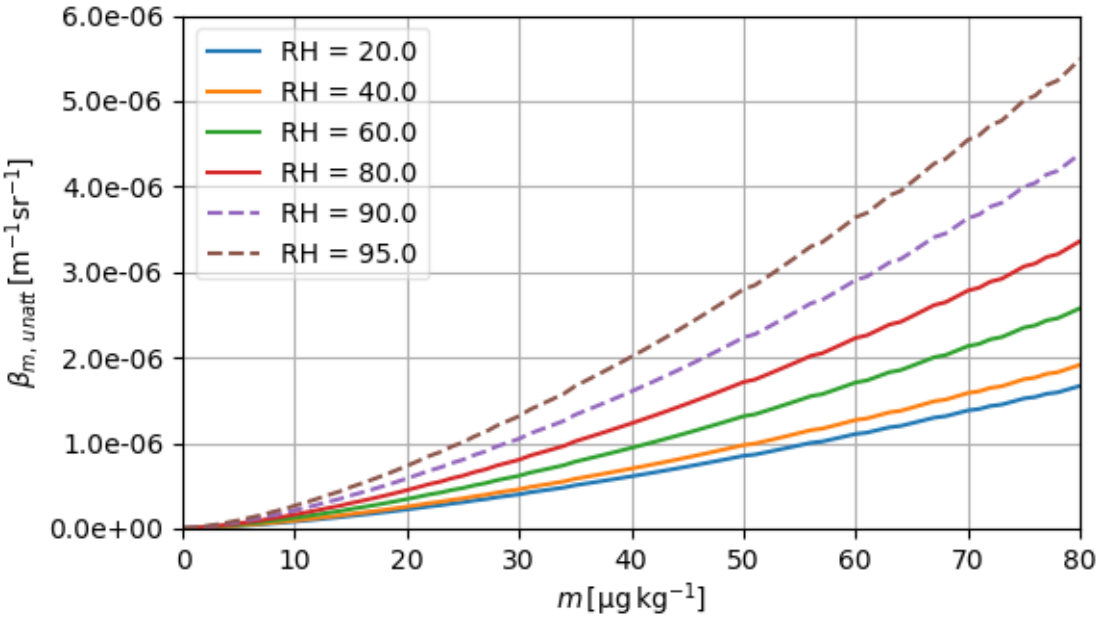

Figure 2-14: Sensitivity of aerFO estimated unattenuated backscatter ($\beta_{m,\ unatt}$) [m^{-1} sr^{-1}] to modelled aerosol mass m [µg kg^{-1}] for different RH [%] (colour). Dashed lines indicate decreased confidence in $\beta_{m,\ unatt}$ because the assumption that no aerosols have reached a critical radius and become cloud condensation nuclei is unlikely to be met.

2.6.5 Sensitivity of water vapour extinction coefficient to lidar central wavelength

The sensitivity of the water vapour extinction coefficient $M_{abs,wv}$ to wavelength across the range 895 – 915 m can be high and it varies (Wiegner and Gasteiger, 2015). Water vapour absorption is low around the wavelength of 1064 nm so can be ignored for sensors operating lasers in that region. $M_{abs,wv}$ for the default settings of the aerFO (section 2.4) is found to vary by a factor of two across the wavelength range 895 - 915 nm (not shown). Nonetheless, as the sensitivity of β_m to $M_{abs,wv}$ is rather low, a maximum difference in β_m of 6.3 % is associated with the $M_{abs,wv}$ values for the wavelength region tested. Again, if the specific central wavelength for each instrument (e.g. measured or manufacturer provided) was available to the aerFO, the uncertainty associated with water vapour absorption could be reduced (Wiegner and Gasteiger, 2015).

2.7 Performance of aerFO on cloud-free days in 2015-2016

Between 5 February 2015 – 31 December 2016, there were 11 entirely cloud-free days with more than one ALC in the LUMO network (Table 2.3) recording data. Of these days, 4 are modelled using the *Best* scheme and 7 using MORUSES (section 2.5.3). One case day was a high pollution event (19/01/16, > 50 µg m^{-3}), and given the importance to accurately forecast aerosol during such events,

it was analysed separately before analysed with the other 10 cases (Appendix 2.C). Using all 11 cases, the impact of the urban surface scheme is discussed (section 2.7.1) before analysing the general performance of aerFO on the cloud-free days (Sect. 2.7.2 and 2.7.3).

2.7.1 Importance of surface scheme in estimating attenuated backscatter

The aerFO can be used as a tool to explore implications of the change in MO urban surface schemes (section 2.5.3) on β_m results. To demonstrate the use of the aerFO in this way, the difference between modelled and observed attenuated backscatter is compared along with sensible heat flux and other UKV model fields. Figure 2-15 shows a comparison between two spring cases (14 May 2015 and 04 May 2016), when the *Best* and MORUSES urban surface schemes were used respectively (section 2.5.3). Figure 2-16 shows the sensible heat flux (Q_H) from the UKV for each grid cell that overlays a ceilometer site, and the sensible heat flux observed from KSSW (Figure 2-8), for both spring cases.

From this case study comparison, it is evident that the mixing layer starts to grow ~3 h later on the day simulated using the *Best* scheme (14 April 2015) compared to the MORUSES example (04 May 2016). Several factors, including synoptic conditions, radiative forcing, and atmospheric stability are likely to explain some day-to-day variations in boundary layer dynamics. However, a systematic delay of the rise in turbulent sensible heat flux, leading to morning transition onset, is found for results obtained with the *Best* scheme in general (King, 2015), and was also the case on 14 April 2015 (Figure 2-16). The reduced vertical exchange predicted by the *Best* scheme is linked to a delayed increase of the modelled sensible heat flux $Q_{H,m}$. As a result, a strong positive bias ($\Delta\beta = \beta_m - \beta_o$) near the surface (Figure 2-15c, left) persisting for the time where delayed mixing layer growth occurs is usually detected for days modelled with the *Best* scheme. As reported by King (2015), the timing of turbulent mixing is clearly improved in the MORUSES scheme so that the positive morning bias in attenuated backscatter near the surface vanishes much earlier (~0900) in the example shown (Figure 2-15i). Given the importance of vertical mixing dynamics for the spatial distribution of scattering aerosols, the accuracy of turbulent processes represented in the surface scheme have direct influence on β_m results. For the UKV, the comparison (Figure 2-15) suggests that the improvements of MORUSES compared to *Best* translate into a better agreement between modelled and observed attenuated backscatter.

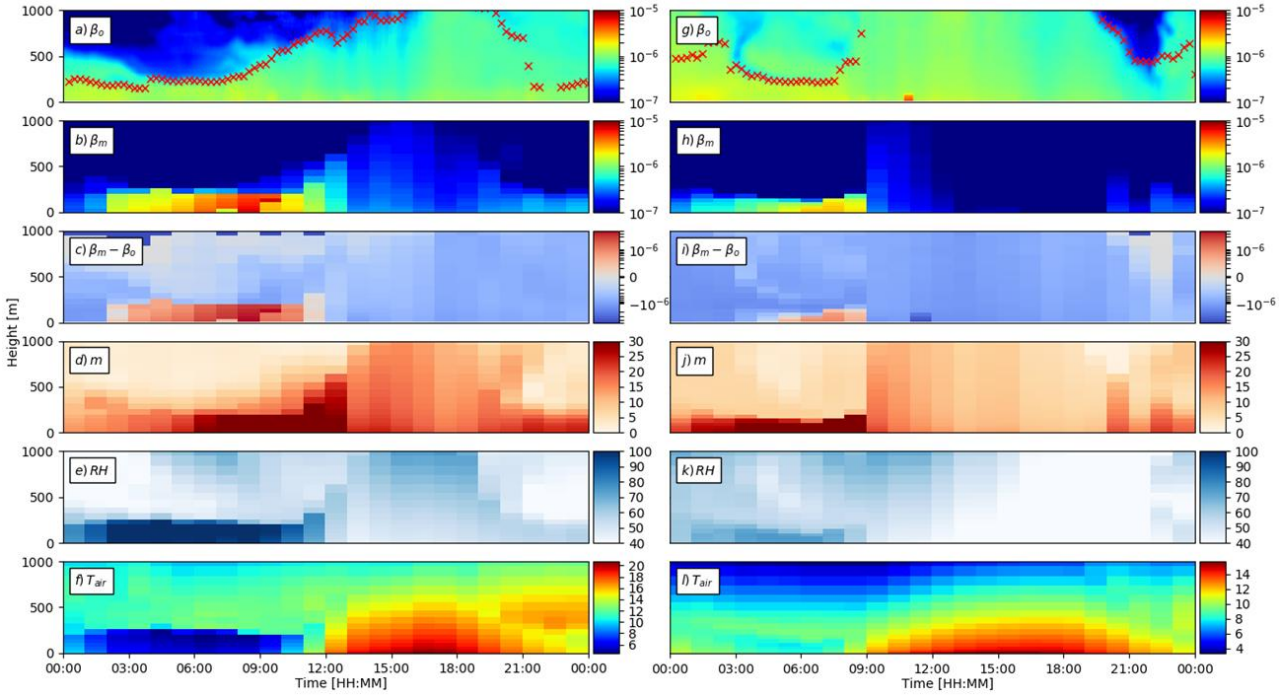

Figure 2-15: Vertical profiles of attenuated backscatter [m^{-1} sr^{-1}] that are (a, g) observed (β_o) with estimated mixing layer height (red crosses, Kotthaus and Grimmond, 2018) and (b, h) forward modelled (β_m) using the aerFO (section 2.4). (c, i) Attenuated backscatter difference (β_m - β_o) calculated using the hourly β_m vertical profile and the vertical profile of β_o nearest in time; (d, j) aerosol mass mixing ratio (m) [$\mu g\ kg^{-1}$]; (e, k) relative humidity (RH) [%] and (f, l) air temperature (T_{air}) [°C] at MR (Figure 2-8), with (a-e)14 April 2015 using the Best urban surface scheme and (g-l) 04 May 2016 using the MORUSES urban surface scheme. Note that (a, b, g, h) are plotted on a log scale and (c, i) colour bar has a range of $\pm 5 \times 10^{-6}$ m^{-1} sr^{-1} on a symmetric log scale.

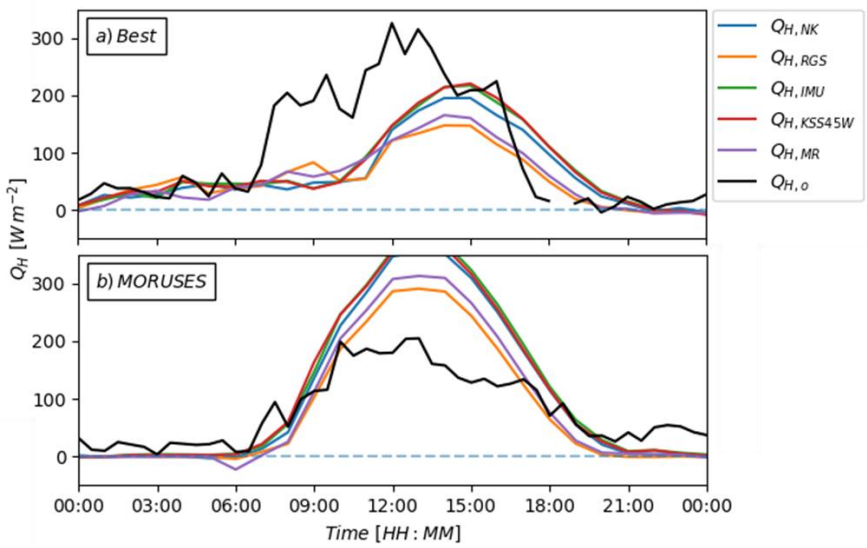

Figure 2-16: Modelled sensible heat flux [W m^{-2}] from the UKV for the grid cells over each ceilometer site (Figure 2-8) and observed sensible heat flux at KSSW ($Q_{H,o}$ black) from a CSAT3 sonic anemometer for: (a) 14 April 2015 (Best urban scheme) and (b) 04 May 2016 (MORUSES urban scheme).

2.7.2 Near-surface evaluation

Near-surface data across the 11 cloud-free days are analysed (β_m, β_o, m_{MURK} and PM_{10}, section 2.5.4). The data includes the *in-situ* PM_{10} measurements and m_{MURK} taken at the nearest vertical model level to the PM_{10} instrument height. β_o is taken at the second range gate (section 2.5.1) and β_m is selected at the model level closest to that height. Generally, β_m is lower than β_o (mean $\beta_m/\beta_o = 0.53$), and the variability of β_m/β_o increases with increasing m and RH (Appendix 2.D).

The difference in UKV bulk aerosol mass mixing ratio and observed PM_{10} (Δm) near the surface (~20 m range) when compared to $\Delta\beta$ has a positive correlation (Figure 2-17, Spearman correlation coefficient $R = 0.46$; Table 2.6 for collocated LAQN/LUMO sites NK and MR, Figure 2-8), and between Δm and $\Delta log10(\beta)$ the correlation is stronger (0.80).

Comparing Δm and $\Delta\beta$, the larger outliers are associated with particularly high model RH ($RH_m >$ ~90 %). As β_m is very sensitive to RH_m in humid conditions (section 2.6.4), any errors in RH_m translate to errors in β_m via errors in $f_{ext,rh}$. Hence, at high levels of RH, uncertainty in humidity can be the dominant factor in $\Delta\beta$, explaining the increased scatter in the relation between $\Delta\beta$ and Δm and a weaker Spearman correlation coefficient R_s. If conditions of $RH_m > 90$ % are excluded, the $\Delta\beta$ and Δm correlation increases R_s to 0.47, while excluding $RH_m > 70$ % conditions increases R_s further to 0.58 (Table 2.6). In addition, other aerFO assumptions may make β_m inappropriate at high RH, leading to greater scatter. For example, $f_{ext,rh}$ does not account for the activation of particles into cloud droplets and the continental aerosol aerFO lidar ratio is incompatible with cloud droplets. This supports the argument that under high RH, errors in RH dominate $\Delta\beta$. However, the impact of sub-sampling at lower RH for $\Delta log10(\beta)$ makes little discernible difference.

At lower RH_m (< ~ 70 %), errors in modelled β_m will be smaller, due to the lower sensitivity of β_m to RH (Figure 2-14). Assuming the settings of the aerFO are appropriate, these statistically significant relations suggest that these instances of $\Delta\beta$ and $\Delta log10(\beta)$ at lower RH_m can be attributed more specifically to errors in m_{MURK}. More generally this also suggests that, $\Delta\beta$ and $\Delta log10(\beta)$ at higher heights could likely be explained by inaccurate m_{MURK} at higher levels when RH_m is low.

Despite their statistically significant relation, some of $\Delta\beta$ (and $\Delta log_{10}(\beta)$) remains unexplained by Δm. Several aspects could play a role. One source of uncertainty lies in the observed β_o in the near range where instrument-related artefacts (reported e.g. for CL31) introduce errors into the attenuated backscatter profiles. Although these artefacts are corrected on average, uncertainty remains (Kotthaus et al., 2016). Some uncertainty may also arise from using observed PM_{10} to compare to m_{MURK}. PM_{10} measures the total mass of aerosols of size < 10 μm whereas m_{MURK} represents aerosols of any size. Although the mass above 10 μm is not always insignificant due to the presence of very coarse particles such as sea salt, the number of particles above 10 μm is often small (Figure 8.11 in Seinfeld and Pandis, 2016). Given that aerFO is based on a series of assumptions regarding the

composition of the bulk aerosol, errors will be introduced into β_m if the aerosol proxy used does not sufficiently represent the real atmospheric composition (both in terms of constituents and their relative proportions). Many other aerosol species may be present in London with variations at seasonal and hourly timescales (Young et al., 2015). Since that the complex index of refraction and hygroscopicity vary with aerosol species, inappropriate simplifications can cause errors in Q_{ext} (section 2.6.2).

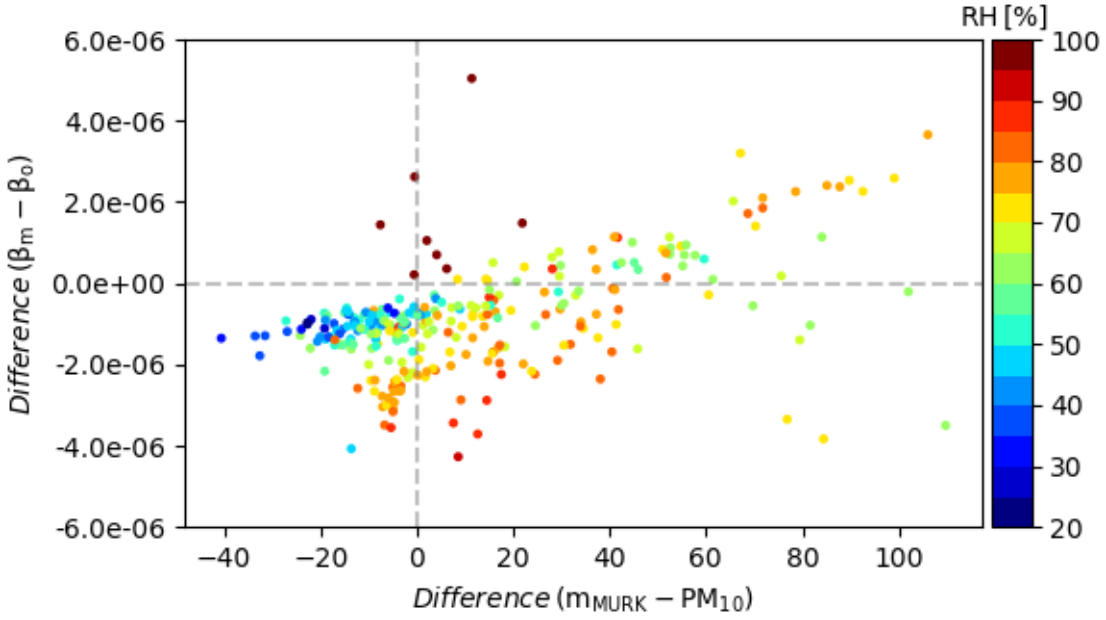

Figure 2-17: Hourly differences in aerosol mixing ratios (modelled m_{MURK} - observed PM_{10} [$\mu g\ m^{-3}$]) versus differences in attenuated backscatter ($\Delta\beta = \beta_m - \beta_o$ [$m^{-1}\ sr^{-1}$]) at MR (Figure 2-8) with UKV relative humidity (colour) indicated. β_m values for 21.7 m agl, which corresponds to the ALC's second range gate. m_{MURK} for 5 m agl which is closest to the PM_{10} observation height. Nearest time of β_o vertical profile is compared to each β_m profile (288 profiles, from 12 days, Table 2.3). Spearman correlation coefficient $R_s = 0.46$, p-value > 0.00.

Table 2.6: Spearman correlation coefficient (R_s) of $\Delta log_{10}(\beta)$ and $\Delta\beta$ against Δm, using different hourly sub-samples defined by model relative humidity (RH_m).

Model RH range included [%]	Sample size [hours]	Spearman correlation coefficient (R_s)	
		$\Delta log_{10}(\beta)$ against Δm	$\Delta\beta$ against Δm
0 – 100	280	0.80	0.46
0 – 90	276	0.81	0.47
0 – 80	238	0.80	0.50
0 – 70	159	0.80	0.58
0 – 60	93	0.67	0.54
0 – 50	53	0.74	0.62
0 – 40	17	0.61	0.70

Figure 2-18 shows the absolute error in attenuated backscatter ($|\Delta\beta| = |\beta_m - \beta_o|$) binned by Δm. $|\Delta\beta|$ tends to decrease as Δm increases from -50 to 75 μg m^{-3}, and is likely compensating a bias in the aerFO. The aerFO attempts to only represent aerosol in the accumulation mode (Table 2.1) as it contains a large proportion of total aerosol mass. When m_{MURK} is high, this leads to higher estimates of N and r_{md} in the aerFO, which in turn increases estimates of β_m. The higher β_m estimates then compensate for the lack of coarse and fine mode aerosol representation in the aerFO. Consequently, future versions of the aerFO could improve estimates of β_m if multiple modes of the aerosol distribution are represented more accurately through estimating r_{md} and N individually.

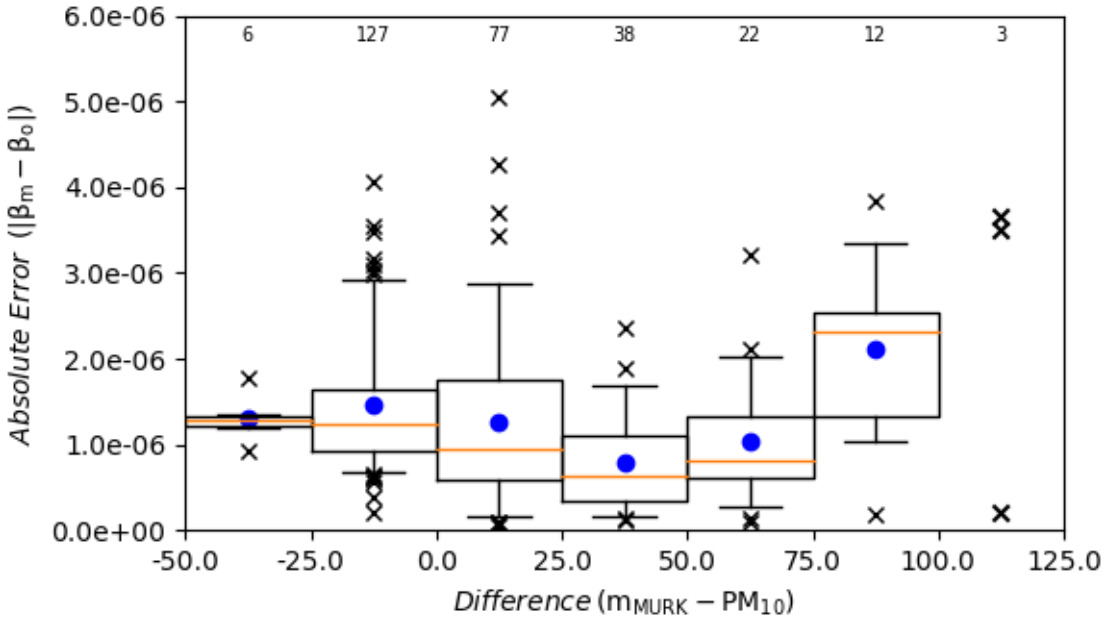

Figure 2-18: Absolute error in attenuated backscatter ($|\beta_m - \beta_o|$ [m^{-1} sr^{-1}]) binned by difference in aerosol mixing ratios, (Δm = modelled m_{MURK} - observed PM_{10} [μg m^{-3}]) at MR (Figure 2-8). Bin size = 25 kg m^{-3}. Sample means (blue circles), median (orange line) and inter-quartile range (box extent), and 5th and 95th percentiles (whiskers), statistical outliers (crosses) shown, with sample sizes above each boxplot. Same data samples used as in Figure 2-17. Boxplot omitted for 100 < $\Delta m \leq$ 125 μg m^{-3} due to the low sample size.

2.7.3 Diurnal variation in agreement between integrated forward modelled attenuated backscatter and observed attenuated backscatter

The differences in vertical profiles of β_m and β_o are assessed in order to identify whether consistent differences occur by time of day. Figure 2-19 shows the median and interquartile range of the absolute error ($|\beta_m - \beta_o|$) at each site (Figure 2-8) by time of day. Four days are analysed when β_o is available across all sites, from 73 m to 10 % below the mixing layer height (estimated from MR CL31 observations, Kotthaus and Grimmond, 2018). This is to ensure that the same number of samples are used at each site despite different measurement heights (Table 2.3), whilst the mixing layer height varies (Figure 2-19 numbers for each hour).

The variation in absolute error for each site is similar across the sites (~ 5 × 10^{-7} m^{-1} sr^{-1}). The generally higher absolute error values at KSS45W and RGS, with older generation hardware and software (Table 2.3), is likely explained by the higher noise levels and increased susceptibility to instrumental artefacts present in observed attenuated backscatter profiles (section 3.1 in Kotthaus et al., 2016). The higher noise and effect of instrumental artefacts in older instrument results also reduced Spearman correlation coefficients (not shown).

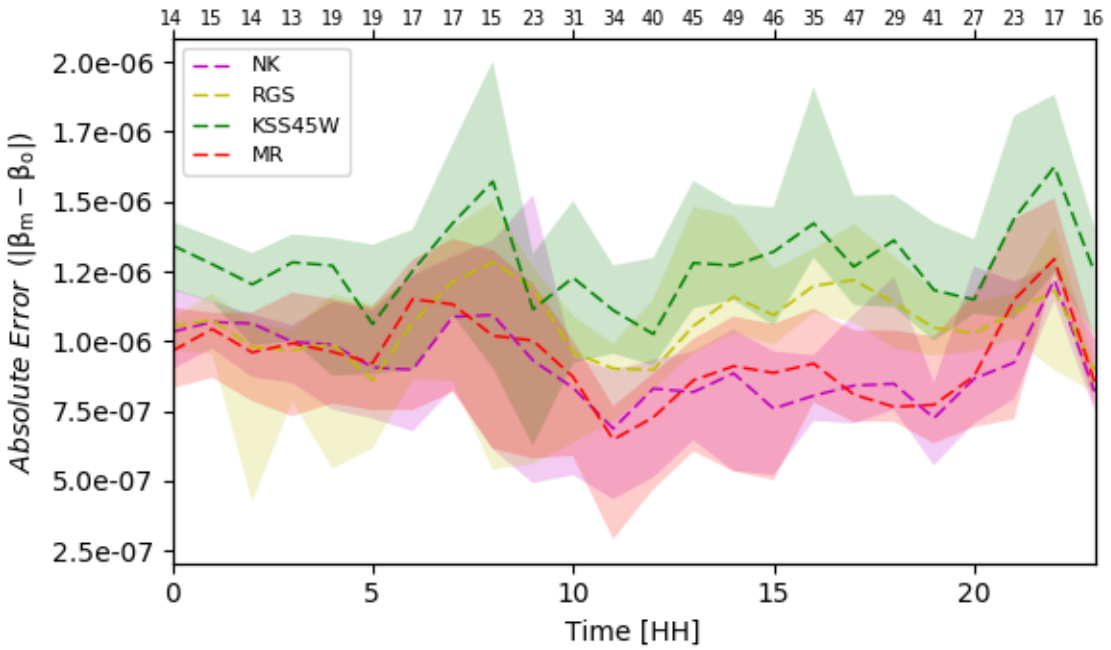

Figure 2-19: Hourly median (dashed) and IQR (shading) of the absolute error $|\beta_m - \beta_o|$ at five ALC sites (Figure 2-8), calculated for four cloud-free sample days (14, 15, 21 April 2015 and 19 January 2016), using hourly β_m vertical profiles and the vertical profile of β_o nearest in time. Profiles of β_m and β_o were compared between 73 m and 10 % below the top of the mixing layer height estimated from Vaisala CL31 observations at MR (Kotthaus and Grimmond, 2018b). Sample size for each hourly composite bin are shown above each hour. Note: sample sizes are the same for each site and are lower when the mixing layer height was lower in the early morning and late evening.

2.8 Conclusions

Following an initial assessment (section 2.3), the aerosol lidar forward operator (aerFO) is further developed (section 2.4) to estimate attenuated backscatter (β_m) for cloud-free conditions, when aerosols are the dominant scatterers, and is evaluated in an urban setting. The aerFO requires an aerosol forecast for input, which can come from a numerical weather prediction (NWP) model. Besides the aerosol field, relative humidity (*RH*) is required. If only a bulk aerosol mass mixing ratio (m) is available from NWP, the aerFO includes estimates of species-specific properties based on aerosol composition assumptions. Here, m_{MURK} and *RH* taken from the Met Office UKV (1.5 km) are used to explore the aerFO and applications.

Assumptions within the aerFO are assessed based on sensitivity studies. The choice of radius range used to calculate N_0 is critical given the sensitivity of β_m to number concentration N. Use of observed

aerosol distribution data could allow more appropriate values for N_0 to be determined and reduce the uncertainty in β_m. There are differences in both the dry extinction efficiency ($Q_{ext,dry}$) and the extinction enhancement factor ($f_{ext,rh}$) for commonly observed urban constituents. Notably, black carbon has very distinct signatures caused by its high absorptivity, low scattering properties, hygroscopicity and physical particle growth. Consequently, accuracy of β_m can be improved by providing aerFO with more accurate aerosol species composition and proportions of the bulk aerosol quantity. If atmospheric constituents are significantly different to those assumed, biases could be inadvertently introduced when assimilating β_o.

Different automatic lidars and ceilometers (ALC) have different laser central wavelengths plus uncertainty (e.g. for Vaisala CL31, available specifications are vague at 905 ± 10 nm). The impact of varying wavelength on estimated optical properties is explored. The sensitivity of $f_{ext,rh}$ to wavelength for the wavelength range 895 – 915 nm (i.e. common for Vaisala ALC) is small (< 1.5%). However, for some aerosol species $Q_{ext,dry}$ is clearly affected by uncertainties in laser wavelength (~ 8 % for ammonium sulphate and ammonium nitrate). The mass absorption of water vapour is found to vary by a factor of two across this wavelength range, translating into an uncertainty in β_m of 6.3 %. Hence, given spectral variations of absorption and extinction even small variations in laser wavelength can be important in this region. If the centre laser wavelength can be specified more accurately in the aerFO settings (e.g. than Vaisala CL31 rather vague 905 ± 10 nm) such variations in $Q_{ext,dry}$ could be accounted for by the aerFO. Manufacturers are encouraged to release this information for each instrument, to permit more accurate estimates of the various wavelength-related variables, including water vapour absorption.

Comparing results for the two most common ALC wavelengths, i.e. 905 nm (e.g. Vaisala) and 1064 nm (e.g. Lufft) reveals large differences in both $Q_{ext,dry}$ and $f_{ext,rh}$. Therefore, different look up tables of aerFO should be used for different laser wavelengths.

Sensitivity studies of unattenuated backscatter ($\beta_{m,unatt}$) to aerFO inputs of m and RH allow attribution from their inaccuracies to be assessed. $\beta_{m,unatt}$ is sensitive to m with a weak exponential relation that increases with increasing m. Overall, RH has a relatively weaker relation with β_m than with m. However, there is a stronger exponential relation between RH and $\beta_{m,unatt}$ at high RH values (>~60%) caused by the exponential growth of particles parameterised by $f_{ext,rh}$. This suggests that RH becomes increasingly important for the accurate estimation of β_m as RH increases. Therefore, the impact of errors in RH will increase as RH approaches 100 %.

Implications of a change in UKV urban surface scheme from the *Best* to MORUSES is explored based on two spring case study days. The mean bias error ($\Delta\beta = \beta_m - \beta_o$) is positive during morning hours at heights below ~250 m for the day modelled using the *Best* scheme. This is partially due to the delayed vertical mixing linked to a delayed morning increase of modelled sensible heat flux (Q_H)

(noted previously for this surface scheme). The delay in vertical mixing limits near-surface NWP aerosol dispersion. With an improved Q_H timing with MORUSES, the bias in $\Delta\beta$ is reduced less in the morning below ~250 m. This highlights the importance of the mixing dynamics within the NWP model to accurately distribute aerosols within the boundary layer and hence obtain representative β_m results.

Comparison of β_m/β_o to RH with respect to m_{MURK} shows that β_m is often less than β_o (mean β_m/β_o = 0.53). The variation in β_m/β_o depends upon the magnitudes and errors of m and RH. Analysis of the relation between Δm (= $m_{MURK} - PM_{10}$) and $\Delta\beta$ found that a positive, statistically significant Spearman correlation exists (R = 0.46), which increases when restricted to periods with lower RH values. The correlation between $\Delta log_{10}(\beta)$ was larger (R = 0.80) but does not increase when restricted to periods with lower RH. Hence, assimilation of β_o into NWP models is expected to have an increasingly effective contribution when atmospheric conditions are drier. The aerFO is still effective at high RH values, despite potentially violating the assumption that the cloud droplets have not yet formed. This could be because $f_{ext,rh}$ is representing well the exponential growth in particle radii at high RH values. A specific parameterisation to incorporate cloud droplets at high RH values would broaden the applicability of the aerFO to instances of near saturation. The absolute error ($|\beta_m - \beta_o|$) between vertical profiles of β_m and β_o show a dependence on noise levels within the observations. Quality of attenuated backscatter measurements can be improved by operating recent hardware, using the most appropriate firmware version and settings, and ensuring careful maintenance (e.g. window cleaning). This finding highlights the need for careful processing of ALC measurements prior to use in model evaluation or data assimilation. Furthermore, detailed knowledge of sensor characteristics can provide useful insights for interpretation.

As boundary layer dynamics are critical for the vertical distribution of aerosol, data assimilation of the entire β_o profile would have a limited effect if the aerosol is a passive tracer that is one-way coupled with model dynamics such as the UKV MURK. However, to account for a potential general bias affecting the whole aerosol profile, an integrated value of β_o could be assimilated.

The bulk aerosol composition is important in accurately estimating the extinction efficiency (Q_{ext}) and therefore β_m. Performance of the aerFO depends on the level of detail provided by the aerosol field(s) available from the NWP model. Some models contain information on various aerosol species, such as the AQUM configuration of the MO Unified model (Savage et al., 2013), but such models usually still run at lower spatial resolution. The high spatial resolution (grid ≤ 1.5 km) MO models (e.g. UKV or MO London model) enable a more detailed representation of the high spatial heterogeneity of urban areas. However, these usually have very simple aerosol schemes (e.g. passive tracer m_{MURK}).

The current default for aerFO includes a constant lidar ratio representative of continental aerosol.

Given the importance of this parameter, a more dynamic approach accounting for variable aerosol composition could significantly improve results. However, this can be challenging in urban settings where the relative contribution of various aerosol types is often highly variable due to the complex mixture of local sources and long-range transport. Nonetheless, the amount of aerosol and its composition can be distinct in an urban area when compared to the surroundings; therefore, a speciated lidar ratio would be more suitable. For example, urban areas tend to have higher proportions of black carbon, which has very specific aerosol properties.

The aerFO has demonstrated utility as a tool to evaluate vertical profiles of aerosol bulk mass mixing ratio and RH from NWP models. Its results can also be used in conjunction with meteorological observations to explore processes related to near-surface aerosol mass variability or horizontal advection of air masses. As aerosol is a tracer for boundary layer dynamics, application of the aerFO has proven very useful to evaluate the performance of urban surface parameterisation schemes and their ability to drive growth of the mixing layer.

Acknowledgements

EW was funded by the NERC Scenario DTP and University of Reading. The LUMO instruments and measurements have been funded by NERC AirPro, NERC ClearfLo, EUf7 BRIDGE and EUf7 emBRACE, H2020 URBANFLUXES, Met Office/Newton Fund CSSP-China, KCL, and University of Reading. Sites are provided by Royal Geographical Society (with IBG), King's College London, and Islington council. PM_{10} data are provided by KCL Environmental Research Group (ERG)/London Air Quality Network (LAQN). Helpful discussions were had with Peter Clark, Jim Haywood, Lucy Neal, Bernard Claxton, Claire Ryder, Nicholas Bellouin, Ben Johnson, James Manners, Stephen Havemann, Joelle Buxmann, Ewan O'Conner, Anthony Illingworth, and Helen Dacre. Helpful comments on the manuscript were made by Sylvia Bohnenstengel. We would also like to acknowledge the invaluable discussions within the TOPROF community (European Cooperation in Science and Technology (COST) action: ES1303). The authors are grateful to Will Morrison, Kjell zum Berge, Christoph Kent, Andy Gabey, Izzy Capel-Timms, David Meyer, Ben Crawford and the many others who maintain the LUMO operations.

Appendix 2.A MURK aerosol and visibility scheme

MURK is calculated from UKV ancillaries using (Neal, 2019):

$$m = 0.33 * SO_2 + 0.15 * NO_2 + 0.34 * VOC \quad (2.A.1)$$

where SO_2 is sulphur dioxide, NO_2 is nitrogen dioxide and VOC are volatile organic compounds. The equation converts these compounds into ammonium sulphate, ammonium nitrate and organic carbon from the species emissions inventories.

The MURK aerosol has a simple flux term as a boundary condition to help distinguish European from Atlantic emissions, as well as small background terms to represent distant sources which are designed to decay exponentially with height (Eq. 6 in Clark et al., 2008). It also adds a source term to cells over the ocean to represent sea salt. Wet and dry deposition are parameterised. MURK has a simple monthly variation with a smooth transition between months (Neal, 2019), as well as a sinusoidal diurnal cycle that peaks at 12 UTC (personal communication, Clark, 09 August 2016). On 3 February 2015, updates were made to the emission ancillaries to take advantage of more recent emission data (internal MO communication). This included using: NAEI (DEFRA, 2016), ENTEC and EMEP (Neal, 2019) datasets. In the MO UKV, emissions are partially constrained using data assimilation of visibility observations. However, this has not always led to an improvement in the forecast of visibility, i.e. if visibility. For instance, if visibility is underestimated and model *RH* is low, the MURK emissions are raised to unrealistically high levels in order to correct the visibility (Claxton, 2013).

The visibility parameterisation was originally developed assuming the aerosol was comprised entirely of ammonium sulphate (Clark et al., 2008), as the inventory used from the mid 1990's was based on sulphur dioxide emission data (Wright, 1997). In order to assess the parameterisation an aircraft observation campaign was undertaken (2006-2007) (Haywood et al., 2008), which obtained number concentration distributions and found that the typical aerosol composition in the UK included sulphates, nitrates and organic carbon. Several parameters of the MURK scheme were modified accordingly, including the standard dry radius and number concentration.

Appendix 2.B Variation of the extinction enhancement factor with respect to wavelength around the central wavelength of a Vaisala CL31

For each wavelength the ratio of $f_{ext,rh}$ to its values at the nominal centre wavelength (λ=905 nm) are determined for three MURK constituents and the bulk aerosol (Figure 2.B.1). The relative differences are generally < 2% but non-linear with relative humidity changes. However, these are mostly artefacts from using LUTs from the radiation code, with discontinuities in the parameterisations with *RH* (section 2.4). Thus, the impact of wavelength variation between individual sensors of the same design can be treated as negligible.

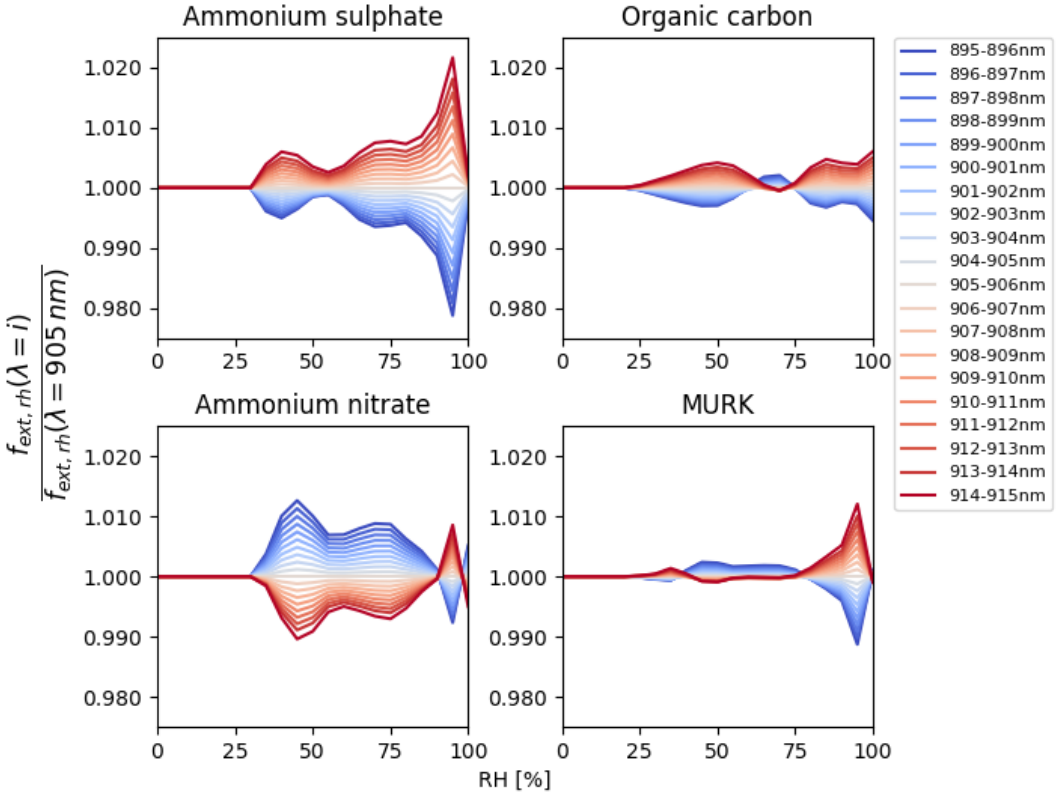

Figure 2.B.1: Extinction enhancement coefficient ($f_{ext,rh}$) variation with RH for different wavelengths (colour) normalized by 905 nm results for MURK and its three constituents. Wavelength ranges from 895 (light green) to 915 (dark blue) nm, with a bandwidth of 1 nm.

Appendix 2.C London high pollution case study (19/01/16)

β_m was estimated using the aerFO and, along with its model input fields of aerosol m_{MURK} and relative humidity *RH*, was evaluated in detail on a day with high pollution (19 January 2016). On this day, daily mean PM_{10} (51.9 μg m^{-3}) exceeded the EU legal limit and World Health Organisation (WHO) recommended limit of 50 μg m^{-3} (EEA, 2015; WHO, 2006) as measured at an urban background measuring site at North Kensington (NK) in London (Figure 2-8). An urban background site is one where the measurements are not strongly influenced by a single source point and are representative of a wider area, typically several kilometres squared (DEFRA, 2019). A high pollution case study was chosen for analysis given that accurate forecasts of aerosol content during high aerosol pollution events are critical for providing accurate air quality warnings, and limiting the negative impact on people's health (Grell and Baklanov, 2011).

Synoptic conditions favoured very light southerly winds, thus, discouraging rapid aerosol transport. Although two stratus-like clouds were present at ~4 km above ground level between ~02:00 and 12:00 UTC, there was a cloud-free atmosphere below this. Thus, the aerFO is applicable and can be used to evaluate conditions in the atmospheric boundary layer.

The aerFO (section 2.4) was evaluated using 21Z UKV forecast aerosol m_{MURK} and RH from 18 January 2016, allowing 3 hours of spin-up time, to provide model attenuated backscatter β_m (Figure 2.C.1b). This UKV forecast was run using the *Best* scheme (Best, 1998) and the aerFO was set to consider the aerosol constituents of m_{MURK} in equal parts (Table 2.2). The observed, post-processed and calibrated β_o (section 2.5.1) from the Vaisala CL31 at the Marylebone Road site (MR), in London (Figure 2-8) is compared to β_m.

Vertical profiles of absolute difference in attenuated backscatter ($\Delta\beta = \beta_m - \beta_{o,}$) on 19 January 2016 are typically ± 5 × 10^{-6} m^{-1} sr^{-1} (Figure 2.C.1c). Furthermore, the sign and magnitude of $\Delta\beta$ vary throughout the day (Figure 2.C.1a). During the day, below ~150 m, $\Delta\beta$ is positive, whereas above ~150 m, $\Delta\beta$ is negative. This positive and negative $\Delta\beta$ pattern is partially due to insufficient boundary layer growth in the UKV compared to reality. If the UKV boundary layer had developed to reach greater heights, the near surface m_{MURK} could be increasingly mixed vertically, modifying m_{MURK} near the surface and aloft, and hence changing $\Delta\beta$. The weak mixed layer growth is likely associated with underestimated modelled sensible heat fluxes in central London. Figure 2.C.2 shows estimated sensible heat flux at two locations in the UKV: MR ($Q_{H.MR}$) and for a grid box overlaying a nearby site at King's College London (KSS45W, $Q_{H.KSS45W}$). Figure 2.C.2 also shows eddy covariance observations of sensible heat flux at KSS45W ($Q_{o,KSS45W}$). Both $Q_{H.MR}$ and $Q_{H.KSS45W}$ are low compared to $Q_{o,KSS45}$, especially during mid-day, despite the relatively high vegetation cover in the MR grid cell from Regent's Park (Figure 2-8). Parks typically have lower sensible heat fluxes compared to urban areas, due to more energy being partitioned into the latent heat flux. However, given that the

modelled flux is very low at both MR and KSS45W grids, the higher vegetation fraction of the MR grid cell likely does not explain the small modelled $Q_{H,MR}$.

Low estimates of sensible heat fluxes might be partially due to inaccuracies in the modelled cloud amount and position. Though 19 January 2016 was mostly clear of cloud, some cloud was present in the morning (Figure 2.C.3). Positive model estimates of cloud volume fraction were estimated as low as 2 km compared to observed cloud at ~3.5 km. The presence of lower cloud in the model could have reduced the surface insolation, therefore increasing stability and leading to weaker boundary layer development. As radiation variables were not archived from the UKV, their impact on the modelled surface energy balance could not be assessed here.

For any NWP model, this analysis suggests that vertical mixing is important for accurately estimating β_m. Insufficient vertical transport of m_{MURK} within the NWP leads to a positive $\Delta\beta$ bias near the surface and negative discrepancies aloft.

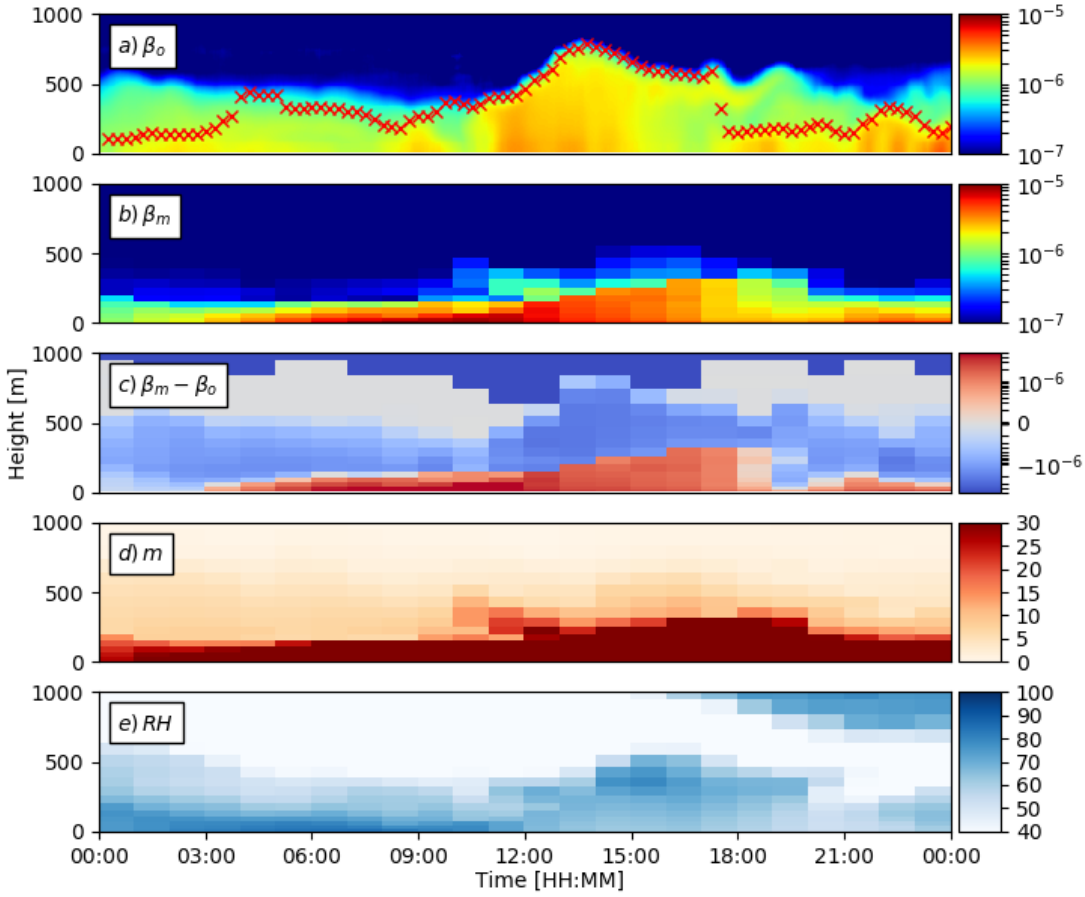

Figure 2.C.1: Vertical profiles of attenuated backscatter [m^{-1} sr^{-1}] that are a) observed (β_o) with estimated mixed layer height (red crosses, Kotthaus and Grimmond, 2017) and b) forward modelled (β_m) with the aerFO (section 2.4). c) Attenuated backscatter absolute difference ($\Delta\beta = \beta_m - \beta_o$) calculated using the hourly β_m vertical profile and the vertical profile of β_o nearest in time; d) UKV aerosol mass mixing ratio (m) [$\mu g\ kg^{-1}$] and e) UKV relative humidity (RH) [%]. All fields are for 19 January 2016 at MR (Figure 2-8). Peak values of m are ~140 $\mu g\ kg^{-1}$. Note that (a-b) are plotted on a log scale and panel (c) colour bar has a range of $\pm 5 \times 10^{-6}$ m^{-1} sr^{-1} on a symmetric log scale.

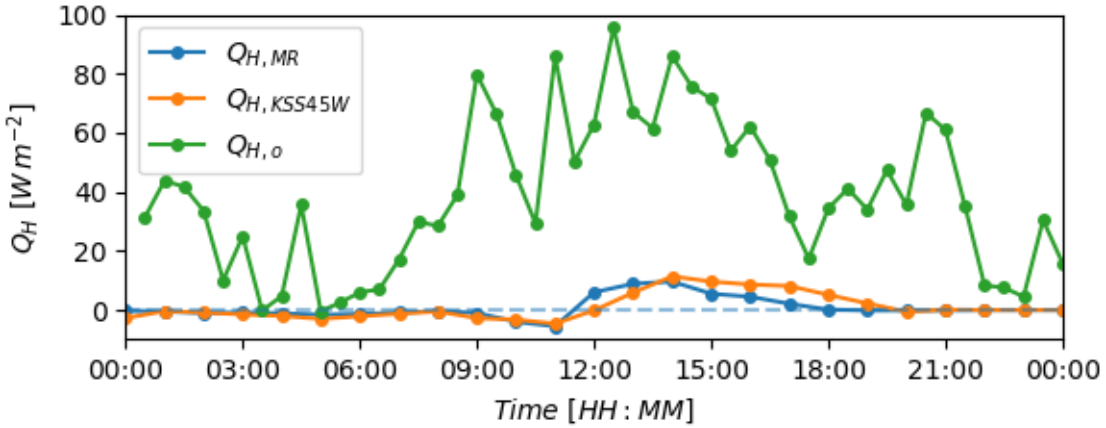

Figure 2.C.2: Sensible heat flux (Q_H) [W m^{-2}] from the UKV at MR ($Q_{H,MR}$) and KSS45W ($Q_{H,KSS45W}$) both at a height of 73.0 m agl, and observed with a CSAT3 sonic anemometer ($Q_{H,o}$) at KSSW (Figure 2-8) for the 19 January 2016.

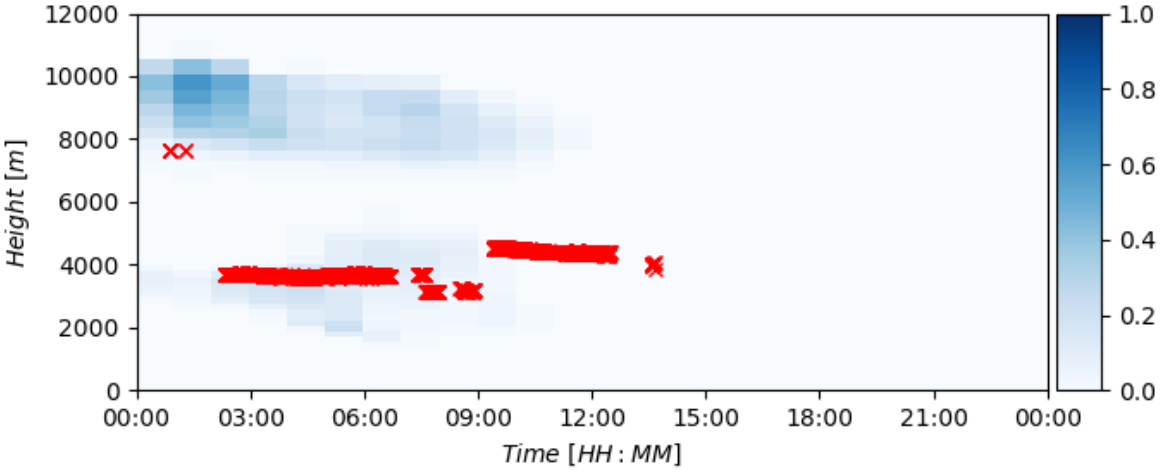

Figure 2.C.3: Cloud volume fraction from the UKV [fraction] (blue) and cloud base height from CL31-C (red crosses) for MR on 19 January 2016. UKV heights are model level heights, CL31-C heights are height above ground.

Appendix 2.D Variation in ratio of modelled to observed attenuated backscatter with respect to modelled relative humidity and aerosol.

The β_m/β_o ratio versus *RH* with m_{MURK} indicated (colour) allows further understanding of how the two attenuated backscatters vary more generally (Figure 2.D.1). β_m/β_o is often low (mean ratio = 0.53 at MR), with the smallest values typically occurring when m_{MURK} was relatively low (< ~ 30 µg kg^{-1}). Lower m_{MURK} will lead to smaller mean dry radii (r_{md}) and mean number concentration (N) estimates, which generate a lower extinction coefficient and β_m. Consequently, this also limits the effect of *RH* as it is parameterised as an extinction enhancement; this means that β_m/β_o with relatively low m_{MURK} values do not increase greatly with increasing *RH* (Figure 2-14). β_m/β_o increases with larger values of m_{MURK} (> ~ 30 µg kg^{-1}), which reflects the larger estimates of r_{md} and N values and the relatively larger impact of extinction enhancement from *RH*. In addition, the increased sensitivity to *RH* at higher m_{MURK} leads to an increased scatter, as *RH* errors have a larger impact on β_m.

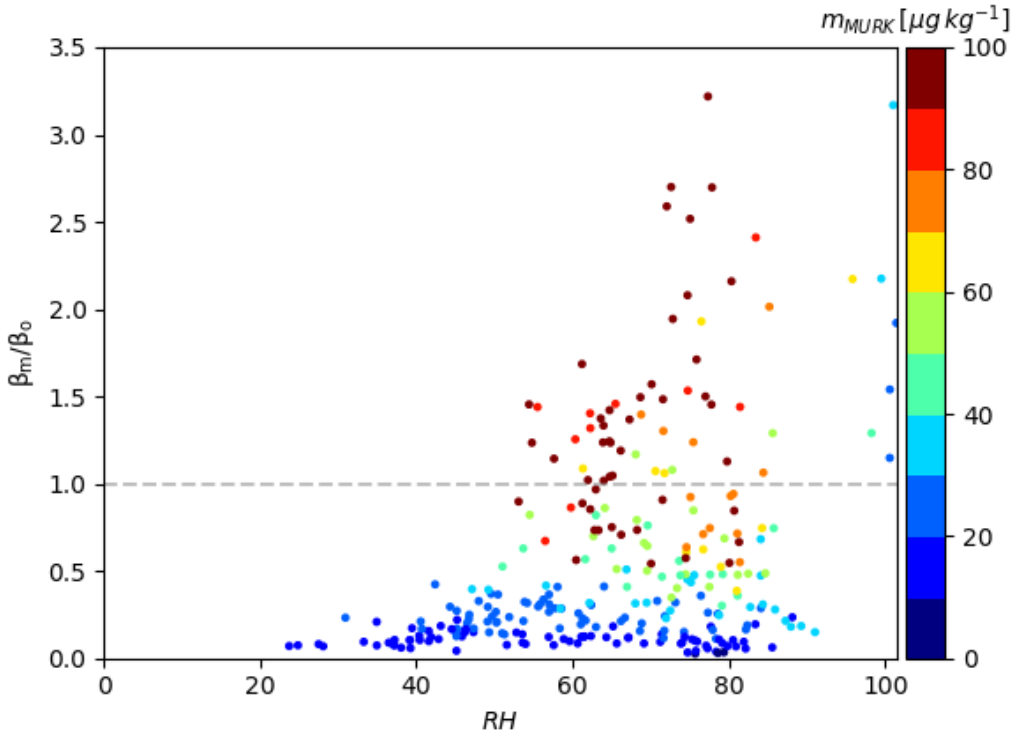

Figure 2.D.1: Ratio of modelled attenuated backscatter to observed attenuated backscatter (β_m / β_o) verses UKV relative humidity (RH), with the UKV aerosol (m, colour) indicated, for MR. β_m values for 21.7 m agl, which corresponds to the ALC's second range gate. m and RH for 5 m agl. Nearest time of β_o vertical profile is compared to each β_m profile (261 profiles, from 12 days, Table 2.3).

Chapter 3 Observed aerosol characteristics to improve forward-modelled attenuated backscatter in urban areas

Abstract

Numerical weather prediction (NWP) models often parameterise aerosols to reduce computational needs, while aiming to accurately capture their impact adequately. Increasingly, aerosols are monitored *in-situ* directly and/or indirectly (e.g. by automatic lidars and ceilometers, ALC). ALC measure the aerosol optical characteristic of attenuated backscatter. This can also be estimated using forward models that combine forecast aerosol and relative humidity to parameterise aerosol physical and optical characteristics. The aerFO is one such forward model, designed to use Met Office NWP model output and parameterisations from the MURK visibility scheme. Given the aerFO-MURK scheme link, assessing the aerFO and its output could therefore be used to inform future developments of the MURK scheme. To identify which parameterised physical and optical aerosol characteristics in the scheme are the most critical in urban settings, aerFO is driven with different *in-situ* aerosol observations at a background site in central London. Estimated attenuated backscatter is then assessed against ALC observations. It is shown that the original MURK scheme parameterisation underestimates the variance of both dry mean volume radius and total number concentration. Representing both the accumulation and coarse mode aerosols in the aerFO reduces the median bias error of estimated attenuated backscatter by 69.1 %. Providing more realistic temporal (monthly to hourly) variability of relative mass for different species leads to little improvement, compared to using monthly climatological means. Numerical experiments show that having more realistic estimates of number concentration is more important than providing more accurate values of the dry mean volume radius for the accumulation mode. Hence, improving the parameterisations for number concentration should be a main focus for further development of the MURK scheme. To estimate aerosol attenuated backscatter, the aerFO requires an extinction-to-backscatter ratio (i.e. the lidar ratio). In addition to forward modelling, the lidar ratio can also be used with ALC attenuated backscatter to calculate aerosol properties estimated in aerosol forecasts. Here, a model is developed that estimates the ratio using *in-situ* observations of the number, size distribution and speciation of aerosol masses. The values of lidar ratio derived at the London background site (14 – 80 sr across selected common lidar wavelengths) compare well to the literature. However, the modelled lidar ratio is unexpectedly correlated to relative humidity. Further, a stronger dependence exists at shorter wavelengths (355 and 532 nm) compared to longer wavelengths (905 and 1064 nm); this is due to the critical relation of lidar wavelength to aerosol size.

3.1 Introduction

Forecasting aerosol characteristics accurately is important, due to their impact on radiation budgets and visibility (Roessler and Faxvog, 1981; Seinfeld and Pandis, 2016; Stull, 1988). Understanding the impact of various aerosol characteristics on radiation forecast accuracy would be beneficial for the development of efficient aerosol parameterisations while minimising additional computational resources, as some aerosol species may impact radiation differently.

Currently, it is computationally expensive to use numerical weather prediction (NWP) models with explicit representations of multiple aerosols characteristics, such as mass by species and total number concentration. NWP aims to balance computational cost and forecast accuracy by including aerosol characteristics that are the most critical for the latter. For example, the ECMWF Integrated Forecasting System (IFS) (native resolution up to ~ 9 km) contains mass mixing ratios for sea salt, desert dust, organic matter, elemental carbon and sulphate separately (Benedetti et al., 2009; ECMWF, 2018).

High-resolution Met Office (MO) NWP models, such as the MO UKV (variable resolution 1.5 km inner, 4 km outer domain) (Tang et al., 2013) only explicitly include the total dry mass of bulk aerosol to minimise computational expense (Clark et al., 2008). All other characteristics in the MO UKV, such as number concentration and mass mean radius, are parameterised from the estimated mass through the 'MURK' visibility scheme, which produces a visibility diagnostic from the aerosol characteristics and humidity. Estimating the multiple aerosol characteristics with greater accuracy could help improve MO operational weather forecasts for visibility, and potentially radiation, if aerosol is later coupled to the radiation scheme.

Increasingly, *in-situ* observations in cities include multi-site campaigns to understand spatio-temporal variations in aerosol characteristics, such as particle mass (*PM*, total mass of particulate matter below a given size), size distributions and composition (Crilley et al., 2017; Hama et al., 2017a; Harrison et al., 2012; Klompmaker et al., 2015; Young et al., 2015). Extended routine aerosol monitoring provides further information about inter-annual variability with a wider range of meteorological conditions (Hofman et al., 2016; Ruths et al., 2014); relative importance of aerosol sources in cities that vary both temporally and spatially with regional background (Abdalmogith and Harrison, 2006); long-range transport (Abdalmogith and Harrison, 2005); and local sources within cities such as transport, construction and fuel burning (Liu et al., 2014; Ye et al., 2013). Interaction between these sources creates a highly diverse aerosol composition with spatial and temporal (hourly to interannual) variations (Stachlewska et al., 2018).

Alongside *in-situ* aerosol observations, automatic lidars and ceilometers (ALC; Wiegner et al., 2014) are being deployed (Flentje et al., 2010b; Illingworth et al., 2007; Met Office, 2017; Osborne et al., 2018; Pappalardo et al., 2014). Further, ceilometers are becoming increasingly sensitive with

measurements and becoming more comparable to sophisticated lidars, such as Raman (Heese et al., 2010; Madonna et al., 2018, 2014; Song et al., 2017). ALC observations have been used to derive information on aerosol and meteorological conditions, including mixed layer height (Kotthaus and Grimmond, 2018b; Wagner and Schäfer, 2015) and, in the absence of hydrometeors (cloud, precipitation and fog), relations between attenuated backscatter (β) and *PM* concentrations (Münkel et al., 2004; You et al., 2016). Calibrated and corrected ALC measurements provide the optical property of β, rather than a direct measurement of physical aerosol characteristics. Thus, aerosol forward operators (FO) are required to relate aerosol physical properties to β. They allow for the estimation of attenuated backscatter (β_m) from both NWP or chemistry transport model output of aerosol parameters for use in model evaluation or in data assimilation (Benedetti and Dabas, 2016; Chan et al., 2018; Charlton-Perez et al., 2016; Geisinger et al., 2017). One such forward operator (hereafter aerFO, section 2.4), uses a bulk mass of dry aerosol (m) [kg kg^{-1}] and *RH* as inputs, to estimate β_m for clear-sky conditions (i.e. without hydrometeors). aerFO, originally built for the MO MURK visibility scheme (Clark et al., 2008), shares some parameterisations with the MURK scheme to estimate aerosol physical properties, but with a revised approach to estimate aerosol optical properties for multiple laser wavelengths.

One important aerosol optical property is the ratio between particle extinction and particle backscatter, called the lidar ratio (S) (Song et al., 2018), and depends on the refractive index and size distribution of the aerosol particles (Müller et al., 2007). S can be used to solve the lidar equation with elastic-backscatter lidars, to recover the particle extinction coefficient (Fernald, 1984; Klett, 1981), or to derive further aerosol properties including aerosol optical depth (AOD) (Seinfeld and Pandis, 2016). S has been calculated both directly from Raman lidars (Müller et al., 2007; Papayannis et al., 2005; Wang et al., 2016) or high spectral resolution lidars (Burton et al., 2012; Hair et al., 2008; Reid et al., 2017), and indirectly from sunphotometers (Cattrall et al., 2005). However, all instruments have inherent limitations. Due to the high cost, very few Raman lidars are available and usually they are not operated continuously. Further, they only achieve complete optical overlap several hundred metres above the surface (Li et al., 2016; Wandinger and Ansmann, 2002) so near-surface aerosols are not well observed. Sunphotometers measure AOD, an integrated value for the total atmospheric column, that cannot differentiate near-surface contributions from those aloft. Hence, to derive S using a sunphotometer requires additional aerosol modelling (with inherent assumptions) without additional instrumentation (Dubovik and King, 2000). Furthermore, few instruments with the capability to derive S are being operated in urban measurement networks because of general constraints such as accessibility, communications and security (Muller et al., 2013). As the lidar ratio in cities is usually lower than that for background continental aerosol types (Müller et al., 2007), applying values representative of rural settings may create uncertainty in urban studies of derived aerosol optical properties. Hence, further analysis of urban-based aerosol

measurements (i.e. beyond Raman lidars or sunphotometers) could help determine lidar ratio values that are more representative of urban settings. *S* has been derived from aerosol observations using Monte-Carlo analysis with assumed aerosol particle size distributions and refractive indices (Barnaba et al., 2007; Dionisi et al., 2018), though without accounting for aerosol speciation. This may introduce uncertainty as aerosol species have widely varying optical properties (section 2.6.2). The Monte-Carlo values are broadly applicable to continental aerosol, but without the temporal variability of aerosol characteristics typical in urban environments (Hama et al., 2017b). A new approach to derive *S* from urban aerosol observations could improve the understanding of *S* variability in cities.

This study is structured in two parts. First, a method is presented to estimate *S* at a range of common lidar wavelengths based on *in-situ* aerosol observations collected at an urban background site. Second, aerFO is used to assess the relative importance of different aerosol properties in the accurate estimation of forward modelled attenuated backscatter (β_m). For this, parameterised aerosol characteristics are successively replaced by more realistic estimates derived from observations. Results can then inform future design of aerosol schemes in MO NWP models with respect to estimating optical properties, by suggesting which aerosol characteristics should be prioritised for inclusion or improvement in the NWP models.

3.2 Methods

An improved version of the forward operator, aerFO (first outlined in section 2.4), is used to model vertical profiles of attenuated backscatter. Improvements to the aerFO look-up-tables (LUT) are based on observations taken at North Kensington (NK) (Figure 3-1) and outlined in section 3.2.3.

3.2.1 Observations

Measurements of attenuated backscatter (β_o), aerosol properties and relative humidity (*RH*) are collected at an urban background site (NK) in the London Air Quality Network (LAQN) (DEFRA, 2018a).

Vertical profiles of β_o are observed with a Vaisala CL31 ALC (Table 3.1), that is part of the London Urban Meteorological Observatory (LUMO; http://micromet.reading.ac.uk/), and has been used in pervious boundary layer studies (Kotthaus et al., 2018, 2016). β_o is corrected for instrument-related background and near-range artefacts (Kotthaus et al., 2016). A centred moving average is applied in time and space, using 25 min windows (101 time steps) and 110 m (11 range gates), to increase the signal-to-noise ratio (SNR) (section 4.2 in Kotthaus et al., 2016). The CL31 reaches complete overlap at 70 m (Kotthaus et al., 2016). Insufficient optical overlap is corrected for internally by the sensor firmware. β_o from the second range gate is used in this study (20 m above ground level), which is

approximately the average of the lowest 7 range gates (70 m). β_o from the first range gate is not used due to its relatively high noise (Kotthaus et al., 2016).

Calibration coefficients were computed for periods using the 'stratus cloud' method, which has been shown to produce stable estimates (±5 % over 2015) (Hopkin et al., 2019). Coefficients produced here are interpolated for days between status conditions, with key changes in window transmission, instrument hardware and software changes accounted for (section 2.5.1). Daily calibration coefficients are applied to the smoothed profiles.

The manufacturer-specified central wavelength of the CL31 is 905 nm. The central wavelength uncertainty is ± 10 nm and full width half maxima (FWHM) is 4 nm (Kotthaus et al., 2016; Wiegner and Gasteiger, 2015). To incorporate this wavelength spectrum into our estimates of aerosol optical properties, we use the FWHM to define a Gaussian weighting function to calculate the dry Mie extinction efficiency ($Q_{ext,dry}$), extinction enhancement factor ($f_{ext,rh}$) with a dependency on *RH*, and the water vapour extinction coefficient ($\sigma_{ext,wv}$) with a dependency on ambient water vapour mass mixing ratio (r_v) (section 3.2.3 gives an updated overview of aerFO variables).

Comparison of aerFO-derived β_m and observed β_o is undertaken using seven clear-sky days when all variables (aerosol variables, backscatter and *RH*) are observed, ensuring a fair comparison between all the aerFO experiments. Clear-sky days were defined as those where the number of observed backscatter profiles with cloud, from each ceilometer between midnight and midnight, was below 5 % . Although this intentionally strict criterion reduces the number of case studies, it minimises the impact of cloud on local radiation budgets and consequently simplifies aerFO analysis. More general application of the aerFO is less stringent on cloud amount, for example forward modelling aerosol backscatter below cloud.

As no *RH* observations are conducted at the NK site, measurements from a different central London LUMO site (KCL; 6.8 km southeast of NK, Figure 3-1) are used instead as input to parameterise the swelling and drying of aerosol particles. To assess how well the KCL *RH* measurements also represent NK, Pearson correlation between two MO UKV NWP model grid boxes containing the respective sites is calculated for hourly *RH* from 30 clear sky days from 2014 to 2015. *RH* at both locations varied between 40 and 95 % and a good correlation (> 0.99) was found (Table 3.1). An additional 2 clear days were present in between 2014 and 2015 but had near saturated conditions and were not included in the sample.

With measurements of the mass of ammonium sulphate and ammonium nitrate aerosol unavailable, we use the CLASSIC aerosol scheme (Bellouin et al., 2011; Martin et al., 2006) to model these with the observed mass of reactants (Table 3.2). This assumes the entire ammonium (NH_4) reacts with sulphate (SO_4) preferentially, and only the remaining NH_4 reacts with nitrate (NO_3). Sea salt mass is calculated from observed chlorine (Cl) mass, assuming all Cl atoms were contained in sea salt.

Ammonium nitrate, ammonium sulphate and sea salt aerosol 15 min mass data are averaged to hourly resolution. To obtain hourly masses of elemental carbon (EC) and organic carbon (OC), the daily masses are linearly interpolated due to the lack of additional information. As a large proportion of EC is black carbon (strongly absorbing particles, Briggs and Long, 2016), the density and shape factor of black carbon have been used from the literature for EC.

Two aerosol particle diameter size ranges are measured at NK (Table 3.1): (i) *Smaller sizes* (0.016 – 0.6 μm): Scanning Mobility Particle Sizer 3081 (SMPS; TSI, 2009) with a Condensation Particle Counter 3775 (CPC; TSI, 2007) operated by King's College London Environmental Research Group (ERG) and National Physics Laboratory (NPL) measures dried particle diameters.

(ii) *Larger sizes* (0.5 – 20 μm): Aerodynamic Particle Sizer 3321 (APS; TSI, 2004) operated by DEFRA. Measurements of ionic components of PM_{10} by an URG-9000B Ambient Ion Monitor (DEFRA, 2018b; URG, 2011) are used to determine ammonium nitrate, ammonium sulphate and sea salt aerosol proportions. PM_{10} is sampled onto filters (Tissuquartz™ 2500 QAT-UP) with a Partisol 2025 sequential air sampler (DEFRA, 2018b; Thermo Fisher Scientific, 2015) and analysed for EC and OC using a Sunset Laboratory thermal–optical analyser according to the QUARTZ protocol (with very similar results to EUSAAR 2: Cavalli et al., 2010) (Beccaceci et al., 2013). Particles are measured at ambient *RH*.

Table 3.1: Observations from North Kensington (NK) (Figure 3-1) operated by LUMO (Kotthaus and Grimmond, 2018b, 2014) *and Environmental Research Group (ERG) LAQN (Mittal et al., 2016). † Particle number distribution data from two instruments (by diameter size) are combined to extend the range of particle sizes used to estimate aerosol optical properties (section 3.2.3). * RH from King's College London (KCL) is used.*

Variables		Instrument	Sampling	Period	Source
Attenuated backscatter β_o		Vaisala CL31	15 s, 10 m	01/01/2014 – 31/12/2015	LUMO NK
PM_{10} by species	Ammonium sulphate $[(NH_4)_2SO_4]$	URG-9000B AIM	15 min	02/02/2011 – 07/02/2018	ERG/ NPL NK
	Ammonium nitrate $[NH_4NO_3]$		15 min		
	Sea salt [NaCl]		15 min		
	Elemental carbon (EC)	Partisol 2025/Sunset	Daily		
	Organic carbon (OC)		Daily		
Particle number distribution †		SMPS 3081 (with CPC 3775)	~0.016 – 0.6 μm, 51 size bins (small)	01/01/2014 - 31/12/2015	ERG/ NPL NK
		APS	~0.5 – 20.0 μm, 52 size bins (large)	01/01/2014 – 31/12/2015	ERG UK
Relative Humidity *RH* *		Vaisala WXT520	5 s	01/01/2014 – 31/12/2015	LUMO KCL

*Table 3.2: Aerosol species modelled: ammonium sulphate, ammonium nitrate and sea salt. Mass calculated from respective reactants observed. † Elemental carbon assumed to be completely hygrophobic. * Efflorescence and deliquescence limits taken from swelling method source. Varutbangkul et al. (2006) found no efflorescence or deliquescence limits for organic carbon. ^ shape factor (χ) used if particles are on the 'dry' branch of the hysteresis curve. See Table 3.1 for aerosols species. Sources: Fi75 Fitzgerald (1975) Sc07 Schkolnik et al. (2007), SP16 Seinfeld and Pandis (2016), Va06 Varutbangkul et al. (2006), Ze06 Zelenyuk et al. (2006), Zh16 Zhang et al. (2016),*

Aerosol species	Estimated from observed reactants	Swelling method applied	Efflorescence * relative humidity [%]	Deliquescence * relative humidity [%]	Dynamic shape factor		Density	
					χ ^	Source	ρ_p [kg m^{-3}]	Source
$(NH_4)_2SO_4$	NH_4, SO_4	Fi75	30	81	1.0	SP16	1770	n/a
NH_4NO_3	NH_4, NO_3	Fi75	30	61	1.0	n/a	1720	n/a
OC	n/a	Va06	n/a	n/a	1.0	Ze06	1100	Sc07
EC	n/a	n/a †	n/a	n/a	1.2	Zh16	1200	Zh16
NaCl	Cl	Fi75	42	75	1.08	SP16	2160	n/a

The SMPS and APS datasets together cover a large range of particle sizes, enabling a more representative estimate of aerosol physical and optical properties to be derived. Unfortunately, the SMPS and APS observe aerosol particle diameters in slightly different ways. The SMPS provides the mobility diameter (D_m) or particle diameter relative to a perfect sphere with the same electrical mobility in a constant electrical field (DeCarlo et al., 2004). The APS provides aerodynamic diameter (D_a) or diameter relative to an aerodynamic property measured (DeCarlo et al., 2004). Therefore, to merge these datasets together, both are first converted to a common particle size, i.e. the volume equivalent diameter (D_v). D_v is based on particle surface area, and chosen as it is more closely related to aerosol optical properties (e.g. extinction coefficient) than D_m or D_a (Seinfeld and Pandis, 2016). The diameter conversions to D_v are done prior to computing particle optical properties based on Mie theory.

The diameter conversion equations require additional aerosol parameters that can vary slightly by species (Table 3.2). For instance, the dynamic shape factor (χ) is needed to convert D_m and D_a to D_v, and the particle density (ρ_p) is also required to convert D_a to D_v. χ describes the sphericity of a particle: $\chi = 1$ for a perfect sphere and $\chi > 1$ for non-spherical particles (Hinds, 1999). Often the Cunningham slip correction factor (C_c) is used when converting between equivalent diameters to account for the no-slip condition for small particles (Seinfeld and Pandis, 2016). However, we assume C_c can be ignored as the contribution by small particles (< ~ 0.07 μm) to the optical properties of total aerosol is negligible (section 3.3.2).

From the merged number distribution, three aerosol mode ranges are defined (D_v: fine < 0.08 μm; accumulation 0.08 – 0.8 μm; coarse > 0.8 μm). These are interpreted based on the mean dV/dlogD

of the dry particle number distribution (where V is the volume) between 1 January 2014 – 31 December 2015 at NK.

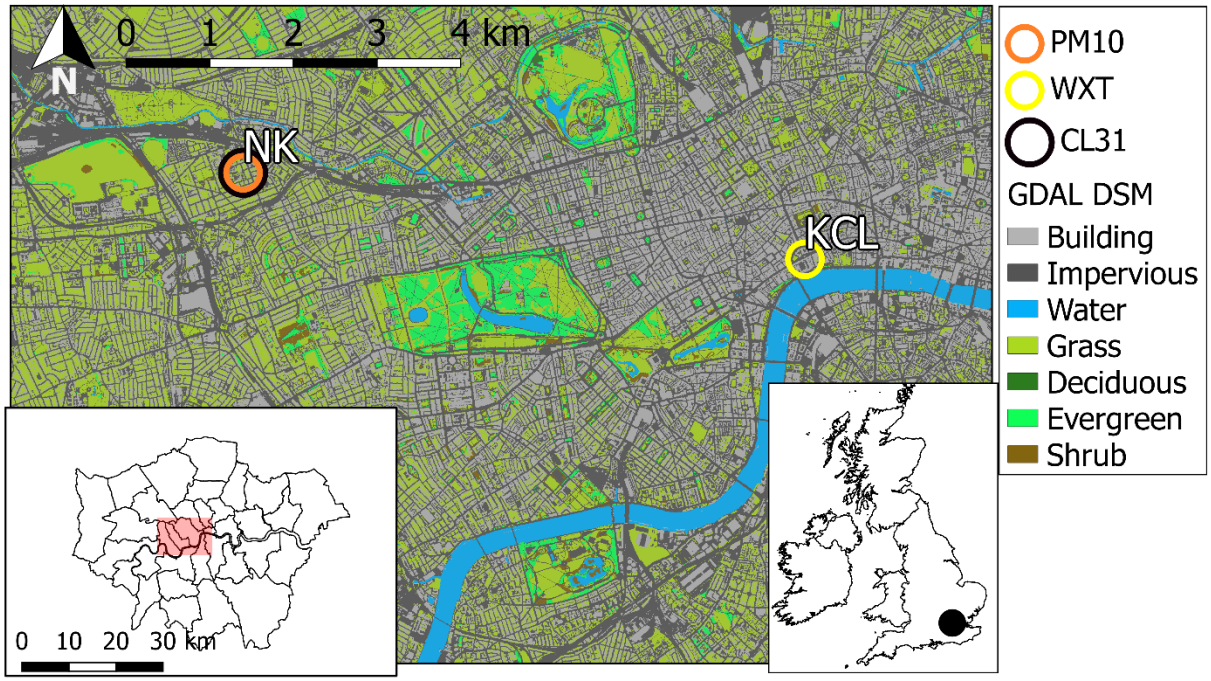

Figure 3-1: LUMO (CL31, WXT) and LAQN measurement sites in central London and surface cover type (data source described in Lindberg and Grimmond, 2011); within Greater London and the British Isles (insets).

3.2.2 Using observations to model the lidar ratio

The extinction-to-backscatter ratio, or lidar ratio S, is a key component of aerFO as it links the modelled extinction to the modelled backscatter. As is common (Doherty et al., 1999; Müller et al., 2007), the aerFO used in Chapter 2 assumed a constant value (S = 60 sr^{-1}, typical of continental aerosol) to estimate β_m from σ_{ext}. Here, a more detailed parameterisation is developed to model S using aerosol observations of particle size distribution, speciated aerosol mass and RH. S in aerFO can then be updated with estimates from the parameterisation.

To estimate S (Figure 3-2), the overall approach is to model the aerosol optical properties based on the aerosol physical properties, including the observed dry number distribution by aerosol species ($N_{aer}(D)$) binned by diameter (D). The extinction and backscatter properties for each aerosol species (Table 3.2) are modelled separately, assuming an external aerosol mixture. Hygroscopic growth and consequential changes in the complex index of refraction of each aerosol species, from additional water, are also parameterised.

Initially, with only non-speciated observations of dry number distribution by diameter ($N_{obs}(D)$) available, $N_{aer}(D)$ are estimated by speciating $N_{obs}(D)$ using observed hourly PM_{10} mass of different aerosol types (m_{aer}, Figure 3-2, part 1). m_{aer} is used with observed aerosol density of each species from the literature (ρ_{aer}, Table 3.2) to create volumetric weights ($V_{weight,aer}$; ratio of each species (V_{aer})

to total aerosol (V_{tot}) volume). The $V_{weight,aer}$ provide *a priori* weights for each bin of $N_{obs}(D)$ to obtain $N_{aer}(D)$ *by* species:

$$N_{aer} = N_{obs}(V_{aer}/V_{tot}) \quad (3.1)$$

After estimating $N_{aer}(D)$ for each species, the hygroscopic effects on aerosol particles are parameterised. As *RH* increases, hydrophilic aerosol particles swell (i.e. increase in size) with water, and the complex index of refraction changes, reaching some state between that of a dry particle and a water droplet (Haeffelin et al., 2016). For hydrophilic species, the parameterised physical growth of each binned diameter follows the CLASSIC aerosol scheme (Bellouin et al., 2011) such that the swollen diameters $D_{wet,aer}$ and dry diameters $D_{dry,aer}$ are calculated using ambient *RH*. In this study, the hygroscopic parameterisations require the observed dry SMPS particle sizes to be swollen to find $D_{wet,aer}$, and the observed humidified APS particle sizes to be dried to find $D_{dry,aer}$. Once $D_{dry,aer}$ and $D_{wet,aer}$ for the merged number distribution are calculated, the physical growth factor g_{aer} is estimated (Figure 3-2 part 2). g_{aer} is used with the size parameter X_{aer} (eqn. 15.6 in Seinfeld and Pandis, 2016) for each size bin and each species, to estimate the mixed complex index of refraction between that of dry aerosol and of pure water ($n_{wet,aer}$) using the volume mixing method (Liu and Daum, 2008). Once $D_{wet,aer}$ and the mixed complex index of refraction are calculated, Mie scattering code is used with the swollen particles to calculate the optical properties: extinction and backscatter efficiencies ($Q_{ext,aer}$, $Q_{back,aer}$), and extinction and backscatter cross-sections ($C_{ext,aer}$, $C_{back,aer}$). The Mie scattering code used was written with a module for Python 2.7 (pymiecoated, version 0.1.1: https://code.google.com/archive/p/pymiecoated/). Following this, the total extinction and backscatter for each aerosol species ($\sigma_{ext,aer}$, $\sigma_{back,aer}$) is calculated using $N_{aer}(D)$ with $C_{ext,aer}$ and $C_{back,aer,}$ respectively. The sum of $\sigma_{ext,aer}$, and $\sigma_{back,aer}$ is used to compute the overall extinction (σ_{ext}) and backscatter (σ_{back}). Finally, *S* is computed as the ratio of σ_{ext} to σ_{back} (Figure 3-2, part 3).

As optical properties of aerosol particles are highly sensitive to particle size (Jacobson, 2005), the number and arbitrary position of diameter bins in the observed number distribution can affect the final modelled $\sigma_{ext,aer}$, $\sigma_{back,aer}$ and *S* results. To minimise the impact of number and arbitrary bin diameter sizes on optical property results, we follow Geisinger et al. (2017 section 2.2.4) in part 2 (Figure 3-2). This increases the total number of diameter bins by a factor of 4, by interpolating between the adjacent diameter bins. X_{aer} and $n_{wet,aer}$ are calculated for each interpolated bin, and the optical properties for each original bin (D_i) are derived as the average of the interpolated bins for that size interval.

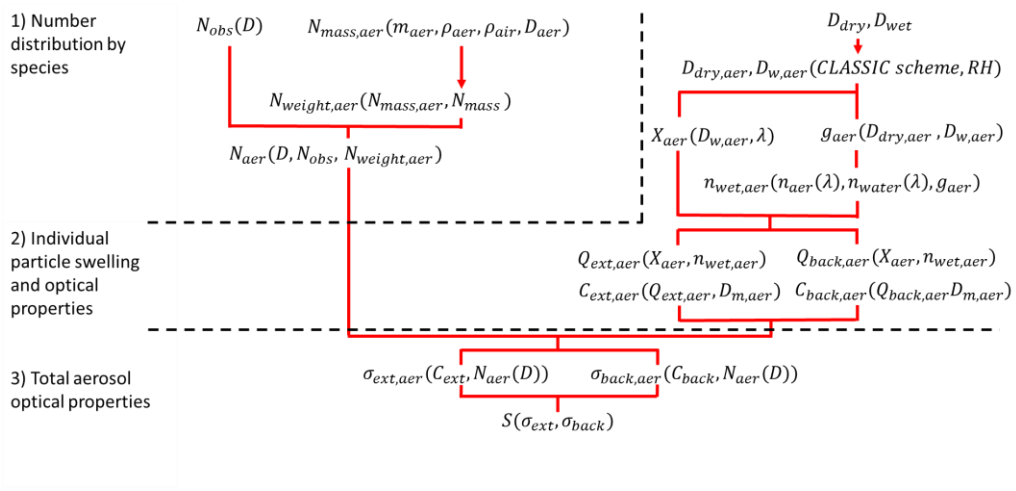

Figure 3-2: Lidar ratio (S) calculation flow chart. The main inputs are diameter bins (D), number distribution by diameter ($N_{obs}(D)$), and aerosol species mass at a given size range (m_{aer}), e.g. PM_{10}, and relative humidity (RH). See Chapter 3 notation and abbreviations table for a full notation list.

For some particles (both pure and mixed species aerosol), particle hygroscopic growth has been observed to exhibit hysteresis (Figure 3-3) with a differing shrinking and growing relation as *RH* changes (Fierz-Schmidhauser et al., 2010; Tang and Munkelwitz, 1994). Thus, these particles have no simple relation between *RH* and their extinction coefficient. For example, many particles that are originally dry, solid and without any water condensation on them, can be defined as being below their efflorescence *RH* limit and on the lower branch (Figure 3-3, red). Particles already swollen through hygroscopic growth that are liquid solutions containing aerosol, can be defined as being above their deliquescent *RH* limit and on the upper branch (Figure 3-3, blue). When ambient conditions change so that *RH* is between these limits, the particles may not shrink or swell until they reach the other *RH* limit (Figure 3-3, grey). Particles originally above the deliquescence *RH* limit may remain relatively large, despite a decrease in *RH* until the critical efflorescence limit is reached, at which point the particle suddenly crystallises and loses all remaining condensed water (Tang and Munkelwitz, 1994). Conversely, dry particles below the efflorescence *RH* limit may stay dry and solid as ambient *RH* increases until reaching the deliquescent *RH* limit, then suddenly swell with water.

Given the importance of particle size on aerosol optical properties, the hysteresis effect was incorporated into the model to estimate *S*. The magnitude of the hysteresis effect for ambient aerosol varies by aerosol species (Figure 5 in Fierz-Schmidhauser et al., 2010). Consequently, we parameterise the hysteresis effect for ammonium sulphate, ammonium nitrate and sea salt separately (Table 3.2; Figure 3-3). Originally dry particles (beginning below efflorescence limit) did not swell and were treated as being in a 'dry' crystalline state (s_1, s_2). Once *RH* exceeds the deliquescence limit (per aerosol type) the parameterised particles are considered to be in a 'wet' liquid state (s_3, s_4) and their growth follows the aerosol type swelling method (Table 3.2). While in the 'wet' state, the growth of particles continues due to swelling until *RH* falls below the efflorescence limit. Particles

then return to a 'dry' state and no swelling method is applied. Hysteresis is not assumed for OC as its efflorescence and deliquescence limits are unclear and the hygroscopic growth has been observed to be smooth without an apparent hysteresis (Varutbangkul et al., 2006).

The magnitude of the hysteresis effect is based on the difference between s_3 and s_2 and varies by aerosol species and dry particle size. More hygroscopic species have the largest hysteresis effect because of the larger difference of g_{aer} between the s_3 and s_2 branches. For instance, dry 1.1 $\times 10^{-8}$ µm ammonium sulphate particles on the s_2 branch can be ~ 30 % smaller compared to swollen ammonium sulphate particles on the upper s_3 branch. Further, dry 1.1 $\times 10^{-8}$ µm sea salt particles can be twice as small compared to swollen sea salt particles on the upper s_3 branch.

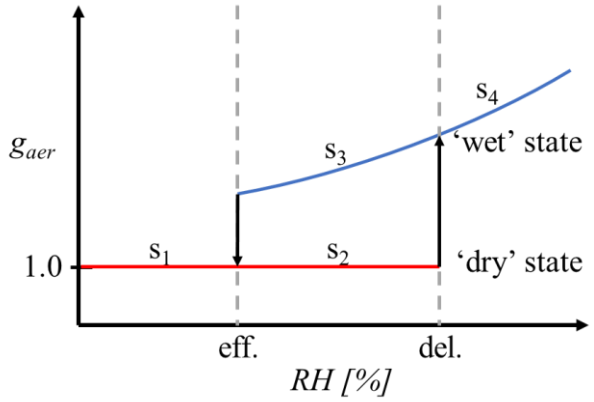

Figure 3-3: Schematic of the particle physical growth factor (g_{aer}) hysteresis with RH. A particle is assumed to be in either a dry (s2) or wet (s3) state when between the efflorescence (eff) and deliquescence (del) RH limits. 'Dry' particles at s1 move to s2. Once the ambient RH reaches the deliquescence RH limit particles move from s2 to the s4 branch. 'Wet' particles at s4 move to s3 and once the ambient RH reaches the efflorescence RH limit the particles have a dry branch response (s1). This is based on observations presented in Table 3.2.

3.2.3 Overview of aerFO

A complete description of the original aerFO was presented in section 2.4. In Chapter 3, an updated version of the aerFO is used with a brief outline presented in this section, including the updates.

aerFO is designed to estimate clear-sky β_m, using bulk mass of dry aerosol (m) [kg kg^{-1}] and RH from NWP as inputs. However, observations can also be used. Clear-sky conditions are defined as sub-saturated conditions based on the input RH, where no hydrometeors, such as cloud, fog or rain impact β_m, and where β_m is dominated by aerosol backscatter.

Initially, aerFO estimates the bulk aerosol physical properties of the aerosol accumulation mode, including dry mean radius (r_{md}) and total number concentration (N) using observed and empirically derived constants (Figure 2-7, part 1). These include the 'standard' climatological mean dry mass of

aerosol (m_0), climatological volume mean radius for the aerosol mode (r_0), climatological mean total number concentration for the aerosol mode (N_0) and a scaling factor (p). Following the MO visibility scheme (Clark et al., 2008) for the parameterisations of N and r_{md}, a fixed geometric standard deviation (d_g) of 1.7 is used for the accumulation mode as calculated from observations. Properties calculated include: aerosol optical (e.g. $Q_{ext,dry}$ and σ_{ext}; Figure 2-7, parts 2 – 3), backscatter and transmission (Figure 2-7, part 4). aerFO can represent multiple aerosol modes (e.g. fine, accumulation and coarse) by calculating $\sigma_{ext,aer}$ (Figure 2-7, parts 1 – 2) for each mode and adding them to obtain the overall $\sigma_{ext,aer}$. The effect of water vapour absorption ($\sigma_{ext,wv}$) is also included in the total extinction coefficient (σ_{ext}).

For computational efficiency, $Q_{ext,dry}$ and $f_{ext,rh}$ (Figure 2-7, part 2) are pre-calculated for aerFO. As the relative composition of aerosols may vary (e.g. by month, Young et al., 2015), a monthly climatology of urban aerosol composition is derived from NK aerosol observations (Table 3.2, section 3.2.1).

Look up tables (LUT) are derived for $f_{ext,rh}$ based on the geometric mean radius (r_g), using SOCRATES (Manners et al., 2015) Mie scattering code within the radiation suite. r_g is estimated from r_{md} using linear regression based on observations for the aerosol accumulation mode (Appendix 3.A). Relations between r_{md} and r_g are derived for each aerosol mode separately (e.g. fine, accumulation and coarse mode in experiment 5; section 3.3.2).

As shown in Chapter 2, N_0 is highly sensitive to the defined size range of the accumulation mode. Therefore, the default aerFO N_0 and the mean dry volume radius (r_0) are re-calculated for NK, from the mean observed particle number distribution for the accumulation mode (Table 3.3) from the SMPS and APS measurements (Table 3.1).

Table 3.3: Default input parameters for aerFO at NK (Figure 3-1) using all available data.

Site (type)	Accumulation radius size range [μm]	N_0 [cm^{-3}]	r_0 [μm]
NK (urban background)	0.04 – 0.4	1311	0.159

3.2.4 Aerosol forward operator experiments

The improvement in β_m estimates when using observed aerosol characteristics (e.g. size, total number) is evaluated using a set of aerFO experiments (Table 3.4). For these we use observations (rather than NWP model forecasts) as inputs to aerFO. Since observations replace parameterised variables, β_m is changed (Figure 2-7, part 1). For comparison, we define a control run (ID = 0, Table 3.4), with physical properties N and r_{md} (Figure 2-7, part 1) estimated based on aerFO parameterisations, with measured PM_{10} mass and RH (Table 3.1). In experiments 1 – 7 (ID = 1-7),

parameterised variables, such as N and r_{md}, are replaced by estimates from aerosol observations (Table 3.4). Experiments 1 – 3 consider the accumulation mode whereas experiments 4 – 7 address multiple aerosol modes (e.g. fine, accumulation and coarse).

For experiments 1-3 we do the following: the previously parameterised N is replaced by observed N for the accumulation mode (ID = 1); the parameterised r_{md} is replaced by observed r_{md} for the accumulation mode (ID = 2); both an observed N and r_{md} are provided (ID = 3) to eliminate the need for a mass estimate as input to aerFO. Each subsequent experiment (ID > 3) uses further observations and accounts for more aerosol modes. In experiment 4, using observed N and r_{md} for two (accumulation and fine) aerosol modes allows assessment of the importance of the fine mode to accurately calculate the aerosol optical properties. To represent multiple modes in aerFO (Figure 2-7, parts 1 -3) each mode is treated separately when obtaining the extinction coefficient. Experiment 5 uses observed N and r_{md} for the fine, accumulation and coarse aerosol modes. Experiment 6 is the same as 5, but with hourly aerosol composition used to calculate hourly varying $Q_{ext,dry}$ and $f_{ext,rh}$ LUT. Experiments 0 – 6 use a mean S (section 3.2.2), whereas 7 is the same as 6, but with hourly estimates of S.

The experiments use seven clear-sky days (in 2014: 06/06; 03/07; and 2015: 07/03; 20/04; 21/04; 04/06; 02/08) when all the required aerosol observational data are available for all the experiments. Although the evaluation of β_m is done at 905 nm, β_m can be computed for other ALC wavelengths (e.g. 355, 532, 1064 nm); this is not assessed as observations at other wavelengths are unavailable.

*Table 3.4: Experiments (ID) assumptions for number concentration (N) and optical properties as input to aerFO with which aerosol characteristics are replaced with observations measured at NK (Figure 3-1). All aerFO experiments use observed RH from KCL. Mode diameter ranges: < 0.08 µm (fine), 0.08 – 0.8 µm (accumulation) and > 0.8 µm (coarse). * control run.*

ID	Aerosol variables estimated based on observations	Aerosol observations used
*0**	$m = PM_{10}$	Particle mass
1	N (accumulation mode)	Particle number distribution (fine and coarse)
2	r_{md} (accumulation mode)	
3	N, r_{md} (accumulation mode)	
4	N, r_{md} (fine and accumulation mode)	
5	N, r_{md} (fine, accumulation and coarse mode)	
6	N, r_{md} (fine, accumulation and coarse mode), $Q_{ext,dry}$, $f_{ext,RH}$	Particle number distribution (fine and coarse) and mass by species
7	N, r_{md} (fine, accumulation and coarse mode), $Q_{ext,dry}$, $f_{ext,RH}$, S	

3.3 Results

3.3.1 Lidar ratio sensitivity to aerosol type and size

Aerosol lidar observations are commonly conducted at typical Raman lidar channels (e.g. 355, 532 nm) or at the wavelengths used by ALC (e.g. 905, 1064 nm). To account for differences in lidar wavelengths, the lidar ratio *S* is estimated from aerosol observations at NK at these four wavelengths (Figure 3-4). Across all wavelengths *S* varies between ~14 – 80 sr. This wide range of *S* and corresponding β (not shown) are similar to those obtained using aerosol characteristics with Monte-Carlo methods (Barnaba et al., 2007; Dionisi et al., 2018). The 905 nm *S* average of 43.1 sr is similar to that measured in urban areas by Raman lidars (Müller et al., 2007; Stachlewska et al., 2018; Wang et al., 2016), but is low compared to that derived using sunphotometer observations (Cattrall et al., 2005).

To date, there is little published research on the relation between *S* and *RH*. These experiments suggest that for these given parameters, *S* generally has a positive linear relation to *RH*. This is strongest at ALC wavelengths 905 and 1064 nm (Pearson correlation of 0.83 and 0.78, respectively). The dependence of *S* on *RH* (Figure 3-4) is linked to the relation between the aerosol size distribution and *S*, and likely to the relative position of the accumulation mode size range (Figure 3-5). For OC and any particles in the 'wet' deliquescent state (section 3.2.2), particles swell or shrink in response to changing *RH*, as water vapour condenses onto or evaporates from them. Consequently, particles move up or down the curve (Figure 3-3). Many particles in the size distribution lie within the accumulation mode, with the smaller half of the accumulation mode largely lying on a monotonically increasing part of the curve, before the first maxima (e.g. ~ 0.2 μm at 355 nm and 0.5 μm at 1064 nm for most aerosol types). Consequently, for the hygrophilic particles on the increasing part of the curve, as *RH* increases and the particles increase in size, many of the particles collectively move up the curve which acts to increase the overall *S* of the combined aerosol.

The variability in the relation between *RH* and *S* is further complicated by the variation in wavelength *S* is estimated for. In the accumulation mode size range (diameter: 0.08 – 0.8 μm), the number of inflection points varies with wavelength for all aerosol species considered (Figure 3-5). At the longer 905 and 1064 nm wavelengths (Figure 3-5c, d), the first local maximum of *S* is associated with larger particles compared to the Raman channels (Figure 3-5a, b) so there are fewer inflection points across the accumulation size range (Figure 3-5, dashed lines). At shorter wavelengths (355, 532 nm) there are more inflection points within the accumulation range causing a more variable response (greater scatter) of *S* to changes in *RH* (Figure 3-4a, b).

Additionally, *S* is influenced by the proportions of aerosol species at some wavelengths and humidities. At 355 (Figure 3-6) and 532 nm (not shown) and at low *RH*, *S* increases with greater relative EC abundance as EC absorbs radiation more efficiently than other aerosol types. However,

when *RH* increases, the relation between EC and *S* is less clear, this could be because the other hygroscopic aerosol types are in the 'wet' deliquescent state, so would have swollen and increased their size. Consequently, the swollen hygroscopic aerosol *S* values increase to a magnitude like or above those of dry (un-swollen) EC. Though the relation between EC and *S* is most pronounced at shorter wavelengths, it is unclear at longer wavelengths including at low *RH* (905, 1064 nm; not shown). In addition to the impact of aerosol proportion on *S*, a clear impact is found for $f_{ext,rh}$ at 355, 532, 905 and 1064 nm (905 nm shown in Appendix 3.B).

Some assumptions made in calculating *S* may reduce *S* variability with changing *RH* (Figure 3-4). Applying Mie theory based on spherical particles might limit variability in *S* for both EC and hygroscopic particles at low *RH*, before the hygroscopic particles become swollen (i.e. pre-spherical). Speciating *N(D)* evenly across all bins based on the total mass of each species available (section 3.2.2) likely does not portray the actual diversity in size distributions. This simplification does not allow individual species to contribute more to one specific mode; for example, sea salt could contribute more to the coarse mode.

In the parameterisation of *S* (Figure 3-2) internal aerosol mixing is not considered. This would lead to a wide range of efflorescence and deliquescence *RH* limits for different particles and hence change the particle physical growth factor with respect to water (g_{aer}). Omitting internal mixing likely means hysteresis-related variability is not fully represented. In laboratory conditions, Fierz-Schmidhauser et al. (2010) found a clear hysteresis effect on enhanced scattering from several pure aerosol species, but in field conditions with ambient aerosol hysteresis it was not always detected. Similarly, Smith et al. (2012) found that varying the proportions of organic material to ammonium sulphate changed the efflorescence and deliquescence *RH* limits of ammonium sulphate greatly, with any apparent *RH* limits eliminated once organic material volume fractions exceeded 0.6.

Evaluation of β_m (section 3.3.2) uses several different parameterisations of *S* (e.g. linear regression with *RH*) to help identify which should be the default in aerFO. Given this evaluation (not shown), as well as the remaining uncertainty and various causes for variability in estimating *S*, the mean *S* for each wavelength is used as the default value in aerFO. The mean *S* for 905 nm is 43.1 sr (median 44.1 sr), calculated from a sample size of 4699 hours (across 17 months in 2014 – 2015).

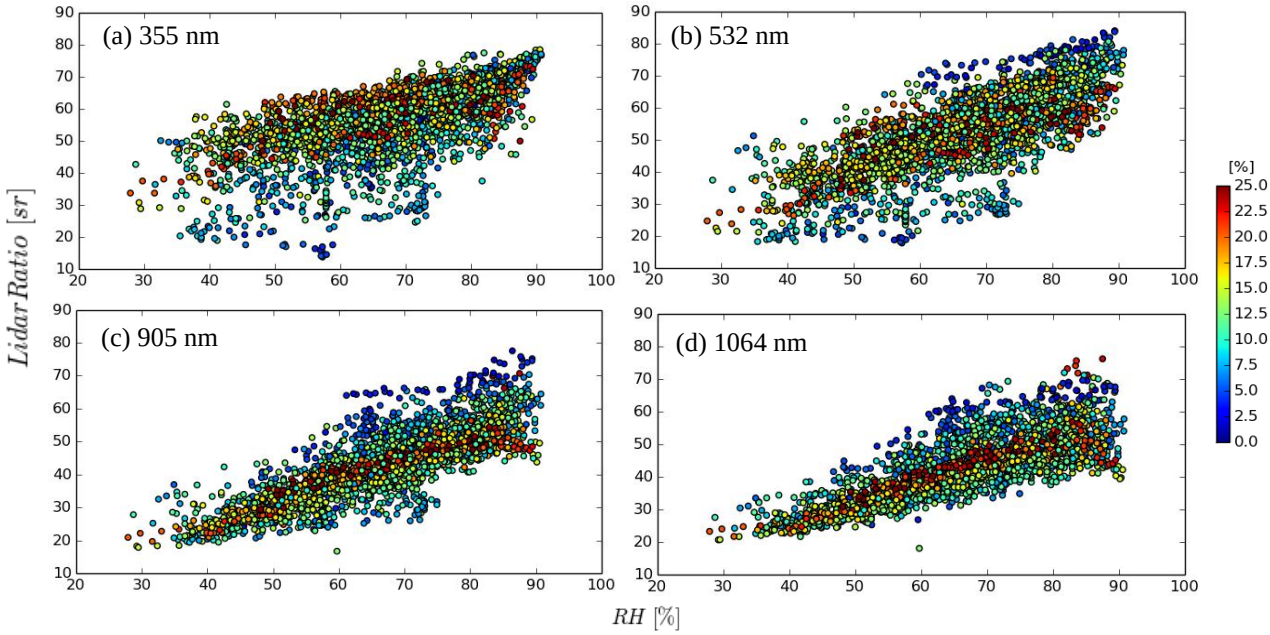

Figure 3-4: Calculated lidar ratio (S, sr, Figure 3-2) as a function of relative humidity (RH) [%] and elemental carbon proportion of total aerosol (%, colour) at four wavelengths that are typical of: Raman lidar (a) 355, (b) 532 nm; and ALC (c) 905 (Vaisala CL31) and (d) 1064 nm (Lufft CH15MK). Calculated hourly using NK (Figure 3-1) data for the period 01/01/2014 – 31/12/2015.

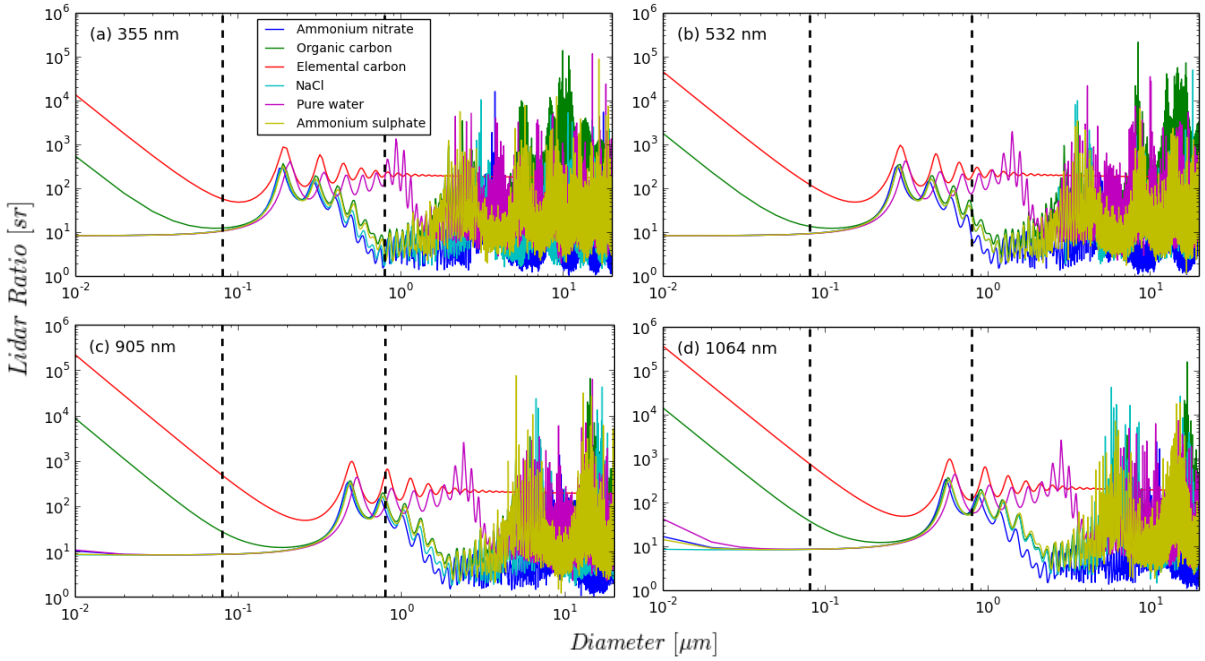

Figure 3-5: Lidar ratio (S, sr) of pure monodisperse aerosol species for four wavelengths (as Figure 3-4) using Mie scattering and assuming perfect sphericity. Accumulation diameter range (~80 – 800 nm) indicated (vertical dashed lines) (Figure 3-1).

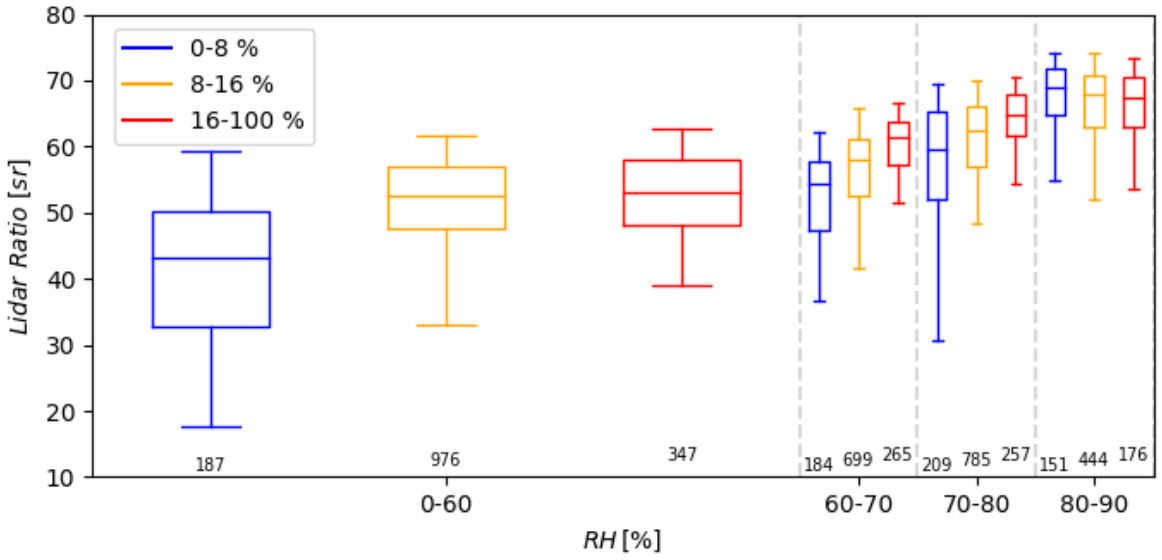

Figure 3-6: Calculated lidar ratio (S, sr) as Figure 3-4, but of relative humidity (%, x-axis) then by relative volume of elemental carbon to total aerosol (colour), for 355 nm. Median (box centre line), inter-quartile range (box edges), 5th and 95th percentiles (whiskers). Sample size below each box and whiskers.

As EC and OC have a relatively high absorption of radiation compared to other ambient aerosols, their presence impacts the bulk aerosol absorption (Esteve et al., 2014). However, the imaginary component of the complex index of refraction (CIR) of these species (which governs absorption), as well as the physical growth factor of OC (g_{OC}) from the absorption of water, are relatively uncertain due to the large number of chemical species present and the extent of external mixing (Esteve et al., 2014). Measurements in areas with distinct aerosol sources (e.g. South American biomass burning (Kotchenruther and Hobbs, 1998; Figure 3 in Thornhill et al., 2018), biogenic sources in North America, (Lowenthal et al., 2009; Taylor et al., 2017)) highlight the vast uncertainty in g_{OC}. Observed water soluble fractions of OC vary with location, between 20 and 65% of total OC (Du et al., 2014; Miyazaki et al., 2006; Sullivan et al., 2004; Zappoli et al., 1999), indicating the hygroscopicity of total OC. Uncertainty in g_{OC} arises at NK as the OC sources differ from previous g_{OC} studies, with urban contributions from cooking and fuel burning, and local vegetation differences. To explore the impact of CIR and g_{OC} uncertainty on *S* variability, five combinations of variables at 905 nm are used (Table 3.5). For experiment 5, an adjusted g_{OC} (growth is ~1/3 of fresh OC, model 2 in Kotchenruther and Hobbs (1998), hereafter KH98) gives the scattering enhancement factor closest to the literature.

Table 3.5: Experiment (ID) settings for calculation of lidar ratio S: physical growth factor (g_{OC}) and imaginary component of the complex index of refraction (CIR) for organic carbon (OC) and elemental carbon (EC). Settings modified from default (shaded). Aged and fresh OC growth factors as per Met Office CLASSIC aerosol scheme (Bellouin et al., 2011). KH98 (Kotchenruther and Hobbs 1998) adjusted growth factor best reproduced the scattering enhancement observed. All values applicable for 905 nm. ID = 0 is control (section 3.3.2).

ID	Physical growth factor (g_{OC})	CIR imaginary component	
Species	*OC*	*OC*	*EC*
0	Aged	0.006	0.700
1	Aged	0.006	0.440
2	Aged	0.010	0.700
3	Aged	0.000	0.700
4	Fresh	0.000	0.700
5	KH98 adjusted	0.006	0.700

The positive *RH* - *S* relation occurs in all experiments (not shown), with the range of *S* mostly unchanged. Sensitivity analyses suggest that the uncertainty of the imaginary component of CIR for OC and EC is small when estimating *S* at NK. Using either fresh or aged OC growth factor has a minor impact. Using KH98 adjusted g_{OC} reduced the inter-quartile range (IQR) and the mean *S* (30.1 sr, not shown). Hence, constraining g_{OC} uncertainty is expected to reduce the lidar ratio uncertainty. This is in agreement with Esteve et al. (2014), who found the physical growth factor of OC to be important for the calculation of the bulk aerosol optical properties. However, their results also indicate that the uncertainty of the imaginary component of CIR should be accounted for, which is not considered critical in the current analysis.

3.3.2 Improvement to forward-modelled attenuated backscatter using observed aerosol characteristics

Multiple aerFO experiments are conducted with different combinations of aerosol observations to identify which aerosol characteristics are most important for estimating β_m (section 3.2.4). Estimates of β_m are compared with observed β_o and summary statistics across all the experiments are shown in Table 3.6. Statistics computed include the Spearman rank correlation (R_S) and Pearson correlation (R_P) between β_m and β_o, as well as a normalised standard deviation (nSD = SD(β_m)/SD(β_o)). β_m is most relevant for the near surface (height level of *in-situ* aerosol measurements), whereas β_o was taken at 20 m above ground level (section 3.2.1).

In the control experiment (ID = 0), PM_{10} is the aerosol mass *m*, used to parameterise *N* and r_{md}. Modelled and observed attenuated backscatter are highly correlated with both R_S and R_P (0.78 and 0.74, respectively) and statistically significant (> 99.9 %). However, β_m is generally underestimated with a median bias error (medBE) of ~ - 1.1 $\times 10^{-6}$ sr. The nSD is low (0.37, i.e. small modelled

standard deviation) as the monotonic function of m (PM_{10}, Figure 3-7) does not account for the large variance observed.

In experiment 1 (ID = 1) observed aerosol number concentration of the accumulation mode replaces the parameterised N, bringing a slight reduction in error. The parameterised aerosol number concentration has slightly underestimated N when $m > \sim 20$ μg kg^{-1}, compared to a locally weighted scatter smoothing function (Cleveland, 1979) through N observations (Figure 3-7a, black (parameterisation) and red (LOWESS) lines). Providing more variable and realistic values of N enhances the modelled variability and increases nSD to 0.68.

For experiment 2, observed dry mean volume radius replaces the parameterised r_{md} causing poorer (cf. ID = 1) correlation coefficients R_S and R_P. Experiment 2 has the lowest nSD (0.32) of all the experiments despite introducing observed variability in r_{md} (Figure 3-7b). These results suggest the variance of β is controlled more by the variation in aerosol number concentrations in the accumulation mode than by the variation of r_{md} at NK. Hence, a more realistic estimate of N is more beneficial to accurately reproducing β_m than adjusting r_{md}. Overall, the level of underestimation of β_m is larger for experiment 2 (cf. ID = 1). This is likely explained by the underestimation of N by aerFO for $m > \sim 20$ μg kg^{-1} (Figure 3-7a).

Experiment 3, with both observed N and r_{md} for the accumulation mode used, still has a negative medBE (~ - 1.1 ×10^{-6} sr). This suggests other characteristics are important to accurately obtain β_m.

In experiment 4, the N and r_{md} observations used are for both fine and accumulation modes. This addition of fine aerosol mode information has little impact on the results as R_S, R_P, medBE and nSD only differ slightly (cf. ID = 3). This is explained by the impact of aerosol size and surface area on extinction. Given that $r_{md} \propto Q_{ext,dry} \propto \sigma_{ext}$, and the average r_{md} of the accumulation mode (0.142 μm) distinctly exceeds that of the fine mode (0.029 μm) by definition, the extinction coefficient (σ_{ext}) of the fine mode is ~100 times smaller than that of the accumulation mode at a wavelength of 905 nm. Although more smaller particles are present, and both N and $f_{ext,rh}(r_{md})$ are greater for the fine mode, the effect of number concentration does not outweigh the size and surface area effects. Hence, we conclude that representing the fine mode N and r_{md} is not critical to accurately estimating β_m at the 905 nm wavelength.

For experiment 5, when N and r_{md} observations for three aerosol modes (fine, accumulation and coarse) are used, β_m is overestimated (medBE ~3.6 ×10^{-7} sr). This demonstrates the strong impact of coarse mode particles (D_v > 0.8 μm) on the bulk optical properties. The underestimation in experiments 0 - 4 is likely caused by the lack of coarse mode aerosol information (cf. medBE ~-1.1 ×10^{-6} sr (ID = 3, 4)). This is likely to be more important at ~900 nm than 300 – 500 nm.

A strong horizontal advection event likely affects this positive bias. On 04 June 2015 a change in wind direction (not shown) is associated with advection of $(NH_4)_2SO_4$ and NH_4NO_3, leading to a 3 – 4-fold increase in NH_4 mass and a 2 – 3 increase in NO_3 and SO_4. This impacts the five hours after the air mass change with a larger nSD (2.1). The assumption that all available NH_4 combines with SO_4 and NO_3 when estimating S (section 3.2.1) might not hold during this period, causing the relative amount of $(NH_4)_2SO_4$ and NH_4NO_3 to be overestimated. The associated increase of $f_{ext,rh}$ would explain some of the overestimation of β_m. Removing the affected period (5 hours) reduces the nSD to 1.56. The medBE also improves from 3.6×10^{-7} to 2.8×10^{-7} sr.

Experiment 6 uses hourly estimates of $Q_{ext,dry}$ and $f_{ext,rh}$ to replace the monthly LUT. The medBE, nSD and correlation statistics are similar to experiment 5. As additional relative mass of different aerosol species information does not improve β_m, the default monthly LUT of $Q_{ext,dry}$ and $f_{ext,rh}$ appear adequate. However, other uncertainties within aerFO may limit its ability to improve β_m. For instance, assuming particles are perfectly spherical when using Mie theory will introduce errors. As particles likely interact and aggregate, their interaction with water and light is altered. Hence, not considering internal mixing in aerFO or its LUT is another source of uncertainty. Measurement uncertainties arise in β_o from imperfect overlap correction and calibration. Height differences between the near-surface *in-situ* aerosol observations and ceilometer β_o also cause uncertainties.

Finally, experiment 7 uses hourly S estimates (section 3.3.1) to ID = 6. The error is greater than those for experiments 5 and 6, with a larger standard deviation and interquartile range in $\beta_m - \beta_o$. Correlations between model results and observations (R_S and R_P) are the lowest across all experiments. The large differences between β_m and β_o is reflected in a high nSD (3.5), but excluding the five hours (see experiment 5 discussion) reduces nSD to 1.83. The reduced overall performance of experiment 7 suggests the method to compute S (section 3.3.1) is unable to provide hourly variability. Specifically, uncertainties in estimating S (section 3.3.1) are likely varying with aerosol as particles interact and age, including through hysteresis, relative aerosol speciation across the size distribution, as well as OC and EC characteristics. Hence, the method is currently limited to providing only the overall behaviour of the lidar ratio.

Table 3.6: aerFO model estimates for experiments (ID) with different aerosol observations (Table 3.4), observed attenuated backscatter and $\beta_m - \beta_o$ evaluation statistics. Metrics: standard deviation (SD), interquartile range (IQR), normalised standard deviation (nSD = SD(β_m)/ SD(β_o)), Spearman rank (R_S) and Pearson (R_P) correlation between β_m and β_o. All correlation coefficients are statistically significant (> 99.9 %). Equal sample size of 106 hours from the seven days for all experiments.

		ID							
		0	1	2	3	4	5	6	7
β_m	Mean	5.6×10^{-7}	7.8×10^{-7}	4.6×10^{-7}	6.2×10^{-7}	6.3×10^{-7}	2.4×10^{-6}	2.4×10^{-6}	3.0×10^{-6}
	Median	5.3×10^{-7}	7.0×10^{-7}	4.6×10^{-7}	5.4×10^{-7}	5.4×10^{-7}	2.1×10^{-6}	2.1×10^{-6}	2.6×10^{-6}
β_o	Mean	1.8×10^{-6}							
	Median	1.6×10^{-6}							
$\beta_m - \beta_o$	Mean	-1.3×10^{-6}	-1.1×10^{-6}	-1.4×10^{-6}	-1.2×10^{-6}	-1.2×10^{-6}	5.3×10^{-6}	5.3×10^{-6}	1.1×10^{-6}
	Median	-1.1×10^{-6}	-1.1×10^{-6}	-1.2×10^{-6}	-1.1×10^{-6}	-1.1×10^{-6}	3.6×10^{-7}	4.0×10^{-7}	8.2×10^{-7}
	SD	4.7×10^{-7}	4.1×10^{-7}	5.2×10^{-7}	4.3×10^{-7}	4.3×10^{-7}	9.6×10^{-7}	9.3×10^{-7}	1.9×10^{-6}
	IQR	5.9×10^{-7}	5.1×10^{-7}	6.5×10^{-7}	6.6×10^{-7}	6.5×10^{-7}	1.1×10^{-6}	9.7×10^{-7}	9.7×10^{-7}
nSD		0.37	0.68	0.32	0.50	0.51	2.14	2.14	3.48
R_S		0.78	0.68	0.66	0.70	0.70	0.72	0.72	0.47
R_P		0.78	0.74	0.62	0.75	0.75	0.74	0.76	0.50

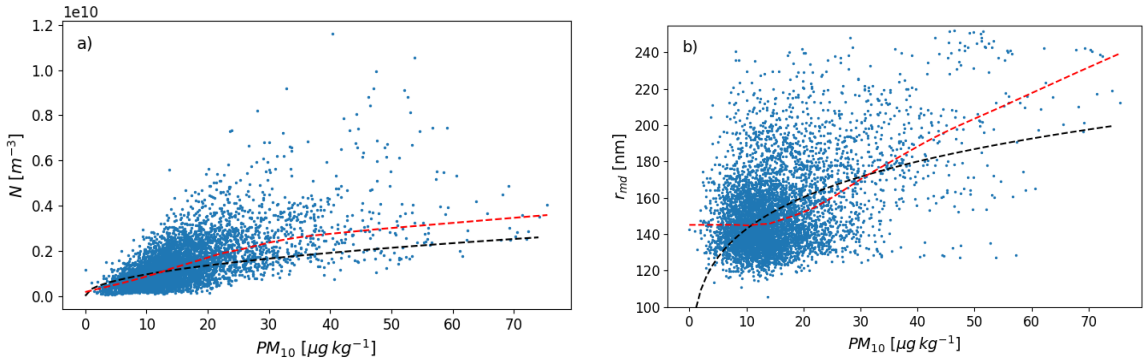

Figure 3-7: Hourly observations for two years (2014 – 2015) at NK (Figure 3-1) of PM_{10} concentrations and aerosol characteristics for the accumulation diameter range (80 – 800 nm): a) total particle number concentration (N) and b) dry volume mean radius (r_{md}) with a locally weighted scatterplot smoothing function (LOWESS, red). The aerFO parameterisations (black) are a function of aerosol mass (m) derived from PM_{10} concentration.

3.4 Conclusions

An updated version of the aerosol forward operator (aerFO, section 3.2.3), first used in Chapter 2, is used in Chapter 3 to quantify the relative importance of accurately describing different aerosol characteristics when estimating attenuated backscatter (β_m) at an urban background site. Observations are used to calculate aerosol variables by aerFO in increasing detail, including total aerosol mass, number concentration and mean volume radius for fine, accumulation and coarse modes. Results

inform the development of the Met Office MURK visibility scheme for application in urban settings, as the aerFO links the output of this scheme to observed aerosol characteristics.

aerFO is used to assess the relative importance of various aerosol characteristics when modelling aerosol optical properties at an urban background site in London, UK. With respect to where efforts should be focused on improving the MURK aerosol scheme, the following conclusions can be drawn which are in line with the data in Table 3.6:

- A more realistic representation of total particle number concentration is the most critical existing factor that could be improved to improve the variability in attenuated backscatter and should be considered a priority in future developments of Met Office NWP aerosol schemes for estimating aerosol optical properties.
- Including representation of the coarse mode (diameter: > 0.8 μm) in addition to the accumulation mode leads to a reasonable improvement in the median difference between β_m and observed attenuated backscatter (β_o), and a clear overall increase in β_m. Given the large sensitivity of β_m to the coarse aerosol, this should be considered a second priority.
- Monotonic parameterisations within aerFO for total particle number concentration and dry volume mean radius from aerosol mass alone lead to an underestimation of the variance of attenuated backscatter by about 2/3. If the variance in optical properties is important, this should be considered a third priority.

The following conclusions were also found, and highlight aspects which should not be a priority:

- At a wavelength of 905 nm the extinction coefficient of the fine mode aerosol (diameter: < 0.08 μm) is ~100 times smaller than that of the accumulation mode, so that more realistic fine mode information has no clear benefit and should not be a priority.
- Using more realistic values of dry volume mean radius had little effect and could not be identified as a priority.
- Constraining aerosol species relative mass has little impact on modelled results and could not be identified as a priority. The low impact may be attributed to a limitation of the aerFO which does not account for internal mixing, but rather uses some simplifications with respect to aerosol hygroscopic and optical properties.

In addition to improvements for the Met Office aerosol scheme, the aerFO needs to represent the coarse mode in addition to the accumulation mode to better estimate the magnitude of β_m.

A new method to estimate the lidar ratio *(S)* from observations is developed that uses relative humidity and *in-situ* aerosol observations. This extends the possibility to calculate the lidar ratio without the need for the rare measurements from Raman lidars or sunphotometers. *S* can be used to derive further aerosol optical properties (e.g. optical depth, extinction) from elastic-backscatter lidars

(e.g. ceilometers), and the ratio can be parameterised before use in forward models to minimise computational expense. For the London background site, modelled *S* varies with wavelength between 14 – 80 sr across four common lidar wavelengths (355, 532, 905 and 1064 nm), with values consistent with the literature. A strong dependence of the lidar ratio on relative humidity is found which varies between common lidar wavelengths; this is a relation not found in the wider literature. The relation found is due to another relation between lidar wavelength and aerosol size.

Adequately representing elemental carbon (EC) is important when estimating *S* at shorter wavelengths (355 and 532 nm) and low relative humidity, given its characteristically high absorption of radiation. The effect of hygroscopic aerosol swelling likely explains why EC is less important at longer wavelengths and at higher relative humidity. Remaining uncertainties when estimating *S* are associated with the assumption of particle sphericity for use with Mie theory, the speciation of the number distribution, and not including the effect of internal aerosol mixing which would lead to errors in the hygroscopic growth of particles, including hysteresis effects. These uncertainties could not be constrained in the current study due to a lack of appropriate observations. Since, the current study is based on aerosol properties and relative humidity observed near the surface, future work should be to assess vertical variations in lidar ratio, which will enable a better comparison to modelled estimates in the literature (Dionisi et al., 2018). Nonetheless, this would require further observations at multiple levels, which could potentially come from instruments on towers or aircraft.

The current study uses a range of very detailed aerosol observations. However, data availability of number concentrations is slightly limited for relatively large particles (diameter: > 0.6 μm). While small particles are a critical research topic due to their impact on health, the fundamental impact of larger particles on bulk optical characteristics of aerosols warrant additional, increased and simultaneous monitoring of number concentrations in both the accumulation and the coarse mode. Furthermore, to estimate optical properties at more and diverse locations, more observation of speciated aerosol mass alongside number size distributions are needed.

Acknowledgements

EW was funded by the NERC Scenario DTP and University of Reading. The instruments and measurements have been funded and operated by LUMO (NERC ClearfLo, NERC TRUC, EPSRC DARE, KCL, UoR, Newton Fund/Met Office CSSP China (SG, SK), EUf7 BRIDGE, EUf7 emBRACE, H2020 UrbanFluxes), LAQN/ERG, MO, DEFRA, STFC, NPL; Paul Quincy, Jordan Tompkins, James Allen, Max Priestman, Anke Finnenkoetter, Adrian Hill, Jordan Tompkins, Paul Agnew, Emma Hopkin, Anthony Illingworth and Helen Dacre; and the vast number of people who help keep the equipment operating on a daily basis. We acknowledge the contribution of one of the authors, Sue Ballard, who sadly passed away before the completion of this paper.

Appendix 3.A Estimation of geometric mean radius from volume mean radius

The dry geometric number mean radius (r_g) is calculated using linear regression and the dry mean volume radius (r_v), as r_g is not estimated by the aerFO (Figure 3.A.1 – Figure 3.A.3). r_g is calculated as (Hinds, 1999):

$$r_g = \frac{(\Pi\, D_i^n)^{1/N_{tot}}}{2} \tag{3.A.1}$$

where D_i is the mid-point particle size in bin i, n is the number of particles in bin i and N_{tot} the total number of particles across all bins. Calculations are done for dry particles.

r_v is calculated as:

$$r_v = \frac{\Sigma\left(dn_i/dlogD_i * dlogD_i * D_i^4\right)}{\Sigma\left(dn_i/dlogD_i * dlogD_i * D_i^3\right)} \tag{3.A.2}$$

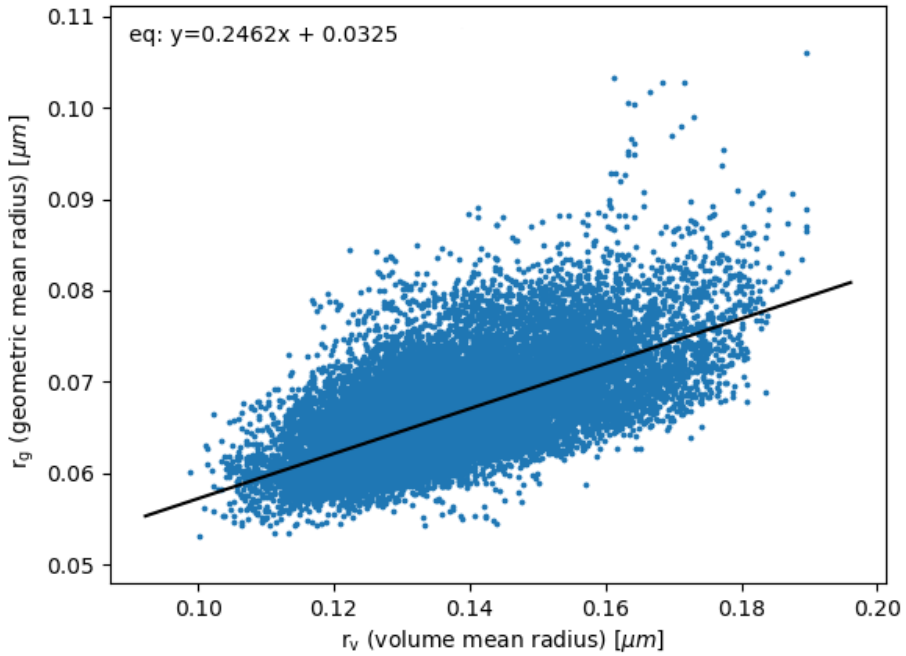

Figure 3.A.1: Observed dry geometric number mean radius (r_g) against observed dry mean volume radius (r_v) for the accumulation mode (radius: 0.04 – 0.4 µm) at North Kensington (NK) 01 January 2014 – 31 December 2015, from combined TSI SMPS and APS data. APS data are 'dried' based on physical growth factors, calculated from observed RH at KCL (section 3.2.1, Figure 3-1). Linear regression (line), Pearson correlation (R_P) = 0.65 and statistically significant > 99 %.

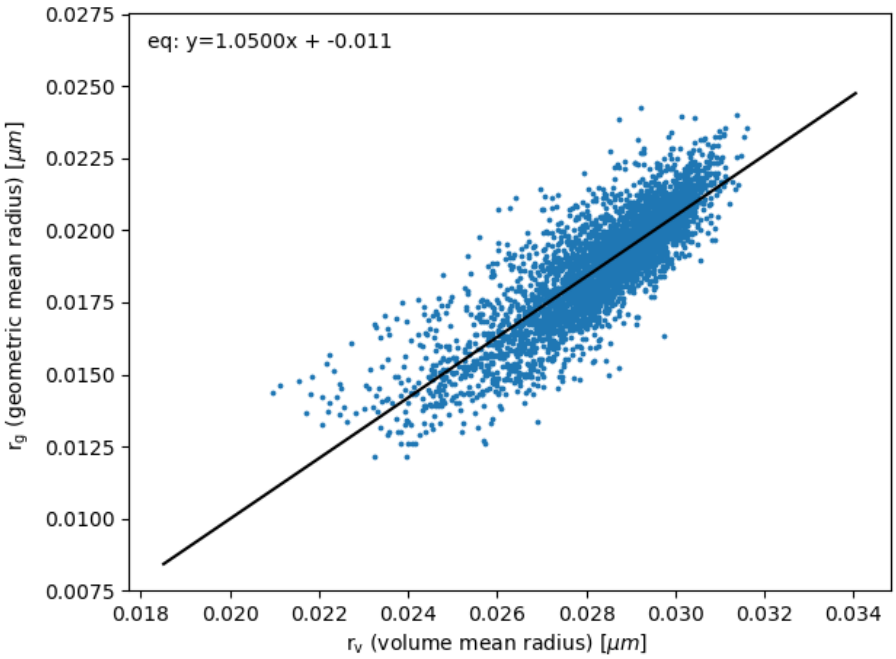

Figure 3.A.2: As Figure 3.A.1 but for fine mode (radii: < 0.04 μm). Pearson correlation (RP) = 0.83 and statistically significant > 99 %.

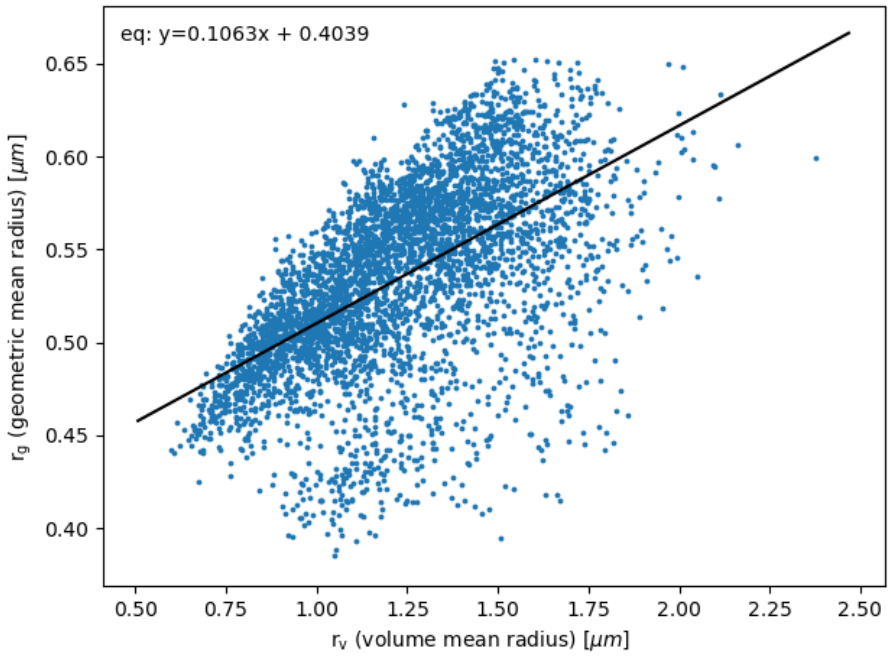

Figure 3.A.3: As Figure 3.A.1 but for the coarse mode (radii:> 0.4 μm). Linear regression (line), Pearson correlation (R_P) = 0.57 and statistically significant > 99 %.

Appendix 3.B Extinction enhancement factor sensitivity to observed aerosol speciation

As bulk aerosol optical properties vary strongly based on the hygroscopic properties and the relative proportion of different aerosol species present (section 2.6, section 3.3.1), it is useful to understand how the common urban aerosol species contribute to extinction enhancement factor $f_{ext,rh}$ calculations.

Daily mean $f_{ext,rh}$ are estimated from hourly observed aerosol species for three UK sites: North Kensington (urban, NK), Chilbolton (rural, Ch) and Harwell (rural, Ha), for a fixed r_{md} = 0.11 μm and a *RH* range of 0 – 100 %. $f_{ext,rh}$ is lower when the relative volumes of EC are greater (Figure 3.B.1). Across all sites the mean relative volume of EC is 7.25 %. The lower $f_{ext,rh}$ values (Fig. A2.1 red, EC relative volume 9 – 100 %) reflect the hygrophobic nature of EC which has a constant $f_{ext,rh}$ of 1.0. Hence, aerosol mixtures with greater volumes of EC have a lower combined $f_{ext,rh}$.

At higher *RH* the reduction in $f_{ext,rh}$ is greatest with respect to aerosol with the larger fraction of EC (Figure 3.B.1). As *RH* increases, $f_{ext,rh}$ increases exponentially for hygroscopic aerosols. However, as $f_{ext,rh}$ for EC remains constant, the exponential increase in $f_{ext,rh}$ of the total aerosol is limited. In addition, the variability of $f_{ext,rh}$ increases with *RH*, due to the increasing variability in $f_{ext,rh}$ between different aerosol species. Greater relative volumes of sea salt lead to higher daily estimates of $f_{ext,rh}$ (not shown) due its highly hygrophilic nature. Greater relative volumes of OC are related to lower $f_{ext,rh}$ values (not shown). Although OC is not strongly hygrophobic, the relatively high-volume fractions (average 41.7 %) make it proportionally important when computing the optical properties of the combined aerosol.

In section 3.3.1, the impact of hysteresis is discussed as it relates to calculations of *S*. The hysteresis effect is not accounted for in the $f_{ext,rh}$ parameterisation here as SOCRATES (which is used, section 2.4) does not include this effect. We recognize that hysteresis should be considered in future computation of LUTs of $f_{ext,rh}$, to provide more realistic extinction enhancement.

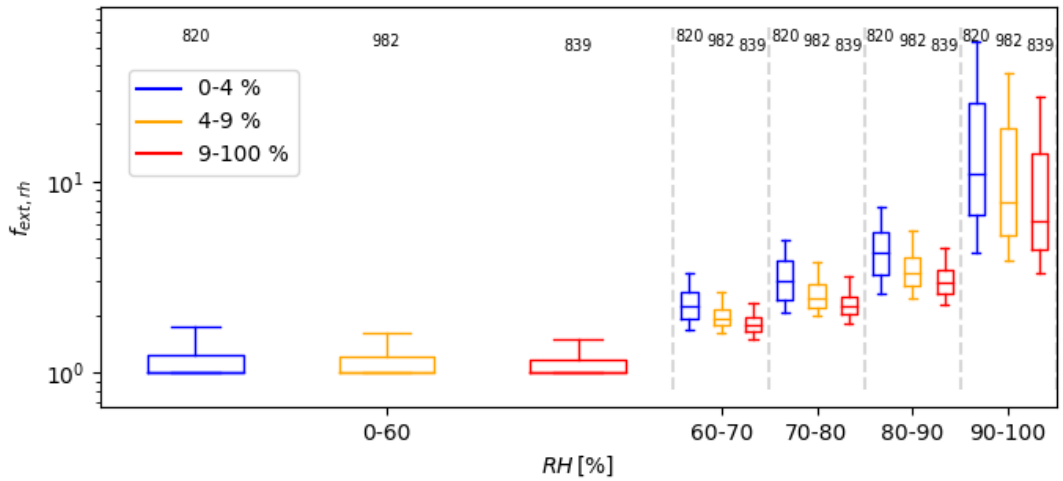

Figure 3.B.1: Modelled daily extinction enhancement factors ($f_{ext,rh}$) stratified by relative humidity (RH) [%] and relative volume of elemental carbon [%] using observed daily mean relative volumes of each aerosol species as input. Median, IQR, 5th and 95th percentile whiskers, and outliers (crosses) shown with sample size. Estimates calculated for 0 – 100 % RH for a radius of 0.11 µm using data from three sites: North Kensington (n=1378), Chilbolton (n=192) and Harwell (n=1071). Note log scale on y-axis.

Chapter 4 Spatial variability of forward modelled attenuated backscatter in clear-sky conditions over a megacity: Implications for observation network design

Abstract

Sensors that measure attenuated backscatter (e.g. automatic lidar and ceilometers (ALC)) provide information on aerosols which impact urban citizen climate and health. To design an observational network for ALC sensors that could contribute to data assimilation in cities, therefore improving prediction of urban weather and air quality, a methodology is needed. In this study, the spatio-temporal patterns of attenuated backscatter are modelled using Met Office numerical weather prediction (NWP) models at two resolutions, 1.5 km (UKV) and 333 m (London Model) for 28 clear-sky days and nights.

Initially, the β_m data area is analysed using S-mode principal component analysis with VARIMAX rotation (Kaiser, 1958). Four to seven empirical orthogonal functions (EOFs) are produced for each model level with common EOFs found across different heights (day and night) in both NWP models. Several EOFs relate strongly to orography, wind and relative aerosol emissions source location suggesting that these critically control the megacity attenuated backscatter spatial variability. In several night-time EOFs potential gravity wave features are found in both NWP models. Increasing the horizontal resolution of the native ancillaries (model input parameters) and improving the urban surface scheme in the LM may enhance the urban signal in the EOFs.

Principal component analysis (PCA) output, with agglomerative Ward cluster analysis (CA), minimises intra-group variance. In the network design process, the number of potential instruments to be deployed can be selected. The UKV and LM CA shape and size results are similar, and strongly related to orography. PCA-CA is a simple, but adaptable methodology, allowing close alignment with observation network design goals.

4.1 Introduction

Automatic lidars and ceilometers (ALC) can measure the attenuated backscatter from aerosol particles. They are used in a wide variety of applications, including to investigate particle mass concentrations (Münkel and Roininen, 2010; You et al., 2016), to measure boundary layer height (Kotthaus and Grimmond, 2018b; Peña et al., 2013; Wang et al., 2012) and sea-breeze dynamics (Lemonsu et al., 2006). As instrument sensitivity has increased, the utility of ALC measured aerosol backscatter has grown (Jin et al., 2018; Madonna et al., 2014), but many studies analyse relative variations or the signal-to-noise ratio of the measurements (Wiegner and Gasteiger, 2015). To relate

the aerosol attenuated backscatter more directly to aerosol physical characteristics, a forward operator (FO) is required.

An aerosol FO estimates the attenuated backscatter using relevant aerosol variables as inputs. Numerical weather prediction (NWP) or chemistry transport models (CTMs) can run a FO, and the estimated attenuated backscatter can be evaluated against ALC observations or the latter can be assimilated (Benedetti and Dabas, 2016; Chan et al., 2018; Charlton-Perez et al., 2016; Geisinger et al., 2017; Chapter 2; Chapter 3). Several of these studies take advantage of ALC networks to provide more spatial information.

ALC networks cover spatial scales spanning cities (Kotthaus and Grimmond, 2018a), countries (Flentje et al., 2010b; Osborne et al., 2018) and continents (Guerrero-Rascado et al., 2016; Illingworth et al., 2015, 2007; Nishizawa et al., 2016; Pappalardo et al., 2014; Welton et al., 2000). However, the ideal spatial distribution of ALC within a network is uncertain and may depend on the spatial scale of features targeted for investigation, whether micro-, meso- or synoptic scale. If instrument spacing is too fine, undesirable redundancy can exist within the network, as multiple-ALC are measuring the same meteorological features. Increasing the distance between instruments can allow instrument re-deployment (or reduced initial capital costs) to measure other meteorological features. Conversely, a sparse network configuration may miss or partially capture meteorological features.

Though practical constraints will impact final deployment locations (e.g. geo-political boundaries or suitability of local infrastructure), an understanding of spatial variability in the meteorological features present in the area is critical. For example, in urban areas a detailed knowledge of atmospheric boundary layer processes is needed.

Network design for weather and hydrological instrumentation has been informed by the analysis of spatial or spatio-temporal variability in meteorological observations using (geo-)statistical techniques (e.g. Bastin et al., 1982; Bayat et al., 2019; Burn and Goulter, 1991), and regional climate models' climatological variability (St-Hilaire et al., 2003). However, the resolution of the latter can be coarse compared to many meso-scale and smaller meteorological sub-hourly processes.

Principal component analysis (PCA) is a technique that reduces a dataset down into a series of orthogonal functions or modes, to find the main patterns of variability (Wilks, 2011). S-mode PCA (or empirical orthogonal function (EOF) analysis) is one variant that focuses on identifying *spatial* patterns in the different eigenvectors and has been used extensively to analyse the variability of many meteorological variables. These include rainfall (Smith and Phillips, 2013; Yu and Lin, 2015); wind (Álvarez-García et al., 2020; Farjami and Hesari, 2020) and temperature (Li et al., 2018). PCA has also been used to explore variables specifically within the boundary layer; for example, boundary layer turbulence (Lin et al., 2008; Wilson, 1996), urban heat island characteristics (Qiao et al., 2018;

Vicente-Serrano et al., 2005) and air quality (Chan and Mozurkewich, 2007; Fleming et al., 2012; Gupta et al., 2018; Henry et al., 1991; Rogula-Kozłowska et al., 2015). Further in-depth reviews of different PCA approaches in meteorology and climatology can be found in Monahan et al. (2009), Schmidt et al. (2019), Wilks (2011), and Zhang and Moore (2015). Some studies follow PCA with cluster analysis (CA) to explore the relations between modes or with other meteorological variables (Beaver and Palazoglu, 2006; Henry et al., 1991) and to identify and group geographical regions together where the variability in the target meteorological variables are similar (Dogruparmak et al., 2014; Neal and Phillips, 2009; Smith and Phillips, 2013). With respect to network design, PCA and CA could be used to identify spatial regions that covary, which could then in turn inform the placement of instrumentation in a network, to capture unique phenomena and maximise network utility.

In this study, our objectives are to (1) identify the main spatial patterns of forward modelled attenuated backscatter (β_m) variability across a metropolitan area using NWP forecasts and PCA; and (2) develop a method using PCA and CA to inform ALC network design constrained by areas of similar β_m from similar meteorological processes. To understand the impact of horizontal resolution two NWP models are used.

4.2 Methods

4.2.1 Numerical weather prediction (NWP) data

Here two configurations of the Met Office Unified Model (Davies et al., 2005) are used: the operational convection-permitting UK regional model (UKV, 1.5 km; Tang et al., 2013), and the experimental London Model (LM , 333 m, Boutle et al., 2016). These have been applied and evaluated for London (Boutle et al., 2016; Lapworth and Osborne, 2016; Chapter 2). The urban surface energy balance scheme within the research LM at the time of this study was JULES/Best 1-tile (Best, 1998; Best et al., 2011, 2006) and within the UKV the JULES/MORUSES 2 tile (Best et al., 2011; Bohnenstengel et al., 2011; Porson et al., 2010). The latter requires more detailed surface information and differentiates between urban canyons and roofs and aims to better estimate surface fluxes. The aerosol emission ancillary was derived from 1km native resolution National Atmospheric Emissions Inventory dataset (Neal, 2019) to produce a dry mass mixing ratio of aerosol. The orography ancillary is from a Digital Terrain Elevation Data (DTED) model with a native 100 m resolution (Boutle et al., 2016). Land cover is separated into 9 different land-use tiles and the ancillary also has a native resolution of 100 m (Boutle et al., 2016; Walters et al., 2017). All ancillaries are rescaled to NWP model resolution for implementation.

Data from both models for 28 clear-sky days (Appendix 4.A) are used to force the aerosol lidar forward operator (hereafter aerFO) to create 3-D hourly fields of attenuated backscatter (β_m) for each

model. Hourly aerFO calculations are conducted from midnight for 24 h using the prior day 21 Z forecast (i.e. 3 h at the start of each model run are not used to allow for any spin up, particularly in the 333m model which is downscaling the UKV). The NWP data are extracted for the *London domain* (Figure 4-1).

Analysis is done for two periods separately to differentiate meteorological processes (e.g. boundary layer depth) for 24 model levels: *day* (sunrise (SR) + 2 h to sunset (SS) -2 h) and *night* (i.e. SS +2 h to SR – 2 h). 24 model levels were chosen (representing the bottom 2075 m above ground level (agl)), to ensure the day-time boundary layer and night-time residual layer were always within the domain analysed, as these layers typically contain most of the aerosol in the atmospheric column (Seinfeld and Pandis, 2016). 2075 m is above typical daytime maximum mixed layer heights observed in London, where the median boundary layer depth is 1704 m, as estimated from ceilometer observations during thermally unstable days following clear nights (Kotthaus and Grimmond, 2018b). Furthermore, 2075 m was also above the highest boundary layer height across the 28 clear-sky cases, following subjective interpretation of β_m from of each case.

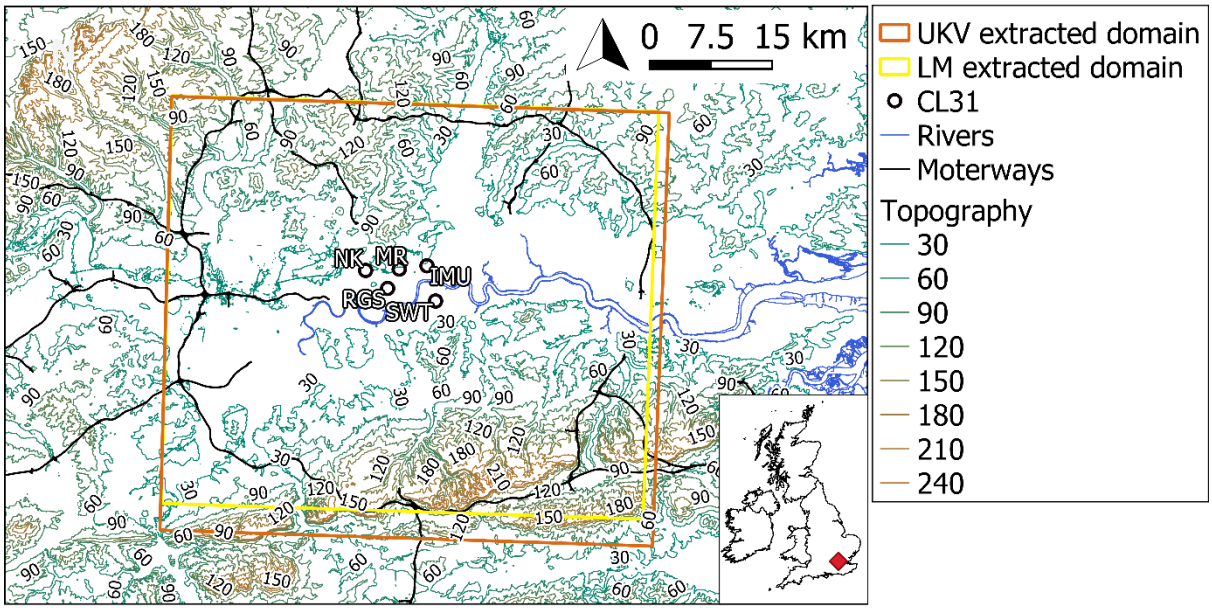

Figure 4-1: London Urban Meteorological Observatory (LUMO) ALC (dots and labels), 'London' grid domain used from the UKV (1.5 km) and London (333 m) model, for the Principal Component Analysis. Contains OS data © Crown copyright and database right (2019).

4.2.2 Overview of the aerosol lidar forward operator (aerFO)

The forward modelled attenuated backscatter is calculated using the aerFO version from Chapter 3 (section 3.2.3), hence an brief overview is provided here. The two primary aerFO input variables are obtained from the NWP output: dry mass of total aerosol (m, kg kg^{-1}) and relative humidity (RH). Model fields of air temperature (T_{air}) and pressure (p_{air}) are also used with the specific water vapour

(q) to compute the water vapour absorption. Both the UKV and LM provide m from the passive aerosol tracer (m_{MURK}) in the MURK visibility scheme, that considers the most common aerosol species in the UK (Neal, 2019).

Firstly, aerFO estimates physical properties of the aerosol in accumulation mode from m_{MURK}, including the dry mean volume radius and total number concentration. This follows the MURK empirical parameterisation based on the climatological mean mass of dry aerosol, mean dry volume radius and total number concentration (Clark et al., 2008). The climatological values and the geometric distribution for the accumulation mode are calculated from aerosol observations at an urban background site (North Kensington, London, UK; Figure 4-1; DEFRA, 2018).

Secondly, the optical properties of the aerosol are calculated, including the extinction efficiencies for the dry aerosol particles and the effect of hygroscopic growth on particle extinction. For computational efficiency, pre-calculated look-up-tables of dry extinction efficiency and hygroscopic growth effects are used. The extinction coefficient calculation includes the effect of water vapour. Thirdly, the backscatter is calculated using a fixed lidar ratio, and finally, the transmission factor is applied to compute attenuated backscatter.

4.2.3 Spatial mode principal component analysis (S-mode PCA)

To explore the nature of backscatter variability in the London model domain (Figure 4-1) PCA is used to extract the most important spatial patterns from the original dataset (x) and identify when these spatial patterns are most prominent in time. As the distribution of β_m is highly positively skewed and therefore not directly appropriate for PCA, it is transformed logarithmically ($\log_{10}(\beta_m)$) to reduce the skewness (Neal and Phillips, 2009).

S-mode PCA is carried out on the covariance matrix (S) of $\log_{10}(\beta_m)$ for each model level separately following Wilks, (2011). PCA produces unit-scaled eigenvectors (e_i where $\|e_i\| = 1$) with paired eigenvalues (λ_i), where $i = 1, \ldots, n$, where n is the number of original spatial variables (x). The eigenvectors are then used to calculate a set of new, uncorrelated variables (the principal components, PCs), that relate each e_i to each x_i (length of e_i is equal to the number of spatial points in x). A PC time series can be interpreted as a series of 'scores' with high PC scores equating to the spatial pattern e_i being more relevant at a given time (length of a PC is equal to the length of time in x). More negative scores indicate the inverse of e_i is more relevant. Here, the covariance matrix is used in the PCA, so each successive PC explains the maximum remaining variability in the original dataset.

To ease physical interpretation of e, VARIMAX rotation (Jolliffe and Cadima, 2016; Kaiser, 1958; Richman, 1986) is performed on a limited number of leading vectors in e. VARIMAX rotation finds a new set of e_i which effectively re-distributes the explained variability across them, so more unique sub-groups can be identified (Neal and Phillips, 2009; Richman, 1986). Here, only the e_i that explains

more than 1 % of the total variability in the original dataset are analysed. This threshold is chosen to balance the need to retain as much of the original variability as possible while limiting the number of e_i that require interpretation. The selection limits the potential creation of multiplets (i.e. sets of e_i that effectively describe the same phenomena) helping physical interpretation (Wilks, 2011).

Before VARIMAX rotation is performed, the e_i are scaled by $\sqrt{\lambda_i}$ ($\|e_i\| = \sqrt{\lambda_i}$) to '*load*' the vectors. The individual elements of scaled eigenvectors ($e_{i,k}$, where $k = 1, \ldots, n$) are hereafter referred to as *loadings*. As some form of scaling with λ_i is required to effectively incorporate the information about the explained variability of e_i, into the newly rotated eigenvectors, we use $\sqrt{\lambda_i}$ (see Table 12.3, in Wilks (2011). However, although VARIMAX rotation can maintain orthogonality of e_i, the scaling used here means e_i are not orthogonal and can be correlated (Wilks, 2011). As the eigenvectors are created with respect to space (and to be consistent with the literature e.g. Jolliffe and Cadima, 2016; Qiao et al., 2018), the retained, rotated, loaded eigenvectors used are hereafter referred to as eigenvector orthogonal functions (EOFs).

Following VARIMAX rotation, rotated PCs are calculated using regression (Field, 2009):

$$PC_i = x'(S^{-1} \cdot EOF_i) \tag{3.B.1}$$

where x' is the original, mean centred dataset and $S^{-1} \cdot EOF_i$ act as PC coefficients for PC_i. As S is often an ill-conditioned matrix and unsuitable for inversion, to calculate S^{-1} the Moore-Penrose pseudo-inverse of S is taken using singular value decomposition (Strang, 1988):

$$S^{-1} = V\,\Sigma^{-1}\,U^T \tag{3.B.2}$$

Spearman correlation coefficients (r) are calculated between EOFs and PCs to find the EOFs multiplets that are to be interpreted as a group.

Comparison of the PCA output generated from the study using Python 2.7 code (Warren, 2019) to that generated from IBM SPSS Statistics 25 program package (Field, 2009), for a test case (air temperature ((11:00 – 18:00 h) at 645 m as an input), found indiscernible differences in all outputs, including the rotated scaled eigenvectors and the resultant PCs. Thus, we assume our PCA code is suitably accurate.

To aid interpretation of each EOF, the relation between each EOF and different meteorological variables is explored. For each EOF, x is subsampled in time twice. When the paired PC scores are (1) above the 90th percentile, and (2) below the 10th percentile. These two data distributions are assumed to be the meteorological conditions that occur when the EOF, and the inverse of the EOF, are most prominent in time. Analysis is further supported by qualitatively comparing two-dimensional cross sections of the meteorological variables when the PC scores are high or low.

4.2.4 Ward cluster analysis (CA)

To inform ALC sensor siting, CA is used with the PCA output to measure the variability in β_m captured by the EOFs. Agglomerative Ward minimum variance CA is chosen here as this clustering method merges variables based on their similarity (Ward, 1963; Wilks, 2011). Across all 24 heights analysed, we use all the *unrotated* eigenvectors (uEOF) that explain more than 1 % of the variability for their respective height as inputs (i.e. to 2075 m model level), to identify the geographic regions of common β_m variability throughout the atmospheric column analysed. For each k grid box, the unrotated eigenvector loadings for that grid box, ($e_{i,k}$) from each uEOF, are combined into new vectors. CA is performed on the new vectors.

The initial k separate groups (each with one variable member) are iteratively paired and merged to reduce the number of clusters, until j specified clusters remain. The clusters to be merged are chosen to minimise the sum of squared distances between all the cluster variable elements and cluster centroids (i.e., merging the two most similar clusters). Unrotated loaded eigenvectors are used in the clustering as they are orthogonal, unlike the rotated loaded eigenvectors. The loading vectors also provide a beneficial weighting to the CA. Eigenvectors that represent more of the original variability across x have larger λ_i, hence larger loading magnitudes are calculated (both positive and negative). Consequently, these vectors have a larger range of values, which effectively gives more weight to those vectors in the CA (Kaufman and Rousseeuw, 2005). The differing weights are beneficial as the more common spatial variability patterns have greater weight in the CA.

As CA allows users to define the number of clusters to create, this can be related to the number of sensors to be deployed in a network. As Ward CA minimises the total variance within groups (Ward, 1963), an instrument per cluster should optimise the balance between representativity of measurements and sensor redundancy. It also allows for operational decisions if a sensor fails (i.e. should one/more be moved) to retain the required information about variability. In our analysis the EOFs are scaled relative to the total variability at each height (i.e. not weighted with respect to other heights). Therefore, all heights are considered equally important when clustered. This can be modified by applying additional weightings.

4.3 Results

4.3.1 Spatial variability of forward modelled attenuated backscatter

The PCA with VARIMAX rotation for the daytime period produced between four and seven EOFs for each height in both models that explained more than 1 % of variability in the original dataset. Broadly, similar EOF patterns are identified across different heights. Examples of common daytime EOF patterns from the UKV (Figure 4-2) and the LM (Figure 4-3) are shown for the 111.7 m agl

model level. At this height, five EOFs are needed to explain this level of β_m variance whereas at other model levels this may differ.

EOF patterns derived from both NWP models are strongly related to wind speed and direction. The highest loadings in each EOF (Figure 4-2a-e, Figure 4-3a-d) are associated with areas where wind has transported emitted aerosol (cf. Figure 4-2f, Figure 4-3f) leading to an increase in aerosol and consequently β_m (Figure 4-4a-b). As many aerosol emission sources are located near the centre of the domain (i.e. the most urbanised area), the highest loadings in each EOF are often located near the domain edges. For example, the highest loadings in EOF_1 are in the north-west of the domain and EOF_1, most prevalent under south/south-easterly winds (Figure 4-5). Although only 28 case study days were used to derive the EOFs, a large range of wind directions and speeds were included (Figure 4-6) and therefore the importance of wind to β_m variability is highly likely at these NWP scales.

A second factor determining the shape of the EOFs is topography. London, located in a river valley, is situated between the Chiltern Hills (north) and North Downs (south and south-east). It has relatively low elevation compared to its surroundings (> 200 m above sea level, Figure 4-2a-e, Figure 4-3a-e). The EOF_1, EOF_2, and EOF_4 spatial patterns appear to represent air advected onto these hills which are forced upward in the north-west, south-east and north-east of the domain respectively. As the air is forced up, the air temperature reduces which increases *RH*. This advected air also increases m_{MURK} locally, where background emissions are typically low. As $\exp(\beta_m)$ is proportional to ($\propto$) *RH* and $\beta_m \propto m_{MURK}$, the advection effectively increases and affects the local variation of attenuated backscatter. For example, if an air parcel with m_{MURK} = 24 μg kg^{-1} is advected onto a hill with background concentration of m_{MURK} = 18 μg kg^{-1}, and subsequently raised adiabatically by 100 m such that *RH* increases from ~80% to 85 %, β_m consequently increases locally by ~1.2×10^{-6} m^{-1} sr^{-1}. This combined orographic effect and advection of aerosol emitted from the major sources (middle of the domain, Figure 4-2f) increases β_m in areas surrounding the city. An example of increased β_m, m_{MURK} and *RH* for EOF_1 in the UKV is shown in Figure 4-4.

The general distribution of loadings in the EOFs across the London domain are similar for both NWP models, as the synoptic winds are the most important factor driving the spatial patterns. While the LM EOFs have more spatial detail in the loading distribution around complex orography from the higher resolution elevation data used (cf. Figure 4-2b, Figure 4-3b), the resolution in the UKV orography is sufficient to generate orographic uplift and the consequential peak in loadings.

EOF_1 and EOF_2 from both NWP models, across each of the model levels analysed, explain a large amount of the total variability, typically between ~25 - 40 % (EOF_1) and ~20 – 40 % (EOF_2), respectively, of the total variability. EOF_1 most frequently has the highest loadings in the north-west and lowest loadings in the south-east of the London domain. EOF_1 typically occurs under southerly/south-easterly winds and the inverse pattern is found during northerly flow, common

daytime wind directions across the 28 case study days (Figure 4-6). However, EOF_2 tends to occur with lower domain mean m_{MURK} and lower mean temperatures from northerly winds, leading to higher *RH* across the domain. EOF_1 and EOF_2 are part of a multiplet that are highly inversely correlated (e.g. r = -0.85 at 111.7 m from UKV), likely because the wind direction that is favourable to one EOF, is unfavourable to the other (Figure 4-5).

For both NWP models, EOF_3 explains ~13 – 25 % of the total variability in β_m and can be interpreted as a response to two processes that increase β_m. In the UKV, EOF_3 has high loadings in the south-west due to relatively larger *RH* (not shown), and high loadings in the centre of the domain due to higher aerosol concentrations over London (not shown). As the VARIMAX rotation has merged the spatial variability from the two processes together, this shows the effects of both processes on β_m are temporally in phase.

For level heights above 471.7 m in the UKV, EOF_3 as in the lower levels is no longer present. Instead, the two processes represented in EOF_3 are represented in two separate EOF patterns at higher level heights. As the two EOFs represent parts of EOF_3, they can be considered EOF_3 sub-patterns (hereafter $EOF_{UKV,3\text{-}1}$ and $EOF_{UKV,3\text{-}2}$). $EOF_{UKV,3\text{-}1}$ contains the peak in loadings in the south-west of the domain from the high *RH* (not shown), while $EOF_{UKV,3\text{-}2}$ contains the high values near the domain centre from the second process involving high aerosol concentrations (not shown). Above 471.7 m, the effect of the two processes on β_m variability must have become more temporally distinct or out of phase, and hence separated into two different EOF patterns as a part of the VARIMAX rotation.

In the LM, the two scenarios of high RH in the south-west and high aerosol in the centre of the domain are more distinct at all heights. EOF_3 from the LM contains the effect of high RH in the south-west of the domain, with possibly some of the high aerosol towards the domain centre (Figure 4-3c). However, the peak in high aerosol in the centre of the domain occurs when $EOF_{LM,5}$ from the LM has the highest PC scores, strongly suggesting $EOF_{LM,5}$ contains this process and is equivalent to $EOF_{UKV,3\text{-}2}$.

Importantly, the complex $EOF_{LM,5}$ (Figure 4-3e) is very similar to $EOF_{UKV,3\text{-}2}$ (not shown) and highly likely to be the first to show urban-rural spatial patterns. $EOF_{LM,5}$ appears to combine two processes:

(1) G*rowth of the urban boundary layer* in the model: m_{MURK} and β_m increases in the centre of the domain (not shown), which could represent when aerosol built-up near the surface during the night and early morning is transported upwards during the day. When this process is most prevalent, $EOF_{LM,5}$ has high PC scores.

(2*) Effect of relatively high and increasing RH in the rural areas surrounding the city centre*: leading to greater β_m surrounding the city (not shown). When this process is prevalent, $EOF_{LM,5}$ has negative PC scores, showing the inverse of $EOF_{LM,5}$ is most relevant.

Additionally, there is some height dependence in the $EOF_{LM,5}$ PC scores above the surface layer which supports the likelihood that the two aforementioned processes are relevant. Above 111.7 m, PC scores tend to vary diurnally (Figure 4-7a): (1) decreasing after sunrise, (2) rapidly increasing during the day, and (3) rapidly decreasing before sunset. This temporal pattern in PC scores reflects: (1) the relatively high *RH* surrounding the city in the residual layer which could be due to lower air temperature the previous night steadily increasing; (2) the increase of aerosol in the city and reduction in rural *RH*, which could be as the boundary layer quickly grows to the level height, entrains the residual layer and vertically transports relatively higher concentrations of near surface aerosol; and (3) the reduction in air temperature and renewed increase of *RH* in the rural boundary layer, potentially following to the rural boundary layer being decoupled from the surface. At greater heights, stage 2 occurs later after sunrise and stage 3 occurs sooner before sunset (Figure 4-7b). This smaller range (shorter rise, lower fall) of PC scores of $EOF_{LM,5}$ could reflect the extra time required for boundary layer mixing processes to reach greater heights and entrain the existing residual layer. Additionally, The PC scores suggest $EOF_{LM,5}$ is most relevant at a height when the aerosol load is relatively low across the domain height analysed (not shown), leading to a relatively larger increase in β_m over the city as relatively higher concentrations of surface aerosol are likely mixed vertically.

$EOF_{LM,5}$ tends to represent a greater proportion of total β_m variability than its UKV counterpart, $EOF_{UKV,3\text{-}2}$. The higher resolution of both the urban characteristics and aerosol emission sources, in the LM ancillaries, may allow the urban effects to be better resolved. The native resolution of the aerosol and emission ancillaries are 1 km. These are coarsened for use in the UKV (to 1.5 km) and interpolated for the LM (to 333 m). Higher resolution LM ancillaries would provide greater variability in β_m, particularly in central London where emission sources and urban characteristics are particularly heterogeneous.

It should also be noted that high wind speeds can strongly affect the presence of EOFs by advecting urban (rural) air over to rural (urban) areas, distorting the diurnal pattern of PC scores.

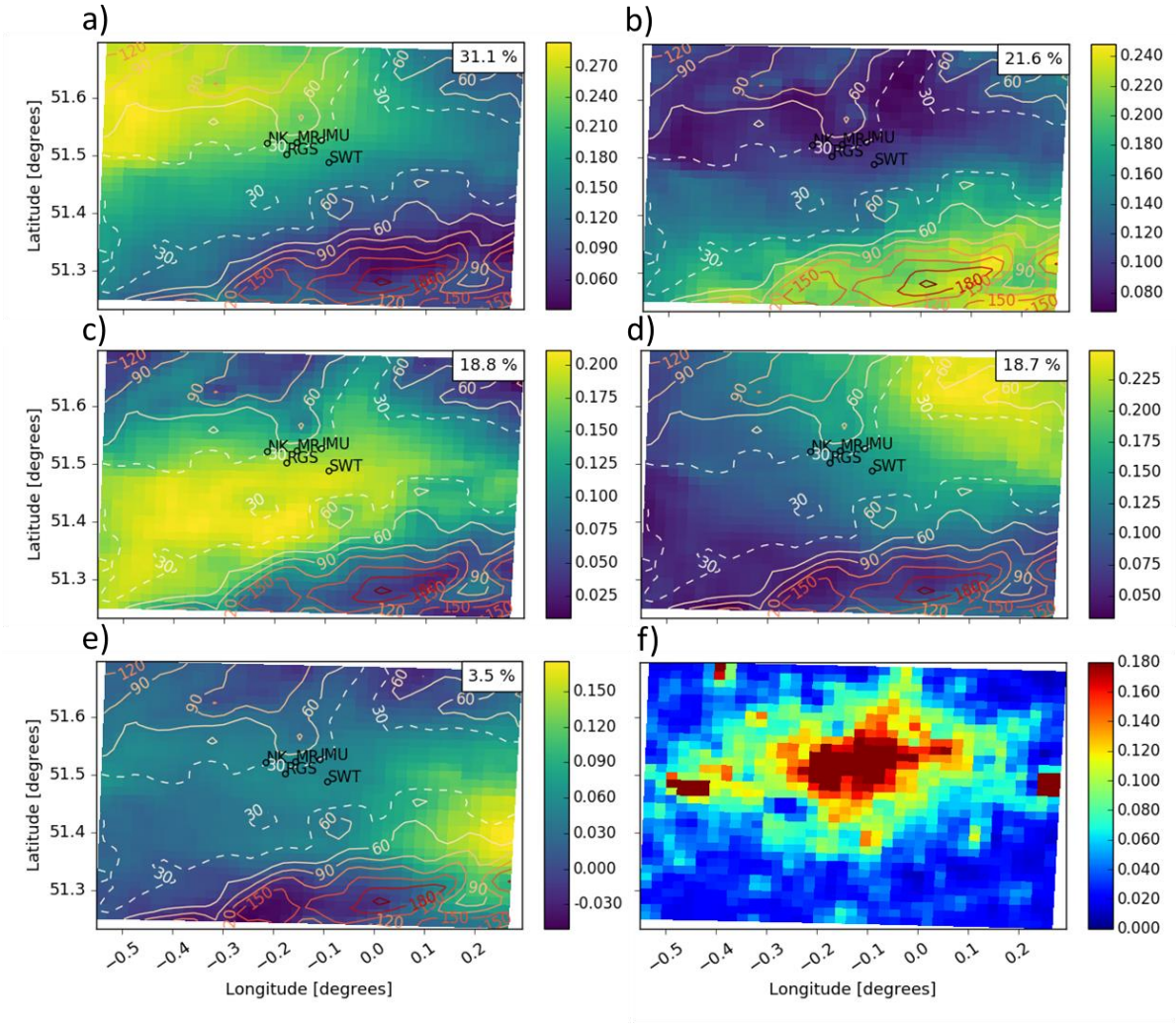

Figure 4-2: Examples of the main loaded, rotated EOF patterns that typically occur across different model heights from the UKV (1.5 km) during the daytime (colour) with topographic contours (lines, 30 m dashed). EOFs explain decreasing variability (% in top right) of the original dataset (111.7 m above ground level) (a) EOF_1 (b) EOF_2 (c) EOF_3 (d) EOF_4 (e) EOF_5 and (f) m_{MURK} background aerosol emissions [μg m^{-2} s^{-1}] climatology for July. LUMO ALC network (Figure 4-1) shown as dots and labels in the domain centre.

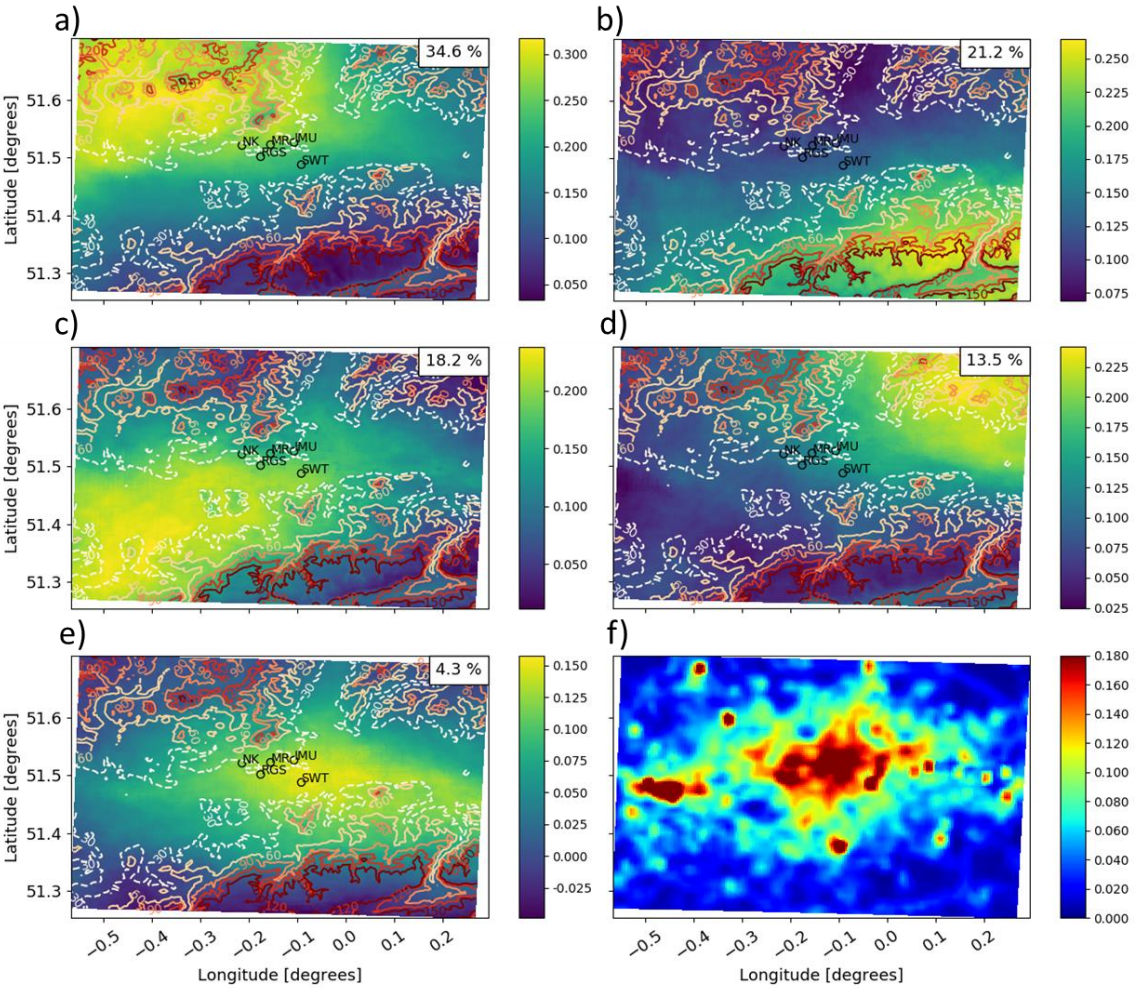

Figure 4-3: As Fig. 2 but for the London Model (LM, 333 m).

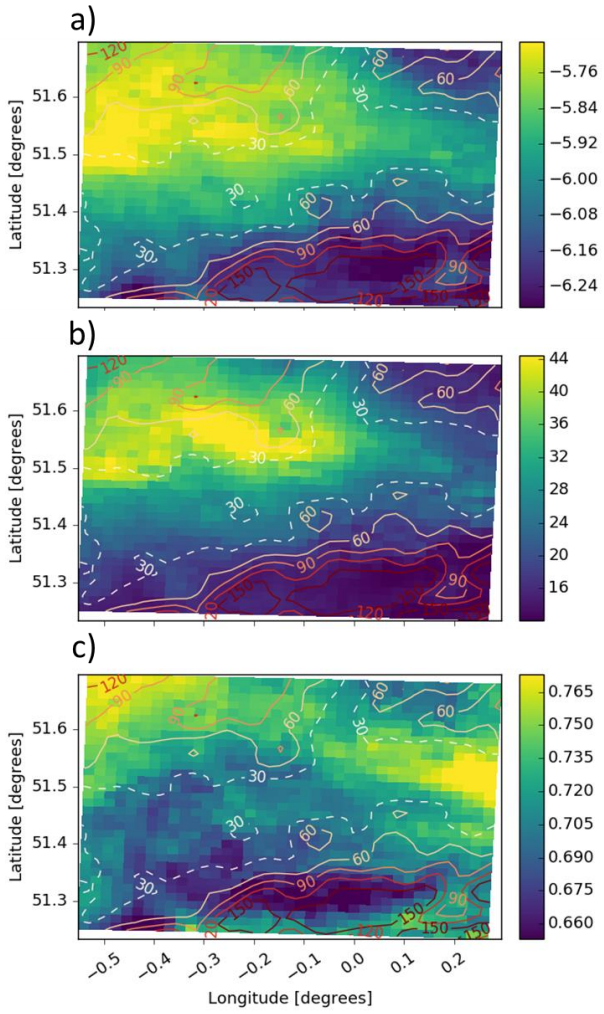

Figure 4-4: Median with respect to time of (a) $\log_{10}(\beta_m)$ [m^{-1} sr^{-1}], (b) m_{MURK} [$\mu g\ kg^{-1}$], and (c) RH [fraction] at model height 111.7 m above ground level across the entire London domain in the UKV (1.5 km) during daytime, when the PC scores for EOF_1 were above the 10^{th} percentile.

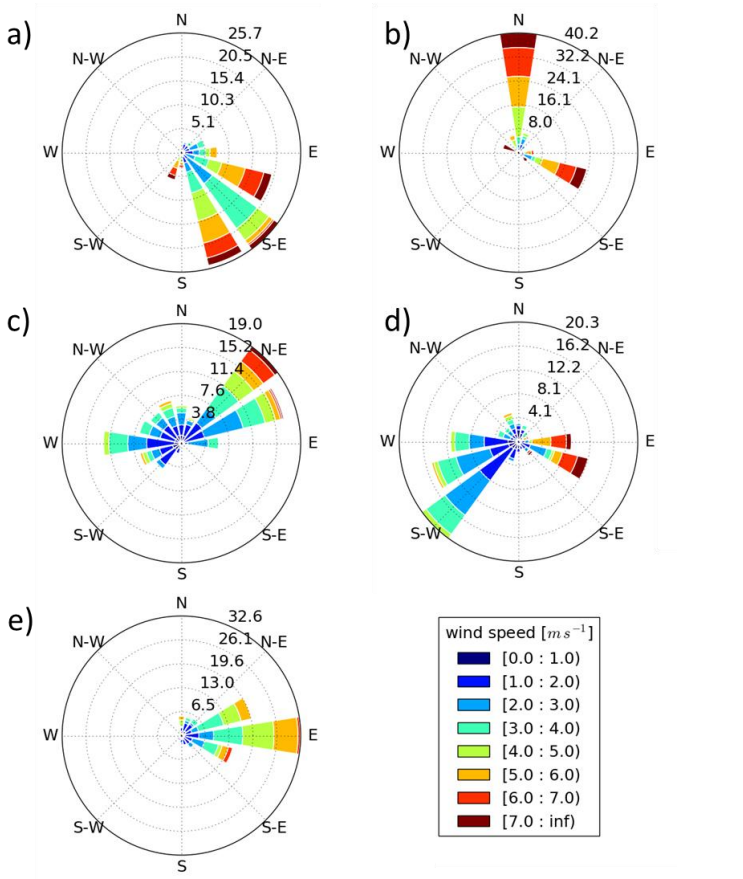

Figure 4-5: Paired daytime wind roses of UKV model wind speed [m s^{-1}] and direction (°) for each rotated EOF at model height 111.7 m (Figure 4-2). Wind speeds are temporally sampled when the scores for each PC were above the 10th percentile, and spatially sampled across the domain (wind characteristics most associated with each EOF). Bins extend from low inclusive and high exclusive (key). Radial axis (%) frequency of occurrence by wind direction bin.

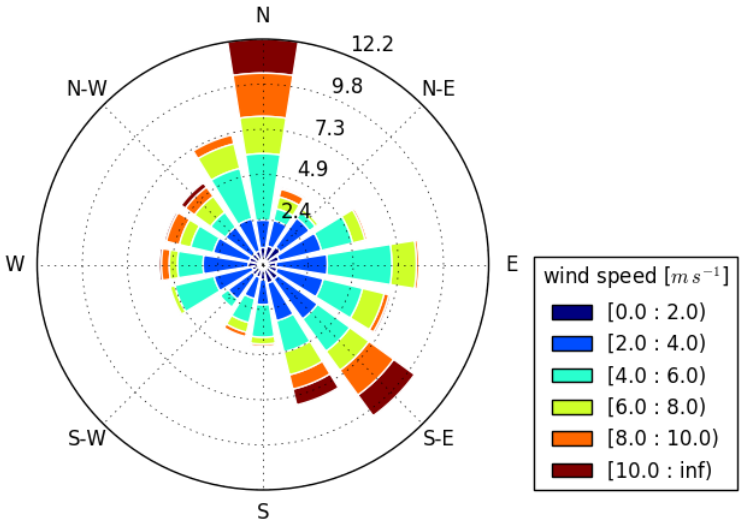

Figure 4-6: Combined day and night time wind rose (28 days) UKV model wind speed [m s^{-1}] and direction (°) at model height 111.7 m across the domain. Bins extend from low inclusive and high exclusive (key). Radial axis (%) frequency of occurrence by wind direction bin. Note different bin sizes compared to Figure 4-5.

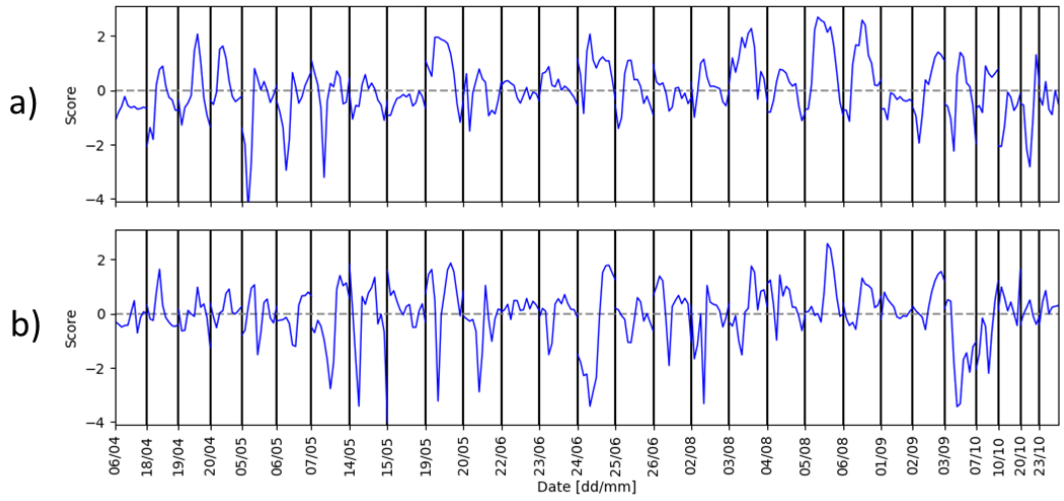

Figure 4-7: Daytime hourly principal component scores for all 28 days $EOF_{LM.5}$ (Figure 4-3e) for a) 261.7 m and c) 741.7 m.

Similar EOFs are derived from the PCA with VARIMAX rotation for the night-time period as for the daytime. Again, EOFs relate strongly to wind speed, wind direction, emission source location, and orography (not shown), with peak loadings in similar positions. However, there were some differences between the daytime and night-time EOFs.

Between model heights 325.0 m and 955.0 m, several EOFs have wavelike patterns that are not present during the day. For example, EOF_2 in both NWP models at 417.7 m (Figure 4-8), which generally occurs under north-westerly winds. Similar wave structures are found in EOFs associated with southerly winds (not shown). These wave patterns likely represent gravity waves produced as north-westerly flow passes over the Chiltern Hills (north west London domain, Figure 4-2f) (Figure 4-8a). Gravity waves can occur in statically stable conditions (Wallace and Hobbs, 2006) and have been found in EOFs derived from large-eddy simulations of the planetary boundary layer under weakly convective conditions (Wilson, 1996). These have been observed associated with the orography surrounding Greater London in both the UKV and LM (e.g. Chiltern Hills and North Downs; Lapworth and Osborne, 2019, 2016). Lidar observations of gravity waves have been shown to cause fluctuations in *RH* (Gibert et al., 2011), which translates to variability in β_m that is subsequently captured in the EOFs.

The presence of waves is visible in the w-wind component when the PC scores for EOF_2 are > 4 (e.g. Figure 4-8c, d). Under clear-sky conditions at night, the boundary layer is more likely to be stable, which can reinforce the presence of gravity waves (Wallace and Hobbs, 2006). The more unstable conditions during the day limits their formation (Figure 4-2, Figure 4-3). However, the frequency and spatial prevalence of these waves in the NWP models might be too large compared to reality,

causing them to be overrepresented in the EOFs. The LM used the simpler 1-tile urban surface scheme, which tends to underestimate sensible heat fluxes in the evening in London (Hertwig et al., n.d.), and in turn would further favour more stable conditions, reinforcing the waves' presence. Additionally, and common to both NWP models, some of the increased stability will also be caused by the underestimation of the anthropogenic heat fluxes over London (Bohnenstengel et al., 2014). Thus, these phenomena may not be as important as the EOFs derived from NWP model data imply.

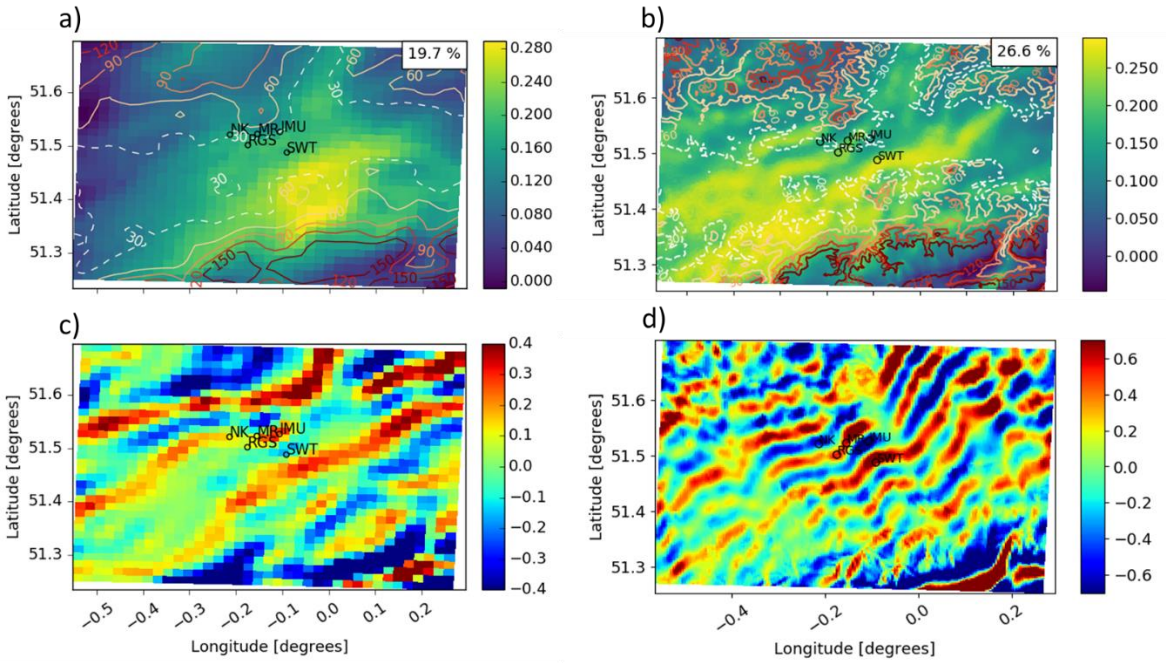

Figure 4-8: Nocturnal EOF_2 at the 471.7 m model level for a) UKV and b) London Model. Horizontal cross section of w-wind component when EOF_2 PC scores were > 4 on 23/10/2018 23:00 (at 471 m) from c) UKV and d) London Model.

The scales of spatial variability in β_m was also explored using semi-variograms, with smaller scale features being more dominant around the mixing layer height during low winds (section Appendix 4.B).

4.3.2 Application of spatial variability to network design

To find geographic regions where the variability in β_m is similar, Ward CA is performed on the unrotated EOFs (Ward, 1963; Wilks, 2011). CA is performed using the unrotated EOFs (uEOF) for day/night and for UKV/LM with the cluster numbers (n) set to 5, 7 and 20. For example, Figure 4-9 shows the clustering when n = 7 clusters as well as the associated distributions of ground elevation heights. Clusters are ordered in descending size (i.e. 1 has the largest area).

For both the UKV and LM, clusters occur over the city centre and in the surroundings for different wind directions, (Figure 4-9). The central LM cluster (7, $C_{LM,7}$) is smaller compared to its UKV

counterpart ($C_{UKV,5}$), which could be due to the higher loadings over central London compared to its surroundings in the first LM uEOF. The clusters have three elevation ranges (Figure 4-9c-d): (1) wide range in the north-west and southeast, (2) small range of low elevations for most of the others, and (3) predominantly higher elevation in the south.

Qualitatively, the boundaries between adjacent clusters follow contour lines of the orography (n=7), especially $C_{UKV,1}$, $C_{UKV,7}$ and $C_{LM,5}$, $C_{LM,6}$ (Figure 4-9). This relation between cluster shape and orography holds with different number of clusters (n=5 or n=20). When n=20 (Figure 4-10) the smaller clusters align with more detailed orographic features (e.g. valley variations and hill tops, $C_{UKV,12}$ in the south-west). With n=20 the relative size of different clusters remains similar, despite large spatial variations in urban heterogeneity and theoretically meso-scale meteorological processes across the London domain.

The similar cluster sizes might reflect both the dominance of larger scale processes, such as the synoptic winds, and the aerosol emission source distribution, in determining the spatial distribution of β_m compared to the smaller scale urban processes. However, the NWP models can only accurately represent the spatial aspect of features several times larger than their grid length resolution without aliasing (e.g. Lean and Clark, 2003). Consequently, the accurate spatial representation of similar and smaller scaled features on cluster shape and size will be strongly constrained by NWP resolution.

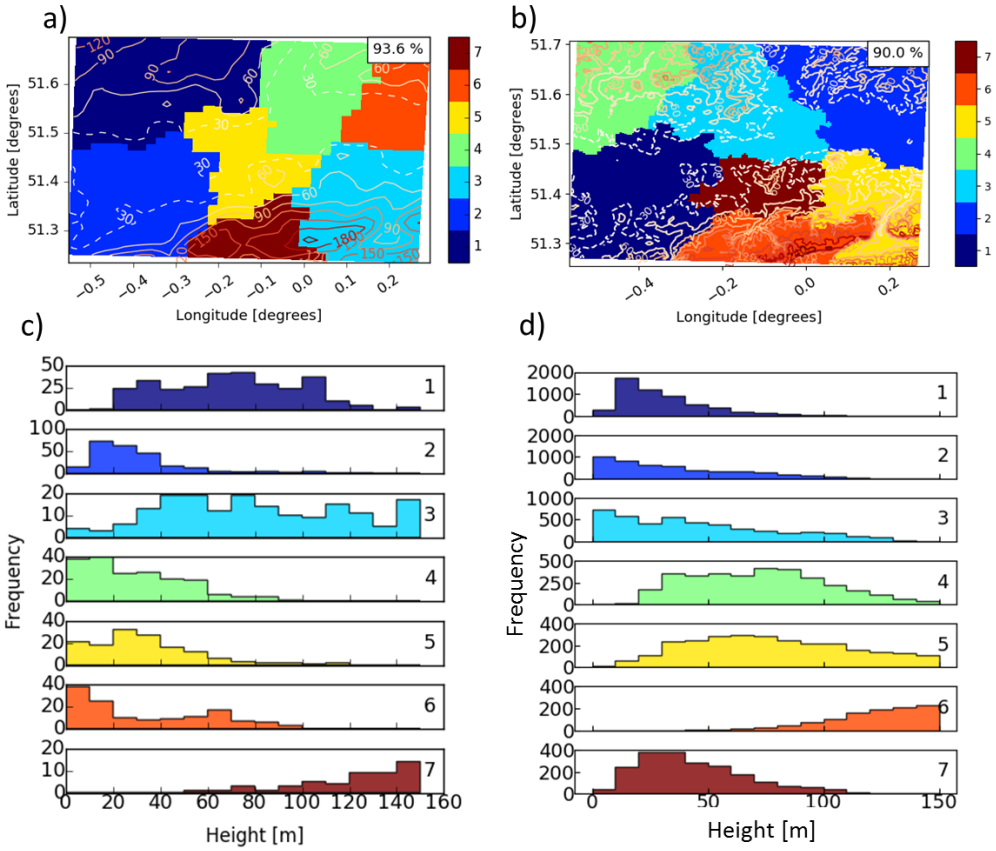

Figure 4-9: Ward cluster analysis (n=7 clusters) (of unrotated EOFs from PCA) of daytime β_m from a) UKV and b) London Model. Cluster analysis groups (colour patches) numbered in descending order such that C_1 has the largest area and C_7 the smallest. Orography contours overlaid (lines), with 30 m orography contour line dashed for clarity. All EOFs included explain ≥ 1% of the variability in modelled attenuated backscatter. Orography (10 m bins) in each cluster for c) UKV and d) LM. Note frequency (y-axis) varies.

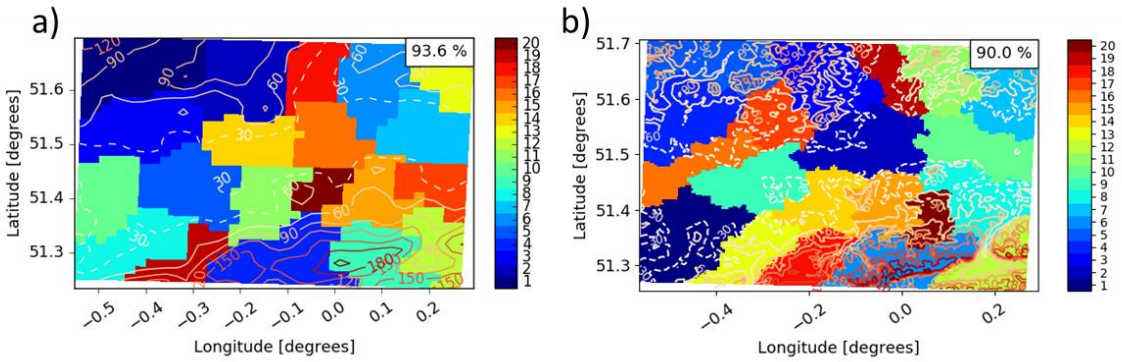

Figure 4-10: As Figure 4-9a,b, but with n=20 clusters.

The nocturnal CA results (Figure 4-11) are similar to the daytime (Figure 4-9b), in terms of cluster shapes and sizes, due to the impacts of wind, relative aerosol emission location and orography on the EOFs, which do not depend highly on time of day (section 4.3.1). $C_{UKV,7}$ does shown a slight elongation in the south-west to north-east direction which qualitatively aligns with the relatively high MURK ancillaries in the central domain. This is likely because of the aerosol build up over time in the surface layer at night strongly affecting the lower model level uEOFs. The elongated shape of $C_{UKV,7}$ could also be partially due to the presence of nocturnal gravity wave features but the effect would likely be limited as the β_m variability due to the waves was small in the uEOFs.

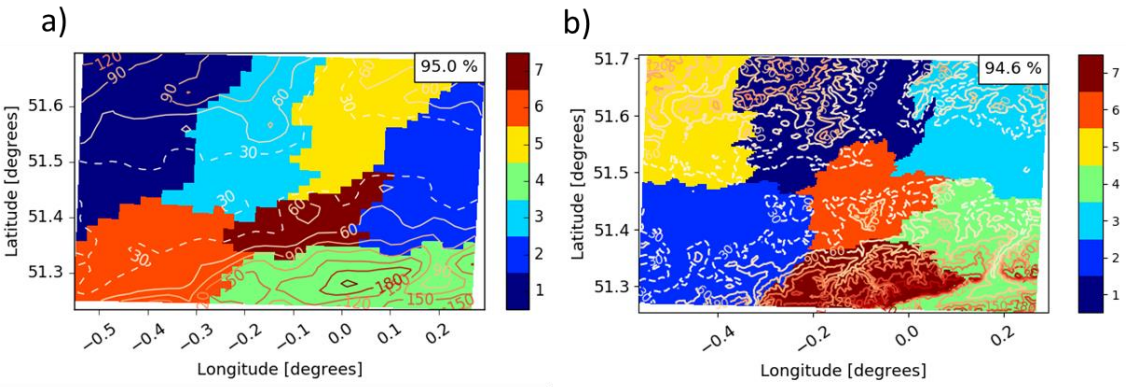

Figure 4-11: As Figure 4-9a,b, but for night.

The CA maps can be used to inform the deployment of ALC instruments. As Ward clustering aims to minimise the variability in β_m within a cluster, one logical approach would be to locate a single ALC instrument in each cluster. Measurements from an instrument located in the daytime cluster 5 would be most representative of the city centre (n=7). At night the cluster shapes and areal extents change slightly, so the most representative instrument location may change but can be understood from the night-time CA maps.

4.4 Conclusions

To assess what drives spatial variations in aerosol (and relative humidity), within the boundary layer, across a region in and around a megacity under cloud-free conditions, a method is developed to identify common spatial patterns of variability. Attenuated backscatter (β_m) is modelled using the aerosol forward operator (aerFO) and is driven by two numerical weather prediction (NWP) at two resolutions: 1.5 km Met Office UKV (1.5 km) and 333 m research London Model (LM) for a domain around Greater London. Principal component analysis (PCA) with VARIMAX rotation is used for two periods (day, night), to create empirical orthogonal functions (EOFs) for each model level.

PCA results are combined with cluster analysis (CA) to identify areas of similar aerosol variability. The CA results could then be used to identify potential locations for automatic lidar and ceilometer (ALC) instruments to reduce redundancy and maximise uniqueness across a network with respect to observing features at spatial scales larger than the model resolution used.

Common EOF patterns are identifiable across different heights during both day and night with similar results for both NWP models. Many EOFs from both models were strongly related to orography, wind and relative aerosol emission source location, suggesting these are the main factors determining the spatial variability in β_m over the study area.

Although the relatively coarse UKV (1.5 km) is sufficient to represent the main spatial patterns of β_m variability, the higher horizontal resolution of the LM can identify more detailed spatial features. The LM includes more detailed ancillaries (e.g. topography, emissions), but some could be even more detailed (e.g. emissions here interpolated from 1 km) to possibly provide greater spatial variations in aerosol and relative humidity patterns. Improved urban energy balance fluxes within

the NWP models, notably anthropogenic heat emissions, may be another factor that could enhance detection of small-scale effects.

It is concluded that:

- The EOF results can be mostly explained by orography, transport of aerosols from source locations (mainly in city centre), and wind direction. As only 28 days (rather than a climatology with all possible directions) are analysed, this impacts the conclusions that can be drawn. Peak loadings are found downwind of the city caused by wind direction dependent advection. EOFs could be quite different if precipitation conditions were analysed, as this would include wet deposition and changed aerosol in the boundary layer.
- Possible gravity waves in the EOFs are identified in the residual layer at night in both NWP models. Vertical oscillations from gravity waves are associated with changes in relative humidity, which causes β_m variations detected in the EOFs.
- CA identifies distinct regions from unrotated EOFs (uEOF, from PCA using the UKV and the LM) that could inform an ALC sensor network design when observing features of larger spatial resolution to the models. The cluster shapes are similar between the LM and UKV, with cluster boundaries following topographic contours.
- No major differences are found in CA results when the cluster number is set to 5, 7 or 20.
- CA results are similar between the day and night periods, despite the wave-like patterns in night-time EOFs.
- The CA-uEOF results with the coarser resolution model are comparable to the higher resolution (~333 m) NWP model.
- In both PCA and CA-uEOF results, the LM may not have resolved as much of the variability in β_m because of low native resolution of the urban geometry and emission ancillaries. Improving the resolution of these may give greater benefit from higher resolution NWP models, through more accurate β_m estimates, and consequentially more representative EOFs and better clustering.
- The identification of distinct regions in CA is mostly constrained by NWP model resolution and the spatial scales of features it can effectively resolve. To inform network design for ALCs to observe smaller spatial features, higher resolution NWP models are needed.

Using the PCA-CA methodology could bring several benefits to improve NWP model forecast accuracy. Firstly, improving the spatial distribution of instruments could result in observing more meteorological processes. This could in turn provide more information for data assimilation to further improve analyses of aerosol and relative humidity. Furthermore, the PCA-CA technique could also be used on sub-samples of NWP data in order to focus on better capturing information for particular

meteorological situations or regions, for example above average aerosol events or the upper extent of the boundary layer where for instance, observations can be sparse in urban areas (Barlow, 2014). In addition, the PCA-CA technique could aid the spatial interpretation of verification statistics. The CA highlights spatial regions where the grid cells covaried most similarly, therefore the spatial applicability of verification statistics using ALC instruments located in identified cluster regions can be better understood.

Acknowledgements

EW was funded by the NERC Scenario DTP and University of Reading. Instruments operated and funded by LUMO (NERC ClearfLo, NERC TRUC, KCL, UoR, EUf7 BRIDGE, H2020 UrbanFluxes), and the vast number of people who help keep the equipment operating on a daily basis and the institutions that provide sites. Funding from EPSRC DARE and Newton Fund/Met Office CSSP China (SG) support are acknowledged. Mapped terrain, river and road data provided by Ordnance Survey (OS).

Thanks also to Anthony Illingworth and Helen Dacre for their helpful advice and comments.

Appendix 4.A Clear Sky days

To select the 28 clear-sky cases daily plots of observed attenuated backscatter from ALC of the London Urban Meteorological Observatory (LUMO, http://micromet.reading.ac.uk/) were reviewed for the period 01/04/2018 to 31/10/2018. The clear days selected (Table 4.A.1) are absent of hydrometeors.

Table 4.A.1: Days in each month of 2018 used for the analysis

April	May	June	July	August	Sept	Oct
06	05	22		02	01	07
18	06	23		03	02	10
19	07	24		04	03	20
20	14	25		05		23
	15	26		06		
	19					
	20					

Appendix 4.B Spatial scales of forward modelled attenuated backscatter from modelled semi-variogram range and sill

4.B.1 Background

The scales of spatial variability in forward modelled attenuated backscatter (β_m) over London is explored by examining semi-variograms (or often just called variograms). The procedures used primarily follow Oliver and Webster (2014).

Semi-variograms, often the first step in kriging (a spatial interpolation technique), describe the spatial autocorrelation of points across a domain (ESRI, 2019; Matheron, 1963; Oliver and Webster, 2014). As points that are geographically close are likely to be more similar than those further apart, the auto-correlation decreases with distance. Semi-variograms can be used to help determine the spatial scales of dominant processes in a domain.

4.B.2 Methods

To construct the semi-variograms the semi-variance (γ) calculated as a function of lagged distance using sampled data:

$$\gamma(h) = \frac{1}{2 * m(h)} \sum_{j=1}^{m(h)} \{z(x_j) - z(x_j + h)\}^2 \quad (4.B.1)$$

where h is the lag distance, m is the number of samples in the current lag, z is the sampled data and x its location. γ effectively describes the similarity of points across the domain, where greater dissimilarity between the points leads to higher values of γ.

A semi-variogram model is fitted (see section 3.1 of Oliver and Webster, 2014 for possible functions) to give a γ that is continuous for all h. Two common properties that the semi-variogram models (Figure 4.B.1) possess are the:

(1) *range:* The distance that γ in the model semi-variogram tends to become more constant (ESRI, 2019) or the scale of the dominant processes responsible for variability across the domain. Points at distances < range should be affected by same processes. Points at distances > range are affected by different processes (i.e. greater dissimilarity).

(2) *sill:* γ at the range.

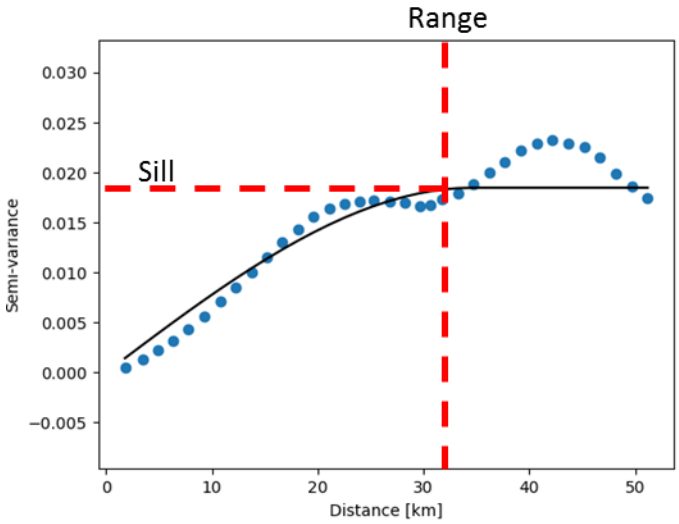

Figure 4.B.1: Semi-variogram produced from forward modelled attenuated backscatter in a north-south strip of the London domain (ref methods section in paper 3) from the UKV at a model level height of 21.7 m, at 15:00 on 07/10/2018. Semi-variance as a function of distance (blue dots) and a spherical semi-variogram model fit (black line). Sill (horizontal dashed red line) and range (vertical dashed red line) are also shown.

Semi-variograms were produced from β_m derived using the aerFO from section 3.2.3, estimated hourly from Met Office UKV (Davies et al., 2005) for each model level to 2075 m agl, for the 28 clear-sky days (Appendix 4.A) using the 21Z forecast from the previous day. Semi-variograms used a fixed lag of 1.5 km (i.e. UKV grid length (blue points, Figure 4.B.1)), as little sensitivity was found to lag choice. Using a spherical semi-variogram model (black line, Figure 4.B.1) fit to the semi-variances the range and sill (red dashed lines, Figure 4.B.1) are obtained. Ordinary Kriging with 'PyKrige' (version 1.4.0, https://pypi.org/project/PyKrige/) Python 2.7 is used to calculate the semi-variance and semi-variograms.

If data are anisotropic or have directional dependencies due to the wind speed and direction, calculated semi-variograms can be interpreted incorrectly. Here, this could lead to an under or overestimation of the range as features are advected across the domain and consequently, incorrect understanding of the scale of dominating meteorological processes. Transforming the data prior to calculating γ can help minimize these effects (section 3.2.4 of Oliver and Webster, 2014). To explore if strong anisotropy exists, several sets of directional semi-variograms are constructed with small subsamples of data, semi-variograms are constructed for model levels 5 and 111.7 m for 90, 180, 270 and 0 °, with anisotropic rescaling, for the 03 August 2019. Differences in the calculated *range* between the semi-variograms at the same time and height are < ~±5 km. Qualitative assessment found very little variation in semi-variogram shape and sill. Therefore, no anisotropic correction is applied in the main analysis.

Prohibitive computing resources would be needed to calculate the semi-variograms for London Model (333 m) as binning based on pair-wise distances between points between all pair-wise

distances for every grid cell are needed. This could possibly be overcome using either parallel processing or by constructing and inserting a function into the semi-variance calculation that takes advantage of the regular grid spacing from the numerical weather prediction models.

4.B.3 Results and discussion

Typically, shorter ranges indicate smaller scale processes dominate the variability of β_m across the domain, whereas smaller sills indicate a lower overall variability across the domain. The range may vary with the mixing layer height (MLH), with morning range for levels around the MLH being variable (< 30 km, Figure 4.B.2). Overnight there is build-up of aerosol near the surface layer, with low winds (case shown) limiting horizontal mixing creating relatively high heterogeneity in aerosol mass concentrations. As the MLH grows, the *range* decreases at higher model levels, possibly indicative of convective processes dominating the β_m variability. After sunset, the mixing layer collapses and the range gradually increases at levels near the top of the residual layer, possibly as the more heterogenous convective processes weaken and light horizontal mixing equalises some of the variation in horizontal mass concentration. As the MLH reduces, the range near the surface layer begins to decrease.

Similarly, the sill also varies with the MLH (Figure 4.B.3). The sill is relatively large in the surface layer, increases with height as the MLH increases, and is high in the surface layer when the mixing layer collapses. This likely reflects the greater horizontal variability in β_m at those heights due to the influence of smaller, meso-scale processes as described above. However, unlike the range, the sill also has a clear and strong relation with the top of the residual layer. The highest sill values occur at those heights. It is possible that variograms at these heights are including air both below and above the boundary layer, with high and low aerosol concentrations respectively, leading to a large amount of variability.

The strength of the relations between the range and sill with the MLH or residual layer are only clear in cases with low daily averaged wind speeds (< ~ 5 m s^{-1} < 2075 m), with mixed clarity at higher wind speeds. As higher wind speeds are typically driven by synoptic scale meteorology (i.e. larger than the domain), it is beyond the maximum range possible in the domain analysed here.

With many of the semi-variance estimates being bi-modal structures (Figure 4.B.1) and few suitable semi-variogram fitting models bi-modal distributions (Oliver and Webster, 2014), subsequent range and sill calculations are poor. Therefore, despite smaller scale processes having an identifiable effect in the calculated semi-variance, the lack of suitable model-fitting hampered application of this method. As the calculated range and sill are effectively averaging across the domain, small scale variations in part of the domain (e.g. adjacent residential and industrial areas), may not be clearly identified if other more common spatial scales exist (urban vs rural differences).

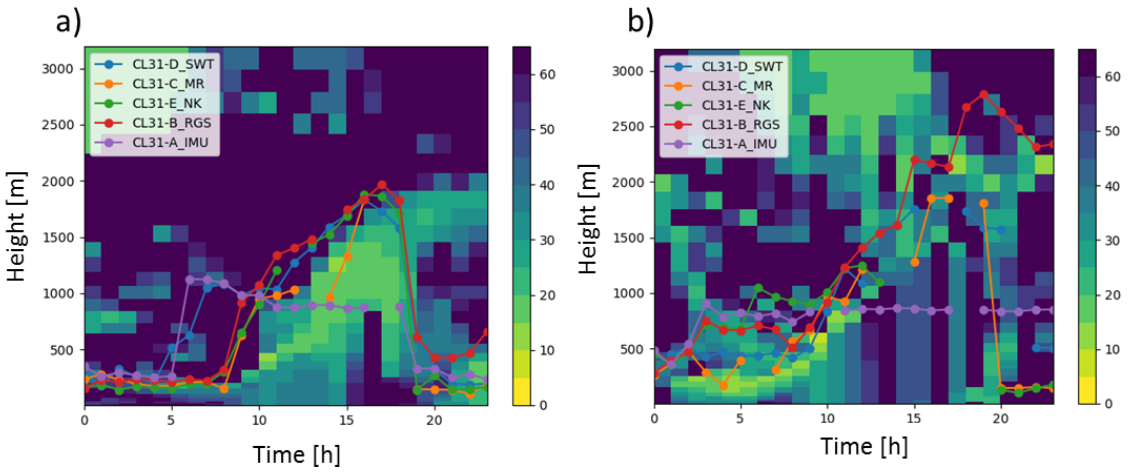

Figure 4.B.2: Semi-variograms range [km] (colour) determined for UKV forecasts of β_m over Greater London with mixing layer height derived from ceilometer observations (circles and lines) (Figure 4-1) for a) 03/09/2018 (mean daily wind speed 2.6 m s^{-1} m < 2075 m) and b) 06/08/2018 (2.9 m s^{-1} < 2075 m).

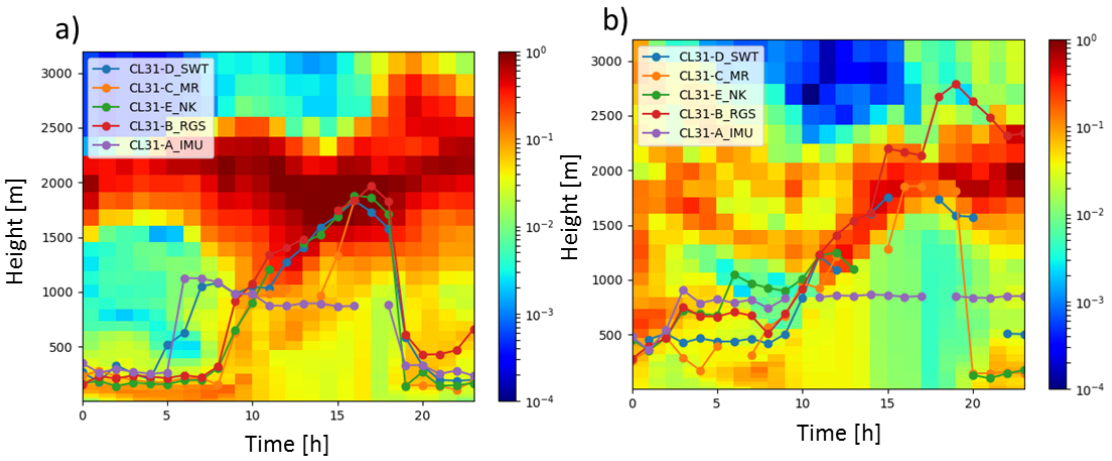

Figure 4.B.3: As Figure 4.B.2 but, showing the sill [m^{-2} s^{-2}] (colour, note logarithmic).

Chapter 5 Conclusions and contributions

5.1 Main contributions

Numerical weather prediction (NWP) models and chemistry transport models (CTM) produce forecasts of aerosol which are becoming increasingly sophisticated. However, uncertainties remain in the forecast accuracy and their evaluation can be challenging. As automatic lidar and ceilometers (ALC) that measure vertical profiles of attenuated backscatter (β_o) are becoming both increasingly sensitive to aerosol and more common in observation networks, they could be used for aerosol forecast evaluation. However, to relate β_o to physical aerosol properties present in aerosol forecasts, a forward operator is required. Thus, the key contribution of this thesis is:

(1) *An aerosol forward operator (aerFO) that can estimate ALC aerosol attenuated backscatter (β_m) in clear-sky conditions (no hydrometeors) using inputs of relative humidity (RH) and aerosol*. The aerFO is novel and built with a wide range of default aerosol options which allow it to be exceptionally flexible, computationally cheap and usable with a wide range of NWP, CTM or observational data as inputs, for different locations and multiple wavelengths. Unlike existing lidar aerosol forward operators (Benedetti et al., 2009; Chan et al., 2018; Geisinger et al., 2017) this is highly flexible, requires less input and is not model specific.

Through development and utilisation of the aerFO, several other key contributions have been made:

(2) *Identification of the importance of different aerosol characteristics to accurately estimate aerosol optical properties.* Sensitivity tests, evaluation of β_m against ALC observations and evaluation of modelled with observed aerosol characteristics, have identified the importance of several aerosol physical characteristics and aerosol speciation in estimating aerosol optical properties.

(3) *Evaluation of NWP parameterisations including the MURK visibility and urban land-surface energy exchanges.* Evaluation of the existing aerosol parameterisations in the MURK scheme provided insight into what aerosol characteristics to prioritise for improvement, in the future development of the scheme for urban areas. Evaluation of β_m against ALC observations demonstrated the importance of the land-surface scheme in the distribution of aerosol and *RH*. aerFO can be used to indirectly assess the impacts of the land surface scheme.

(4) *Improve understanding of the spatio-temporal variability of β_m over a megacity (London, UK).* Common spatial patterns in βm variability are identified in the boundary layer, in two high resolution NWP models using principal component analysis (PCA). These are found to relate to orography, wind and emission source location.

(5) *Develop a method to inform ALC network design.* Cluster analysis (CA) of PCA output can produce maps to identify homogeneous regions of βm variability. Using these can inform ALC network design, for example to allow better management of measurement redundancy (or not) between instruments.

5.2 Conclusions

Although the aerFO can be driven using *RH* and a wide range of aerosol data for different locations, in this thesis, the Met Office (MO) NWP models and greater London are the focus.

To understand the relative importance of different input and parameterised variables required to estimate β_m is highly challenging given the large and varied complexity in aerosol characteristics. This is made clear through sensitivity studies of modelled aerosol physical characteristics on optical characteristics, evaluation of β_m against β_o from the ALC network in London and estimating the particle extinction-to-backscatter ratio (lidar ratio, *S*) from aerosol observations. Despite the complexity of aerosol variation, several conclusions are drawn:

- Sensitivity studies of β_m to aerFO inputs shows that β_m has an exponential relation with *RH* and β_m becomes increasingly sensitive to *RH* as *RH* increases (especially when $RH > \sim 60\%$) due to the hygroscopic growth of aerosol particles. RH becomes increasingly important in accurately estimating β_m at high *RH*, and errors in other related meteorological variables such as air temperature, can become important as an indirect effect on β_m estimates. Additionally, aerFO β_m estimates are still valid at high *RH* even when cloud droplets may have formed, and the clear-sky assumption violated. This is because of the exponential relation between the parameterised particle extinction enhancement factor from aerosol hygroscopic growth of water vapour ($f_{ext,rh}$) and *RH*.
- Through sensitivity studies of modelled aerosol based on aerosol observations it can be concluded that that aerosol optical properties of the dry particle extinction efficiency (Q_{ext}) and $f_{ext,rh}$ can vary greatly for different aerosol species found in urban areas, including ammonium sulphate, ammonium nitrate, sea salt, black carbon. Black carbon is especially important to consider in urban areas given its relatively high abundance (Liu et al., 2014), and very strong absorption and hygrophobic properties. Additionally, organic carbon is important because of its hydrophobic properties. Therefore, accurate estimates of aerosol optical properties are needed to represent the different aerosol species in the bulk aerosol. The current lack of information on the particle size distribution for each aerosol species and accurate bulk aerosol speciation, create large uncertainties in Q_{ext} estimates, with additional uncertainties in the hysteresis in aerosol swelling that impact estimates of $f_{ext,rh}$.
- Studies of the sensitivity of aerosol measurement to wavelength used concluded that aerosol optical properties can vary largely between different common ALC wavelengths (905 cf.

1064 nm), as well as within the range of uncertainty in wavelength for single ALC instruments (Vaisala CL31 905 ± 10 nm). Vitally, Q_{ext} and $f_{ext,rh}$ should be estimated separately for different ALC models (Vaisala CL31 cf. Lufft CHM15K), and ideally for different ALC instruments as well, once the wavelength variation of the instruments are better known. The correction for water vapour absorption can vary greatly within the range of some ALC wavelengths (e.g. Vaisala CL31: 905 ± 10 nm). It would be highly advantageous for ALC manufacturers to provide the central wavelength and wavelength variation for each instrument, in order to reduce the wavelength uncertainty in estimates of β_m.

- The coarse and accumulation aerosol modes (diameter: > 0.08 μm) are very important to estimate aerosol optical characteristics in urban areas. Although previous MURK visibility studies found the accumulation mode to be important for total aerosol optical properties (Clark et al., 2008; Haywood et al., 2008), the coarse mode is also very important and should be appropriately represented. Fine mode aerosol (diameter: < 0.08 μm) is less important in estimating σ_{ext} or β_m in urban areas because of their small size, despite their large number. However, it is possible that the importance of different aerosol properties in accurately estimating aerosol optical characteristics may vary between urban and non-urban areas.
- It is possible to derive S where Raman lidar or sunphotometer data are unavailable using *in-situ* aerosol observations and aerosol modelling. However, its derivation using aerosol observations is challenging. S estimated from *in-situ* observations at the urban background site of North Kensington in London is found to vary between 14 – 80 sr (mean = 43 sr). The results are comparable to urban Raman lidar or sunphotometer based data (Müller et al., 2007; Stachlewska et al., 2018; Wang et al., 2016). Seemingly, S has a strong dependence on RH, but this dependence may be masked by other uncertainties (e.g.) related to accurate aerosol speciation, particle size distribution by species, and internal mixing. The amount of elemental carbon is found to be highly influential on the estimation of S. This is likely to be especially important in cities, given its relative abundance.

Evaluation of β_m using MO UKV data to force aerFO was undertaken for: (1) 12 clear-sky days (2015 -16) with five CL31 ceilometers (Table 2.3), and (2) 7 clear-sky days (2014 - 15) with one CL31 (North Kensington). From this several conclusions are drawn related to NWP parameterisations and evaluation against observations:

- The monotonic parameterisations of the total number concentration (N) and mean dry volume radius (r_{md}) as a function of m used in the MURK visibility scheme (Clark et al., 2008), allows adequate estimate of the mean N and r_{md} in London, but largely underestimates the variance of both by ~ 2/3. This is despite using derived climatological means from observations in London in the scheme. However, using N from observations improves the β_m

variability (more so than improving r_{md} values). Improving N in the MURK scheme should be a development priority as this will enhance estimates of aerosol optical characteristics.

- The NWP urban surface scheme is extremely important to accurately estimate β_m, due to the indirect effect of vertical mixing. Better representation of surface fluxes can improve vertical mixing and consequently the aerosol distribution, preventing build-up of aerosol near the surface (especially in the morning after sunrise). Additionally, RH estimates can be improved as the underestimation of air temperature in this surface layer is reduced. Spring case studies with the (1) older Best 1-tile scheme and (2) newer MORUSES scheme, found that using MORUSES improved vertical mixing and sensible heat flux in the morning, reducing the mean bias error ($\Delta\beta = \beta_m - \beta_o$) below ~250 m by more than half. Therefore, urban NWP modelling studies involving aerosol (and RH) require a good urban land-surface scheme (i.e. improvements to this will benefit numerous applications).
- Evaluation of β_m with β_o is highly dependent on the quality of ALC observations. Instrumental noise, artefacts and incomplete overlap are well-researched in the literature (Bachour and Perez-Astudillo, 2016; Kotthaus et al., 2016). These effects increase differences between β_m and β_o. This additional uncertainty can complicate analyses as β_m - β_o differences add uncertainties into the evaluation, making it more difficult to discern errors in β_m due to errors in the aerFO inputs variables (want to identify) and instrumental error (do not want to identify). Operating ALC with the latest firmware and hardware can reduce instrumental error and lead to a fairer assessment of β_m. Additionally, this would also improve the quality of potential, future data assimilation of ALC observations.

Principal component analysis (PCA) using 28 clear-sky days of $\log_{10}(\beta_m)$ data derived from the UKV (1.5 km) and London Model (LM, 333 m) has common spatial and temporal patterns in β_m variability over Greater London. Cluster analysis (CA) of the PCA output to identify homogenous regions of β_m variability, leads to the following conclusions relevant to ALC network design:

- Common spatial and temporal patterns in β_m variability identified from the NWP data are largely driven by orography, winds and relative aerosol emission source locations during both the day and night. The coarser UKV resolves the same broad patterns as the LM, but the LM resolves some higher resolution spatial variations in β_m (e.g. small river valleys). This suggests NWP models at ~1.5 km can adequately capture large-scale variation in β_m over a megacity. However, higher resolution ancillaries (e.g. orography, emissions) in the LM would likely allow more urban features to be identified (e.g. growth of the boundary layer over the city). More detailed ancillaries or even higher resolution NWP models could provide more spatially detailed urban spatial patterns.
- The PCA of NWP output is highly flexible and could be used for larger areas (e.g. entire country) if processes being targeted were to inform national ALC networks. CA of the PCA

output allowed spatial maps to be created with homogenous regions of β_m variability. These could be used to identify potential locations for ALC instruments to be deployed to measure the type of variability found in the NWP models.

- Like the PCA output, cluster shape and size from the CA are related to orography and emission location with little difference between the UKV and LM (day, night). CA with more clusters (e.g. 20) have boundaries that follow the topographic contours more closely. Despite large-scale, wavelike features being apparent in nocturnal spatial patterns, they had little impact on the cluster shape, suggesting they are of low importance.
- Identification of distinct clusters is mostly constrained by NWP model resolution and the spatial features they can effectively resolve. Higher resolution NWP models are likely needed to further refine cluster shapes or better account for smaller scale features.

5.3 Limitations and recommendations for future work

As the developed aerFO is unsuitable for conditions when hydrometeors are present (e.g. fog, cloud or rain), parameterising these would allow greater utility. Reintegrating more of the original forward operator from Charlton-Perez et al. (2015) would be a sensible starting point.

The aerFO currently only parameterises the accumulation aerosol mode (diameter range, D, 0.08 – 0.8 μm) effects but the coarse mode (D = 0.8 – 10 μm) can contribute strongly to β_m for 905 nm (Chapter 3) and may be important for other ALC wavelengths. Therefore, the coarse mode should be included into future aerFO versions, possibly by estimating the total number concentration and dry mean radius of the coarse mode in a similar way to the accumulation mode (i.e., using climatological mean number mean concentration and dry mean radii for each).

The aerFO parameterises aerosol swelling effects on optical properties in a single coefficient ($f_{ext,rh}$) that can be turned on or off (Chapter 2). Kotthaus et al. (2018) found discrepancies between the urban boundary layer height derived from ALC and doppler lidars and hypothesised it could possibly be due to aerosol swelling altering the vertical gradients of attenuated backscatter from the two instrument types differently. The aerFO could be used with $f_{ext,rh}$ turned on and off to help explore this hypothesis for both types of lidars.

Although the analysis is undertaken for a typical European city (London), the generalisation of conclusions to other urban areas contains some uncertainty, notably, for cities with large differences in land use, urban characteristics and climate to London. Ideally, in other cities ALC instruments and networks could be used to evaluate β_m to generalise the analyses.

With ALC networks, such as LIDARNET (Adam et al., 2015) and within EARLINET (Pappalardo et al., 2014), attenuated backscatter data are stored using consistent data processing and calibration, making measurements more comparable. The aerFO could be used to understand variations in

attenuated backscatter and to evaluate aerosol forecasts across multiple cities, rural areas, countries or continents. As many ALC networks can monitor and track aerosol layers including volcanic ash (Emeis et al., 2011; Pappalardo, 2013), aerFO could be used to forward model forecasts of volcanic ash events in near real time given aerFO's low computational cost and ability to estimate backscatter for instruments with different wavelengths. The aerFO could also be used for data assimilation of observed attenuated backscatter to improve aerosol forecasts and possibly other processes if aerosols become coupled with other variables in the forecast models (i.e. aerosols intercepting radiation thereby modifying surface sensible heat fluxes). Given that boundary layer depth in NWP models can be different to reality, and the importance of vertical mixing (section 2.7.1), correcting total volumetric mass of aerosol in the vertical atmospheric profile should be undertaken before correcting the mass as a function of height.

Spatial variations in β_m explored at two model resolutions over London (1.5 km and 333 m) allowed indirect evaluation of aspects of the NWP models, including the urban surface scheme (section 2.7.1). Using the higher resolution NWP models being developed, the aerFO could be used to evaluate aspects of these models, including the extent of vertical mixing and indirect effects of the urban surface scheme.

aerFO is designed to work with a wide array of aerosol inputs, such as from chemical transport models (e.g. Met Office Numerical Atmospheric-dispersion Modelling Environment (NAME, Maryon et al., 1991), or air quality models (e.g. Air Quality Unified Model (AQUM, Savage et al., 2013)). These could benefit from being evaluated using ALC data (e.g. LIDARNET network) given the observations' high temporal and vertical resolution.

References

Abdalmogith, S.S., Harrison, R.M., 2006. An analysis of spatial and temporal properties of daily sulfate, nitrate and chloride concentrations at UK urban and rural sites. J. Environ. Monit. 8, 691–699. https://doi.org/10.1039/b601562j

Abdalmogith, S.S., Harrison, R.M., 2005. The use of trajectory cluster analysis to examine the long-range transport of secondary inorganic aerosol in the UK. Atmos. Environ. 39, 6686–6695. https://doi.org/10.1016/j.atmosenv.2005.07.059

Abdullahi, K.L., Delgado-Saborit, J.M., Harrison, R.M., 2013. Emissions and indoor concentrations of particulate matter and its specific chemical components from cooking: A review. Atmos. Environ. 71, 260–294. https://doi.org/10.1016/j.atmosenv.2013.01.061

Adam, M., Horseman, A., Turp, M., Buxmann, J., Sugier, J., 2015. Towards an operational lidar network across the UK. Geophys. Res. Abstr. 17, 11455.

Alvarado, M.J., Lonsdale, C.R., MacIntyre, H.L., Bian, H., Chin, M., Ridley, D.A., Heald, C.L., Thornhill, K.L., Anderson, B.E., Cubison, M.J., Jimenez, J.L., Kondo, Y., Sahu, L.K., Dibb, J.E., Wang, C., 2016. Evaluating model parameterizations of submicron aerosol scattering and absorption with in situ data from ARCTAS 2008. Atmos. Chem. Phys. 16, 9435–9455. https://doi.org/10.5194/acp-16-9435-2016

Álvarez-García, F.J., Fresno-Schmolk, G., OrtizBevia, M.J., Cabos, W., RuizdeElvira, A., 2020. Reduction of aggregate wind power variability using Empirical Orthogonal Teleconnections: An application in the Iberian Peninsula. Renew. Energy 159, 151–161. https://doi.org/https://doi.org/10.1016/j.renene.2020.05.153

Angelini, F., Barnaba, F., Landi, T.C., Caporaso, L., Gobbi, G.P., 2009. Study of atmospheric aerosols and mixing layer by LIDAR. Radiat. Prot. Dosimetry 137, 275–279. https://doi.org/10.1093/rpd/ncp219

Ansmann, A., Tesche, M., Groß, S., Freudenthaler, V., Seifert, P., Hiebsch, A., Schmidt, J., Wandinger, U., Mattis, I., Müller, D., Wiegner, M., 2010. The 16 April 2010 major volcanic ash plume over central Europe: EARLINET lidar and AERONET photometer observations at Leipzig and Munich, Germany. Geophys. Res. Lett. 37, 1–5. https://doi.org/10.1029/2010GL043809

Aptowicz, K.B., Pinnick, R.G., Hill, S.C., Pan, Y.L., Chang, R.K., 2006. Optical scattering patterns from single urban aerosol particles at Adelphi, Maryland, USA: A classification relating to particle morphologies. J. Geophys. Res. Atmos. 111, 1–13. https://doi.org/10.1029/2005JD006774

AQEG, 2005. Particulate Matter in the United Kingdom. DEFRA, London, UK.

Arshinov, Y.F., Bobrovnikov, S.M., Zuev, V.E., Mitev, V.M., 1983. Atmospheric temperature measurements using a pure rotational Raman lidar. Appl. Opt. 22, 2984–2990. https://doi.org/10.1364/AO.22.002984

Atkinson, R.W., Kang, S., Anderson, H.R., Mills, I.C., Walton, H.A., 2014. Epidemiological time series studies of PM $_{2.5}$ and daily mortality and hospital admissions: a systematic review and meta-analysis. Thorax 69, 660–665. https://doi.org/10.1136/thoraxjnl-2013-204492

Bachour, D., Perez-Astudillo, D., 2016. Ceilometer signal performance with height in high aerosol loads. J. Atmos. Solar-Terrestrial Phys. 138–139, 215–226. https://doi.org/10.1016/j.jastp.2015.12.005

Barlow, J.F., 2014. Progress in observing and modelling the urban boundary layer. Urban Clim. 10, 216–240. https://doi.org/10.1016/j.uclim.2014.03.011

Barnaba, F., Tafuro, A.M., De Tomasi, F., Perrone, M.R., 2007. Observed and simulated vertically resolved optical properties of continental aerosols over southeastern Italy: A closure study. J.

Geophys. Res. Atmos. 112, 1–16. https://doi.org/10.1029/2006JD007926

Bastin, G., Lorent, B., Duqué, C., Gevers, M., 1982. Optimal Estimation of the Average Areal Rainfall Over a Basin and Optimal Selection of Rain-Gauge Locations. IFAC Proc. Vol. 15, 1571–1576. https://doi.org/10.1016/S1474-6670(17)63223-3

Bayat, B., Hosseini, K., Nasseri, M., Karami, H., 2019. Challenge of rainfall network design considering spatial versus spatiotemporal variations. J. Hydrol. 574, 990–1002. https://doi.org/10.1016/j.jhydrol.2019.04.091

Beaver, S., Palazoglu, A., 2006. Cluster analysis of hourly wind measurements to reveal synoptic regimes affecting air quality. J. Appl. Meteorol. Climatol. 45, 1710–1726. https://doi.org/10.1175/JAM2437.1

Beccaceci, S., Mustoe, C., Butterfield, D., Tompkins, J., Sarantaridis, D., P., Q., Brown, R., Green, D., Fuller, G., Tremper, A., Font Font, A., Jones, A., 2013. Airborne Particulate Concentrations and Numbers in the United Kingdom (phase 3) - Annual report 2012. Teddington, U.K.

Beevers, S., Carslaw, D., Westmoreland, E., 2009. Air pollution and emissions trends in London. King's Coll. London Inst. Transp. Stud. 1–47.

Bellouin, N., Rae, J., Jones, A., Johnson, C., Haywood, J., Boucher, O., 2011. Aerosol forcing in the Climate Model Intercomparison Project (CMIP5) simulations by HadGEM2-ES and the role of ammonium nitrate. J. Geophys. Res. Atmos. 116, 1–25. https://doi.org/10.1029/2011JD016074

Bender, F.A.M., Frey, L., McCoy, D.T., Grosvenor, D.P., Mohrmann, J.K., 2019. Assessment of aerosol–cloud–radiation correlations in satellite observations, climate models and reanalysis. Clim. Dyn. 52, 4371–4392. https://doi.org/10.1007/s00382-018-4384-z

Benedetti, A., Dabas, A., 2016. Assessment of the necessary developments to assimilate Aeolus and EarthCARE aerosol profile products, Technical Note 1 of ESA contract number 4000116106.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R.J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J.W., Kinne, S., Mangold, A., Razinger, M., Simmons, A.J., Suttie, M., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J.W., Razinger, M., Schulz, M., Serrar, S., Simmons, A.J., Sofiev, M., Suttie, M., Tompkins, A.M., Untch, A., 2009. Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling. J. Geophys. Res. 114, D06206. https://doi.org/10.1029/2008JD011235

Best, M.J., 1998. A Model to Predict Surface Temperatures. Boundary-Layer Meteorol. 88, 279–306. https://doi.org/10.1023/A:1001151927113

Best, M.J., Grimmond, C.S.B., Villani, M.G., 2006. Evaluation of the urban tile in MOSES using surface energy balance observations. Boundary-Layer Meteorol. 118, 503–525. https://doi.org/10.1007/s10546-005-9025-5

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., Harding, R.J., 2011. 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

Bohnenstengel, S.I., Belcher, S.E., Aiken, A., Allan, J.D., Allen, G., Bacak, A., Bannan, T.J., Barlow, J.F., Beddows, D.C.S., Bloss, W.J., Booth, A.M., Chemel, C., Coceal, O., Di Marco, C.F., Dubey, M.K., Faloon, K.H., Fleming, Z.L., Furger, M., Gietl, J.K., Graves, R.R., Green, D.C., Grimmond, C.S.B., Halios, C.H., Hamilton, J.F., Harrison, R.M., Heal, M.R., Heard, D.E., Helfter, C., Herndon, S.C., Holmes, R.E., Hopkins, J.R., Jones, A.M., Kelly, F.J., Kotthaus, S., Langford, B., Lee, J.D., Leigh, R.J., Lewis, A.C., Lidster, R.T., Lopez-Hilfiker, F.D., McQuaid, J.B., Mohr, C., Monks, P.S., Nemitz, E., Ng, N.L., Percival, C.J., Prévôt, A.S.H., Ricketts, H.M.A., Sokhi, R., Stone, D., Thornton, J.A., Tremper, A.H., Valach, A.C., Visser, S., Whalley, L.K., Williams, L.R., Xu, L., Young, D.E., Zotter, P., Beddddows, D.C.S., Blossss, W.J.,

Booth, A.M., Chemel, C., Coceal, O., Di Marco, C.F., Dubey, M.K., Faloon, K.H., Flemiming, Z.L., Furger, M., Gietl, J.K., Graves, R.R., Green, D.C., Grimmimmimmond, C.S.B., Halios, C.H., Hamiamiamilton, J.F., Harrisison, R.M., Heal, M.R., Heard, D.E., Helfter, C., Herndon, S.C., Holmes, R.E., Hopkins, J.R., Jones, A.M., Kelly, F.J., Kotthaus, S., Langford, B., Lee, J.D., Leigh, R.J., Lewisis, A.C., Lidsidsidster, R.T., Lopez-Hilfiker, F.D., McQuaidaid, J.B., Mohr, C., Monks, P.S., Nemimitz, E., Ng, N.L., Percival, C.J., Prévôt, A.S.H., Ricketts, H.M.A., Sokhi, R., Stone, D., Thornton, J.A., Tremper, A.H., Valach, A.C., Vississsisser, S., Whalley, L.K., Williamsiamsiamsiams, L.R., Xu, L., Young, D.E., Zotter, P., 2015. Meteorology, Air Quality, and Health in London: The ClearfLo Project. Bull. Am. Meteorol. Soc. 96, 87–95. https://doi.org/10.1175/BAMS-D-12-00245.1

Bohnenstengel, S.I., Evans, S., Clark, P.A., Belcher, S.E., 2011. Simulations of the London urban heat island. Q. J. R. Meteorol. Soc. 137, 1625–1640. https://doi.org/10.1002/qj.855

Bohnenstengel, S.I., Hamilton, I., Davies, M., Belcher, S.E., 2014. Impact of anthropogenic heat emissions on London's temperatures. Q. J. R. Meteorol. Soc. 140, 687–698. https://doi.org/10.1002/qj.2144

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M.V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S.K., Sherwood, S., Stevens, B., Zhang, X.Y., Zhan, X.Y., 2013. Clouds and Aerosols. Clim. Chang. 2013 Phys. Sci. Basis. Contrib. Work. Gr. I to Fifth Assess. Rep. Intergov. Panel Clim. Chang. 571–657. https://doi.org/10.1017/CBO9781107415324.016

Boutle, I.A., Finnenkoetter, A., Lock, A.P., Wells, H., 2016. The London Model: Forecasting fog at 333 m resolution. Q. J. R. Meteorol. Soc. 142, 360–371. https://doi.org/10.1002/qj.2656

Boyouk, N., Léon, J.-F., Delbarre, H., Augustin, P., Fourmentin, M., 2011. Impact of sea breeze on vertical structure of aerosol optical properties in Dunkerque, France. Atmos. Res. 101, 902–910. https://doi.org/10.1016/j.atmosres.2011.05.016

Briggs, N.L., Long, C.M., 2016. Critical review of black carbon and elemental carbon source apportionment in Europe and the United States. Atmos. Environ. 144, 409–427. https://doi.org/10.1016/j.atmosenv.2016.09.002

Brito, J., Carbone, S., Monteiro Dos Santos, D.A., Dominutti, P., De Oliveira Alves, N., Rizzo, L., Artaxo, P., 2018. Disentangling vehicular emission impact on urban air pollution using ethanol as a tracer. Sci. Rep. 8, 1–10. https://doi.org/10.1038/s41598-018-29138-7

Burn, D.H., Goulter, I.C., 1991. An approach to the rationalization of streamflow data collection networks. J. Hydrol. 122, 71–91. https://doi.org/10.1016/0022-1694(91)90173-F

Burton, S.P., Ferrare, R.A., Hostetler, C.A., Hair, J.W., Rogers, R.R., Obland, M.D., Butler, C.F., Cook, A.L., Harper, D.B., Froyd, K.D., 2012. Aerosol classification using airborne High Spectral Resolution Lidar measurements-methodology and examples. Atmos. Meas. Tech. 5, 73–98. https://doi.org/10.5194/amt-5-73-2012

Campbell Scientific, 2013. CS135 Ceilometer.

Cattrall, C., Reagan, J., Thome, K., Dubovik, O., 2005. Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations. J. Geophys. Res. D Atmos. 110, 1–13. https://doi.org/10.1029/2004JD005124

Cavalli, F., Viana, M., Yttri, K.E., Genberg, J., Putaud, J.-P., 2010. Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the EUSAAR protocol. Atmos. Meas. Tech. 3, 79–89. https://doi.org/10.5194/amt-3-79-2010

Chan, K.L., Wiegner, M., Flentje, H., Mattis, I., Wagner, F., Gasteiger, J., Geiß, A., 2018. Evaluation of ECMWF-IFS (version 41R1) operational model forecasts of aerosol transport by using ceilometer network measurements. Geosci. Model Dev. 11, 3807–3831. https://doi.org/10.5194/gmd-11-3807-2018

Chan, T.W., Mozurkewich, M., 2007. Application of absolute principal component analysis to size distribution data: Identification of particle origins. Atmos. Chem. Phys. 7, 887–897. https://doi.org/10.5194/acp-7-887-2007

Charlton-Perez, C., O'Connor, E., Ballard, S., Adam, M., Klugmann, D., Cox, O., 2015. Review of ceilometer operator design. Met Office.

Charlton-Perez, C., Simonin, D., Ballard, S., Hopkin, E., Illingworth, A., Kotthaus, S., Westbrook, C., Grimmond, C.S.B., 2016. Suitability of Ceilometer Observations for DA, Data Assimilation and Ensembles Science Report 10. Met Office.

Chen, C.H., Chan, C.C., Chen, B.Y., Cheng, T.J., Leon Guo, Y., 2015. Effects of particulate air pollution and ozone on lung function in non-asthmatic children. Environ. Res. 137, 40–48. https://doi.org/10.1016/j.envres.2014.11.021

Chen, W., Tang, H., Zhao, H., 2015. Diurnal, weekly and monthly spatial variations of air pollutants and air quality of Beijing. Atmos. Environ. 119, 21–34. https://doi.org/10.1016/j.atmosenv.2015.08.040

Cheng, S., Wang, G., Lang, J., Wen, W., Wang, X., Yao, S., 2016. Characterization of volatile organic compounds from different cooking emissions. Atmos. Environ. 145, 299–307. https://doi.org/10.1016/j.atmosenv.2016.09.037

Chudnovsky, A., Lyapustin, A., Wang, Y., Tang, C., Schwartz, J., Koutrakis, P., 2013. High resolution aerosol data from MODIS satellite for urban air quality studies. Cent. Eur. J. Geosci. 6, 17–26. https://doi.org/10.2478/s13533-012-0145-4

Clark, P.A., Harcourt, S.A., Macpherson, B., Mathison, C.T., Cusack, S., Naylor, M., 2008. Prediction of visibility and aerosol within the operational Met Office Unified Model. I: Model formulation and variational assimilation. Q. J. R. Meteorol. Soc. 134, 1801–1816. https://doi.org/10.1002/qj.318

Clarke, A.D., Owens, S.R., Zhou, J., 2006. An ultrafine sea-salt flux from breaking waves: Implications for cloud condensation nuclei in the remote marine atmosphere. J. Geophys. Res. Atmos. 111, 1–14. https://doi.org/10.1029/2005JD006565

Claxton, B.M., 2013. Parameters controlling the aerosol hydration scheme within the UKV visibility parametrization. Met Off. Cardingt. Tech. Note No. 85 1–35.

Cleveland, W.S., 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. J. Am. Stat. Assoc. 74, 829. https://doi.org/10.2307/2286407

Clough, S.A., Iacono, M.J., Moncet, J.-L., 1992. Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor. J. Geophys. Res. 97, 15761. https://doi.org/10.1029/92JD01419

Clough, S.A., Shephard, M.W., Mlawer, E.J., Delamere, J.S., Iacono, M.J., Cady-Pereira, K., Boukabara, S., Brown, P.D., 2005. Atmospheric radiative transfer modeling: A summary of the AER codes. J. Quant. Spectrosc. Radiat. Transf. 91, 233–244. https://doi.org/10.1016/j.jqsrt.2004.05.058

Crilley, L.R., Lucarelli, F., Bloss, W.J., Harrison, R.M., Beddows, D.C., Calzolai, G., Nava, S., Valli, G., Bernardoni, V., Vecchi, R., 2017. Source apportionment of fine and coarse particles at a roadside and urban background site in London during the 2012 summer ClearfLo campaign. Environ. Pollut. 220, 766–778. https://doi.org/10.1016/j.envpol.2016.06.002

Davies, T., Cullen, M.J.P., Malcolm, A.J., Mawson, M.H., Staniforth, A., White, A.A., Wood, N., 2005. A new dynamical core for the Met Office's global and regional modelling of the atmosphere. Q. J. R. Meteorol. Soc. 131, 1759–1782. https://doi.org/10.1256/qj.04.101

Dawson, K.W., Meskhidze, N., Josset, D., Gassó, S., 2015. Spaceborne observations of the lidar ratio of marine aerosols. Atmos. Chem. Phys. 15, 3241–3255. https://doi.org/10.5194/acp-15-3241-2015

DeCarlo, P.F., Slowik, J.G., Worsnop, D.R., Davidovits, P., Jimenez, J.L., 2004. Particle Morphology and Density Characterization by Combined Mobility and Aerodynamic Diameter Measurements. Part 1: Theory. Aerosol Sci. Technol. 38, 1185–1205. https://doi.org/10.1080/027868290903907

DEFRA, 2019. Site environment types [WWW Document]. URL https://uk-air.defra.gov.uk/networks/site-types (accessed 7.10.19).

DEFRA, 2018a. Site Information for London N. Kensington (UKA00253) [WWW Document]. URL https://uk-air.defra.gov.uk/networks/site-info?site_id=KC1#startcontent (accessed 10.8.18).

DEFRA, 2018b. Particle Numbers and Concentrations Network [WWW Document]. URL https://uk-air.defra.gov.uk/networks/network-info?view=particle (accessed 10.8.18).

DEFRA, 2016. National Atmospheric Emissions Inventory [WWW Document]. URL http://naei.defra.gov.uk/about/index (accessed 11.10.16).

Dionisi, D., Barnaba, F., Diémoz, H., Di Liberto, L., Gobbi, G.P., 2018. A multi-wavelength numerical model in support to quantitative retrievals of aerosol properties from automated-lidar-ceilometers and test applications for AOT and PM10 estimation. Atmos. Meas. Tech. Discuss. 1–47. https://doi.org/10.5194/amt-2018-79

Dockery, D., Pope III, C.A., 1994. Acute Respiratory Effects of Particulate Air Pollution. Annu. Rev. Public Health 15, 107–132. https://doi.org/10.1146/annurev.publhealth.15.1.107

Dogruparmak, S.C., Keskin, G.A., Yaman, S., Alkan, A., 2014. Using principal component analysis and fuzzy c–means clustering for the assessment of air quality monitoring. Atmos. Pollut. Res. 5, 656–663. https://doi.org/10.5094/APR.2014.075

Doherty, S.J., Anderson, T.L., Charlson, R.J., 1999. Measurement of the lidar ratio for atmospheric aerosols with a 180° backscatter nephelometer. Appl. Opt. 38, 1823. https://doi.org/10.1364/AO.38.001823

Du, Z., He, K., Cheng, Y., Duan, F., Ma, Y., Liu, J., Zhang, X., Zheng, M., Weber, R., 2014. A yearlong study of water-soluble organic carbon in Beijing I: Sources and its primary vs. secondary nature. Atmos. Environ. 92, 514–521. https://doi.org/10.1016/j.atmosenv.2014.04.060

Dubovik, O., King, M.D., 2000. A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements. J. Geophys. Res. Atmos. 105, 20673–20696. https://doi.org/10.1029/2000JD900282

ECMWF, 2018. Operational configurations of the ECMWF Integrated Forecasting System (IFS) [WWW Document]. URL https://www.ecmwf.int/en/forecasts/documentation-and-support (accessed 12.18.18).

Edwards, J.M., Slingo, A., 1996. Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. Q. J. R. Meteorol. Soc. 122, 689–719. https://doi.org/10.1002/qj.49712253107

EEA, 2015. Air quality in Europe-2015 report. https://doi.org/10.2800/62459

Elliot, A.J., Smith, S., Dobney, A., Thornes, J., Smith, G.E., Vardoulakis, S., 2016. Monitoring the effect of air pollution episodes on health care consultations and ambulance call-outs in England during March/April 2014: A retrospective observational analysis. Environ. Pollut. 214, 903–911. https://doi.org/10.1016/j.envpol.2016.04.026

Emeis, S., 2011. Surface-Based Remote Sensing of the Atmospheric Boundary Layer, Atmospheric and Oceanographic Sciences Library. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-90-481-9340-0

Emeis, S., Forkel, R., Junkermann, W., Schäfer, K., Flentje, H., Gilge, S., Fricke, W., Wiegner, M., Freudenthaler, V., Gro, S., Ries, L., Meinhardt, F., Birmili, W., Münkel, C., Obleitner, F.,

Suppan, P., 2011. Measurement and simulation of the 16/17 April 2010 Eyjafjallajökull volcanic ash layer dispersion in the northern Alpine region. Atmos. Chem. Phys. 11, 2689–2701. https://doi.org/10.5194/acp-11-2689-2011

Emeis, S., Schäfer, K., Münkel, C., Friedl, R., Suppan, P., 2012. Evaluation of the Interpretation of Ceilometer Data with RASS and Radiosonde Data. Boundary-Layer Meteorol. 143, 25–35. https://doi.org/10.1007/s10546-011-9604-6

ESRI, 2019. Understanding a semivariogram: The range, sill, and nugget [WWW Document]. URL https://pro.arcgis.com/en/pro-app/help/analysis/geostatistical-analyst/understanding-a-semivariogram-the-range-sill-and-nugget.htm (accessed 7.2.19).

Esteve, A.R., Highwood, E.J., Morgan, W.T., Allen, G., Coe, H., Grainger, R.G., Brown, P., Szpek, K., 2014. A study on the sensitivities of simulated aerosol optical properties to composition and size distribution using airborne measurements. Atmos. Environ. 89, 517–524. https://doi.org/10.1016/j.atmosenv.2014.02.063

Farjami, H., Hesari, A.R.E., 2020. Assessment of sea surface wind field pattern over the Caspian Sea using EOF analysis. Reg. Stud. Mar. Sci. 35, 101254. https://doi.org/https://doi.org/10.1016/j.rsma.2020.101254

Fernald, F.G., 1984. Analysis of atmospheric lidar observations: some comments. Appl. Opt. 23, 652. https://doi.org/10.1364/AO.23.000652

Field, A., 2009. Discovering statistics using IBM SPSS statistics. sage., 5th Editio. ed. SAGE, London, UK.

Fierz-Schmidhauser, R., Zieger, P., Wehrle, G., Jefferson, A., Ogren, J.A., Baltensperger, U., Weingartner, E., 2010. Measurement of relative humidity dependent light scattering of aerosols. Atmos. Meas. Tech. 3, 39–50. https://doi.org/10.5194/amt-3-39-2010

Fitzgerald, J.W., 1975. Approximation Formulas for the Equilibrium Size of an Aerosol Particle as a Function of Its Dry Size and Composition and the Ambient Relative Humidity. J. Appl. Meteorol. 14, 1044–1049. https://doi.org/10.1175/1520-0450(1975)014<1044:AFFTES>2.0.CO;2

Fleming, Z.L., Monks, P.S., Manning, A.J., 2012. Review: Untangling the influence of air-mass history in interpreting observed atmospheric composition. Atmos. Res. 104–105, 1–39. https://doi.org/10.1016/j.atmosres.2011.09.009

Flentje, H., Claude, H., Elste, T., Gilge, S., Köhler, U., Plass-Dülmer, C., Steinbrecht, W., Thomas, W., Werner, A., Fricke, W., 2010a. The Eyjafjallajökull eruption in April 2010 - Detection of volcanic plume using in-situ measurements, ozone sondes and lidar-ceilometer profiles. Atmos. Chem. Phys. 10, 10085–10092. https://doi.org/10.5194/acp-10-10085-2010

Flentje, H., Heese, B., Reichardt, J., Thomas, W., 2010b. Aerosol profiling using the ceilometer network of the German Meteorological Service. Atmos. Meas. Tech. Discuss. 3, 3643–3673. https://doi.org/10.5194/amtd-3-3643-2010

Flynn, C., 2004. Vaisala ceilometer (model CT25K) handbook, Vaisala.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D.W., Haywood, J., Lean, J., Lowe, D.C., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schultz, M., Van Dorland, R., 2007. Changes in atmospheric constituents and in radiative forcing. Cambridge University Press, Cambridge, United Kingdom.

Francis, P.N., 2012. The direct assimilation of surface cloud reports in the UKV model. Satellite Applications Technical Memo 9, Met Office.

Gantt, B., Kelly, J.T., Bash, J.O., 2015. Updating sea spray aerosol emissions in the Community Multiscale Air Quality (CMAQ) model version 5.0.2. Geosci. Model Dev. 8, 3733–3746. https://doi.org/10.5194/gmd-8-3733-2015

Geisinger, A., Behrendt, A., Wulfmeyer, V., Strohbach, J., Förstner, J., Potthast, R., 2017. Development and application of a backscatter lidar forward operator for quantitative validation of aerosol dispersion models and future data assimilation. Atmos. Meas. Tech. 10, 4705–4726. https://doi.org/10.5194/amt-10-4705-2017

Gibert, F., Arnault, N., Cuesta, J., Plougonven, R., Flamant, P.H., 2011. Internal gravity waves convectively forced in the atmospheric residual layer during the morning transition. Q. J. R. Meteorol. Soc. 137, 1610–1624. https://doi.org/10.1002/qj.836

Goldsmith, J.E.M., Bisson, S.E., Ferrare, R.A., Evans, K.D., Whiteman, D.N., Melfi, S.H., 1994. Raman Lidar Profiling of Atmospheric Water Vapor: Simultaneous Measurements with Two Collocated Systems. Bull. Am. Meteorol. Soc. 75, 975–982. https://doi.org/10.1175/1520-0477(1994)075<0975:RLPOAW>2.0.CO;2

Gosse, S.F., Wang, M., Labrie, D., Chylek, P., 1997. Imaginary part of the refractive index of sulfates and nitrates in the 0.7-2.6-micron spectral region. Appl. Opt. 36, 3622–3634.

Goto, D., Badarinath, K.V.S., Takemura, T., Nakajima, T., 2011. Simulation of aerosol optical properties over a tropical urban site in India using a global model and its comparison with ground measurements. Ann. Geophys. 29, 955–963. https://doi.org/10.5194/angeo-29-955-2011

Grell, G., Baklanov, A., 2011. Integrated modeling for forecasting weather and air quality: A call for fully coupled approaches. Atmos. Environ. 45, 6845–6851. https://doi.org/10.1016/j.atmosenv.2011.01.017

Grell, G.A., Knoche, R., Peckham, S.E., McKeen, S.A., 2004. Online versus offline air quality modeling on cloud-resolving scales. Geophys. Res. Lett. 31, 6–9. https://doi.org/10.1029/2004GL020175

Guerrero-Rascado, J.L., Landulfo, E., Antuña, J.C., de Melo Jorge Barbosa, H., Barja, B., Bastidas, Á.E., Bedoya, A.E., da Costa, R.F., Estevan, R., Forno, R., Gouveia, D.A., Jiménez, C., Larroza, E.G., da Silva Lopes, F.J., Montilla-Rosero, E., de Arruda Moreira, G., Nakaema, W.M., Nisperuza, D., Alegria, D., Múnera, M., Otero, L., Papandrea, S., Pallota, J.V., Pawelko, E., Quel, E.J., Ristori, P., Rodrigues, P.F., Salvador, J., Sánchez, M.F., Silva, A., 2016. Latin American Lidar Network (LALINET) for aerosol research: Diagnosis on network instrumentation. J. Atmos. Solar-Terrestrial Phys. 138–139, 112–120. https://doi.org/10.1016/j.jastp.2016.01.001

Gupta, S., Gadi, R., Sharma, S.K., Mandal, T.K., 2018. Characterization and source apportionment of organic compounds in PM 10 using PCA and PMF at a traffic hotspot of Delhi. Sustain. Cities Soc. 39, 52–67. https://doi.org/10.1016/j.scs.2018.01.051

Haeffelin, M., Laffineur, Q., Bravo-Aranda, J.A., Drouin, M.A., Casquero-Vera, J.A., Dupont, J.C., De Backer, H., 2016. Radiation fog formation alerts using attenuated backscatter power from automatic lidars and ceilometers. Atmos. Meas. Tech. 9, 5347–5365. https://doi.org/10.5194/amt-9-5347-2016

Hair, J.W., Hostetler, C.A., Cook, A.L., Harper, D.B., Ferrare, R.A., Mack, T.L., Welch, W., Izquierdo, L.R., Hovis, F.E., 2008. Airborne High Spectral Resolution Lidar for profiling aerosol optical properties. Appl. Opt. 47, 6734. https://doi.org/10.1364/AO.47.006734

Hama, S.M.L., Cordell, R.L., Kos, G.P.A., Weijers, E.P., Monks, P.S., 2017a. Sub-micron particle number size distribution characteristics at two urban locations in Leicester. Atmos. Res. 194, 1–16. https://doi.org/10.1016/j.atmosres.2017.04.021

Hama, S.M.L., Cordell, R.L., Monks, P.S., 2017b. Quantifying primary and secondary source contributions to ultrafine particles in the UK urban background. Atmos. Environ. 166, 62–78. https://doi.org/10.1016/j.atmosenv.2017.07.013

Harrison, R.M., Dall'Osto, M., Beddows, D.C.S., Thorpe, A.J., Bloss, W.J., Allan, J.D., Coe, H., Dorsey, J.R., Gallagher, M., Martin, C., Whitehead, J., Williams, P.I., Jones, R.L., Langridge,

J.M., Benton, A.K., Ball, S.M., Langford, B., Hewitt, C.N., Davison, B., Martin, D., Petersson, K.F., Henshaw, S.J., White, I.R., Shallcross, D.E., Barlow, J.F., Dunbar, T., Davies, F., Nemitz, E., Phillips, G.J., Helfter, C., Di Marco, C.F., Smith, S., 2012. Atmospheric chemistry and physics in the atmosphere of a developed megacity (London): An overview of the REPARTEE experiment and its conclusions. Atmos. Chem. Phys. 12, 3065–3114. https://doi.org/10.5194/acp-12-3065-2012

Haywood, J., Boucher, O., 2000. Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review. Rev. Geophys. 38, 513–543. https://doi.org/10.1029/1999RG000078

Haywood, J., Bush, M., Abel, S., Claxton, B., Coe, H., Crosier, J., Harrison, M., Macpherson, B., Naylor, M., Osborne, S., 2008. Prediction of visibility and aerosol within the operational Met Office Unified Model. II: Validation of model performance using observational data. Q. J. R. Meteorol. Soc. 134, 1817–1832. https://doi.org/10.1002/qj.275

Haywood, J, Francis, P., Dubovik, O., Glew, M., Holben, B., 2003. Comparison of aerosol size distributions, radiative properties, and optical depths determined by aircraft observations and Sun photometers during SAFARI 2000. J. Geophys. Res. 108. https://doi.org/Artn 8471 Doi 10.1029/2002jd002250

Haywood, Jim, Osborne, S.R., Francis, P.N., Keil, A., Formenti, P., Andreae, M.O., Kaye, P.H., 2003. The mean physical and optical properties of regional haze dominated by biomass burning aerosol measured from the C-130 aircraft during SAFARI 2000. J. Geophys. Res. 108, 8473. https://doi.org/10.1029/2002jd002226

He, J., Chen, S., Zhang, Y., Guo, P., Chen, H., 2019. Accurate inversion of tropospheric bottom temperature using pure rotational Raman lidar in polluted air condition. Opt. Commun. 452, 88–94. https://doi.org/10.1016/j.optcom.2019.07.030

Heese, B., Flentje, H., Althausen, D., Ansmann, A., Frey, S., 2010. Ceilometer lidar comparison: Backscatter coefficient retrieval and signal-to-noise ratio determination. Atmos. Meas. Tech. 3, 1763–1770. https://doi.org/10.5194/amt-3-1763-2010

Henry, R.C., Wang, Y., Gebhart, K.A., 1991. The relationship between empirical orthogonal functions and sources of air pollution. Atmos. Environ. Part A. Gen. Top. 25, 503–509. https://doi.org/10.1016/0960-1686(91)90322-X

Hertwig, D., Grimmond, S., Hendry, M., Saunders, B., Wang, Z., Jeoffrion, M., Vidale, P.L., McGuire, P.C., Bohnenstengel, S.I., Ward, H.C., Kotthaus, S., n.d. Urban signals in high-resolution weather and climate simulations: The role of urban land-surface models. In preperation.

Hinds, W.C., 1999. Aerosol Technology: Properties, Behaviour and Measurement of Airborne Particles. Wiley, New York, USA.

Hofman, J., Staelens, J., Cordell, R., Stroobants, C., Zikova, N., Hama, S.M.L., Wyche, K.P., Kos, G.P.A., Van Der Zee, S., Smallbone, K.L., Weijers, E.P., Monks, P.S., Roekens, E., 2016. Ultrafine particles in four European urban environments: Results from a new continuous long-term monitoring network. Atmos. Environ. 136, 68–81. https://doi.org/10.1016/j.atmosenv.2016.04.010

Hood, C., Stocker, J., Johnson, K., Carruthers, D., MacKenzie, I., Doherty, R., Vieno, M., 2018. Air quality simulations for London using a coupled regional-to-local modelling system. Atmos. Chem. Phys. 18, 11221–11245. https://doi.org/10.5194/acp-18-11221-2018

Hopkin, E., Illingworth, A., Charlton-Perez, C., Westbrook, C., Ballard, S., 2017. Calibration of the Met Office Ceilometer Network using the Cloud Method. EMS Annu. Meet. Abstr. 14.

Hopkin, E., Illingworth, A.J., Charlton-Perez, C., Westbrook, C.D., Ballard, S., 2019. A robust automated technique for operational calibration of ceilometers using integrated backscatter from totally attenuating liquid cloud. Atmos. Meas. Tech. 12, 4131–4147.

https://doi.org/10.5194/amt-2018-427

Illingworth, A.J., Cimini, D., Gaffard, C., Haeffelin, M., Lehmann, V., Löhnert, U., O'Connor, E.J., Ruffieux, D., 2015. Exploiting existing ground-based remote sensing networks to improve high-resolution weather forecasts. Bull. Am. Meteorol. Soc. 96, 2107–2125. https://doi.org/10.1175/BAMS-D-13-00283.1

Illingworth, A.J., Hogan, R.J., O'Connor, E.J., Bouniol, D., Brooks, M.E., Delanoë, J., Donovan, D.P., Eastment, J.D., Gaussiat, N., Goddard, J.W.F., Haeffelin, M., Klein Baltinik, H., Krasnov, O.A., Pelon, J., Piriou, J.M., Protat, A., Russchenberg, H.W.J., Seifert, A., Tompkins, A.M., van Zadelhoff, G.J., Vinit, F., Willen, U., Wilson, D.R., Wrench, C.L., 2007. Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations. Bull. Am. Meteorol. Soc. 88, 883–898. https://doi.org/10.1175/BAMS-88-6-883

Jacobson, M.Z., 2005. Fundamentals of Atmospheric Modeling, 2nd Ed. ed. Cambridge University Press, Cambridge, UK.

Janisková, M., Mahfouf, J.-F., Morcrette, J.-J., 2002. Preliminary studies on the variational assimilation of cloud–radiation observations. Q. J. R. Meteorol. Soc. 128, 2713–2736. https://doi.org/10.1256/qj.01.192

Janisková, M., Stiller, O., 2010. Development of strategies for radar and lidar data assimilation. Reading, UK.

Janisková, M., Stiller, O., Michele, S. Di, Forbes, R., Morcrette, J., Ahlgrimm, M., Bauer, P., 2010. QuARL – Quantitative Assessment of the Operational Value of Space-Borne Radar and Lidar Measurements of Cloud and Aerosol Profiles. ESA Contract Report.

Ji, D., Li, L., Pang, B., Xue, P., Wang, L., Wu, Y., Zhang, H., Wang, Y., 2017. Characterization of black carbon in an urban-rural fringe area of Beijing. Environ. Pollut. 223, 524–534. https://doi.org/10.1016/j.envpol.2017.01.055

Jin, Y., Kai, K., Kawai, K., Nagai, T., Sakai, T., Yamazaki, A., Uchiyama, A., Batdorj, D., Sugimoto, N., Nishizawa, T., 2015. Ceilometer calibration for retrieval of aerosol optical properties. J. Quant. Spectrosc. Radiat. Transf. 153, 49–56. https://doi.org/10.1016/j.jqsrt.2014.10.009

Jin, Y., Sugimoto, N., Shimizu, A., Nishizawa, T., Kai, K., Kawai, K., Yamazaki, A., Sakurai, M., Wille, H., 2018. Evaluation of ceilometer attenuated backscattering coefficients for aerosol profile measurement. J. Appl. Remote Sens. 12, 1. https://doi.org/10.1117/1.jrs.12.042604

Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. Philosphical Trans. R. Soc. A 374. https://doi.org/10.1098/rsta.2015.0202

Jose, S., Niranjan, K., Gharai, B., Rao, P.V.N., Nair, V.S., 2016. Characterisation of Absorbing Aerosols Using Ground and Satellite Data at an Urban Location, Hyderabad. Aerosol Air Qual. Res. 16, 1427–1440. https://doi.org/10.4209/aaqr.2014.09.0220

Kaiser, H.F., 1958. The varimax criterion for analytic rotation in factor analysis. Psychometrika 23, 187–200. https://doi.org/10.1007/BF02289233

Kaufman, L., Rousseeuw, P., 2005. Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New Jersey.

King's College London, 2016. Air quality networks – beyond compliance assessment [WWW Document]. URL http://www.kcl.ac.uk/lsm/research/divisions/aes/research/ERG/measurement/air-quality-networks/index.aspx (accessed 8.25.16).

King, R., 2015. Comparison of UKV with MORUSES and JULES in Urban Areas, Met Office. Met Office.

Klett, J.D., 1981. Stable analytical inversion solution for processing lidar returns. Appl. Opt. 20, 211–220. https://doi.org/10.1364/AO.20.000211

Klompmaker, J.O., Montagne, D.R., Meliefste, K., Hoek, G., Brunekreef, B., 2015. Spatial variation of ultrafine particles and black carbon in two cities: Results from a short-term measurement campaign. Sci. Total Environ. 508, 266–275. https://doi.org/10.1016/j.scitotenv.2014.11.088

Knipping, E.M., Dabdub, D., 2003. Impact of chlorine emissions from sea-salt aerosol on coastal urban ozone. Environ. Sci. Technol. 37, 275–284. https://doi.org/10.1021/es025793z

Koenig, J.Q., 2000. Health Effects of Particulate Matter, in: Health Effects of Ambient Air Pollution. Springer US, Boston, MA, pp. 115–137. https://doi.org/10.1007/978-1-4615-4569-9_10

Kotchenruther, R.A., Hobbs, P. V, 1998. Humidification factors of aerosols from biomass burning in Brazil. J. Geophys. Res. Atmos. 103, 32081–32089. https://doi.org/10.1029/98JD00340

Kotthaus, S., Grimmond, C.S.B., 2018a. Atmospheric boundary-layer characteristics from ceilometer measurements. Part 1: A new method to track mixed layer height and classify clouds. Q. J. R. Meteorol. Soc. 144, 1525–1538. https://doi.org/10.1002/qj.3299

Kotthaus, S., Grimmond, C.S.B., 2018b. Atmospheric boundary-layer characteristics from ceilometer measurements. Part 2: Application to London's urban boundary layer. Q. J. R. Meteorol. Soc. 144, 1511–1524. https://doi.org/10.1002/qj.3298

Kotthaus, S., Grimmond, C.S.B., 2014. Energy exchange in a dense urban environment – Part I: Temporal variability of long-term observations in central London. Urban Clim. 10, 261–280. https://doi.org/10.1016/j.uclim.2013.10.002

Kotthaus, S., Halios, C.H., Barlow, J.F., Grimmond, C.S.B., 2018. Volume for pollution dispersion: London's atmospheric boundary layer during ClearfLo observed with two ground-based lidar types. Atmos. Environ. 190, 401–414. https://doi.org/10.1016/j.atmosenv.2018.06.042

Kotthaus, S., O'Connor, E., Münkel, C., Charlton-Perez, C., Haeffelin, M., Gabey, A.M., Grimmond, C.S.B., 2016. Recommendations for processing atmospheric attenuated backscatter profiles from Vaisala CL31 ceilometers. Atmos. Meas. Tech. 9, 3769–3791. https://doi.org/10.5194/amt-9-3769-2016

Kucbel, M., Corsaro, A., Švédová, B., Raclavská, H., Raclavský, K., Juchelková, D., 2017. Temporal and seasonal variations of black carbon in a highly polluted European city: Apportionment of potential sources and the effect of meteorological conditions. J. Environ. Manage. 203, 1178–1189. https://doi.org/10.1016/j.jenvman.2017.05.038

Kuniyal, J.C., Guleria, R.P., 2019. The current state of aerosol-radiation interactions: A mini review. J. Aerosol Sci. 130, 45–54. https://doi.org/10.1016/j.jaerosci.2018.12.010

Lapworth, A., Osborne, S.R., 2019. Gravity-Wave Drag in the Stable Boundary Layer Over Moderate Terrain. Boundary-Layer Meteorol. 171, 175–189. https://doi.org/10.1007/s10546-018-00422-3

Lapworth, A., Osborne, S.R., 2016. Evidence for gravity wave drag in the boundary layer of a numerical forecast model: a comparison with observations. Q. J. R. Meteorol. Soc. 142, 3257–3264. https://doi.org/10.1002/qj.2909

Lean, H.W., Barlow, J.F., Halios, C.H., 2019. The impact of spin-up and resolution on the representation of a clear convective boundary layer over London in order 100 m grid-length versions of the Met Office Unified Model. Q. J. R. Meteorol. Soc. https://doi.org/10.1002/qj.3519

Lean, H.W., Clark, P.A., 2003. The effects of changing resolution on mesocale modelling of line convection and slantwise circulations in FASTEX IOP16. Q. J. R. Meteorol. Soc. 129, 2255–2278. https://doi.org/10.1256/qj.02.57

Lemonsu, A., Bastin, S., Masson, V., Drobinski, P., 2006. Vertical structure of the urban boundary layer over Marseille under sea-breeze conditions. Boundary-Layer Meteorol. 118, 477–501. https://doi.org/10.1007/s10546-005-7772-y

Lewis, E.R., 2008. An examination of Köhler theory resulting in an accurate expression for the equilibrium radius ratio of a hygroscopic aerosol particle valid up to and including relative humidity 100%. J. Geophys. Res. Atmos. 113, 1–17. https://doi.org/10.1029/2007JD008590

Li, J., Liu, X., Yuan, L., Yin, Y., Li, Z., Li, P., Ren, G., Jin, L., Li, R., Dong, Z., Li, Y., Yang, J., 2015. Vertical distribution of aerosol optical properties based on aircraft measurements over the Loess Plateau in China. J. Environ. Sci. 34, 44–56. https://doi.org/10.1016/j.jes.2015.01.021

Li, Jian, Li, C., Zhao, Y., Li, Jing, Chu, Y., 2016. Geometrical constraint experimental determination of Raman lidar overlap profile. Appl. Opt. 55, 4924–4928. https://doi.org/http://dx.doi.org/10.1364/AO.55.004924 1.

Li, Q., Chen, P., Sun, L., Ma, X., 2018. A global weighted mean temperature model based on empirical orthogonal function analysis. Adv. Sp. Res. 61, 1398–1411. https://doi.org/https://doi.org/10.1016/j.asr.2017.12.031

Lin, C.-L., Xia, Q., Calhoun, R., 2008. Retrieval of Urban Boundary Layer Structures from Doppler Lidar Data. Part II: Proper Orthogonal Decomposition. J. Atmos. Sci. 65, 21–42. https://doi.org/10.1175/2007JAS2329.1

Liu, D., Allan, J.D., Young, D.E., Coe, H., Beddows, D., Fleming, Z.L., Flynn, M.J., Gallagher, M.W., Harrison, R.M., Lee, J., Prevot, A.S.H., Taylor, J.W., Yin, J., Williams, P.I., Zotter, P., 2014. Size distribution, mixing state and source apportionment of black carbon aerosol in London during winter time. Atmos. Chem. Phys. 14, 10061–10084. https://doi.org/10.5194/acp-14-10061-2014

Liu, D., Taylor, J.W., Crosier, J., Marsden, N., Bower, K.N., Lloyd, G., Ryder, C.L., Brooke, J., Cotton, R., Marenco, F., Blyth, A., Cui, Z., Estelles, V., Gallagher, M., Coe, H., Choularton, T.W., 2017. Aircraft and ground measurements of dust aerosols over the West Africa coast in summer 2015 during ICE-D and AER-D. Atmos. Chem. Phys. Discuss. 1–32. https://doi.org/10.5194/acp-2017-984

Liu, Y., Daum, P.H., 2008. Relationship of refractive index to mass density and self-consistency of mixing rules for multicomponent mixtures like ambient aerosols. J. Aerosol Sci. 39, 974–986. https://doi.org/10.1016/j.jaerosci.2008.06.006

Lopez, P., 2002. Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes. Q. J. R. Meteorol. Soc. 128, 229–257. https://doi.org/10.1256/00359000260498879

Lowenthal, D., Zielinska, B., Mason, B., Samy, S., Samburova, V., Collins, D., Spencer, C., Taylor, N., Allen, J., Kumar, N., 2009. Aerosol characterization studies at Great Smoky Mountains National Park, summer 2006. J. Geophys. Res. Atmos. 114, 1–11. https://doi.org/10.1029/2008JD011274

Lu, C.H., Da Silva, A., Wang, J., Moorthi, S., Chin, M., Colarco, P., Tang, Y., Bhattacharjee, P.S., Chen, S.P., Chuang, H.Y., Juang, H.M.H., McQueen, J., Iredell, M., 2016. The implementation of NEMS GFS Aerosol Component (NGAC) Version 1.0 for global dust forecasting at NOAA/NCEP. Geosci. Model Dev. 9, 1905–1919. https://doi.org/10.5194/gmd-9-1905-2016

Lufft, 2016. Cloud Height Sensor CHM 15k User Manual. Fellbach, Germany.

Madonna, F., Amato, F., Vande Hey, J., Pappalardo, G., 2014. Ceilometer aerosol profiling vs. Raman lidar in the frame of INTERACT campaign of ACTRIS. Atmos. Meas. Tech. Discuss. 7, 12407–12447. https://doi.org/10.5194/amtd-7-12407-2014

Madonna, F., Rosoldi, M., Lolli, S., Amato, F., Vande Hey, J., Dhillon, R., Zheng, Y., Brettle, M., Pappalardo, G., 2018. Intercomparison of aerosol measurements performed with multi-wavelength Raman lidars, automatic lidars and ceilometers in the framework of INTERACT-II campaign. Atmos. Meas. Tech. 11, 2459–2475. https://doi.org/10.5194/amt-11-2459-2018

Manners, J., Edwards, J.M., Hill, P., Thelen, J., 2015. SOCRATES Technical Guide Suite Of Community Radiative Transfer codes based on Edwards and Slingo. Exeter, UK.

Marenco, F., Johnson, B., Langridge, J.M., Mulcahy, J., Benedetti, A., Remy, S., Jones, L., Szpek, K., Haywood, J., Longo, K., Artaxo, P., 2016. On the vertical distribution of smoke in the Amazonian atmosphere during the dry season. Atmos. Chem. Phys. 16, 2155–2174. https://doi.org/10.5194/acp-16-2155-2016

Marley, N.A., Gaffney, J.S., 2006. Introduction to Urban Aerosols and Their Impacts. ChemInform 37, 1–22. https://doi.org/10.1002/chin.200650274

Martin, G.M., Ringer, M.A., Pope, V.D., Jones, A., Dearden, C., Hinton, T.J., 2006. The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model (HadGEM1). Part I: Model description and global climatology. J. Clim. 19, 1274–1301. https://doi.org/10.1175/jcli3636.1

Martucci, G., Milroy, C., O'Dowd, C.D., 2010. Detection of Cloud-Base Height Using Jenoptik CHM15K and Vaisala CL31 Ceilometers. J. Atmos. Ocean. Technol. 27, 305–318. https://doi.org/10.1175/2009JTECHA1326.1

Maryon, R.H., Smith, F.B., Conway, B.J., Goddard, D.M., 1991. The U.K. nuclear accident model. Prog. Nucl. Energy 26, 85–104. https://doi.org/10.1016/0149-1970(91)90043-O

Matheron, G., 1963. Principles of geostatistics. Econ. Geol. 58, 1246–1266. https://doi.org/10.2113/gsecongeo.58.8.1246

Mauderly, J.L., Chow, J.C., 2008. Health Effects of Organic Aerosols. Inhal. Toxicol. 20, 257–288. https://doi.org/10.1080/08958370701866008

McKendry, I.G., van der Kamp, D., Strawbridge, K.B., Christen, A., Crawford, B., 2009. Simultaneous observations of boundary-layer aerosol layers with CL31 ceilometer and 1064/532 nm lidar. Atmos. Environ. 43, 5847–5852. https://doi.org/10.1016/j.atmosenv.2009.07.063

Melfi, S.H., Whiteman, D., 1985. Observation of Lower-Atmospheric Moisture Structure and its Evolution Using a Raman Lidar. Bull. Am. Meteorol. Soc. 66, 1288–1292. https://doi.org/10.1175/1520-0477(1985)066<1288:OOLAMS>2.0.CO;2

Met Office, 2019. Numerical weather prediction models [WWW Document]. URL https://www.metoffice.gov.uk/research/approach/modelling-systems/unified-model/weather-forecasting (accessed 9.6.19).

Met Office, 2017. LIDARNET- LCBR Network Backscatter visualisation [WWW Document]. URL https://www.metoffice.gov.uk/public/lidarnet/lcbr-network.html (accessed 10.25.17).

Mittal, L., Baker, T., Fuller, G., 2016. London Air Quality Network Summary Report 2014. Kings College London, London, UK.

Miyazaki, Y., Kondo, Y., Takegawa, N., Komazaki, Y., Fukuda, M., Kawamura, K., Mochida, M., Okuzawa, K., Weber, R.J., 2006. Time-resolved measurements of water-soluble organic carbon in Tokyo. J. Geophys. Res. Atmos. 111. https://doi.org/10.1029/2006JD007125

Monahan, A.H., Fyfe, J.C., Ambaum, M.H.P., Stephenson, D.B., North, G.R., 2009. Empirical Orthogonal Functions: The Medium is the Message. J. Clim. 22, 6501–6514. https://doi.org/10.1175/2009JCLI3062.1

Morcrette, J.J., Beljaars, A., Benedetti, A., Jones, L., Boucher, O., 2008. Sea-salt and dust aerosols in the ECMWF IFS model. Geophys. Res. Lett. 35, 1–5. https://doi.org/10.1029/2008GL036041

Muller, C.L., Chapman, L., Grimmond, C.S.B., Young, D.T., Cai, X.-M., 2013. Toward a Standardized Metadata Protocol for Urban Meteorological Networks. Bull. Am. Meteorol. Soc. 94, 1161–1185. https://doi.org/10.1175/BAMS-D-12-00096.1

Müller, D., Ansmann, A., Mattis, I., Tesche, M., Wandinger, U., Althausen, D., Pisani, G., 2007. Aerosol-type-dependent lidar ratios observed with Raman lidar. J. Geophys. Res. Atmos. 112, 1–11. https://doi.org/10.1029/2006JD008292

Münkel, C., Emeis, S., Müller, W.J., Schaefer, K.P., 2004. Aerosol concentration measurements with a lidar ceilometer: results of a one year measuring campaign. Proc. SPIE--the Int. Soc. Opt. Eng. 5235, 486–496. https://doi.org/10.1117/12.511104

Münkel, C., Eresmaa, N., Räsänen, J., Karppinen, A., 2007. Retrieval of mixing height and dust concentration with lidar ceilometer. Boundary-Layer Meteorol. 124, 117–128. https://doi.org/10.1007/s10546-006-9103-3

Münkel, C., Roininen, R., 2010. Automatic monitoring of boundary layer structures with ceilometer, in: TECO-2010. pp. 2–8.

Myhre, G., Lund Myhre, C., Samset, B.H., Storelvmo, T., 2013. Aerosols and their Relation to Global Climate and Climate Sensitivity, Nature Education Knowledge.

Neal, L., 2019. Air Quality Emissions Data. Met Office, Exeter, UK.

Neal, R.A., Phillips, I.D., 2009. Summer daily precipitation variability over the East Anglian region of Great Britain. Int. J. Climatol. 29, 1661–1679. https://doi.org/10.1002/joc.1826

Nishizawa, T., Sugimoto, N., Matsui, I., Shimizu, A., Higurashi, A., Jin, Y., 2016. The Asian Dust and Aerosol Lidar Observation Network (AD-NET): Strategy and Progress. EPJ Web Conf. 119, 19001. https://doi.org/10.1051/epjconf/201611919001

Noh, Y.M., 2014. Single-scattering albedo profiling of mixed asian dust plumes with multiwavelength raman lidar. Atmos. Environ. 95, 305–317. https://doi.org/10.1016/j.atmosenv.2014.06.028

O'Connor, E.J., Illingworth, A.J., Hogan, R.J., 2004. A Technique for Autocalibration of Cloud Lidar. Am. Meteorol. Soc. 21, 777–786. https://doi.org/10.1175/1520-0426(2004)021<0777:ATFAOC>2.0.CO;2

Oke, T.R., 1978. Boundary Layer Climates, 2nd ed. Routledge, Cambridge, UK.

Oke, T.R., Mills, G., Christen, A., Voogt, J.A., 2017. Urban Climates. Cambridge University Press. https://doi.org/10.1017/9781139016476

Oliver, M.A., Webster, R., 2014. Catena A tutorial guide to geostatistics : Computing and modelling variograms and kriging. Catena 113, 56–69. https://doi.org/10.1016/j.catena.2013.09.006

Osborne, M., Adam, M., Buxmann, J., Sugier, J., Marenco, F., Haywood, J., 2018. Saharan dust and biomass burning aerosols during ex-hurricane Ophelia : validation of the new UK lidar and sun-photometer network.

Palacios-Penã, L., Jiménez-Guerrero, P., Baró, R., Balzarini, A., Bianconi, R., Curci, G., Christian Landi, T., Pirovano, G., Prank, M., Riccio, A., Tuccella, P., Galmarini, S., 2019. Aerosol optical properties over Europe: An evaluation of the AQMEII Phase 3 simulations against satellite observations. Atmos. Chem. Phys. 19, 2965–2990. https://doi.org/10.5194/acp-19-2965-2019

Pandis, S.N., Skyllakou, K., Florou, K., Kostenidou, E., Kaltsonoudis, C., Hasa, E., Presto, A.A., 2016. Urban particulate matter pollution: a tale of five cities. Faraday Discuss. 189, 277–290. https://doi.org/10.1039/C5FD00212E

Papayannis, A., Balis, D., Amiridis, V., Chourdakis, G., Tsaknakis, G., Zerefos, C., Castanho, a. D. a., Nickovic, S., Kazadzis, S., Grabowski, J., 2005. Measurements of Saharan dust aerosols over the Eastern Mediterranean using elastic backscatter-Raman lidar, spectrophotometric and satellite observations in the frame of the EARLINET project. Atmos. Chem. Phys. Discuss. 5, 2075–2110. https://doi.org/10.5194/acpd-5-2075-2005

Pappalardo, G., Amodeo, A., Apituley, A., Comeron, A., Freudenthaler, V., Linné, H., Ansmann, A., Bösenberg, J., D'Amico, G., Mattis, I., Mona, L., Wandinger, U., Amiridis, V., Alados-

Arboledas, L., Nicolae, D., Wiegner, M., 2014. EARLINET towards an advanced sustainable European aerosol lidar network. Atmos. Meas. Tech. 7, 2389–2409. https://doi.org/10.5194/amt-7-2389-2014

Pappalardo, G. et al., 2013. Four-dimensional distribution of the 2010 Eyjafjallajökull volcanic cloud over Europe observed by EARLINET. Atmos. Chem. Phys. 13, 4429–4450. https://doi.org/10.5194/acp-13-4429-2013

Peña, A., Gryning, S.E., Hahmann, A.N., 2013. Observations of the atmospheric boundary layer height under marine upstream flow conditions at a coastal site. J. Geophys. Res. Atmos. 118, 1924–1940. https://doi.org/10.1002/jgrd.50175

Poltera, Y., Martucci, G., Collaud Coen, M., Hervo, M., Emmenegger, L., Henne, S., Brunner, D., Haefele, A., 2017. PathfinderTURB: an automatic boundary layer algorithm. Development, validation and application to study the impact on in-situ measurements at the Jungfraujoch. Atmos. Chem. Phys. 17, 10051–10070. https://doi.org/https://doi.org/10.5194/acp-17-10051-2017

Porson, A., Clark, P.A., Harman, I.N., Best, M.J., Belcher, S.E., 2010. Implementation of a new urban energy budget scheme in the MetUM. Part I: Description and idealized simulations. Q. J. R. Meteorol. Soc. 136, 1514–1529. https://doi.org/10.1002/qj.668

Povey, A.C., Grainger, R.G., Peters, D.M., Agnew, J.L., 2014. Retrieval of aerosol backscatter, extinction, and lidar ratio from Raman lidar with optimal estimation. Atmos. Meas. Tech. 7, 757–776. https://doi.org/10.5194/amt-7-757-2014

Qiao, Z., Wu, C., Huang, N., Xu, X., Sun, Z., Sun, X., 2018. Spatio-temporal structure of the urban thermal environment in Beijing based on an empirical orthogonal function. J. Spat. Sci. 63, 297–310. https://doi.org/10.1080/14498596.2018.1488632

Redington, A.L., Witham, C.S., Hort, M.C., 2016. Source apportionment of speciated PM10 in the United Kingdom in 2008: Episodes and annual averages. Atmos. Environ. 145, 251–263. https://doi.org/10.1016/j.atmosenv.2016.09.047

Reid, J.S., Kuehn, R.E., Holz, R.E., Eloranta, E.W., Kaku, K.C., Kuang, S., Newchurch, M.J., Thompson, A.M., Trepte, C.R., Zhang, J., Atwood, S.A., Hand, J.L., Holben, B.N., Minnis, P., Posselt, D.J., 2017. Ground-based High Spectral Resolution Lidar observation of aerosol vertical distribution in the summertime Southeast United States. J. Geophys. Res. 122, 2970–3004. https://doi.org/10.1002/2016JD025798

Richman, M.B., 1986. Rotation of Principle Components. J. Climatol. 6, 293–335.

Rodríguez, S., Van Dingenen, R., Putaud, J.-P., Dell'Acqua, A., Pey, J., Querol, X., Alastuey, A., Chenery, S., Ho, K.-F., Harrison, R., Tardivo, R., Scarnato, B., Gemelli, V., 2007. A study on the relationship between mass concentrations, chemistry and number size distribution of urban fine aerosols in Milan, Barcelona and London. Atmos. Chem. Phys. 7, 2217–2232. https://doi.org/10.5194/acp-7-2217-2007

Roessler, D.M., Faxvog, F.R., 1981. Visibility in absorbing aerosols. Atmos. Environ. 15, 151–155. https://doi.org/10.1016/0004-6981(81)90006-8

Rogula-Kozłowska, W., Majewski, G., Czechowski, P.O., 2015. The size distribution and origin of elements bound to ambient particles: a case study of a Polish urban area. Environ. Monit. Assess. 187. https://doi.org/10.1007/s10661-015-4450-5

Ruths, M., von Bismarck-Osten, C., Weber, S., 2014. Measuring and modelling the local-scale spatio-temporal variation of urban particle number size distributions and black carbon. Atmos. Environ. 96, 37–49. https://doi.org/10.1016/j.atmosenv.2014.07.020

Ryder, C.L., Marenco, F., Brooke, J.K., Estelles, V., Cotton, R., Formenti, P., McQuaid, J.B., Price, H.C., Liu, D., Ausset, P., Rosenberg, P.D., Taylor, J.W., Choularton, T., Bower, K., Coe, H., Gallagher, M., Crosier, J., Lloyd, G., Highwood, E.J., Murray, B.J., 2018. Coarse-mode mineral

dust size distributions, composition and optical properties from AER-D aircraft measurements over the tropical eastern Atlantic. Atmos. Chem. Phys. 18, 17225–17257. https://doi.org/10.5194/acp-18-17225-2018

Sabbagh-Kupelwieser, N., Horvath, H., Szymanski, W.W., 2010. Urban aerosol studies of PM1 size fraction with reference to ambient conditions and visibility. Aerosol Air Qual. Res. 10, 425–432. https://doi.org/10.4209/aaqr.2010.02.0009

Savage, N.H., Agnew, P., Davis, L.S., Ordonez, C., Thorpe, R., Johnson, C.E., O'Connor, F.M., Dalvi, M., 2013. Air quality modelling using the Met Office Unified Model (AQUM OS24-26): model description and initial evaluation. Geosci. Model Dev. 6, 353–372. https://doi.org/10.5194/gmd-6-353-2013

Schafer, K., Emeis, S., Schrader, S., Torok, S., Alfoldy, B., Osan, J., Pitz, M., Munkel, C., Cyrys, J., Peters, a, Sarigiannis, D., Suppan, P., 2011. A measurement based analysis of the spatial distribution, temporal variation and chemical composition of particulate matter in Munich and Augsburg. Meteorol. Zeitschrift 20, 47–57. https://doi.org/10.1127/0941-2948/2011/0498

Schäfer, K., Wagner, P., Emeis, S., Jahn, C., Münkel, C., Suppan, P., Ecology, L., Gmbh, V., 2012. Mixing layer height and air pollution levels in urban area. Proc. SPIE 8534, 1–10. https://doi.org/10.1117/12.974328

Schkolnik, G., Chand, D., Hoffer, A., Andreae, M.O., Erlick, C., Swietlicki, E., Rudich, Y., 2007. Constraining the density and complex refractive index of elemental and organic carbon in biomass burning aerosol using optical and chemical measurements. Atmos. Environ. 41, 1107–1118. https://doi.org/10.1016/j.atmosenv.2006.09.035

Schmidt, O.T., Mengaldo, G., Balsamo, G., Wedi, N.P., 2019. Spectral Empirical Orthogonal Function Analysis of Weather and Climate Data. Mon. Weather Rev. 147, 2979–2995. https://doi.org/10.1175/MWR-D-18-0337.1

Schoenfelder, J., Brown, M., Pettibone, A., Stanier, C., 2009. Evaluation of the Vaisala CL31 ceilometer as a tool for boundary layer characterization within carbon cycle studies.

Seinfeld, J.H., Pandis, S.N., 2016. Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. Wiley, Haboken.

Sharma, S., Vaishnav, R., Shukla, M. V., Kumar, Prashant, Kumar, Prateek, Thapliyal, P.K., Lal, S., Acharya, Y.B., 2016. Evaluation of cloud base height measurements from Ceilometer CL31 and MODIS satellite over Ahmedabad, India. Atmos. Meas. Tech. 9, 711–719. https://doi.org/10.5194/amt-9-711-2016

Shiraiwa, M., Ueda, K., Pozzer, A., Lammel, G., Kampf, C.J., Fushimi, A., Enami, S., Arangio, A.M., Fröhlich-Nowoisky, J., Fujitani, Y., Furuyama, A., Lakey, P.S.J., Lelieveld, J., Lucas, K., Morino, Y., Pöschl, U., Takahama, S., Takami, A., Tong, H., Weber, B., Yoshino, A., Sato, K., 2017. Aerosol Health Effects from Molecular to Global Scales. Environ. Sci. Technol. 51, 13545–13567. https://doi.org/10.1021/acs.est.7b04417

Smith, M.J., Phillips, I.D., 2013. Winter daily precipitation variability over the East Anglian region of Great Britain and its relationship with river flow. Int. J. Climatol. 33, 2215–2231. https://doi.org/10.1002/joc.3594

Smith, M.L., Bertram, A.K., Martin, S.T., 2012. Deliquescence, efflorescence, and phase miscibility of mixed particles of ammonium sulfate and isoprene-derived secondary organic material. Atmos. Chem. Phys. 12, 9613–9628. https://doi.org/10.5194/acp-12-9613-2012

Song, X., Zhai, X., Liu, L., Wu, S., 2017. Lidar and Ceilometer Observations and Comparisons of Atmospheric Cloud Structure at Nagqu of Tibetan Plateau in 2014 Summer. Atmosphere (Basel). 8, 9. https://doi.org/10.3390/atmos8010009

Song, Y., Zhang, B., Shi, G., Li, S., Di, H., Yan, Q., Hua, D., 2018. Correlation between the lidar ratio and the Ångström exponent of various aerosol types. Particuology 40, 62–69.

https://doi.org/10.1016/j.partic.2017.12.002

St-Hilaire, A., Ouarda, T.B.M.J., Lachance, M., Bobée, B., Gaudet, J., Gignac, C., 2003. Assessment of the impact of meteorological network density on the estimation of basin precipitation and runoff: A case study. Hydrol. Process. 17, 3561–3580. https://doi.org/10.1002/hyp.1350

Stachlewska, I.S., Samson, M., Zawadzka, O., Harenda, K.M., Janicka, L., Poczta, P., Szczepanik, D., Heese, B., Wang, D., Borek, K., Tetoni, E., Proestakis, E., Siomos, N., Nemuc, A., Chojnicki, B.H., Markowicz, K.M., Pietruczuk, A., Szkop, A., Althausen, D., Stebel, K., Schuettemeyer, D., Zehner, C., 2018. Modification of local urban aerosol properties by long-range transport of biomass burning aerosol. Remote Sens. 10, 1–28. https://doi.org/10.3390/rs10030412

Stier, P., Seinfeld, J.H., Kinne, S., Boucher, O., 2007. Aerosol absorption and radiative forcing. Atmos. Chem. Phys. 7, 5237–5261. https://doi.org/10.5194/acp-7-5237-2007

Strang, G., 1988. Linear algebra and its applications, 3rd Editio. ed. Harcourt Brace Jovanovich, San Diego.

Stull, R.B., 1988. An Introduction to Boundary Layer Meteorology. Kluwer Academic Publishers, Dordlecht, Netherlands.

Sugimoto, N., Matsui, I., Shimizu, A., Nishizawa, T., Hara, Y., Xie, C., Uno, I., Yumimoto, K., Wang, Z., Yoon, S.-C., 2008. Lidar network observations of tropospheric aerosols, in: Singh, U.N., Asai, K., Jayaraman, A. (Eds.), Lidar Remote Sensing for Environmental Monitoring IX. Noumea, New Caledonia. https://doi.org/10.1117/12.806540

Sullivan, A.P., Weber, R.J., Clements, A.L., Turner, J.R., Bae, M.S., Schauer, J.J., 2004. A method for on-line measurement of water-soluble organic carbon in ambient aerosol particles: Results from an urban site. Geophys. Res. Lett. 31, 14–17. https://doi.org/10.1029/2004GL019681

Tang, I.N., Munkelwitz, H.R., 1994. Water activities, densities, and refractive indicies of aqueous sulfates and sodium nitrate droplets of atmospheric importance. J. Geophys. Res. 99, 18801–18808.

Tang, M., Huang, X., Lu, K., Ge, M., Li, Y., Cheng, P., Zhu, T., Ding, A., Zhang, Y., Gligorovski, S., Song, W., Ding, X., Bi, X., Wang, X., 2017. Heterogeneous reactions of mineral dust aerosol: Implications for tropospheric oxidation capacity. Atmos. Chem. Phys. 17, 11727–11777. https://doi.org/10.5194/acp-17-11727-2017

Tang, Y., Lean, H.W., Bornemann, J., 2013. The benefits of the Met Office variable resolution NWP model for forecasting convection. Meteorol. Appl. 20, 417–426. https://doi.org/10.1002/met.1300

Taylor, N.F., Collins, D.R., Lowenthal, D.H., McCubbin, I.B., Gannet Hallar, A., Samburova, V., Zielinska, B., Kumar, N., Mazzoleni, L.R., 2017. Hygroscopic growth of water soluble organic carbon isolated from atmospheric aerosol collected at US national parks and Storm Peak Laboratory. Atmos. Chem. Phys. 17, 2555–2571. https://doi.org/10.5194/acp-17-2555-2017

Thermo Fisher Scientific, 2015. Partisol 2025i Sequential Air Sampler/ Partisol 2025i-D Dichotomous Sequential Air Sampler Instruction Manual [WWW Document]. URL https://assets.thermofisher.com/TFS-Assets/LSG/manuals/EPM-manual-Partisol-2025i-2025iD.pdf (accessed 10.11.18).

Thornhill, G.D., Ryder, C.L., Highwood, E.J., Shaffrey, L.C., Johnson, B.T., 2018. The effect of South American biomass burning aerosol emissions on the regional climate. Atmos. Chem. Phys. 18, 5321–5342. https://doi.org/10.5194/acp-18-5321-2018

Toon, O.B., Pollack, J.B., Khare, B.N., 1976. The optical constants of several atmospheric aerosol species: Ammonium sulfate, aluminum oxide, and sodium chloride. J. Geophys. Res. 81, 5733–5748. https://doi.org/10.1029/JC081i033p05733

TOPROF, 2020. TOPROF (COST Action ES1303) [WWW Document]. URL

http://www.toprof.imaa.cnr.it/ (accessed 7.31.20).

TSI, 2009. Series 3080: Electrostatic Classifiers [WWW Document]. URL http://cires1.colorado.edu/jimenez-group/Manuals/SMPS_3080_manual.pdf (accessed 10.10.18).

TSI, 2007. Model 3775 Condensation Particle Counter: Operation and Service Manual [WWW Document]. URL http://dustmonitors.ru/d/68562/d/cpc-3775r.pdf (accessed 10.10.18).

TSI, 2004. Model 3321 Aerodynamic Particle Sizer® Spectrometer: Instruction Manual [WWW Document]. URL http://www.wmo-gaw-wcc-aerosol-physics.org/files/aps_3321.pdf (accessed 10.10.18).

United Nations, 2018. World Urbanisation Prospects: The 2018 Revision, Key Facts. New York, USA.

URG, 2011. Ambient Ion Monitor: Operation Manual URG-9000B [WWW Document]. URL http://www.urgcorp.com/testupdate/library/manuals/URG-9000B Operations Manual.pdf (accessed 10.11.18).

Uzan, L., Egert, S., Alpert, P., 2016. Ceilometer evaluation of the eastern Mediterranean summer boundary layer height-first study of two Israeli sites. Atmos. Meas. Tech. 9, 4387–4398. https://doi.org/10.5194/amt-9-4387-2016

Vaisala, 2017. CL51 Ceilometer for High-Range Cloud Height Detection Technical data [WWW Document]. URL https://www.vaisala.com/sites/default/files/documents/CL51-Datasheet-B210861EN-D.pdf (accessed 4.26.18).

Vaisala, 2006. Ceilometer CL31 Measurement Performance - Test Report.

Van Tricht, K., Gorodetskaya, I. V., Lhermitte, S., Turner, D.D., Schween, J.H., Van Lipzig, N.P.M., 2014. An improved algorithm for polar cloud-base detection by ceilometer over the ice sheets. Atmos. Meas. Tech. 7, 1153–1167. https://doi.org/10.5194/amt-7-1153-2014

Varutbangkul, V., Brechtel, F.J., Bahreini, R., Ng, N.L., Keywood, M.D., Kroll, J.H., Flagan, R.C., Seinfeld, J.H., Lee, a., Goldstein, a. H., 2006. Hygroscopicity of secondary organic aerosols formed by oxidation of cycloalkenes, monoterpenes, sesquiterpenes, and related compounds. Atmos. Chem. Phys. Discuss. 6, 1121–1177. https://doi.org/10.5194/acpd-6-1121-2006

Vicente-Serrano, S., Cuadrat-Prats, J., Saz-Sánchez, M., 2005. Spatial patterns of the urban heat island in Zaragoza (Spain). Clim. Res. 30, 61–69. https://doi.org/10.3354/cr030061

Wagner, P., Schäfer, K., 2015. Influence of mixing layer height on air pollutant concentrations in an urban street canyon. Urban Clim. 1–13. https://doi.org/10.1016/j.uclim.2015.11.001

Wallace, M.J., Hobbs, V.P., 2006. Atmospheric Science: An Introductory Survey, 2nd ed. Academic Press, London, UK.

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Luigi Vidale, P., Wilkinson, J., Willett, M., Woolnough, S., Xavier, P., 2017. The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. Geosci. Model Dev. 10, 1487–1520. https://doi.org/10.5194/gmd-10-1487-2017

Wandinger, U., Ansmann, A., 2002. Experimental Determination of the Lidar Overlap Profile with Raman Lidar. Appl. Opt. 41, 511–514. https://doi.org/10.1364/AO.41.000511

Wang, W., Gong, W., Mao, F., Pan, Z., Liu, B., 2016. Measurement and study of lidar ratio by using a Raman lidar in central China. Int. J. Environ. Res. Public Health 13. https://doi.org/10.3390/ijerph13050508

Wang, Z., Cao, X., Zhang, L., Notholt, J., Zhou, B., Liu, R., Zhang, B., 2012. Lidar measurement of planetary boundary layer height and comparison with microwave profiling radiometer observation. Atmos. Meas. Tech. 5, 1965–1972. https://doi.org/10.5194/amt-5-1965-2012

Ward, H., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58, 236–244.

Warren, E.L., 2019. ImproveNetworks [WWW Document]. GitHub Repos. URL https://github.com/elliottwarren/improveNetworks (accessed 7.29.19).

Weast, R.C. (Ed.), 1977. CRC Handbook of Chemistry and Physics, 58th Edition, 58th Ed. ed. CRC Press.

Welton, E.J., Campbell, J.R., Berkoff, T.A., Valencia, S., Spinhirne, J.D., Holben, B., Tsay, S., Goddard, N., Flight, S., 2000. The NASA micro-pulse lidar network (MPLNET): Co-location of lidars with aeronet sunphometers and related earth science applications.

WHO, 2006. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: global update 2005: summary of risk assessment. Geneva World Heal. Organ. 1–22. https://doi.org/10.1016/0004-6981(88)90109-6

Wiegner, M., Gasteiger, J., 2015. Correction of water vapor absorption for aerosol remote sensing with ceilometers. Atmos. Meas. Tech. 8, 3971–3984. https://doi.org/10.5194/amt-8-3971-2015

Wiegner, M., Madonna, F., Binietoglou, I., Forkel, R., Gasteiger, J., Geiß, A., Pappalardo, G., Schäfer, K., Thomas, W., 2014. What is the benefit of ceilometers for aerosol remote sensing? An answer from EARLINET. Atmos. Meas. Tech. 7, 1979–1997. https://doi.org/10.5194/amt-7-1979-2014

Wiegner, M., Mattis, I., Pattantyús-Ábrahám, M., Antonio Bravo-Aranda, J., Poltera, Y., Haefele, A., Hervo, M., Görsdorf, U., Leinweber, R., Gasteiger, J., Haeffelin, M., Wagner, F., Cermak, J., Komínková, K., Brettle, M., Münkel, C., Pönitz, K., 2019. Aerosol backscatter profiles from ceilometers: Validation of water vapor correction in the framework of CeiLinEx2015. Atmos. Meas. Tech. 12, 471–490. https://doi.org/10.5194/amt-12-471-2019

Wilks, D.S., 2011. Statistical Methods in the Atmospheric Sciences, 3rd ed. Academic Press, Oxford; Walktham, MA.

Willén, U., Crewell, S., Baltink, H.K., Sievers, O., 2005. Assessing model predicted vertical cloud structure and cloud overlap with radar and lidar ceilometer observations for the Baltex Bridge Campaign of CLIWA-NET. Atmos. Res. 75, 227–255. https://doi.org/10.1016/j.atmosres.2004.12.008

Wilson, K.D., 1996. Empirical Orthogonal Function Analysis of the Weakly Convective Atmospheric Boundary Layer. Part I: Eddy Structures. J. Atmos. Sci. 53, 801–823. https://doi.org/10.1175/1520-0469(1996)053<0801:EOFAOT>2.0.CO;2

Wright, B., 1997. A New Visibility Analysis / Forecast System for Nimrod, Forecasting Research Technical Report No. 222.

Xing, Y.F., Xu, Y.H., Shi, M.H., Lian, Y.X., 2016. The impact of PM2.5 on the human respiratory system. J. Thorac. Dis. 8, E69–E74. https://doi.org/10.3978/j.issn.2072-1439.2016.01.19

Yan, D., Wu, S., Zhou, S., Tong, G., Li, F., Wang, Y., Li, B., 2019. Characteristics, sources and health risk assessment of airborne particulate PAHs in Chinese cities: A review. Environ. Pollut. 248, 804–814. https://doi.org/10.1016/j.envpol.2019.02.068

Ye, X., Tang, C., Yin, Z., Chen, J., Ma, Z., Kong, L., Yang, X., Gao, W., Geng, F., 2013. Hygroscopic growth of urban aerosol particles during the 2009 Mirage-Shanghai Campaign. Atmos. Environ. 64, 263–269. https://doi.org/10.1016/j.atmosenv.2012.09.064

Yin, H., Pizzol, M., Xu, L., 2017. External costs of PM2.5 pollution in Beijing, China: Uncertainty analysis of multiple health impacts and costs. Environ. Pollut. 226, 356–369.

https://doi.org/10.1016/j.envpol.2017.02.029

You, W., Zang, Z., Zhang, L., Zhang, M., Pan, X., Li, Y., 2016. A nonlinear model for estimating ground-level PM10 concentration in Xi'an using MODIS aerosol optical depth retrieval. Atmos. Res. 168, 169–179. https://doi.org/10.1016/j.atmosres.2015.09.008

Young, D.E., Allan, J.D., Williams, P.I., Green, D.C., Flynn, M.J., Harrison, R.M., Yin, J., Gallagher, M.W., Coe, H., 2015. Investigating the annual behaviour of submicron secondary inorganic and organic aerosols in London. Atmos. Chem. Phys. 15, 6351–6366. https://doi.org/10.5194/acp-15-6351-2015

Yu, H.-L., Lin, Y.-C., 2015. Analysis of space–time non-stationary patterns of rainfall–groundwater interactions by integrating empirical orthogonal function and cross wavelet transform methods. J. Hydrol. 525, 585–597. https://doi.org/https://doi.org/10.1016/j.jhydrol.2015.03.057

Zappoli, S., Andracchio, A., Fuzzi, S., Facchini, M.C., Gelencsér, A., Kiss, G., Krivácsy, Z., Molnár, A., Mészáros, E., Hansson, H.C., Rosman, K., Zebühr, Y., 1999. Inorganic, organic and macromolecular components of fine aerosol in different areas of Europe in relation to their water solubility. Atmos. Environ. 33, 2733–2743. https://doi.org/10.1016/S1352-2310(98)00362-8

Zelenyuk, A., Cai, Y., Imre, D., 2006. From Agglomerates of Spheres to Irregularly Shaped Particles: Determination of Dynamic Shape Factors from Measurements of Mobility and Vacuum Aerodynamic Diameters. Aerosol Sci. Technol. 40, 197–217. https://doi.org/10.1080/02786820500529406

Zéphoris, M., Holin, H., Lavie, F., Cenac, N., Cluzeau, M., Delas, O., Eideliman, F., Gagneux, J., Gander, A., Thibord, C., 2005. Ceilometer observations of aerosol layer structure above the Petit Lubéron during ESCOMPTE's IOP 2. Atmos. Res. 74, 581–595. https://doi.org/10.1016/j.atmosres.2004.06.014

Zhang, Y., Zhang, Q., Cheng, Y., Su, H., Kecorius, S., Wang, Z., Wu, Z., Hu, M., Zhu, T., Wiedensohler, A., He, K., 2016. Measuring the morphology and density of internally mixed black carbon with SP2 and VTDMA: New insight into the absorption enhancement of black carbon in the atmosphere. Atmos. Meas. Tech. 9, 1833–1843. https://doi.org/10.5194/amt-9-1833-2016

Zhang, Z., Moore, J.C., 2015. Chapter 6 - Empirical Orthogonal Functions, in: Zhang, Z., Moore, J.C.B.T.-M. and P.F. of C.C. (Eds.), . Elsevier, Boston, pp. 161–197. https://doi.org/https://doi.org/10.1016/B978-0-12-800066-3.00006-1

Zhao, D., Xin, J., Gong, C., Wang, X., Ma, Yongjing, Ma, Yining, 2018. Trends of Aerosol Optical Properties over the Heavy Industrial Zone of Northeastern Asia in the Past Decade (2004–15). J. Atmos. Sci. 75, 1741–1754. https://doi.org/10.1175/JAS-D-17-0260.1

Zhao, H., Mao, J., Zhou, C., Gong, X., 2018. A method of determining multi-wavelength lidar ratios combining aerodynamic particle sizer spectrometer and sun-photometer. J. Quant. Spectrosc. Radiat. Transf. 217, 224–228. https://doi.org/10.1016/j.jqsrt.2018.05.030