Atmospheric Research 104-105 (2012) 1–39

Contents lists available at SciVerse ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmos

Review: Untangling the influence of air-mass history in interpreting observed atmospheric composition

Zoë L. Fleming [a,*], Paul S. Monks [b], Alistair J. Manning [c]

[a] National Centre for Atmospheric Science (NCAS), Department of Chemistry, University of Leicester, Leicester, LE1 7RH, UK
[b] Department of Chemistry, University of Leicester, Leicester, LE1 7RH, UK
[c] Met Office, Exeter, EX1 3PB, UK

ARTICLE INFO

Article history:
Received 7 April 2011
Received in revised form 17 September 2011
Accepted 18 September 2011

Keywords:
Air mass
Trajectory
Dispersion model
Composition

ABSTRACT

Is wind direction an adequate marker of air mass history? This review looks at the evolution of methods for assessing the effect of the origin and pathway of air masses on composition change and trends. The composition of air masses and how they evolve and the changing contribution of sources and receptors are key elements in atmospheric science. Source–receptor relationships of atmospheric composition can be investigated with back trajectory techniques, tracing forward from a defined geographical origin to arrive at measurement sites where the composition may have altered during transport.

The distinction between the use of wind sector analysis, trajectory models and dispersion models to interpret composition measurements is explained and the advantages and disadvantages of each are illustrated with examples. Historical uses of wind roses, back trajectories and dispersion models are explained as well as the methods for grouping and clustering air masses. The interface of these methods to the corresponding chemistry measured at the receptor sites is explored. The review does not detail the meteorological derivation of trajectories or the complexity of the models but focus on their application and the statistical analyses used to compare them with in situ composition measurements. A newly developed methodology for analysing atmospheric observatory composition data according to air mass pathways calculated with the NAME dispersion model is given as a detailed case study. The steps in this methodology are explained with relevance to the Weybourne Atmospheric Observatory in the UK.

Contents

* Corresponding author at: National Centre for Atmospheric Science (NCAS), Department of Chemistry, University of Leicester, Leicester, UK. Tel.: +44 116 223 1289.
E-mail address: zf5@le.ac.uk (Z.L. Fleming).

0169-8095
doi:10.1016/j.atmosres.2011.09.009

1. Introduction

At a rudimentary level, clean, polluted, background and industrially influenced are all terms used to describe the composition of an air mass. These are clearly subjective in this context but in some way reflect an assessment of the air mass composition and its history. There has been a historical push in particular to classify air masses that are representative of direct anthropogenic influences nearby. In particular, this has been driven by a need to quantify the impact of the perturbation of continuing anthropogenic emissions on the overall state of the atmosphere.

Primary pollutants can originate from anthropogenic source regions (cities, industry, roads etc.) or stochastic events (biomass burning, volcanoes). Source–receptor relationships investigate composition over a receptor region produced by emission changes within a source region (Fiore et al., 2009; HTAP, 2007). Source–receptor relationships of atmospheric composition have been extensively investigated to look at the relationships between emission, transport and in situ measurements and reveal the influences that pollutant emissions have not only on the local area but also on regions far from the source. Trends in meteorological variability can explain composition trends at some sites as well as the influence of transport pathways. In this review, methods for following source–receptor relationships on regional and global scales by using in situ measured composition data are explored.

The chemical and physical composition of an air mass is inherently related to its path through the atmosphere and in order to get the maximum information out of long term time series of composition measurements, data are often divided according to air mass history. Atmospheric composition measurements have been interpreted using wind speed and direction measurements as a marker of air mass history for many years but current science requires better attribution than that available from using wind direction.

Atmospheric flow can be viewed in two possible fields of reference: Eulerian or Lagrangian. Eulerian modelling uses a fixed reference system (latitude/longitude and elevation) whilst Lagrangian models use a reference system, which follows the average atmospheric motion. Trajectory models use a set of meteorological fields from within the domain of influence and dispersion models are one step on from trajectory models, in that the complexity of turbulence is included. A further complexity to transport-only or tracer models comes with the addition of emissions and chemistry of the atmospheric species making these Chemistry Transport Models (CTMs).

Using back trajectories it becomes possible to examine source–receptor relationships and the timescales of long-range and local transport and its effect on the observed composition. In short-range transport, the airflow pathway is more influenced by emission source areas than in long-range transport, where various exchange and mixing processes (e.g. deposition and advection), physical losses and chemistry have more influence on the composition at the receptor location. The frequency and type of short-term pollution events can be tracked back to their source and seasonal and long term trends can be studied and compared to seasonal and long term air mass transport patterns. For example, Moody et al. (1989) carried out an analysis of atmospheric transport recognition techniques and found that as much as 30% of chemical variability in the troposphere can be related to transport. Therefore, segregating chemical time series into periods receiving different air masses is important. Early work by Draxler and Taylor (1982) explained long-range transport of pollutants by running trajectory models from a series of vertical layers as fractions of the boundary layer and investigated the effect of wind shear on pollutants represented by instantaneous puffs of particles and showed how wind shear must be incorporated into models to explain the dispersion.

Long term measurement stations at specific locations around the globe provide a vast amount of data on the chemical composition specific to their location. There are hundreds of permanent long term stations that form part of various national (e.g. AURN in UK, PAES in France, EPA's network in the US) or international (e.g. WMO's GAW (Global Atmospheric Watch), and the European AirBase and EMEP) networks as

discussed in Laj et al. (2009). Measurement stations are generally denoted as urban, rural or marine as they are strategically positioned to sample representative segments of air masses, with the closest known pollution source (i.e. roads, industry, cities) or biogenic and natural emissions (e.g. oceans, forests and peat bogs). Background or "baseline" stations are located in geographical locations where they sample and detect any long term trends in the background atmosphere and determine the extent and effect of long range transport bringing pollution to the site. Much of the analysis of ground based composition data is accompanied by studies of what is known as "climatological pathways" with the aim of mapping the probability of hypothetical air masses reaching the station and identifying the emission sources and influences on the site.

Seasonal variations in composition at each station reflect changing meteorological patterns, temperatures, emission patterns, chemistry and physical loss processes within a year. Investigating long term trends (inter-annual or seasonal between years) is now possible, with many stations having continuous measurements for over 30 years. Intra-seasonal and short term changes in composition are very specific to the location of the station and can often be explained by tracing back the air mass history during these events. Short-term measurement campaigns from aircrafts provide a vital picture of the three dimensional composition of the atmosphere and have been described as mobile in situ measurement platforms. Tropospheric–stratospheric exchange processes and advection, boundary layer height, convection, temperature and humidity and their relation to tropospheric composition have been investigated in these experiments. The airborne campaigns are often planned with the idea of investigating specific pollution plumes and use prediction tools (often run using trajectory models) to enable flights to traverse particular air masses and follow composition changes within a plume's evolution (Blake et al., 1993). Aircraft campaigns provide a spatial scale and a representativeness to an area as opposed to a fixed point for ground based platforms but by their very nature they are temporally limited. To some extent this can be expanded by long term networks of aircraft measurements e.g. MOZAIC (Cammas et al., 2009), CARIBIC (Brenninkmeijer et al., 2007) and the HIPPO (HIAPER Pole-to-Pole Observations) Carbon Cycle and Greenhouse Gases Study (http://hippo.ucar.edu/).

Methods to attribute composition changes according to trajectory or dispersion models that do not include emissions are not subject to the errors and assumptions that would propagate from the emission inventories and can result in biased source attribution results. Chemistry Transport Models have many more assumptions in them, so even though both the inventories and the chemistry schemes within the models have been vastly improved, there is still an important role for using trajectories or dispersion models to interpret measurements.

This review overviews previous studies in which atmospheric composition measurements have been analysed with respect to air mass history. It explains the techniques used and looks at the classification of the various methods. A step by step explanation of a new technique that classifies the regional influences of a site is described (for the Weybourne Atmospheric Observatory in Norfolk, UK) as a detailed exemplar of the review topic.

2. Outline of methodologies used to assess air mass history and its influence on observed composition

2.1. In situ wind direction as a proxy for air mass history

Meteorological parameters have been used to derive statistically distinct meteorological regimes as a primary technique. *In situ* wind direction measurements at a given point have been extensively used to trace the direction of air arriving at a given site but this clearly does not take into account the synoptic scale of the flow field. Wind roses have been used to show the distribution of wind influences at a particular station and divide the composition data into the corresponding wind sectors. Trace gas data at Mace Head in Ireland from measurements in 1996 and 1997 were separated into five wind directions in order to isolate "clean" (Atlantic) and polluted (European) air masses as shown in Fig. 1a (Salisbury et al., 2002). The wind rose method often tracks local wind influences (the last 2 or 3 h before reaching the station), but in the longer term it can often be misleading. For example, local coastal sea-breeze effects can be different to the general circulation and the synoptic scale windfield.

Radar wind profilers for surface winds and radiosonde data for vertical profiles were used in the MILAGRO campaign in Mexico City (de Foy et al., 2008) to give an extensive picture of the meteorological periods of the campaign and cluster analysis of the wind data was used to assign hourly air mass clusters for the whole campaign (Fig. 1b) which were linked to composition measurement time series. During a study at La Réunion island, Bhugwant et al. (2001) used a sectorised wind analysis (shown in Fig. 1c) to confirm regional contamination of combustion by-products (and higher black carbon levels) during particular seasons of the year in the marine boundary layer on the island and confirmed this by comparing with the 5 day back trajectories for these periods. The influence of the meteorological component on the observed ozone and NO_2 trends was studied at an urban site in the Athens basin (Varotsos et al., 2003). Seasonal wind-roses, derived from both trajectory and meteorological data showed the air-transport effect on the air pollution of the Athens basin and cross correlations between surface ozone and the frequency of the air transport during different seasons were calculated. Burley and Bytnerowicz (2011) used monthly wind roses to study the ozone levels in the Californian White Mountains and then compared this with back trajectory directions to isolate the origin air during high or low ozone periods.

Many wind rose studies measure wind direction in either 16 direction sectors (22.5° each) (Fig. 2) or 36 sectors (10° each) (Fig. 1c). Droppo and Napier (2008) describe an algorithm to interconvert and standardise analysis using various meteorological datasets with different directional sectors. Daily averaged wind roses of SO_2 and Particulate Matter (PM) at an industrial harbour computed using a Power-Ridge Pollutant (PRP) rose computational scheme (using ordinary least squares regression, outlier handling and weighted averages) (Fig. 1d) were compared to half-hourly wind roses of these pollutants and showed that using daily averaged PRP values was as effective as using half-hourly wind roses (Cosemans et al., 2008).

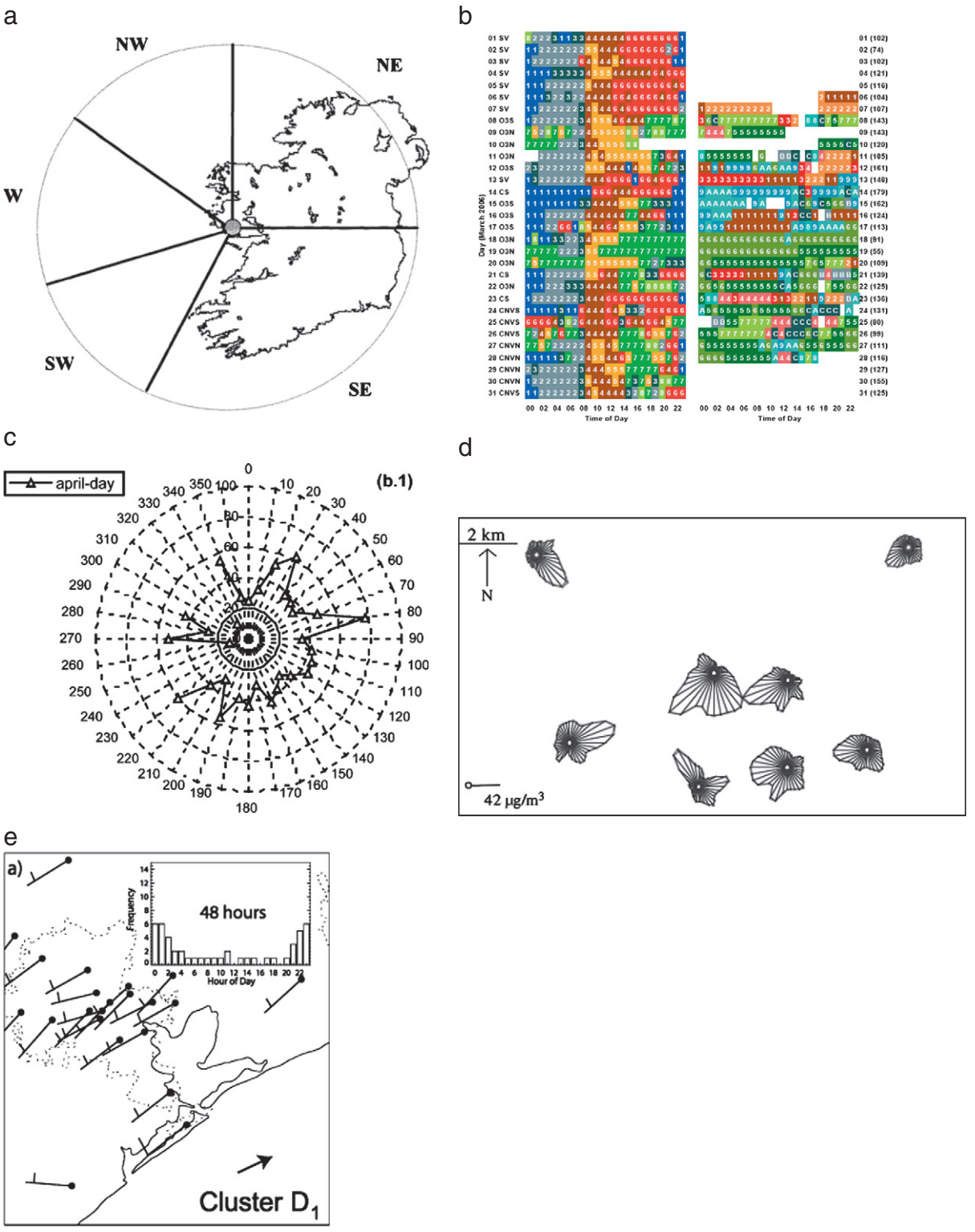

Fig. 1. Visualisation of wind data and separation into wind sectors: a) Mace Head wind sector divisions used for a range of trace gas measurements in 1996 and 1997. b) Meteorological episodes derived from surface wind measurements arranged as a timeseries meteorological flags during the MILAGRO campaign in Mexico City. c) Wind roses of Black Carbon levels from various directions for a month at La Reunion Island. d) High-resolution pollutant roses (known as Power-Ridge Pollutant (PRP) roses) for SO_2 arriving at the Antwerp harbour area. e) The use of wind barbs representing the wind direction and speed used to show one of the 16 meteorological regime clusters for studying ozone exceedances in Houston, Texas.
a): Taken from Salisbury et al. (2002). b): Taken from de Foy et al. (2008). c): Taken from Bhugwant et al. (2001). d): Taken from Cosemans et al. (2008). e): Taken from Darby (2005)© American Meteorological Society. Reprinted with permission.

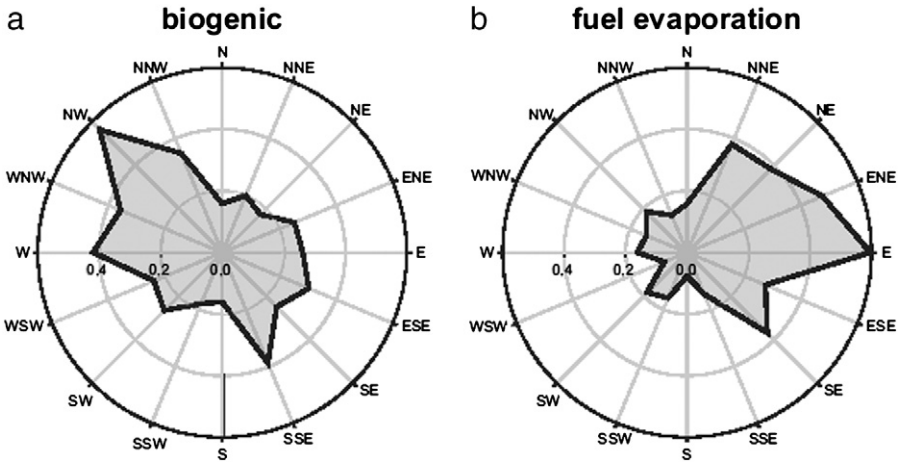

Fig. 2. Wind roses of the 75th percentile PM_{10} levels from biogenic and fuel evaporation sources (out of a total of 8 different sources separated by positive Matrix Factorisation (PMF) analysis) at Houston, Texas.
Taken from Leuchner and Rappengluck (2010).

Many statistical techniques have been used with wind speed and direction, humidity, temperature, pressure and cloud cover parameters to classify distinct meteorological regimes and periods. Crutcher et al. (1986) pioneered the use of cluster analysis to better elucidate the dependence of air quality on meteorology. Other studies include cluster analysis of ozone in St Louis, Missouri (Altshuller, 1986), Birmingham, Alabama (Eder et al., 1994), Houston, Texas (Davis et al., 1998) and ozone exceedances at Houston, Texas (Darby, 2005). The Darby (2005) example is shown in Fig. 1e, where wind barbs are used to represent wind speed and direction to show the wind direction in one air mass type separated by cluster analysis.

Neural networks have been used to remove the meteorological variability from datasets to discern temporal and spatial trends in response to changing precursor emissions. Turias et al. (2006) introduced a neural network approach to classify surface winds that could be used to improve air pollution forecasts. Long term ozone measurements and a suite of meteorological measurements have been analysed with neural networks (Gardner and Dorling, 2001; Gardner and Dorling, 1999, 2000) as well as NOx and PM (Kukkonen et al., 2003). The use of artificial neural networks in interpreting air quality data is discussed in a comprehensive review by Gardner and Dorling (1998).

Beaver and Palazoglu (2006) have used Principal Component Analysis (PCA) to cluster ozone measurements in San Francisco with wind measurements. PCA and Positive Matrix Factorization (PMF) (see Section 4.1.4 for similar technique with back trajectories) were used to interpret the elemental composition and sources of aerosols arriving at Dunkirk, France (Alleman et al., 2010), associating each sample with distinctive emission sources. Hart et al. (2006) have analysed ozone exceedance events over 10 years in Sydney, Australia by using a suite of meteorological measurements combined with PCA and cluster analysis to classify days into low and high ozone days. Oanh et al. (2005) have looked at SO_2 levels in the Mae Moh valley in the northern Thailand with PCA clustering of synoptic meteorological conditions.

The Conditional Probability Function (CPF) that represents the probability of an air mass arriving at a receptor site has been used on wind data in many studies (see Section 4.1.2 for its applications in trajectory residence time analysis). Wind direction and speed were used for sites measuring Black Carbon (BC) in New York state (Venkatachari et al., 2006) in order to identify likely locations of local point sources of BC, with CPF plots showing from which directions around the sites the highest 25% BC levels occur. Extensive wind rose analyses deriving CPF of the concentration of the species of interest from various directions around the measurement station have been carried out: e.g. to attribute PM levels to wind sectors at three sites in Ontario, Canada (Chan and Mozurkewich, 2007), to study sources of PM, O_3, NO, NO_2, CO and SO_2 arriving at Erfurt, Germany (Yue et al., 2008), to study VOC levels in Beijing (Song et al., 2008), to study $PM_{2.5}$, SO_2, CO, and O_3 in Rochester, NY, USA (Kasumba et al., 2009), to analyse PM_{10} levels arriving in Daejeon City, China (Lim et al., 2010) and to analyse VOC levels in Houston during the TexAQS-II campaign (Leuchner and Rappengluck, 2010) (as shown in Fig. 2 with wind roses for two of the source types derived from PMF analysis).

The general weakness of using pollution roses is that one cannot assume that the wind direction measured at a point is consistent with the synoptic scale flow. The turbulent and synoptic nature of wind always leads to changes in the wind direction over a region and this is not shown from local or point wind direction measurements.

2.1.1. Meteorological data as synoptic weather patterns to classify composition

The classical large scale weather classification widely used in Europe is the Grosswetterlagen system, originally conceived by Baur et al. (1944), recently updated by Gerstengabe et al. (1999) and since maintained by the German Weather Service (DWD). The 29 Hess and Brezowsky Grosswetterlagen (HB-GWL) regimes can be viewed as readily identifiable large-scale circulation patterns involving the whole of Europe and the North-East Atlantic, with their primary focus on central Europe. Spichtinger et al. (1996) divided 3 years of ozone, NO and meteorological measurements in Munich into 3 GWL. Linear multiple regression analysis was

Table 1
Overview of studies dealing with using back trajectories or dispersion models to segregate composition measurements.

Reference	Location	Species measured	Trajectory or dispersion model	Cluster analysis methods	Residence time analysis	Chemistry modelling
Aalto et al. (2002)	Pallas, Finland	CO_2, O_3, SO_2, aerosols	TRADOS trajectory		CF	
Abdalmogith and Harrison (2005)	Belfast and Harwell, UK	PM_{10}	HYSPLIT trajectory	k-means		
Amodio et al. (2008)	Bari, Italy	PM	Trajectory			
Apadula et al. (2003)	Plateau Rosa, Monte Cimone and Zugspitze	CO_2	TRAIET trajectory		CF	
Ashbaugh et al. (1985)	Grand Canyon NP, USA	PM	Trajectory		CF, PSCF	
Baker (2010)	Birmingham, Harwell, UK	O_3, NO_x, PM, CO, SO_2, benzene	HYSPLIT trajectory	k-means		
Begum et al. (2005)	Philadelphia, US	PM	HYSPLIT trajectory		PSCF	
Biegalski and Hopke (2004)	Burnt Island, Canada	As, In, Sb, Se, Sn, Zn	AES trajectory		PSCF	
Borge et al. (2007)	Athens, Madrid and Birmingham	PM	HYSPLIT trajectory	k-means (2-stage)		
Brankov et al. (1998)	3 stations in north east USA	O_3, PM	HYSPLIT trajectory	k-means		
Burley and Ray (2007)	Yosemite NP, USA	O_3	HYSPLIT trajectory		PSCF	
Burley and Bytnerowicz (2011)	White mountains, California, US	O_3	HYSPLIT trajectory			
Cape et al. (2000)	Mace Head, Ireland	O_3	UGAMP trajectory	Hierarchical		
Cardenas et al. (1998)	Weybourne, UK	CO, O_3, VOC	Trajectory			
Carvalho et al. (2010)	Lamas d'Olo, Portugal	O_3	HYSPLIT trajectory	k-means	Hierarchical	
Chang et al. (2011)	Arctic Ocean	Aerosols	FLEXPART dispersion, HYSPLIT trajectory		Footprint Emission Sensitivity, PSCF	**PMF**
Cheng and Lin (2001)	Lamont, Oklahoma, US	Aerosols	HYSPLIT trajectory		PSCF	
Choi et al. (2010)	Sukmo Island, Korea	VOC	HYSPLIT trajectory		PSCF	**PMF**
Cohen et al. (2010)	Hanoi, Vietnam	PM	HYSPLIT trajectory		CPF	**PMF**
Colette et al. (2005)	European stations and MOZAIC (aircraft)	O_3	FLEXPART dispersion			**MVA**
Cui et al. (2011)	Jungfraujoch, Switzerland	O_3	LAGRANTO trajectory		Residence time	
Crawford et al. (2007)	Cape d'Aguilar, Hong Kong	Aerosols	HYSPLIT trajectory		PSCF	**PMF**
Crawford et al. (2009)	Cape d'Aguilar, Hong Kong	Aerosols	HYSPLIT trajectory	k-means		
Davis et al. (2010)	Virginia, USA	O_3	HYSPLIT trajectory	k-means (2-stage)		
de Foy et al. (2009)	Mexico City	O_3	FLEXPART dispersion	k-means (2-stage)		**PCA**
Delcloo and De Backer (2008)	Uccle, Belgium	O_3	APTRA trajectory	k-means		
Derwent et al. (2010)	UK	O_3	NAME dispersion			PTM
Doddridge et al. (1994)	Mace Head, Ireland	O_3, CO	NMC trajectory		Residence time	
Dogan et al. (2008)	Antalya, Turkey	Aerosols	ECMWF trajectory		PSCF	**PMF**
Dorling and Davies (1995)	3 Norwegian and 3 UK stations	sulphate and precipitation chemistry	Trajectory	k-means		
Du and Rodenburg (2007)	Camden, NJ, USA	PCB	HYSPLIT trajectory		PSCF	**PMF**
Dueñas et al. (2011)	Málaga, Spain	Aerosols, ^{7}Be and ^{210}Pb	HYSPLIT trajectory	k-means		
Dvorska et al. (2009)	Kosetice, Czech Republic	Persistant Organic Pollutants	HYSPLIT trajectory		Concentration loadings per country	
Ebinghaus et al. (2011)	Mace Head, Ireland	Mercury	NAME dispersion			
Eneroth et al. (2003)	Ny Alesand, Svaalbard	CO_2	HYSPLIT trajectory	Hierarchical		
Eneroth et al. (2007)	Zeppelin, Ny Alesand, Svaalbard	O_3, VOC, Mercury	HYSPLIT trajectory	Hierarchical		
Escudero et al. (2011)	La Castanya, Spain	PM, TSP	HYSPLIT trajectory		Residence time	
Evans et al. (2000)	Mace Head, Ireland	O_3, NOx, CO	UGAMP trajectory			CiTTyCAT
Fenneteaux et al. (1999)	Porspoder, France	O3, VOC, PAN, and NO*x*	NMC trajectory			
Font et al. (2011)	Northern Spain (aircraft)	CO_2	FLEXPART dispersion		Footprint Emission Sensitivity	**PCA**

Forster et al. (2001)	Mace Head and Europe	O_3, aerosols	FLEXTRA trajectory			FLEXPART with CO tracer
Fox and Ludwick (1976)	Quillayute, Washington	Pb	Trajectory			
Gebhart et al. (2005)	Big Bend NP, TX, USA	PM	HYSPLIT, ATAD and CAPITA trajectory		CF and PSCF	
Gebhart et al. (2011)	Rocky Mountain NP, USA	SO_2, NH_3, NO, NO_2	HYSPLIT trajectory		CF and PSCF	
Gheusi et al. (2011)	Pic du Midi, France	O_3	FLEXPART dispersion			
Gogoi et al. (2009)	Dibrugarh, India	Aerosol Optical Depth	HYSPLIT trajectory	Hierarchical		
Gregory et al. (1996) and Merrill (1996)	Pacific (aircraft)	O_3, CO, NO_{xy}, VOCs, minerals, CFCs	NMC trajectory			
Gros et al. (2004)	Atlantic Ocean (ship)	O_3, CO, propane	FLEXTRA trajectory			MATCH-MPIC
Güllü et al. (2005)	Antalya, Turkey	Aerosols	ECMWF trajectory		PSCF	**Factor analysis**
Hains et al. (2008)	Mid-Atlantic USA (aircraft)	O_3, CO, SO_2	HYSPLIT trajectory			
Halse et al. (2011)	86 European background stations	Persistant Organic Pollutants	FLEXPART dispersion		Footprint Emission Sensitivity	
Han et al. (2005)	New York state	Mercury	HYSPLIT trajectory + dispersion		PSCF	
Han et al. (2009)	Daihai, Mongolia	PM	HYSPLIT trajectory			**PCA**
Harrison et al. (2000)	Weybourne, UK	NO_x and NO_y, NH_3, NH_4^+	trajectory			
Harrison et al. (2006)	Birmingham, UK	NO_{xy}, OH	NAME dispersion (forward with chemistry)			
Hirdman et al. (2010a), Hirdman et al. (2010b)	Zeppelin, Alert, Barrow (+ Summit for b)	O_3, BC, aerosols	FLEXPART dispersion		Footprint emission sensitivity	
Hirdman et al. (2009)	Zeppelin, Svaalbard	Mercury	FLEXPART dispersion		Footprint emission sensitivity	
Huang et al. (2010)	Alert, Greenland	BC	HYSPLIT trajectory	k-means		
Hwang and Hopke (2007)	Kalmiopsis, Oregon, USA	$PM_{2.5}$	HYSPLIT trajectory		PSCF	**PMF**
Im et al. (2008)	Istanbul, Turkey	O_3, VOC, NO, NO_2	HYSPLIT trajectory			
Junker et al. (2004)	Orogrande, New Mexico, US	BC	HYSPLIT trajectory		Regional	
Junker et al. (2009)	4 stations in Taiwan and China	O_3, CO, SO_2, NO_x, PM	HYSPLIT trajectory		Regional	
Kaiser et al. (2007)	5 Alpine stations	O_3, CO, NO_x	FLEXTRA trajectory		PSCF, RCF	
Kang et al. (2006)	Seoul, Korea	PM, HNO_3, HONO, SO_2	HYSPLIT trajectory		PSCF	
Karaca and Camci (2010)	Istanbul, Turkey	PM	HYSPLIT trajectory	SOM		
Karaca et al. (2009)	Istanbul, Turkey	PM	HYSPLIT trajectory		PSCF	
Kassomenos et al. (2010)	Athens, Greece	PM_{10}	HYSPLIT trajectory	k-means, Hierarchical, SOM		
Kim and Kim (2008)	Gosan, Korea	PM	HYSPLIT trajectory			
Kocak et al. (2009)	Erdemli, Turkey	PM_{10}	HYSPLIT trajectory		PSCF	**PMF**
Kuhn et al. (2010)	Ellesmere island, Canada	Aerosols	FLEXTRA trajectory		Footprint emission sensitivity	Emission inventory
Law et al. (2010)	West Africa (aircraft)	O_3, CO, CO_2, NO_{xy}	UGAMP trajectory		RDF	
Lawler et al. (2009)	Cape Verde and Atlantic Ocean	O_3, NO_x, Chlorine species	Trajectory			MISTRA
Lee and Ashbaugh (2007a, 2007b, 2007c)	Grand Canyon NP, USA	SO_2	HYSPLIT trajectory		CPF, PSCF, MURA	
Lee et al. (2009)	Cape Verde	NO_x, NO_y	ECMWF trajectory			
Lefohn et al. (2011)	Yellowstone NP and 10 other US sites	O_3	LAGRANTO trajectory			
Lewis et al. (2007)	Atlantic Ocean (aircraft)	CO, PAN, alkanes	UGAMP trajectory			
Lin et al. (2004)	Taiwan	SO_2	Trajectory		CF	Emission inventory
Lupu and Maenhaut (2002)	3 sites in Scandinavia and 1 in Israel	PM, BC	HYSPLIT trajectory		PSCF, CF	
Makra et al. (2010, 2011)	Thessaloniki, Szeged, and Hamburg	Pollen counts and PM respectively	HYSPLIT trajectory	k-means (Mahalanobis, Convhull)		
Malcolm et al. (2000), Malcolm and Manning (2001)	UK and European sites	PM	NAME dispersion			
Manning et al. (2003)	Mace Head, Ireland	CFCs, CH_4 and NO	NAME dispersion			

(continued on next page)

Table 1 *(continued)*

Reference	Location	Species measured	Trajectory or dispersion model	Cluster analysis methods	Residence time analysis	Chemistry modelling
Martin et al. (2011)	Manchester, UK	PM	ECMWF trajectory	k-means		
McConnell et al. (2008)	West Africa (aircraft)	O_3, Ca and Al (dust), aerosols	NAME dispersion			
Merrill and Moody (1996), Merrill (1994)	Barbados, Bermuda, Mace Head, Tenerife	–	NMC trajectory		CPF	
Methven et al. (2001)	Mace Head, Ireland	O_3	UGAMP trajectory		CPF	CiTTyCAT
Methven et al. (2003)	Atlantic and Arctic Oceans (aircraft)	O_3	UGAMP trajectory		RDF	CiTTyCAT
Methven et al. (2006)	Atlantic Ocean (aircraft)	O_3, CO, NO_y	UGAMP trajectory		RDF	CiTTyCAT
Methven et al. (2006)	Atlantic Ocean (aircraft)	O_3, CO, NO_{xy}, VOCs,	UGAMP trajectory, FLEXPART dispersion		RDF	
Moody et al. (1989)	Virginia, Bermuda, Cape point, Amsterdam island	Precipitation chemistry	ATAD trajectory	Hierarchical		GAMBIT
Moy et al. (1994)	Shenandoah NP, Virginia, US	O_3, CO, NO_y	HYSPLIT trajectory	Hierarchical		
Müller et al. (2010)	Cape Verde	PM, Ca, K, Fe	HYSPLIT trajectory			
Nyanganyura et al. (2008)	Rukomechi, Zimbabwe	PM	HYSPLIT trajectory	Non-hierarchical		
Occhipinti et al. (2008)	North Carolina, USA	PM, nitrogen precipitation	HYSPLIT trajectory		Regional	
Park et al. (2008)	Incheon, Korea	PM	HYSPLIT trajectory		PSCF, CPF	
Paris et al. (2010)	Siberia (aircraft)	O_3, CO_2, CO	FLEXPART dispersion	k-means	Footprint emission sensitivity	
Pekney et al. (2006)	Pittsburgh, USA	PM	HYSPLIT trajectory		PSCF	**PMF**
Pochanart et al. (2001)	Arosa, Switzerland	O_3	Trajectory		Residence time	
Pochanart et al. (2003)	Mondy, Siberia	O_3, CO	NIES trajectory		Regional	
Poirot and Wishinski (1986)	Vermont	SO_2	Trajectory		Residence time	
Poirot et al. (2001)	Underhill (Vermont), USA	PM	CAPITA trajectory		PSCF, IP	**PMF**, **UNMIX**
Poissant (1999)	St. Lawrence River valley, Canada	Mercury	AES trajectory		PSCF	
Polissar et al. (1999, 2001a)	Barrow, Alaska	CN, BC, aerosols	CMDC trajectory		PSCF	
Polissar et al. (2001a)	Underhill (Vermont), USA	Na, Br, Ca, Mg, BC, Sulphate	CAPITA trajectory		PSCF	
Pongkiatkul and Kim Oanh (2007)	Bangkok, Thailand	O_3, PM	HYSPLIT trajectory	k-means	PSCF	
Real et al. (2007)	Canada, Arctic, Atlantic (aircraft)	O_3, NOx, CO, PAN	UGAMP trajectory			CiTTyCAT
Real et al. (2008)	Atlantic Ocean (aircraft)	O_3, CO, NO_{xy}, VOCs, aerosols	FLEXTRA trajectory			CiTTyCAT
Real et al. (2010)	Africa and Atlantic (AMMA)	O_3, CO, NO_x	FLEXTRA trajectory			CiTTyCAT
Reddy et al. (2010)	Anantapur, India	O_3	HYSPLIT trajectory			
Reidmiller et al. (2009)	Mount Bachelor, WA, USA	O_3, CO, Mercury	HYSPLIT trajectory		Residence time	GEOS-CHEM
Riccio et al. (2007)	Naples, Italy	O_3, PM	HYSPLIT trajectory	k-means		**PCA**
Robinson et al. (2011)	Borneo	CO, aerosols, halocarbons	ECMWF trajectory	Hierarchical		CF
Rodriguez et al. (2011)	Izaña, Canary islands	Aerosols	HYSPLIT trajectory		MCAR	
Rozwadowska et al. (2010)	Hornsund, Svaalbard	Aerosol Optical Thickness	HYSPLIT trajectory	k-means		
Ryall et al. (2001)	Mace Head, Ireland	CFC	NAME dispersion			
Ryall et al. (2002)	UK	PM	NAME dispersion			
Salvador et al. (2008)	Campisabalos, Spain	PM, aerosols	FLEXTRA trajectory	k-means		
Salvador et al. (2010)	Schauinsland (Germany), Puy de Dôme (France), Sonnblick (Austria)	PM, aerosols	FLEXTRA trajectory	k-means	RCF	
Schelfinger and Kaiser (2007)	Austrian stations	O_3, CO, NO_x	FLEXTRA trajectory		PSCF, RCF	Emission inventory
Schichtel et al. (2006)	Big Bend NP, Texas	PM	CAPITA trajectory		IP, PSCF	
Schmale et al. (2011)	Greenland (aircraft)	Aerosols	OFFLINE and LAGRANTO trajectory FLEXPART dispersion			Emission inventory
Schwarz et al. (2008)	Prague, Czech Republic	PM	HYSPLIT trajectory			

Seibert and Frank (2004)	Stockholm, Sweden	^{137}Cs	FLEXPART dispersion		Residence time	
Shan et al. (2009)	Jinan, China	O_3	HYSPLIT trajectory			
Sharma et al. (2004)	Alert, Greenland	PM, aerosols	CMC trajectory	k-means	CF	
Sharma et al. (2006)	Alert, Barrow, Arctic	PM, aerosols	CMC trajectory	k-means	CF	
Simmonds et al. (2004)	Mace Head, Ireland	O_3	NAME dispersion			
Simmonds et al. (1997)	Mace Head, Ireland	O_3, CO	Trajectory		Regional	
Solberg et al. (1996)	Zeppelin, Svaalbard	O_3	EMEP trajectory		Regional	
Solberg et al. (1997)	Zeppelin and Norwegian stations	O_3	EMEP trajectory		Regional	
Solberg et al. (2008)	European stations	–	FLEXTRA trajectory		Residence time	OSLO CTM EMEP model
Stohl (1996)	European stations	sulphate	FLEXTRA trajectory		RCF	
Stohl et al. (2000)	Jungfraujoch, Sonnblick, Zugspitze, Mt. Cimone	O_3	FLEXPART dispersion			FLEXPART tracers
Stohl et al. (2001)	MOZAIC aircraft	O_3	FLEXTRA trajectory		Regional	
Stohl et al. (2002)	Mace Head, Ireland	–	FLEXPART dispersion	Retroplume clusters		
Stohl et al. (2004)	Zeppelin, Svaalbard	Halocarbons CO_2, O_3, CO, Mercury	FLEXPART dispersion			FLEXPART tracers
Stohl et al. (2006)	(Aircraft)	Aerosol Optical Depth	FLEXPART dispersion		Footprint emission sensitivity	
Stohl et al. (2007)	Zeppelin, Svaalbard	O_3, CO, aerosol	FLEXPART dispersion		Footprint emission sensitivity	FLEXPART tracers
Stohl et al. (2009)	Various stations around the world	HFC	FLEXPART dispersion			Inversion model
Strong et al. (2010)	14 UK rural sites	O_3	HYSPLIT trajectory			ELMO
Tarasick et al. (2010)	North American stations	O_3 (sondes)	HYSPLIT trajectory		Trajectory-mapping	
Tarasova et al. (2009)	Jungfraujoch (Switzerland), Kislovodsk (Russia)	O_3	LAGRANTO trajectory			
Taubman et al. (2006)	US/Atlantic Ocean (aircraft)	O_3, CO, SO_2	HYSPLIT trajectory	Hierarchical		
Toledano et al. (2009)	El Arenosillo, Spain	Aerosol Optical Depth	Trajectory	k-means	Regional	
Traub et al. (2003)	Mediterranean (aircraft)	O_3, NOxy, CO, CO_2, HCHO, CH_4, VOCs	FLEXTRA trajectory		Regional	
Tscherwenka et al. (1998)	Sonnblick, Austria	SO_2	FLEXTRA trajectory		Residence time	
Tuzson et al. (2011)	Jungfraujoch, Switzerland	CO_2	FLEXPART dispersion		Residence time	
Vasconcelos et al. (1996)	Grand Canyon NP, USA	CH_3CCl_3	ATAD trajectory		CF	
Virkkula et al. (1999)	Sevettijärvi, Finland	Aerosols	TRADOS trajectory		Regional	
Walker et al. (2009)	Birmingham, UK	CO, O_3, NO_x, VOC	HYSPLIT and ECMWF trajectory			PTM
Wang et al. (2004)	Hong Kong, China	O_3, CO	HYSPLIT trajectory	Hierarchical		
Wang et al. (2010)	Beijing, China	O_3, CO	HYSPLIT trajectory			
Weiss-Penzias et al. (2004)	Cheeka Peak, WA, USA	O_3, CO	HYSPLIT trajectory		Residence time	GEOS-CHEM
Weiss-Penzias et al. (2006)	Mount Bachelor, WA, USA	O_3, CO, Mercury	HYSPLIT trajectory		Residence time	
Wimolwattanapun et al. (2011)	Bangkok, Thailand	PM	HYSPLIT trajectory		PSCF	**PMF**
Wolfe et al. (2007)	Mount Bachelor, WA USA	O_3, CO, PAN	HYSPLIT trajectory		Residence time	
Wotawa and Kroger (1999)	central Europe	NO_{xy}	FLEXTRA trajectory		CF	IMPO model
Wotawa et al. (2000)	11 stations in the Alps	O_3	FLEXTRA trajectory		CF, PSCF	
Wu et al. (2009)	Beijing, China	PM, NH_3, acidic gases	HYSPLIT trajectory		PSCF	
Xia et al. (2007)	Beijing, China	Aerosol Optical Depth	HYSPLIT trajectory	k-means (fuzzy c-means)		
Xiao et al. (2010)	Tibetan plateau	PCBs, PAH	HYSPLIT Trajectory			
Xie and Berkowitz (2007)	Houston, Texas, USA	VOCs	Trajectory derived from winds		PSCF, CPF	

(continued on next page)

Table 1 *(continued)*

Reference	Location	Species measured	Trajectory or dispersion model	Cluster analysis methods	Residence time analysis	Chemistry modelling
Xu et al. (2006)	US National Parks	PM	HYSPLIT trajectory		Residence time	
Xue et al. (2011)	Mt Waliguan, China	O_3, NO_x, NOy	HYSPLIT trajectory	Hierarchical		
Yan et al. (2008)	Shangdianzi, China	Aerosol optical depth	trajectory	k-means		
Zhu et al. (2011)	Beijing, China	PM	HYSPLIT trajectory	k-means	PSCF	

If the trajectory type is not specified, it is just left as "Trajectory".
If the Residence Time Analysis section and Cluster Analysis section are both left blanks then the methods were not standard and visual divisions or the respective author's own methods have been used.
Abbreviations in the cluster analysis, residence time and modelling columns are explained in Sections 4.2, 4.1 and 2.4 but a summary is included here; CF (Concentration Field), CPF (Conditional Probability Function), PSCF (Potential Source Contribution Function), IP (Incremental Probability), RCF (Redistribution Concentration Field), RDF (Reverse Domain Filling).
Source apportionment models (e.g. PMF (Positive Matrix Factorisation) and PCA (Principal Component Analysis)) are included in the chemistry modelling column in bold as well as whether emission inventories were used to validate and test.

performed on each weather type to reveal how the ozone and NO concentrations are explained in terms of meteorological parameters, showing how there is a better correlation in the cyclonic compared to anticyclonic conditions, which shows how the increasing significance of photochemistry in cyclonic periods.

The classification of daily weather types (DWT) on a daily basis for the British Isles between 1861 and 1997 has been carried out by Lamb (1972) and further developed by Jenkinson and Collison (1977), where the basic flow direction and level of anticyclonicity or cyclonicity over an area were determined. The grid over which the objective DWTs are calculated can be moved to different locations and a new catalogue, appropriate for the new location, can be constructed. Ohare and Wilby (1995) used the Lamb weather types to show that ozone levels from a variety of measurement stations in the UK were strongly correlated with the prevailing weather type.

Meteorological measurements at Szeged, Hungary have been separated according to Factor Analysis (reduction of the dimensionality of the meteorological dataset) and objectively grouped into days with similar weather conditions and then compared to CO, NO, NO_2, SO_2, O_3 and Total Suspended Particles (TSP) measurements over 4 years (Makra et al., 2006). Factor analysis of meteorological data, followed by cluster analysis was applied to air masses arriving in Athens over 4 years and corresponding air pollutant measurements (Sindosi et al., 2003). Helmis et al. (2003) studied the connection of atmospheric circulation to trans-boundary air pollution by using a circulation-to-environment approach where 14 synoptic scale patterns were distinguished over Athens and SO_2, NOx and O_3 data were compared between the synoptic periods and additional modelling linked this to inflow and outflow of the Athens basin. Demuzere and van Lipzig (2010) used linear regression methods to explain O_3 and PM variations and found that classifying according to the automated "Lamb" weather type prior to the regression analysis was superior to just using the linear regression. Comparison of objective air mass types and the "Peczely" weather types to classify daily pollution levels over the Carpathian Basin (using 12 meteorological and eight pollutant parameters) has been conducted for a four year period (Makra et al., 2009).

2.2. Trajectory models

A modelled trajectory is an estimate of the transport pathway of an infinitesimally small air parcel and an estimate of the centreline of an advected air mass subject to vertical and horizontal dispersion. Back trajectories trace the path of a polluted air parcel backward in time and have long been used to track the history and pathway of air parcels arriving at a specific location since they were first developed in the 1940s by Petterssen (1940). The first trajectories were one dimensional lines calculated using recorded meteorological fields in a model dealing with the combination of wind field influences on the air. Computational advances in the 1960s allowed isentropic analysis and trajectory calculations to be performed graphically on computers (Danielsen and Bleck, 1967). These back trajectories followed the path of release backwards on an Eulerian grid, where the flow of particles is depicted as a function of a fixed position and time. More information can be extracted from two-dimensional trajectories by colour-coding their pathways by height and by including markers corresponding to time of travel.

Fox and Ludwick (1976) carried out one of the first studies using back trajectories to identify source regions of atmospheric pollution. They compared levels of lead at Quillayute, Washington using 1 day back trajectories and found that the land areas where source regions and the oceanic air yielded low levels of lead. Ashbaugh et al. (1985) carried out pioneering work on identifying sources of sulphur and corresponding statistical associations between air mass history and above average concentrations. Air quality monitoring records and back trajectories were used to identify those regions from which high sulphur concentrations were most likely to arrive.

A review of the types and uses for back trajectories and the associated errors and probabilities within them has been provided by Stohl (1998), in which it is stated that there can be serious misinterpretations of a flow situation (represented as linear air mass movements) if the magnitude of the errors cannot be estimated. Owing to errors and assumptions in the wind fields used to calculate the trajectories, the uncertainty of trajectories increases with time along the path. Seibert (1993) has investigated the accuracy of trajectories. Kahl (1993) investigated the errors within the calculation of trajectories by calculating trajectories at multiple levels and at regular grid intervals around a site to assess the extent of vertical and horizontal wind shears. The errors and issues associated with back trajectories (uncertainty arising from interpolation of sparse meteorological data, assumptions regarding vertical transport, observational errors, sub-grid-scale phenomenon, turbulence, convection, evaporation, and condensation) are explained in Polissar et al. (1999). Taking into account uncertainties, back trajectories are often better suited to large scale circulation studies, such as shown in the detailed global trajectory study on air parcel circulation between the troposphere and stratosphere carried out by Jackson et al. (2001).

Examples of some of the community's commonly used trajectory models are listed in Table 1. The FLEXTRA model has been used extensively and is described by Stohl et al. (1995). The HYbrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Draxler and Rolph, 2003) has been used extensively in many studies as illustrated in Table 1. The UK Universities' Global Atmospheric Modelling Programme (UGAMP) trajectory model has been used to complement many field campaigns (e.g. Cape et al., 2000; Methven et al., 2006). The Centre for Air Pollution Impact and Trends Analysis CAPITA Monte Carlo (CMC) model (Schichtel et al., 2006) uses meteorological wind fields to advect the particles in three dimensional space, whilst the intense vertical mixing that takes place within the atmospheric boundary layer is simulated using a Monte Carlo technique, which evenly distributes the particles between the surface and the mixing height. Other trajectory models include APTRA (Delcloo and De Backer, 2008) obtained from the ECMWF (European Centre for Medium-Range Weather Forecasts), the LAGRANTO model (Tarasova et al., 2009), the TRADOS model (long-range Trajectory, Dispersion and Dose Model of the Finnish Meteorological Institute) (Virkkula et al., 1999) and trajectory models run directly from meteorological datasets such as the ECMWF three dimensional isentropic model (Dogan et al., 2008), NMC (US National Meteorological

Centre) (Merrill, 1994; Merrill and Moody, 1996), CMDC (Climate Monitoring and Diagnostics Laboratory) (Polissar et al., 1999, 2001b) and NIES (Russian National Institute for Environmental Studies) (Pochanart et al., 2003) trajectory models.

A thorough comparison of various trajectory models with a number of meteorological parameters derived through different systems has been carried out by Gebhart et al. (2005). The 1980s era Atmospheric Transport and Dispersion (ATAD) model was compared to two of the more recent models; HYSPLIT and the CAPITA CMC model. Various data sets (ATAD, BRAVO, MM5 and GDAS) were used as meteorological fields. Tests on data at a field station in Texas showed that there was evidence for systematic differences between the average results of different back trajectory models. Depending on the meteorological input data used in the calculations, the trajectories could be as different by as much as 180° during certain episodes.

Multiple trajectories are used to simulate the air mass history, since the air parcels arriving at a site could have followed different trajectories owing to turbulent atmospheric mixing and advective processes. Multiple trajectories also provide a measure of the uncertainty in the air mass transport pathway and the clustering of individual trajectories discussed later shows how grouping trajectories reduces uncertainty. Cabello et al. (2008) reported on how there are large differences in clustering trajectories when using different NCEP met data (either $1^\circ \times 1^\circ$ FNL data or $2.5^\circ \times 2.5^\circ$ global reanalysis data), and the type of meteorological data rather than the clustering method makes the biggest difference to defining the different clusters.

Example back trajectory analyses are shown in Fig. 3 for Mace Head, Ireland (Cape et al., 2000), Weybourne, UK (Cardenas et al., 1998), the Cape Verde Atmospheric Observatory (Müller et al., 2010) and the TOR station at Porspoder, France (Fenneteaux et al., 1999).

2.3. Particle dispersion models

More complex methods to calculate air mass pathways and air mass footprints have mostly come from Lagrangian Particle Dispersion Models (LPDM) that follow the chaotic pathways of air parcels as probability distributions. Lagrangian methods involve plotting the position of an individual parcel through time as it follows the average atmospheric motion, giving the pathline of the parcel within a specific period of time. The Lagrangian dispersion concept is more accurate because individual particles move independently from each other and can thus carry additional information and the advection scheme is more accurate than those of trajectory models as it attempts to capture turbulence, which causes a more probabilistic and realistic growth in the volume of influence.

The 2-D maps created for backward runs illustrate which geographical regions have influenced the air arriving at a site. These can be seen as four-dimensional as opposed to quasi-one-dimensional back trajectories, making them an interesting extension to conventional trajectory models, allowing a more realistic representation of transport in the planetary boundary layer, where turbulence is important.

Stohl et al. (2002) explains the added accuracy and regional spread of dispersion models and how they can be a replacement for simple back trajectory models for comparing atmospheric composition measurements at a given site. Flesch et al. (1995) describes the parameters within Lagrangian Stochastic models and their use in estimating atmospheric emissions and source–receptor relationships.

The HYSPLIT 4 model can be run with puffs instead of particles with an added dispersion scheme added to the initial trajectory computation (Cohen et al., 2004; Han et al., 2005). This was combined with emissions to extract source–receptor relationships in the area.

Typical models used in many studies include the FLEXPART and NAME models. The FLEXPART model (Stohl et al., 2005, 1998) (an example is shown in Fig. 7), has been used for a variety of research purposes and for emergency preparedness. The model is usually driven by ECMWF meteorological input data. FLEXPART evolved from the FLEXTRA back trajectory model but represents transport and dispersion by calculating the 3-D trajectories of a multitude of particles.

The UK Met. office's NAME (*N*umerical *A*tmospheric Dispersion *M*odelling *E*nvironment) (Jones et al., 2007; Ryall and Maryon, 1998) dispersion model described in detail in the case study in Section 5 can be used to give a footprint of a site on scales of hours to years. The model was developed in the late 1980s following the Chernobyl accident (and originally named Nuclear Accident Model) to give emergency response dispersion predictions for nuclear incidents (Maryon et al., 1991). NAME is usually run with the Met Office's operational global Numerical Weather Prediction NWP model, the Unified Model (Cullen, 1993) meteorological data on a variety of regional or global scales.

Manning et al. (2003, 2011) and Ryall et al. (2001) have used NAME to calculate the distribution of air masses arriving at Mace Head, Ireland, from different Atlantic Ocean and European regions and by combining with composition measurements at the station, they have derived emission inventories for Europe. Ebinghaus et al. (2011) used the NAME model to separate out baseline (non-polluted) air arriving at Mace Head and looked at monthly mercury levels over a 13 year period and found there to be a statistically significant downward trend. In combination with satellite imagery and observational data from Mace Head the NAME model was used to investigate the origin of high Particulate Matter concentrations over the British isles during March 2000 and it showed that the most likely origin of the episode was long range transport of dust from the Sahara region of North Africa and not volcanic ash from an Icelandic volcano (Ryall et al., 2002). Ryall and Maryon (1998) tested NAME using the ETEX database (European Tracer Experiment database of 168 ground-level sampling stations in Western and Eastern Europe) to assess its suitability to predict the overall spread and timing of a pollutant plume across Europe. The NAME model has also been used in conjunction with satellite measurements (Hewitt, 2010) to develop a methodology to investigate regional scale carbon budgets. Polson et al. (2011) used NAME to carry out inversions of aircraft measurements over the UK to derive mapped emissions of the UK which compared well with the UK national emission inventory (NAEI).

2.4. Chemistry transport models

Chemistry Transport Models (CTM) combine meteorological fields, emissions and physical atmospheric processes with chemistry schemes and reaction kinetics to describe the

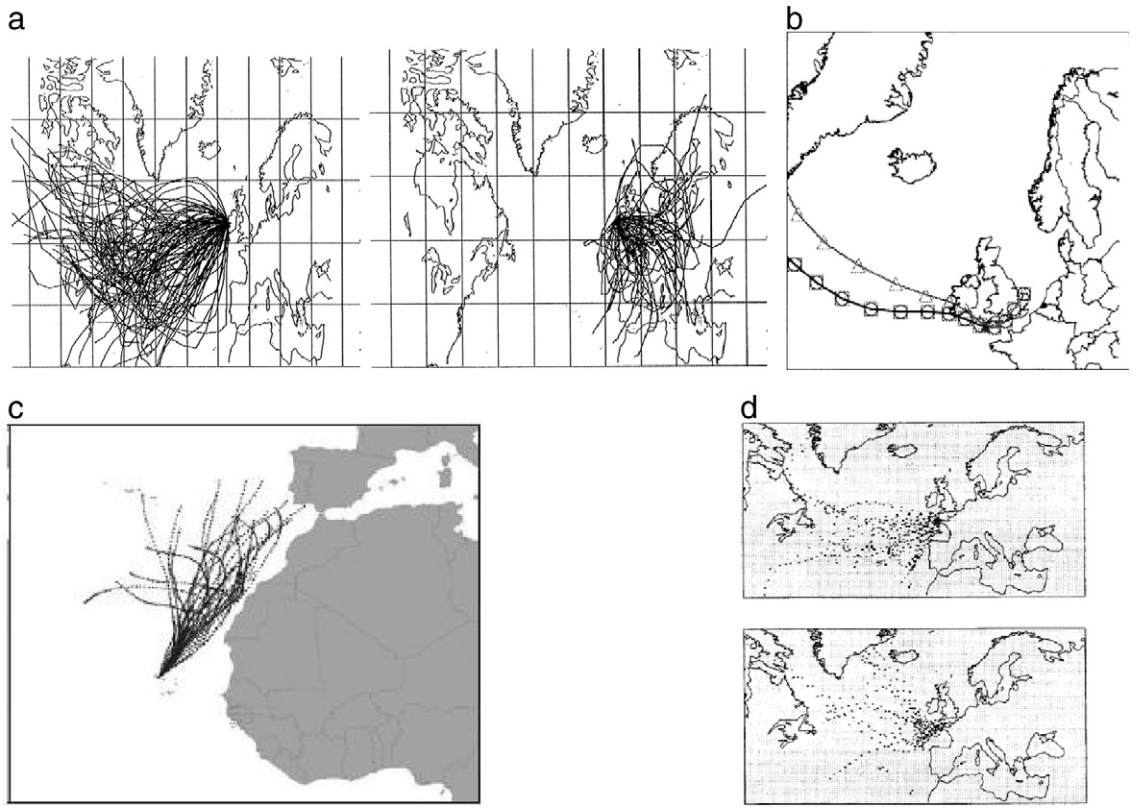

Fig. 3. Trajectory examples: a) Westerly and Easterly Mace Head, Ireland 6 hourly 5 day back trajectory clusters. b) Weybourne, UK 6 hourly trajectories arriving during 1993. c) 4 day backward Cape Verde Observatory daily trajectories from mid May to mid June 2007. d) Back trajectories for the TOR station at Porspoder, France for summers 1992–1995 associated with 10th and 90th percentile ozone levels respectively.
a): Taken from Cape et al. (2000). b): Taken from Cardenas et al. (1998). c): Taken from Müller et al. (2010). d):Taken from Fenneteaux et al. (1999).

chemistry and physical transformations within air masses. The NAME dispersion model has an option to add a chemistry scheme to the tracer transport calculations but is not as complete as a CTM. In this work we define CTMs as ones that by default include emission inventories chemistry and transport schemes.

Emissions databases and chemistry reaction scheme databases are incorporated into models in order to quantitatively predict composition during transport. CTMs contain a large number of uncertainties as do the emission databases that are used within them. They have been used in parallel with many of the studies detailed in this review to test how well they compare with the observed composition from various emission sources and influences and to add value to the trajectory analysis but CTMs are only mentioned in this review if they have been used to compare to simple trajectory or dispersion models.

Wind data have been used as input for many CTMs such as in the distance-weighted wind roses that were combined with measured ozone as input for the Edinburgh Lancaster Model for Ozone (ELMO) model. The model was used to predict ground level ozone concentrations in the UK (Strong et al., 2006) and it predicts the higher ozone episodes (98th percentile) well. ELMO-2 used the HYSPLIT trajectory model as input and was seen to reproduce ozone episodes and diurnal cycles at several UK monitoring sites during summer 1995 (Strong et al., 2010). In the study by Gilliam et al. (2006), the Pennsylvania State University/National Center for Atmospheric Research Fifth Generation Mesoscale Meteorological Model (MM5) was used with seasonal subsets of pseudo-trajectories derived from radar wind profiler data and from simulated wind fields as input to provide an estimate of model errors in terms of wind transport. They found that any inconsistencies in the meteorology were passed on to the air quality model. CO and SO_2 transport was well simulated with both MM5 and WRF (Weather Research and Forecast) meteorological parameters over the Mexico City basin during the MILAGRO study (de Foy et al., 2009). de Foy et al. (2009) used FLEXPART back trajectories, in combination with measured air pollutant concentrations to identify potential source regions and the WRF and MM5 models were used to evaluate the simulated trajectories by comparing the potential source regions with emission inventories.

A Photochemical Trajectory Model (PTM) was used with emissions databases to simulate atmospheric composition, during the Pollution in the Urban Midlands Atmosphere (PUMA) campaign in Birmingham (Walker et al., 2009). The model described the photochemical ozone formation as well as inorganic and organic aerosol formation in northwestern Europe and in this case followed advected air masses from source regions in Europe to the receptor location in the UK. The model was tested against 3 day back trajectories to test the model's ability to pull out the origin of all the pollutant episodes and it was not able to trace back the pollutant

transport during conditions of slow moving anticyclonic conditions. This example shows how trajectory studies should be carried out independently to trajectory chemical modelling studies as in some cases the model cannot interpret all periods.

The Lagrangian chemistry transport model CiTTyCAT (Cambridge Tropospheric Trajectory model of Chemistry and Transport) (Evans et al., 2000) has been used in many studies investigating atmospheric composition. CiTTyCAT simulates chemical transformation following trajectories with a photochemistry scheme that includes the degradation of some hydrocarbons, a representation of the spread of surface emissions into the boundary layer (using emission inventories) and dry deposition. The chemical initial conditions at the trajectory origin are defined by interpolating concentrations from the TOMCAT CTM, which calculates the abundances of chemical species in the troposphere and stratosphere. CiTTyCAT was seen to accurately simulate 70% of the variance in the relationship between chemical composition at Mace Head during field measurements in 1996 and the origin of the resolved flow when compared to time series of trajectory-origin-averaged measured ozone (Strong et al., 2006). Different processes influencing the evolution of pollutant levels in a trans-Atlantic plume have been analysed with the CiTTyCAT model and average trace gas concentrations and their correlations (O_3/CO and NOy/CO) calculated to study the factors governing ozone production (Real et al., 2008).

The GEOS-CHEM 3-D global CTM (using assimilated meteorological data compiled at the NASA Global Modeling and Assimilation Office (GMAO)) was used to calculate a long range transport component to air masses arriving at Cheeka Peak Observatory (western US) from over Asia from trajectory analysis (Weiss-Penzias et al., 2004). The CO produced from Asian biomass burning, and Asian and European fossil fuel and biofuel sources was calculated and compared with measured CO. Reidmiller et al. (2009) used back trajectory analysis coupled with ground-based measurements from the Mount Bachelor Observatory (western US) to confirm GEOS-CHEM simulations, suggesting a significant change in long range transport between 2005 and 2006, owing to changing patterns of long range transport Asian air masses arriving at the site. Investigations of ozone and CO in biomass burning plumes were investigated using the GEOS-CHEM model during the AMMA campaign in West Africa campaign (Real et al., 2010) and the modelled mixing ratios at each successive day along the trajectory are shown in Fig. 4, showing how modelled CO decreases downwind at all heights whereas O_3 decreases downwind only in the lowest trajectory.

Back trajectories calculated for analysis of cruise ship measurements in the Atlantic Ocean and were compared to the MATCH-MPIC (Model of Atmospheric Transport and Chemistry-Max Planck Institute for Chemistry version) model (Gros et al., 2004). MATCH is a global atmospheric offline model, driven by 3-D meteorological parameters with a CH_4–CO–HOx–NOx "background" chemistry, a simplified representation of isoprene and other VOC chemistry and emissions of CO and VOCs from energy and industrial activities taken from the EDGAR inventory (Olivier et al., 1996) emissions database.

The MISTRA 1-D Lagrangian chemistry model has been used to simulate multiphase halogen cycling mechanisms and compared with the observed association between Cl_2 and pollutants at the Cape Verde islands. The model data was compared to the halogen chemistry during particular air transport pathways and found Cl to be involved in CH_4, DMS and O_3 cycling (Lawler et al., 2009).

Another model to quantify source–receptor relationships was the Gaussian trajectory transfer-coefficient model (GTx), which has been used to model PM at Taichung City and Taipei City, Taiwan (Tsuang, 2003; Tsuang et al., 2003a, 2003b). It was able to simulate the daily variation of PM concentrations at these sites and was used to determine the source of particulates and dust from long range transport.

3. Applications of trajectory data to interpret composition observations

Many studies have been carried out to relate variability in chemical observations to variations in synoptic-scale circulation. A summary of previous research studies that combine trajectory or dispersion modelling with chemistry datasets from long term measurement stations and field campaigns is presented in Table 1. Examples of the techniques used, the composition measurements that were analysed, the locations, the type of cluster analysis or sector analysis techniques used (described in Section 4) and if chemistry modelling was used to complement the study are listed for each study.

3.1. Characterisation of long-term in situ ground based measurements

Studies at the Northern hemispheric background measurement site of Mace Head, Ireland used trajectories to as-

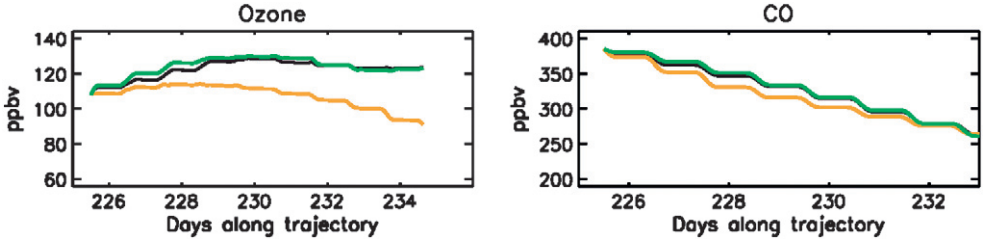

Fig. 4. Chemistry modelling along a trajectory: GEOS-CHEM model simulations of ozone and CO evolution in a biomass burning plume arriving into Western Africa from South America during the AMMA aircraft campaign. Black represents the reference trajectory from the flight altitude, green a higher level trajectory and orange a lower level trajectory.
Taken from Real et al. (2010).

sign each 6 hourly air mass arriving at the site to one of eight 45° sectors (centred on north, northeast, and so on to northwest) (Simmonds et al., 1997). If all four trajectories within a day lay within the same sector, it was classified to that sector and if no such allocation was possible for a particular day, then that day was unclassified. Additional trajectory analysis sorted the data into the specific regions of USA and Canada, Greenland and Iceland, Europe, and southerly latitudes and with both of these sector analyses, so-called background or polluted conditions were separated for analysis of ozone and CO levels. The methodology has been shown to be robust for the determination of background trends. The Zeppelin station in the Arctic has also been classified with the same 8 sectors surrounding the site (see Fig. 8d) and a transport sector allocated to each trajectory if at least 50% of the last 24 h of the trajectory was closer than 850 km from the station (Solberg et al., 1996). Spring low ozone episodes were found to originate from a Westerly/Northerly Arctic oceanic direction.

The air arriving at Porspoder in Brittany, France was split into three kinds of oceanic air masses (North Atlantic northern and southern latitudes and North American continent) and seasonal variations of PAN, VOCs, O_3 and NOx according to air mass types over a period of 4 years were studied (Fenneteaux et al., 1999). Corresponding back trajectories for the 10th and 90th percentiles (see Fig. 3d) and average levels of the species in winter and summer showed that there were strong seasonal and regional influences on these species. Air arriving at the remote mountain site of Mondy in Siberia was classified into 4 transport pathways (Europe, Siberia, High-latitude, and south-west air masses) and CO and O_3 levels at the site were then averaged for each trajectory type to reveal how European air masses had the highest O_3 and CO levels (Pochanart et al., 2003).

Various studies on many US National Parks and wilderness areas (through the IMPROVE sites) used residence time analysis of back trajectories to find source–receptor relationships and the source type and origin of high air-borne pollutants (Hopke and Allan, 2009). The online Combined Aerosol Trajectory (CATT) tool, uses the IMPROVE sites and 5 day back trajectories to colour weight the trajectories to individual or aggregated sites, in order to analyse long term patterns in pollution transport to these areas (Poirot and Allan, 2009).

Brankov et al. (1998) analysed the ozone footprint at Whiteface Mountain in New York state by correlating the short-term component of its ozone time-series with ozone time series at neighbouring stations. The short-term (weather-related) component was separated from the long-term (climate-related) and seasonal components embedded in ozone time-series data with the Kolmogorov filter. They carried out the correlations between each trajectory cluster type and the short-term component of ozone and found it to be correlated with that measured at a number of other sites lying within a cluster envelope, lagged by up to 3 days. This form of analysis tested the hypothesis of whether when ozone was transported from a certain direction, the time-lagged correlation increases in that direction. The distance at which the time-lagged inter-site correlations reach a maximum is expected to be proportional to the distance the air mass can travel in that time and would suggest the transport of ozone pollution to the location of concern and reveal the spatial and temporal scales involved.

In a review of recent aerosol studies in Europe it was found that 11% of all studies used back trajectory methods to cluster aerosol levels according to their origin and transport pathways (Viana et al., 2008). Rozwadowska et al. (2010), Salvador et al. (2008, 2010) and Sharma et al. (2006) show some of the latest aerosol source–receptor studies using back trajectory classification at a range of sites in Europe and the Arctic.

3.1.1. Vertical transport studies at ground based stations

Tarasova et al. (2009) classified the vertical as well as geographical origin of air masses arriving at two mountain stations (Jungfraujoch and Kislovodsk) by using potential vorticity, altitude along the trajectory and boundary layer height to discriminate different vertical source areas, as well as the usual classification using the horizontal coordinates of back trajectories. This methodology allowed classification of the air masses according to their contact with the free troposphere and the stratosphere and showed how ozone levels varied significantly depending on their vertical pathway to the mountain tops. CO_2 measurements at the Jungfraujoch mountain station were compared using the FLEXPART dispersion model and the model was also used to calculate the residence times of the air masses in the boundary layer and relate this to CO_2 exchange processes (Tuzson et al., 2011). Delcloo and De Backer (2008) separated the trajectory analysis clustering of 32 years of trajectories arriving in Uccle, Belgium by elevation, into planetary boundary layer and free troposphere origins to understand how ozone levels have varied in both high and low elevation air masses. Lefohn et al. (2011) used trajectories to isolate periods of stratosphere-tropospheric exchange at several rural US sites that these periods often coincided with ozone exceedance periods, especially at the high elevation site in Yellowstone National park. Gheusi et al. (2011) used 3 day back trajectories to calculate the time spent in the boundary layer before arriving at the high altitude site of Pic du Midi in France in order to study the influences on ozone levels.

Tarasick et al. (2010) ran forward and back trajectory from a series of North American ozone sonde stations and linked the ozone sonde measurements at the receptor sites to the grid points of the trajectory pathways to obtain ozone maps. The trajectory-mapped ozone values showed reasonable agreement to the actual soundings and also compared well with MOZAIC aircraft measurement profiles and surface station data.

3.2. Investigating long term trends and seasonality in composition measurements

Understanding the influence of transport patterns on long-term trends is essential to interpreting their changing chemical and physical signatures. Seasonal cluster analysis of trajectory types has been used to reveal the percentage of each trajectory type arriving in each season as well as the ozone trends in each season in Uccle, Belgium (Delcloo and De Backer, 2008). Monthly averaged CO and O_3 for different trajectory types were calculated for a station in Siberia (Pochanart et al., 2003). Junker et al. (2009) studied 12 years of O_3, CO, SO_2, PM and

NO_x measurements at four measurement stations in Taiwan and China and separated the composition by their air mass history to show long term trends and also seasonal variations. Abdalmogith and Harrison (2005) carried out a seasonal cluster analysis of PM levels at Harwell and Belfast in the UK and found small variations in the direction of the clusters and the levels and type of PM observed between the seasons. Solberg et al. (1997) found the background ozone to have a small seasonal variation and a spring maximum as opposed to the summer maximum and large seasonal variation for European polluted air masses at Birkenes in Norway (Fig. 5a). Simmonds et al. (1997) looked at the relative ozone contribution from each of the source regions to the observed spring maximum at Mace Head. This was obtained by subtracting the mean ozone concentration for the Northern Hemisphere Marine Mid-Latitude Background (NHMLB) (34.8 ppbv at the time) from 5 years of ozone monthly means. These monthly differences were shown as ozone excesses or deficits relative to the NHMLB ozone concentration and displayed marked seasonal differences between Atlantic, European and American air masses in Fig. 5b.

Long term records of composition variations separated by corresponding air mass history are rare due to the fact that running back trajectories on hourly or even daily timescales for over 10 years is very computer-intensive and the meteorological data needed for the model may have changed over the years. Trends in air mass climatology for a 40 year trajectory dataset have been investigated by Shadbolt et al. (2006) and discussed in Section 4.1.2.

Jorba et al. (2004) presented a study on the air arriving in Barcelona spanning July 1997–June 2002 showing the seasonal influence on climatology (Fig. 6a). A similar monthly distribution of regional influences was calculated for 15 years at both Alert and Barrow in the Arctic (Sharma et al., 2006) as shown in the monthly distribution of sector influences in Fig. 6b. Furthermore, Sharma et al. (2006) used a geometric time variation model to describe the temporal variation of 6-hourly Black Carbon (BC) concentrations, including a long-term trend, long-term cycles, seasonal variation, and an autoregressive component that described short-term temporal correlations. The long term trends in both winter and summer varied greatly between the different sectors, especially at Barrow.

Pochanart et al. (2001) found that there were strong seasonal differences in the relationship between residence times of air masses over Europe and ozone levels in Arosa, Switzerland and that ozone concentration depends significantly on European residence times in spring and summer. Eneroth et al. (2003) carried out a cluster analysis of 10 years of trajectories arriving at Ny-Ålesund on a monthly basis, to study inter and intra-annual variations of CO_2 and found that there were seasonal differences in the prevalence of each trajectory type that lead to varying CO_2 levels but they found no conclusive linkage between CO_2 levels and transport pathways. They found that the Arctic CO_2 observations were influenced by synoptic scale atmospheric flow patterns and there were correlations between the North Atlantic Oscillation (NAO)-index and the number of trajectories in each trajectory cluster.

It is clear that for many of these studies the ability to separate the variability, mainly driven by the meteorology from the underlying emission/pre-curser emission trend is a key area for further work or both scientific and policy relevance.

3.3. Analysing the air mass history for aircraft based measurements

Combining trajectory and dispersion studies of air mass histories with aircraft measurements helps to build a 3-D picture of air mass movement and transport. This picture is often achieved by calculating back trajectories and dispersion pathways from multiple elevations along the aircraft flight track.

Back trajectories at 1, 2 and 3 km were run during the RAMMPP aircraft campaign over the mid-Atlantic US (Taubman et al., 2006) in order to pick out any variations in atmospheric circulation patterns in the lower atmosphere and identify the impacts on regional transport. Vertical composition measurements were combined with these back trajectories to investigate the history of the air during summer ozone pollution episodes and the role of transport to the boundary layer.

Hains et al. (2008) calculated average ozone levels in vertical bins of the atmosphere and transposed this on the vari-

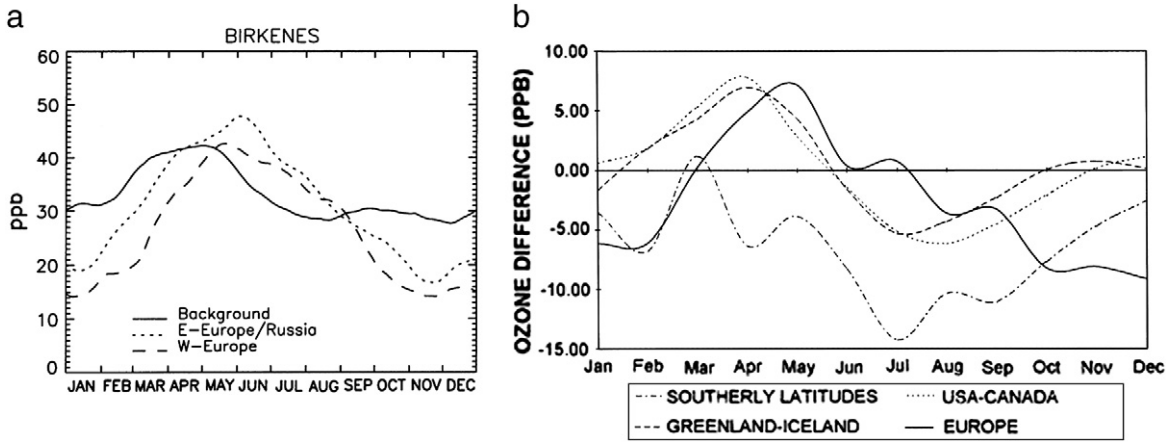

Fig. 5. Seasonal composition cycles in different trajectory types: a) Seasonal O_3 cycles at Birkenes, Norway in 3 air mass sector types derived from trajectories. b) The difference between the Northern Hemisphere Mid Latitude background (NHMLB) average ozone at Mace Head for the various sectors around the site. a): Taken from Solberg et al. (1997). b): Taken from Simmonds et al. (1997).

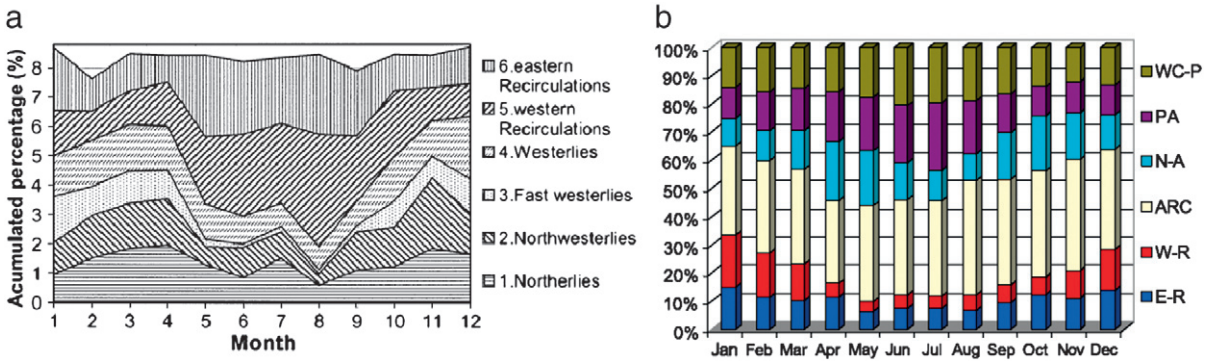

Fig. 6. Seasonality of trajectory types: a) Monthly % distribution of sectors for all 3565 back trajectories arriving in Barcelona over 5 years. b) Monthly distribution of trajectory types representing various sectors (Russia, Europe, America, Arctic in various combinations) influencing the Barrow Arctic station over 15 years 67 and subsequently used for interpreting Black Carbon measurements.
a): Taken from Jorba et al. (2004) © American Meteorological Society. Reprinted with permission. b): Taken from Sharma et al. (2006).

ous trajectory types that were observed during flights over the north-east US in the INTEX-NA campaigns in 1993–2003. An 80 km radius circle was drawn around each hourly point on the trajectory and the NOx emissions from that area were compared to the ozone level at the corresponding altitude bin.

Clustering of air mass composition from the MOZAIC dataset (long-term composition measurements from an aircraft network) using the FLEXPART trajectory model was compared to a multivariate analysis technique that classifies ozone-rich layers observed in tropospheric profiles according to their origin by season (Colette et al., 2005). The ozone multivariate analysis technique was found to underestimate the long-range transport that the trajectory studies identified.

Reverse Domain Filling (RDF) is a technique that defines regions of similar air mass origin with the use of back trajectories and is used to plan flight paths through regions of interest. This method followed on from a technique developed by Sutton et al. (1994) for producing high resolution stratospheric N_2O data from low resolution satellite data using data assimilation of high resolution wind fields which has subsequently been used in many studies for filling in the spatial domain of tracer data from trajectories. Stohl et al. (2001) constructed global ozone climatologies based on MOZAIC aircraft data in both the troposphere and stratosphere by attributing measured ozone values to the entire path of each of the 8 day back trajectories, which allowed a regional identification of the origin of the different ozone levels.

Methven et al. (2003) used RDF trajectories arriving on a high-resolution three-dimensional grid (RDF3D) to simulate air mass structure accurately by colouring arrival grid points according to the specific humidity (or potential vorticity) at the origin of each trajectory. Back trajectories were calculated from every point on the 3-D grid from a reference time near the anticipated flight time and for each RDF forecast trajectories were integrated backwards in time for 3 days, using a combination of ECMWF forecasts and interpolating specific humidity from the forecasts to the origin of each trajectory. The flights were targeted at regions where there were neighbouring air masses with distinct origins. Chemical measurements during the flights were overlaid onto the RDF maps and variations in composition were compared with the pre-selected air mass types to see if composition did vary within the different air mass types. Hydrocarbon measurements during the IGAC Lagrangian 2K4 experiment in July 2004 (Methven et al., 2006) were analysed according to calculated back trajectories and RDF was used to identify Lagrangian matches between flight segments from different aircraft. Fig. 7 shows examples of horizontal and vertical RDF plots used to plan flights that intercept biomass burning plumes during this study. This was the first experiment aiming to take measurements that were linked by trajectories over intercontinental distances through the free troposphere, where vertical motion is important and was described as a "pseudo-Lagrangian experiment". Results from the FLEXPART model, run with CO tracers were also used to confirm matches.

Following on from this form of analysis, Real et al. (2008) isolated the chemistry of an anthropogenic pollutant plume transported across the North Atlantic at low altitudes. Similar use of flight data and back trajectories was used to track Alaskan wild fire plume transport during the ICARTT aircraft campaign (Real et al., 2007) and study the evolution of the composition and ozone formation a few days from emission. Schmale et al. (2011) used the OFFLINE and LAGRANTO trajectory models as well as the FLEXPART dispersion model with EDGAR emissions to study the aerosol composition of a variety of air masses over Greenland during the POLARCAT (Polar Study using Aircraft, Remote Sensing, Surface Measurements and Models, of Climate, Chemistry, Aerosols, and Transport) summer campaign in Greenland.

During the AMMA campaign in West Africa RDF back trajectories were used to split the area over which flights had passed into boxes representing the 4 main regions of recent (10 day) origin (Law et al., 2010). Vertical differences in composition and the occurrence of convection were included in the separation and interpolation of the measured composition data.

4. Methods for deriving classifications of air mass pathways in association with composition at the receptor site

In order to analyse the association between trajectories and concentrations of various species in air arriving at a site, a multitude of methods to carry out trajectory classifications has been devised. These can generally be split into two different methodological groups. The first is to sort air masses by designated air mass sectors, representing a different

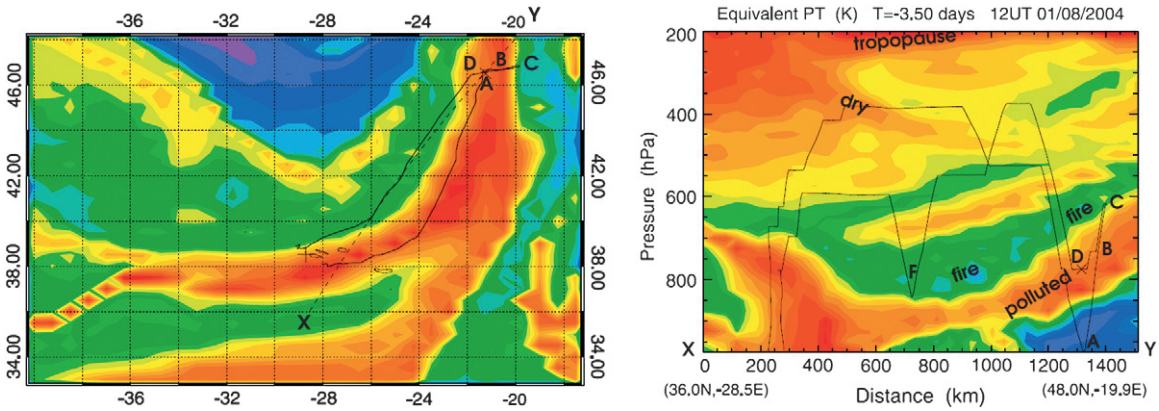

Fig. 7. Reverse Domain Filling (RDF) using FLEXPART plots used in planning the ICARTT flights over the Atlantic and the flight tracks marked in black (different colours represent different source humidities, indicating varying source types).
Taken from Methven et al. (2006).

influence on composition. Relationships between atmospheric composition and air mass origin are often analysed this way, by isolating the highest species concentrations or the exceedance levels and finding the main sectors of influence during those polluted periods. The second is to cluster the trajectories using a mathematical technique and then to analyse the concentrations at the receptor site for each trajectory classification to see whether each classification is chemically distinct. These fall broadly into geographical sector classification and statistical classification (cluster analysis) and are described in detail in the following sections.

4.1. Geographical sector classification

The method of assigning trajectories according to regional influence and comparing the composition in those regions was developed by Ashbaugh (1983). The assignment of the grid boxes is done either from prior knowledge of source regions (e.g. for western US influences in Fig. 8a (Weiss-Penzias et al., 2004) where the region with high CO from the MODIS satellite was chosen to locate an Asian box) or by assigning geographical limits (e.g. for Siberian and European air masses (Paris et al., 2010) and (Salvador et al., 2008) respectively in Fig. 8b and c) or as an objective division of the radius around the site (e.g. such as that done for Svalbard (Solberg et al., 1996) in Fig. 8d).

There are a wide variety of ways to combine the regional history and residence times of the air masses with the composition measurements and an attempt to classify the main types is shown below.

4.1.1. Residence time analysis

Residence time analysis is a qualitative source attribution technique (Ashbaugh, 1983) which generates a probability density function identifying the likelihood that an air mass will traverse a given region *en route* to the site of interest over a given time period. Air parcels that travel quickly through a pollutant source region have less time to accumulate pollutants than air parcels which remain in the source region for a long time.

Many studies look at residence times of air masses passing through polluted areas, such as measurements at the high altitude station of Arosa in Switzerland, classifying them according to the influence of European regional pollution (Pochanart et al., 2001) and the same at the Jungfraujoch observatory where the 20 year O_3 trend in the European air masses shows a large increasing trend in winter concentrations and a smaller increasing summer trend, compared to air from other areas. The *baseline* air showed an increase for the first 10 years and no trend thereafter (Cui et al., 2011). Solberg et al. (2008) calculated residence times of air masses over a central European domain arriving at many European measurement stations and linked this with the potential to form high ozone levels, such as the European heat-wave of 2003.

Back trajectories from Mount Bachelor and Cheeka Peak Observatories on the west coast of the US have been run during a variety of measurement periods to track how long each trajectory particle spends in an “East Asian box” (see Fig. 8a) in order to estimate the magnitude of pollution transport from Asia (Weiss-Penzias et al., 2004, 2006; Wolfe et al., 2007). The number of trajectories passing through the box was weighted by the average amount of time each trajectory spends in the box and the trace gas measurements at their time of arrival showed the effect of long range transport of Asian emissions to the west coast of America.

To deal with back trajectories that reside over two or more defined regions influencing the Mediterranean (from aircraft measurements), Traub et al. (2003) defined that the residence time of the air parcel above their four defined regions had to be above a critical residence time (2.75 days), by which time it was thought that the air mass had adopted the chemical characteristics of that region. Studies at the Zeppelin station in Svalbard (Solberg et al., 1996) allocated a transport sector to a trajectory if at least 50% of the last 24 h of the trajectory were closer than 850 km from the station as shown by the radius in Fig. 8d and used this allocation to understand the origin of air masses causing ozone depletion events. In a study of nitrogen in precipitation and Particulate Matter measurements in North Carolina's large agricultural corridor, Occhipinti et al. (2008) used residence time analysis of back trajectories to find air masses in which a minimum of 50% of the rainfall would have transited either marine or agricultural source regions.

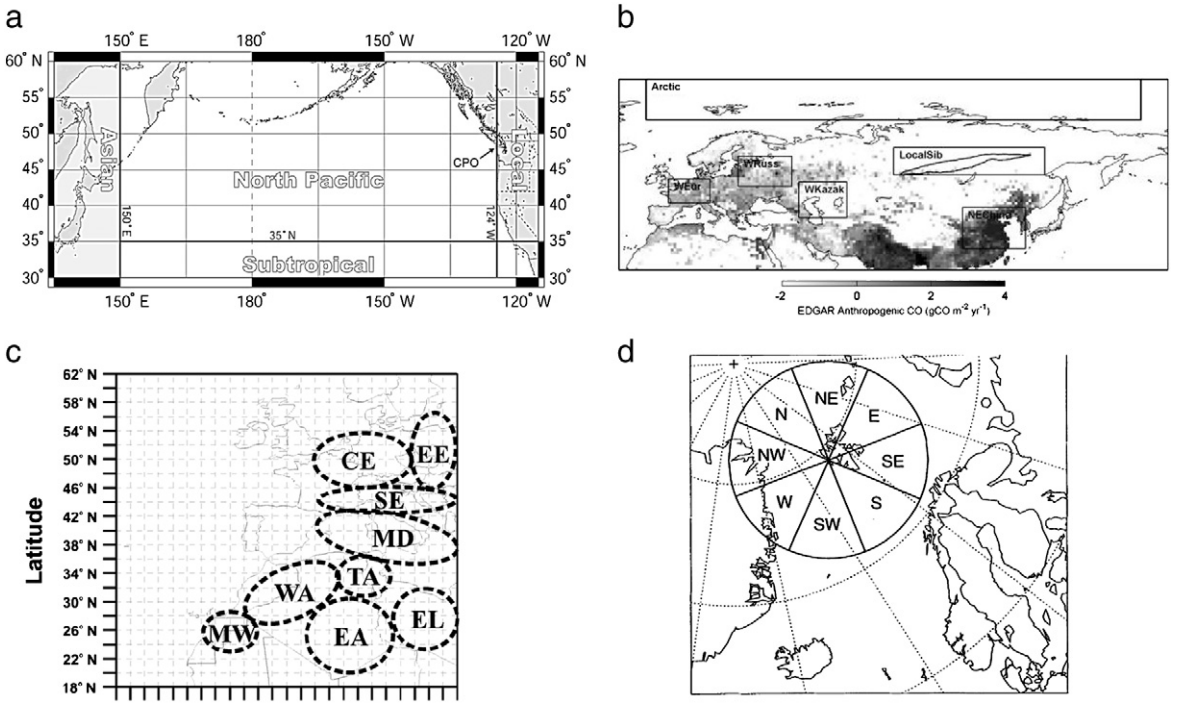

Fig. 8. Regional emission zones and defining sector influences: a) Cheeka Peak Observatory (CPO), WA regions of influence used for classifying trajectory types. b) Source regions assigned for allocating influences from dispersion modelling of Siberian influences for analysis of aircraft measurements. c) European and north African potential emission zones used for classifying European and north African PM transport to Madrid, Spain. d) Zeppelin, Svalbard transport sectors with which to allocate trajectories. Allocation to each sector is done if at least 50% of the last 24 h of the trajectory has travelled less than 850 km.
a): Taken from Weiss-Penzias et al. (2004). b): Taken from Paris et al. (2010). c): Taken from Salvador et al. (2008). d): Taken from Solberg et al. (1996).

To understand the relationship between local and longer range transport, two scales of regional divisions can be constructed. For the IMPROVE sites in various US National Parks and wilderness areas source attribution of ammonium sulphate concentrations was calculated by dividing the area around each station into four quadrants and everything outside the site's US state into one of six larger regions (Xu et al., 2006). Particulate Matter levels from Saharan and non-Saharan air masses arriving at Castanya in Spain were analysed by dividing the Saharan region into 11 source areas and the Iberian peninsula into 3 receptor regions and linking the source regions with the 3 Spanish receptor regions (Escudero et al., 2011).

An example of how residence time analysis is used for investigating long term air mass climatology (which could be used to compare long term composition change) is shown in a 40 year study of airflow trajectories and residence time calculations for the lower peninsula of Michigan (Shadbolt et al., 2006). Monthly air mass climatology anomalies were plotted as standard deviations from the grid cell mean value, where positive anomalies depicted airflow corridors that had more trajectories than average, and negative values had fewer trajectories than average, illustrating a large seasonal variation.

4.1.2. Probability fields from residence time analysis

More advanced forms of residence time analysis superimpose grid cells over specific regions of interest and calculate a probability function for each grid cell representing the probability of an air mass arriving at the receptor site, after having been observed to reside in this specific geographical region. Pioneering work by Ashbaugh et al. (1985) plotted the probability on a map, showing the influences on the air arriving at Grand Canyon National Park as shown in Fig. 9a, indicating which regions would potentially contribute to high sulphur levels and Fig. 9b shows the actual source Contribution Function (CF, described in Eq. (3)), which are often different to each other.

The seasonal variation and climatological pathway of airflow to various stations around the Atlantic Ocean were studied by Merrill (1994) in order to interpret their trace gas measurements. The residence times in each regionally gridded area for a given season were shown as a cumulative probability field and the areas of high probability indicated that trajectories which subsequently reached the site had spent more time in that geographical area. The same technique was used to help interpret composition measurements in the Pacific Exploratory Mission-West A (PEM-West A) flight experiment (Merrill, 1996) and ozone profiles above Bermuda (Merrill et al., 1996). Doddridge et al. (1994) carried out flow climatology studies to show the probability of the air over surrounding areas arriving at Mace Head and assigned a probability contour plot to represent the flow over a 3 month period over 2 years and revealed a change from predominantly Atlantic airflow to European anticyclonic Atlantic air between the years.

Two commonly used probabilities are the Potential Source Contribution Function (PSCF) (see Fig. 10) and the Incremental Probability (IP). The Incremental Probability (IP) is the *difference*

between the sorted and the everyday residence time and identifies regions that are more or less likely to be traversed during periods of high or low concentrations compared to an average day. The Potential Source Contribution Function (PSCF) is the *ratio* of the high concentration probability divided by the everyday probability and identifies locations more likely to be upwind if receptor concentrations are high, associating these upwind regions with emissions that contribute to impacts at the site or areas where secondary formation is enhanced. Scheifinger and Kaiser (2007) explain and compare PSCF and the Concentration Field (CF) and Redistribution Concentration Field (RCF) methods.

The terms Conditional Probability Function (CPF) and PSCF are used interchangeably in the literature but CPF if mostly used to define a directional source apportionment technique and can be derived from wind direction data alone, not just trajectory data. Wimolwattanapun et al. (2011) used CPF on wind direction data to define the direction of PM from a variety of source types and PSCF on trajectory data to derive maps of the areas bringing PM levels from the same source types.

The Conditional Probability that a given factor contribution from a given wind direction will exceed a predetermined threshold criterion is derived as:

$$\mathrm{CPF} = \frac{m\theta\Delta}{n\theta\Delta} \tag{1}$$

where $m\Delta\theta$ is the number of occurrences from wind sector $\Delta\theta$ where the source contributions are over a certain concentration threshold, and $n\Delta\theta$ is the total number of occurrence from this wind sector and $\Delta\theta$ is the size of the wind sector (e.g. 45°).

The PSCF is as a Conditional Probability describing the spatial distribution of probable geographical source locations inferred by using trajectories arriving at the sampling site. A trajectory endpoint lies in a single grid cell of latitude–longi-

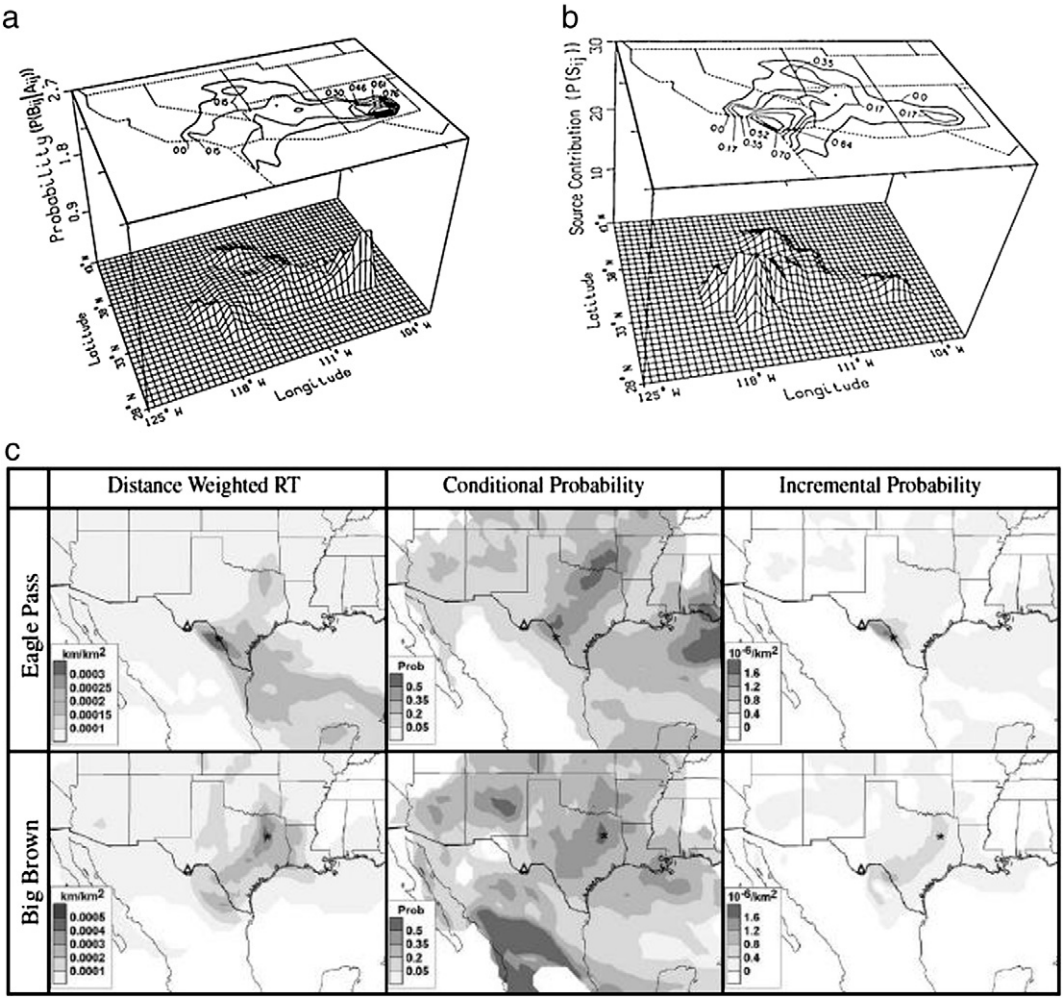

Fig. 9. Conditional Probability maps: a) Conditional Probability Function (CPF) for the highest *potential* for contributing to high sulphur levels arriving at Grand Canyon National Park (highest from New Mexico and south California). b) The source Contribution Function (CF) shows that the region that *actually* contributes most to high sulphur levels is southern California. c) Conditional Probability techniques (residence time, Conditional Probability and Incremental Probability) used for Big Bend National Park, Texas (small triangle). Tracer release from 2 sites (Eagle Pass and Big Brown) were combined with 5 day back trajectories used to study the upper 20th percentile particulate sulphur sources to the National Park.
a): Taken from Ashbaugh et al. (1985). b): Taken from Ashbaugh et al. (1985). c): Taken from Schichtel et al. (2006).

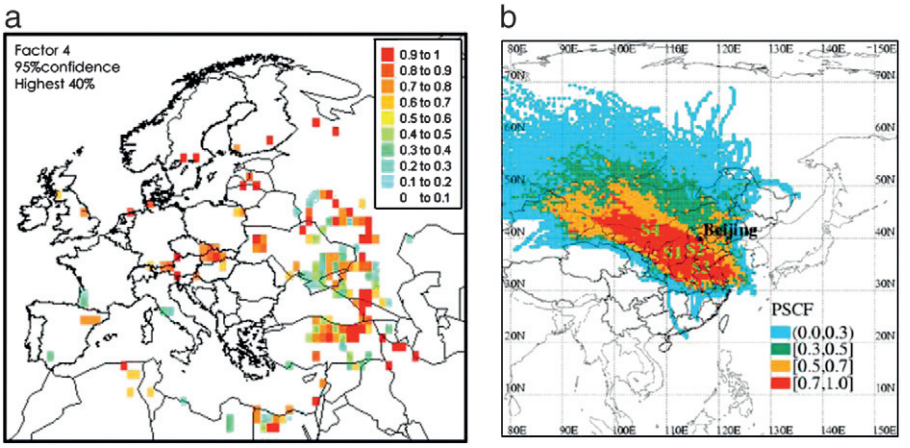

Fig. 10. Potential Source Contribution Factor (PSCF) analysis. a) PSCF map for aerosol levels at Antalya, Turkey. The highest 40% of each of 4 Factor scores from factor analysis for source apportionment of the aerosol were selected as polluted trajectories. b) PSCF map of potential sources of PM_{10} in Beijing.
a): Taken from Güllü et al. (2005). b): Taken from Zhu et al. (2011).

tude coordinates, (i,j). The PSCF value is then defined as the probability that an air parcel that passed through the ij^{th} cell had a high concentration (e.g. 90th percentile) upon arrival at the trajectory endpoint:

$$PCSF_{ij} = \frac{m_{ij}}{n_{ij}} \tag{2a}$$

n_{ij} is the total number of air masses falling into the ij^{th} cell, and m_{ij} is the number of segment trajectory endpoints in the ij^{th} cell on the days that had a source contribution greater than the criterion value.

$$\text{Or} \quad \text{PSCF}_{ij} = \frac{\text{residence time of air parcel}}{\text{Residence time of air parcel above threshold concentration}} \tag{2b}$$

Sources of particulate sulphur measurements at Big Bend National park, Texas were traced using PSCF analysis and IP to segregate the 20th and 80th percentile sulphur levels according to regions with the results from two tracer release sites as shown in Fig. 10 and the mathematical steps in the probability calculations are explained in detail in Schichtel et al. (2006). Kaiser et al. (2007) calculated a PSCF where they divided the sum of the residence times of concentrations over the 75th percentile concentration in each grid by the sum of those concentrations below the 25th percentile. Han et al. (2005) used PSCF to attribute high Mercury levels at 3 rural sites in New York state to nearby coal-fired power stations.

Poirot et al. (2001) stated that only qualitative indications of predominant transport patterns can be obtained from probability and residence time analyses as the techniques can be highly sensitive to the subjective metrics used to define high pollution episodes or the scale of the gridded domain, but agreed that the resulting maps indicating source regions can be a very powerful tool for understanding the air quality influences on a station.

Polissar et al. (1999, 2001a) used PSCF analysis to study the origin of aerosol in the Arctic and found that long-range transport of anthropogenic aerosol to the Arctic is more effective in winter and spring than in the summer. Geographical Information System (GIS) software has shown to be useful for residence time analysis, such as the TrajStat software that has been developed by Wang et al. (2009) to compute PSCF analysis with back trajectories.

It could be seen that there are trajectory statistical methods where residence time is not weighted by the concentration at the receptor point (e.g. PSCF) and methods, which do that (e.g. the Concentration Field) (Scheifinger and Kaiser, 2007). Seibert et al. (1994) dealt with uncertainties in the CPF by calculating a logarithmic mean concentration for each grid cell and a corresponding confidence interval and then smoothing the Concentration Field (CF) imposing the restriction that the values must be kept within the confidence interval. Therefore significant variations were preserved whilst the insignificant ones were removed. Various tests to check the statistical significance of attributing high CFC levels at Big Bend National park (Vasconcelos et al., 1996) have also been carried out.

The Concentration Field (CF) represents the logarithmic mean concentration for each grid cell is calculated according to

$$log\left(\bar{C}_{ij}\right) = \frac{1}{\sum_{l=1}^{M} \tau_{ijl}} + \sum_{n=l}^{M} \log(c_l)\tau_{ijl} \tag{3}$$

where i and j are the indices of the horizontal grid, l the index of the trajectory, M the total number of trajectories, c_l the concentration observed on arrival of the trajectory l and τ_{ijl} the time spent in grid cell i,j by trajectory l. A high value of C_{ij} means that, on average, air parcels passing over cell i,j result in high concentrations at the receptor site.

Potential source regions give an idea of which directions/areas pollutants are coming from but when assigning source regions by probability analysis the concentrations measured at the receptor locations are attributed equally to all segments of the related trajectory, whilst in reality emissions just take place in some segments. To account for this, an iterative scheme was developed by Stohl (1996), redistributing the measured concentrations along the trajectories according to the estimated Concentration Field from the previous iteration. The Redistribution Concentration Field method (RCF)

(Stohl, 1996) aims at extracting more information from the data, a step on from the Conditional Probability method where the concentration measured at the receptor sites is attributed with equal weight to all segments of the trajectory. Pollution sources are usually concentrated in "hot spots" so probes into concentrations at a smaller scale are needed. The concentration values along each trajectory are iteratively re-weighted according to the ratio of the concentration of that grid cell to the mean concentration of all grid cells along the path of that particular trajectory. The results are reported on maps where each grid square is assigned a weighted concentration of the component under study.

Kaiser et al. (2007) weighted the residence time of each trajectory arriving at four Alpine stations with the deviation of the actual concentration from the 3-monthly running mean so to infer potential pollutant source regions. Lin et al. (2004) used the Concentration Field to locate grids with high SO_2 emissions and describe a method from this combined with emission inventories for locating the influential pollution sources and estimating the contribution of the source within a particular geophysical region, for a specific category of emission inventory. Scheifinger and Kaiser (2007) did a validation experiment where they compared CF, RCF and PSCF methods using SO_2 measurements and emission inventories and found that on a small scale the trajectory methods work well but on a European scale these methods did not performing well.

Emission maps for NO_y have been derived over Europe from Redistributed Concentration Field (RCF) derivations (Wotawa and Kroger, 1999) and are shown in Fig. 11b and Wotawa et al. (2000) compared CPF and RCF for the 90th percentile value of O_3 for various stations in the Alps. Apadula et al. (2003) used the RCF method (as a model they call the Identification of Sources of greenhouse Gases Plus (ISOGASP) source–receptor model) to identify the sources and sinks of CO_2 to Plateau Rosa, Monte Cimone and Zugspitze stations and tested the reliability of the method by forcing sources and sinks. Salvador et al. (2010) used the RCF method to study transport pathways of Particulate Matter and aerosols to various locations in Europe and extensive examples of this method are also shown in Aalto et al.'s (2002) characterisation of the origin of CO_2, O_3, SO_2 and aerosols arriving at Pallas in Finland.

During the Rocky Mountain Atmospheric Nitrogen and Sulphur Study (RoMANS), a Concentration Field type analysis was done using the Trajectory Mass Balance (TrMB) Model (Eq. (4)) was used to probe source–receptor relationships.

$$C_{it} = \sum_{J=1}^{j} Q_{ijt} T_{ijt} N_{jt} \tag{4}$$

where the concentration, C and number of trajectory endpoints, N were known and QT (where the emission rate, Q and factor used to account for chemical transformation, T) was calculated. Hourly measured atmospheric NO_x, SO_2 and NH_3 concentrations were compared with residence times in various regions to determine correlations between measured and modelled chemistry and the transport within and into Colorado (Gebhart et al., 2011).

The mean concentration of a particular species arriving at the receptor corresponding to the various pathways and regions it passed over is a useful diagnostic. Methven et al. (2001) studied the back trajectories for air arriving at Mace Head, Ireland using the technique of "Origin Averaging" by calculating a climatological density of origin (the region where the trajectories originate) and assigning a corresponding composition concentration for each area of origin, identifying chemical air masses associated with different ozone levels. Potential source regions of dust and aerosols arriving at Tenerife in the Canary Islands were studied using Median Concentrations at Receptor (MCAR) plots which represent the median aerosol concentrations at the Izaňa station, when air masses passed above each grid box (Rodriguez et al., 2011). Mean concentration loadings from surface sources of Persistent Organic Pollutants for each grid cell within the pathway of Kosetice in the Czech Republic were calculated from the composition of the air masses that pass over each grid (Dvorska et al., 2009). Centres of gravity in defined sectors were determined to quantitatively compare mean loads in particular countries and the relative contribution of these countries to air pollution at the site.

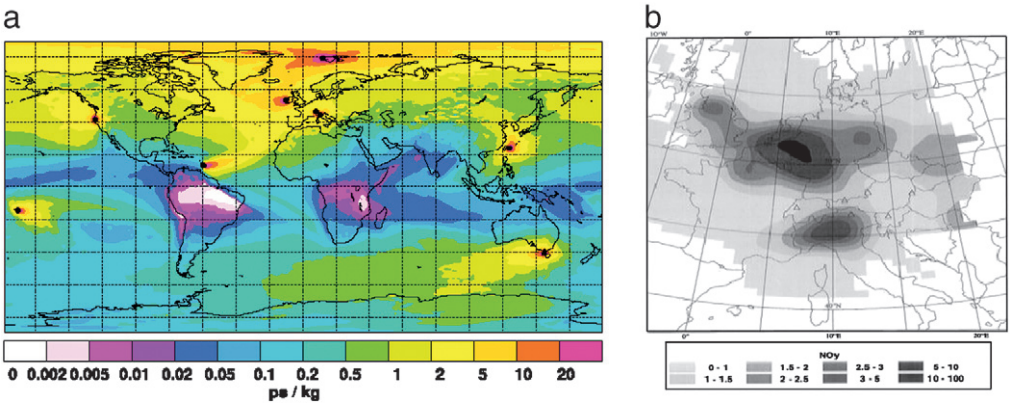

Fig. 11. Footprint emission sensitivity maps: a) footprint emission sensitivity map of hydrofluorocarbons (in picoseconds per kilogramme) obtained from FLEXPART 20 day backward calculations (January 2005–March 2007). b) NOy emission (from 0 to 100×10^{-10} kgm^{-2} s^{-1}) map over Europe derived from the Concentration Field calculated from trajectory statistics and the EMEP emission inventory.
a): Taken from Stohl et al. (2009). b): Taken from Wotawa and Kroger (1999).

4.1.3. Footprint emission sensitivity

In situ greenhouse gas measurement data from three global networks (from nine measurement sites) were combined with back trajectories to extract emission information from global observed concentration increases over a baseline (that was objectively determined by the inversion algorithm) (Stohl et al., 2009). The plot of footprint emission sensitivity for this study is shown in Fig. 11a for 20 day backward runs, showing the areas around the world that are most sampled from this station network. An NOy emission map over Europe has been derived from the Concentration Field calculated from trajectory statistics and the EMEP emission inventory (Wotawa and Kroger, 1999) and is shown in Fig. 11b.

The FLEXPART model was used to analyse transport pathways from potential flux regions towards Siberia as shown in Fig. 8b. Ten day back trajectories released along the aircraft flight track were calculated and the data were grouped according to common transport properties with cluster analysis and this was used to investigate to which extent footprints can explain the air mass chemical composition (Paris et al., 2010). The footprints (relative residence times below 300 m) were calculated using Potential Emission Sensitivity (PES), complemented by 10-day averaged relative contributions from the stratosphere to explain the source of the atmospheric composition along the flight track. Halse et al. (2011) used FLEXPART backward runs to plot the footprint emission sensitivity (the residence time of air masses per grid cell normalised by the volume) for persistent organic pollutants (POPs) from passive air samplers at 86 European background sites. This was also multiplied by emission inventories to calculate emission contributions to the European background.

Average footprint emission sensitivities were calculated around Arctic measurement stations using FLEXPART to study the origin of higher levels of aerosols, black carbon and ozone at Zeppelin, Alert, Barrow (Hirdman et al., 2010a) and the same stations and Summit (Hirdman et al., 2010b). The 10% highest and lowest measured species concentrations were selected to calculate the average emission sensitivity for that data subset. The highest and lowest 10% emission sensitivities (S_P) and the total emission sensitivity (S_T) peak near the observatory (emission sensitivities decrease with distance from the station) so this bias was removed by calculating a relative fraction, R_p:

$$R_p = L/M^* S_P/S_T \quad (5)$$

where M is the number of measured concentrations and L = M/10 highest or lowest concentrations. If the measured species were completely unrelated to air mass transport then the data subset and full dataset would look the same and the fraction would be 0.1 but if it was greater than 0.1, the cell would be a potential source (Hirdman et al., 2010b). The regions used in the footprint analysis, the average contribution of each region over a year and their annual average variation over 20 years are shown in Fig. 12 (Hirdman et al., 2010a).

FLEXPART backward runs at 5 different elevations were used to assess the surface influence on CO_2 measurements from an aircraft over Spain (Font et al., 2011). The horizontal, vertical and temporal extent of the Regional Potential Surface Influence (RPSI) residence time on atmospheric CO_2 mixing ratios was calculated and a principal component analysis

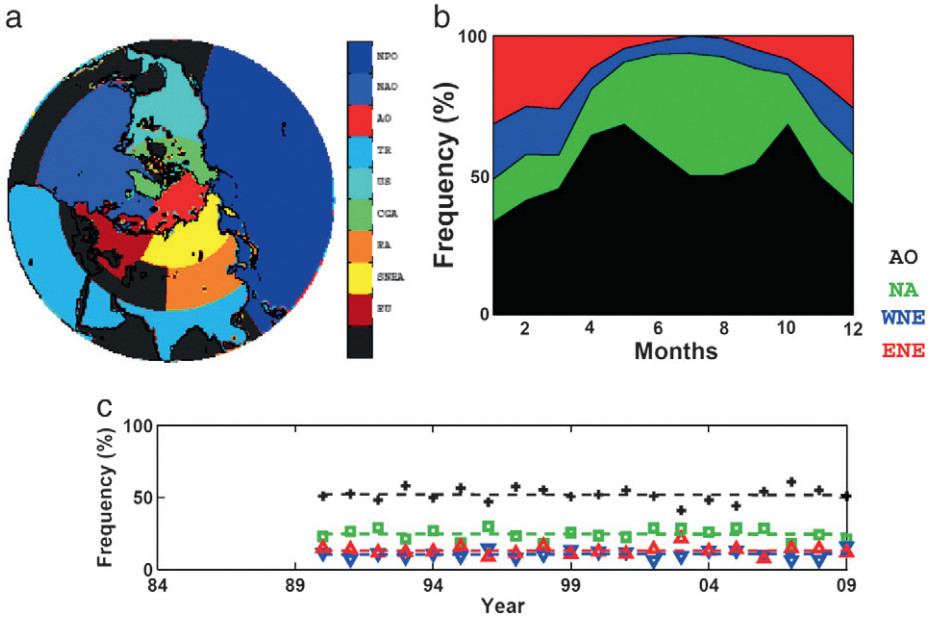

Fig. 12. Source region classifications for Zeppelin, Spitzbergen. a) Regions for clustering of the footprint emission sensitivities, b) monthly mean distribution of the trajectory types (AO: Arctic Ocean, NA: North America, WNE: Western Northern Eurasian cluster, ENE: Eastern Northern Eurasian), c) annual mean trajectory type distribution between 1990 and 2009.
Taken from (Hirdman et al. (2010a)

was carried out on the resulting residence times. Kuhn et al. (2010) looked at the transport of pollutant plumes from Russian and Alaskan forest fires by plotting the footprint emission sensitivities related to aerosol measurements at Ellesmere Island in the Canadian high Arctic and also used combined the footprint analysis with the EDGAR emissions database to isolate the sources of high aerosols.

4.1.4. Combining trajectory studies with source apportionment models

Many studies, as shown in Section 2.1 for wind measurements, combine the regional divisions from trajectory studies with the results from source apportionment models. Positive Matrix Factorisation (PMF) is a multivariate mathematical model that has been used for source–receptor modelling that aims to determine the major sources of a sampled atmospheric species. PMF analyses are used in regulatory studies to assess pollution sources and divide the data into common sources such as various industrial activities or different types of fuel burning.

Chemical Mass Balance (CMB) source apportionment studies are used when the number and nature of the sources in the regions are known and the unknown is the mass contribution of each source to each sample, which is estimated using regression. Song et al. (2008) explains and compares the use of CMB, PMF and UNMIX models for interpreting VOC measurements in Beijing, Watson et al. (2008) for PM in the US Supersites programme and Hopke et al. (2005) for PM in Washington, DC and Phoenix, AZ). In most cases there is a lack of source-specific emission information and changes in the emitted species during transport and therefore factor analysis assigns the sources by combining source contributions with source profiles in matrix multiplication. PCA pulls out a set of components that explain as much as possible of the total variance of the atmospheric species. The EPA's UNMIX model uses a computationally intensive algorithm to estimate the number of sources that can be seen above the noise level in the data and from the estimated number of sources, uses PCA to reduce the dimensionality of the data space.

In an extensive study of the transport of Particulate Matter to a station in Vermont (Poirot et al., 2001) used both PSCF and residence time (Incremental Probability) analysis to help interpret and complement the results from multivariate mathematical models (Positive Matrix Factorization and UNMIX) which had identified seven common sources with different corresponding PM levels. PMF and UNMIX are used when source profiles are not known and are a form of factor analysis that is different from the traditional Principal Component Analysis (PCA, see Section 4.2.3). Various maps showing the influence of nearly the whole of North America for different sources of PM (e.g. coal, woodsmoke, oil, soil, smelting etc.) were constructed and compared. Other examples of PMF studies that incorporate PSCF studies of actual sampled data are shown in Pekney et al. (2006) and Kocak et al. (2009) for Particulate Matter in Pittsburgh and Turkey respectively, Du and Rodenburg (2007) for PCBs in New Jersey, Dogan et al. (2008) for aerosols in Turkey, Choi et al. (2010) for VOCs in Korea. PSCF analysis of trajectories separated by the source apportionment technique of varimax-rotated factor analysis was used for aerosol measurements at Antalya, Turkey (Güllü et al., 2005), highlighting the areas that resulted in high aerosol levels at the site and the map for one factor is shown in Fig. 10a.

4.2. Cluster analysis techniques and other statistical techniques to group air mass histories

Clusters are groups with similar distributions, in the case of back trajectories, similar directions and lengths or a combination of trajectory pathways and composition. Cluster analysis provides an objective means of clustering trajectories whilst giving information about the history of the air mass and the air pollution climatology of a site, helping to determine source–receptor relationships.

Cluster analysis is a multivariate statistical technique that groups individual trajectories of an ensemble into a smaller number of clusters, meaning that the errors in the individual trajectories tend to average out. Some examples of derived mean trajectories with which to carry out composition comparisons are shown in Fig. 13 and a description of the techniques of the various techniques used in this area of research are described in the sections below.

Kassomenos et al. (2010) reviews three of the commonly used cluster analysis techniques and their dependence on arrival height, with examples of PM_{10} from trajectory clusters

Fig. 13. Cluster analysis of trajectories: a) mean trajectory pathways for Alert in January (1990–2005), 10 days backwards with frequency of occurrence in % for each cluster shown. b) Mean trajectory pathways for Zeppelin (5 day backwards) Trajectory lengths in the different clusters were 1, 3180 km; 2, 3920 km; 3, 3880 km; 4, 3200 km; 5, 4850 km; 6, 4330 km; 7, 4330 km and 8, 3880 km. c) Cluster mean trajectories arriving at the mountain site of Dibrugarhin India in the pre-monsoon season.
a): Taken from Huang et al. (2010). b): Taken from Eneroth et al. (2007). c): Taken from Gogoi et al. (2009).

around Athens. These were a hierarchical, non-hierarchical (k-means) and an artificial neural network known as Self Organising Maps (SOM). They recommended that a range of clustering techniques should be preferably used over one type.

There are two different types of clustering algorithms, namely hierarchical and non-hierarchical clustering.

4.2.1. Non-hierarchical clustering methods

Non-hierarchical clustering, sometimes known as partitional clustering attempts to directly decompose the data set into a set of disjoint clusters by minimising the measure of dissimilarity in the trajectories within each cluster, whilst maximising the dissimilarity of different clusters.

The k-means procedure is the most commonly used form of non-hierarchical clustering for trajectory climatology studies. The k-means procedure is an iterative algorithm that uses a specified number of clusters, k, to partition the data by comparing each object to the arithmetic mean of all the members of each of the k clusters (cluster centres). The selection of the optimal number of clusters that best describes the different air flow patterns is performed by computing the percentage change in within-cluster variance, as a function of the number of clusters (Dorling et al., 1992). The assignment of members (trajectories) to a given group (cluster) is carried out by minimising the internal variability within the group of trajectories and maximising the external variability between different groups based on the trajectory co-ordinates. It uses the Root Mean Square Deviation (RMSD) of all individual clusters from their cluster mean trajectory against the number of clusters retained until a "break" is reached, indicating that two clusters have been merged which are unacceptably different. Alternatively, if a threshold percentage change in RMSD is exceeded at any particular point in the clustering process, this is also taken as an indication that an optimum number of clusters have been reached. The k-means clustering method is often quoted as the Dorling method in climatological clustering research and is well suited for large databases because of its relatively small computational requirements (see for example Dorling and Davies, 1995). Non-hierarchical clustering requires that the number of clusters is already known and that the objects are distributed between those. This algorithm is widely used in cases where a priori information on the nature of the measurements is available.

Examples of studies using the k-means algorithm include clustering 15 years of Particulate Matter concentrations at Alert (Huang et al., 2010) in Fig. 13a, monthly average ozone and mercury to the Arctic (Eneroth et al., 2007) where trajectory clusters for Zeppelin and their average length over 5 days is shown in Fig. 13b and comparing composition between the 6 main clusters for Alert and Barrow (Sharma et al., 2006).

Various methods for calculating the distance between the clusters with which to define the final clusters have been used in the literature, with RMSD and Euclidean distance used most frequently. The Mahalanobis distance metric clusters back trajectories by gathering the extreme trajectory positions belonging to a cluster and then enclosing and creating the smallest convex hull with minimum volume covering the backward trajectories of the clusters (Makra et al., 2010).

Fuzzy mean c-clustering is a technique that is very similar to k-means but each trajectory has a degree of belonging to several clusters, as in fuzzy logic, rather than belonging completely to one cluster. Fuzzy c-means uses an iterative algorithm to determine the grade of membership of each trajectory in each cluster, with 0 being no membership and 1 indicating full membership and in between is partial membership. Each trajectory was assigned to a single cluster for which it has the largest value. Xia et al. (2007) used this technique to derive seasonal clusters of trajectories for air arriving in Beijing, China, which managed to separate aerosol into physically distinct groups, explaining 47% of the variance. However, this method could not separate fast from slowly moving trajectories.

Self Organising Maps (SOM) is considered an advanced approach of clustering (a type of Artificial Neural Network) that can produce reliable segregation even in difficult cases (Karaca and Camci, 2010). They operate similarly to k-means, but instead of using a number of clusters they utilise a grid of nodes with predetermined shape and size. This grid iteratively adjusts to the data until it maps as close as possible their structure in space. The obtained nodes (or clusters) are also organised in a 2-D grid so that similar clusters are placed near each other. In that way, clustering is performed following a structured approach, in contrast with the unstructured k-means approach.

Two-stage clustering can further refine the clustering analysis. Borge et al. (2007) re-analysed the short trajectories from 3 European cities with unclear directionality that were derived from an initial clustering that to create a further discrimination between them. Davis et al. (2010) also used this 2-stage technique on a series of stations in Virginia, USA using the distances between the horizontal and vertical trajectory endpoints and the station. Polluted air masses over Athens (Markou and Kassomenos, 2010) were studied by applying a second clustering method (using the Haversine formula, great-circle distance between two points) to separate the clusters already obtained by k-means clustering based on the length of their cluster-mean trajectories, allowing them to distinguish between short slow moving and long fast moving trajectories.

4.2.2. Hierarchical clustering methods

The purpose of the hierarchical clustering is to join objects into successively larger clusters, using some measure of similarity or distance, by constructing clusters within clusters. All the classified objects are considered at each step of the hierarchical clustering and the process is determined by the construction of an agglomeration tree. This approach is usually used when the number of clusters is unknown. Kalkstein et al. (1987) has compared three hierarchical clustering procedures (Ward's, average-linkage and centroid) for climatological studies of back trajectories and showed that the average-linkage method was the most appropriate.

Hierarchical clustering partitions data following a series of steps either by grouping or by separating the objects one by one in each step. The two closest clusters are merged in each step, starting the procedure with singleton clusters and ending with a single cluster that contains all the objects. There are a number of different techniques to measure the distance (or the similarity) between the clusters, which may lead to different subsets.

Ozone measurements at Mace Head (Cape et al., 2000) for 3 years were divided into four 3-month periods and the derived trajectory clusters were used to classify the ozone (these are shown in Fig. 3b as examples of linear trajectories).

The squared distances (kilometres north and east and elevation from ground level (as pressure)) for each arrival time were calculated between pairs of trajectories, one trajectory from each cluster. Starting with each trajectory as a cluster, all possible pairs are evaluated and the two clusters with the smallest average distance between their members were joined. The procedure is designed to minimise within cluster variance and maximise between cluster variance. Fig. 14 from Cape et al. (2000) shows how the R^2, Route Mean Squared (RMS) distance and number of clusters changes as the number of clusters by iterations increases, a method commonly used to calculate the optimum number of clusters. Robinson et al. (2011) used a similar cluster analysis method to cluster the trajectories arriving at the Bukit Atur site in Borneo during 2 month long periods in 2008 and found corresponding aerosol, halocarbon levels in each cluster and this was backed up with residence time analysis of the source regions, showing different signatures in terrestrial and marine sectors.

Aerosol Optical Depth measurements in Northern India were analysed using Cluster Spatial Variance (SPVAR) and their frequency of occurrence is shown in Fig. 13c (Gogoi et al., 2009). SPVAR is the sum of the squared distances between the endpoints of the cluster's component trajectories and the mean of the trajectories in that cluster. A combination of trajectory pairs was used to calculate the SPVAR and the 4 derived clusters. The Total Spatial Variance (TSV), the sum of all the SPVAR, is calculated and the pairs of clusters were combined (with the lowest increase in TSV (which is initially zero)). At each iteration, one more trajectory is joined to a cluster. The iterations are continued until the last two clusters are combined. The iterative step just before the large increase in the change of TSV gives the final number of clusters. Trajectories arriving in Hong Kong have been clustered using this method for analysis with CO and O_3 measurements (Wang et al., 2004) as well as trajectories arriving into Lamas d'Olo in Portugal (Carvalho et al., 2010).

Ward's method is a type of hierarchical cluster analysis that uses the sum of squares of the distance of each trajectory from the cluster's mean trajectory and has been used for classifying trajectory types at various Atlantic Ocean sites (Moody et al., 1989) and to study 10 years of tracer levels at Svalbard (Eneroth et al., 2003). The distances between each 5 day trajectory at every hourly time step along the trajectories were calculated and the spatial variance between two trajectories was quantified as the sum of all squared distances. The smaller the distances, the more similar were the trajectories and they were grouped together until the spatial variance increases rapidly.

4.2.3. Principal component analysis

For the purposes of data dimension reduction in large datasets, Principal Component Analysis (PCA) has been used to group trajectories. Riccio et al. (2007) examined the role exerted by meteorology on air quality through the classification of atmospheric circulation patterns as a function of air mass origin for 10 years of back trajectories arriving into Naples. They used 116,896 trajectories, embedded in a 144 dimensional space. It was found that the first eight components, i.e. the reduction of the (116896×144) data matrix to a (116896×8) matrix, explained almost the total (>98%) portion of initial variance, without sacrificing accuracy and without significantly affecting the classification procedure, but with a large speed-up in computation.

4.2.4. Significance tests between air mass types and composition

Once clusters have been determined, the average composition seen at the receptor site corresponding to these trajectories needs to be investigated. There are a number of ways of statistically testing whether the clusters have distinct composition and are different from one cluster to the other.

For nonparametric data, the Kruskal–Wallis test (Miller, 1981) has been used in many studies e.g. Salvador et al. (2008) and Sharma et al. (2006, 2004) for non-normally distributed data to identify whether the median species concentrations were different between all the air mass sectors. If the Kruskal–Wallis test leads to the rejection of the null hypothesis and, thus, to the conclusion that not all samples are identical, it is appropriate to use a multiple comparison procedure to find out which clusters were different from the others. For example, the Dunn test was used as a multiple comparison procedure after the Kruskal–Wallis test in a study of ozone levels in few stations in Northeast USA (Brankov et al., 1998). The Spearman rank-order correlation coefficients were used in Han et al. (2005) and the Kendall τ and Pear-

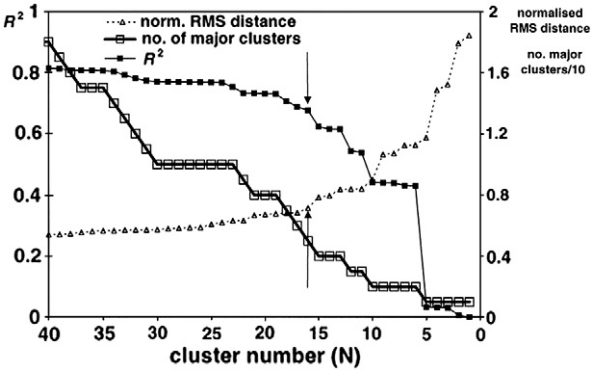

Fig. 14. Changes in the normalised RMS distance between clusters, and the total variance (R^2) as a function of the number of clusters for the daily back trajectories at Mace Head 1995 to 1997. Step changes in these statistics represent logical points for defining an optimum number of clusters to retain in the analysis, in this case, 5 major trajectory clusters.
Taken from Cape et al. (2000).

son's correlation coefficient were used in Paris et al. (2010) to test for significant differences in the chemistry between the clusters or between trajectory and dispersion model clusters as in Han et al. (2005).

For normally-distributed data Analysis of Variance (ANOVA) has been used extensively to detect whether the mean composition is significantly different between the clusters. Borge et al. (2007) used an analysis of variance analysis on 24 hour mean NOx and PM_{10} to test if the cluster averaged concentrations for each pollutant were statistically significant. Occhipinti et al. (2008) used ANOVA statistics to study the influence of agricultural areas on PM and nitrogen deposition and Makra et al. (2006) and Xia et al. (2007) used ANOVA analysis to test for differences in composition between various air mass clusters in the Carpathian basin and Beijing China respectively.

It is useful to display the composition in each cluster with a box and whisker plot, with the mean, percentiles and confidence limits. The confidence limits for the average composition of each trajectory cluster are often very large and overlap between the clusters but when other source apportionment techniques are carried out alongside this, such as time-lagged correlation (Brankov et al., 1998), the difference in the average composition levels is shown to be distinct. Moy et al. (1994) carried out the student *t*-test on O_3, CO and NO_y data for trajectory clusters arriving in Virginia to find statistically significant differences between the mean composition levels in several of the clusters.

4.2.5. Cluster analysis on dispersion models

Cluster analysis has been carried out on dispersion models on a number of occasions. One such study used a k-means cluster technique based on FLEXPART dispersion model footprints where regions were chosen prior to clustering analysis (regions of specific sources or sinks relevant to the site) (Paris et al., 2010) as shown in Fig. 8c. Stohl et al. (2002) used Mace Head FLEXPART runs to compare the accuracy of classical trajectory techniques against dispersion models by comparing representative single trajectories and trajectory clusters (retroplumes). Particles were released in each dispersion run and their position data was used for deriving a condensed model output. Cluster analysis was used as a semi-objective method applied to best characterise the position and shape of the entire retroplume, by calculating the retroplume centroid, followed by on-line cluster analyses of the particle positions at selectable time intervals that minimises the root-mean-square distance between the particles of each of the clusters and their respective retroplume cluster centroids, and maximises the distance between the cluster centroids. As the clustering is performed independently each

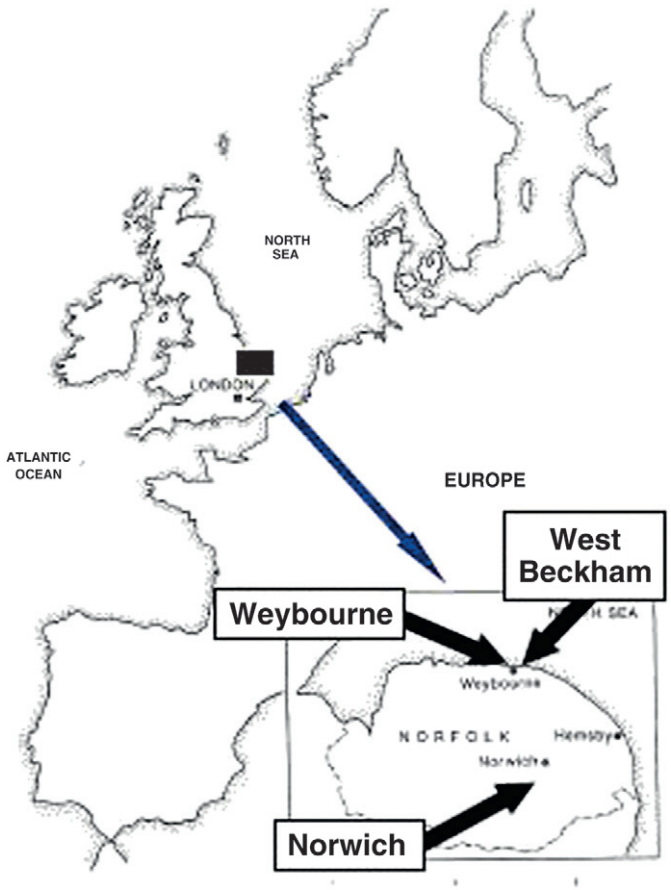

Fig. 15. The Weybourne Atmospheric Observatory.

time, subsequent retroplume cluster centroids do not lie on a trajectory and thus cannot be connected by a line in a trajectory plot.

PSCF multi-receptor (MURA) residence time probability analysis has been used to simulate sulphur concentrations in the South West US (Lee and Ashbaugh, 2007a) on HYSPLIT 4 backwards runs with puffs of particles rather than individual particles. Back trajectory Conditional Probability analysis uses a single receptor at a time (see Section 4.1.3) whereas the MURA method uses several receptors at once in order to detect sources with greater accuracy. The MURA method designates potential source regions by counting trajectories for each grid cell and then examines them to see how often each region affects each receptor. Lee and Ashbaugh (2007b) developed the single receptor forward Conditional Probability (SIRA) method, which is a Conditional Probability method with the second step of the MURA method added to it. A high CP indicates a higher probability that that location contains a source or is on the pathway to the source. To calculate the SIRA back trajectories are divided into 1 hour segments and a grid is superimposed on the area of interest and the number of trajectory segment points located in each grid cell was counted for both sample days and high incident days. Lee and Ashbaugh (2007c) looked at the impact of running trajectories at various elevations on the MURA method and found that it was best to run ensembles of trajectories in the MURA method so as to average out most of the biases found from different trajectory starting heights.

5. Case study: using the NAME model for classification of air mass types and corresponding composition variations at a site

In order to illustrate the use of dispersion models to untangle the regional influences of an atmospheric observatory, a step by step description of a new methodology that has been developed using the NAME model output is detailed for an observatory on the UK North Sea coast as the case study.

The Weybourne Atmospheric Observatory (52°57′N 1°07′E) on the North Sea coast (Fig. 15) is at a strategic location for receiving a variety of Atlantic, Arctic, European, UK and North Sea air masses. In previous work, air masses have been divided according to wind direction or manually classified trajectories (Cardenas et al., 1998; Penkett et al., 2007) but because of the rapidly changing wind directions and the variety of influences close by, a more detailed technique is required.

In this case study the NAME model was run in backwards mode, 10 days backwards in time at 3 hourly intervals (related to the timescale of the Unified Model meteorological fields) for 4 years of Weybourne station data (2006–2009). The particles were released from the height of the station's tower where the instruments sample (10 m). All instances when the particles were near to the ground (0–100 m) were recorded to indicate when surface emissions from different geographical regions (marine or land) will have been picked up by the air mass and transported to the observation station. The horizontal spatial resolution was 0.25° × 0.25° for the 10 day regional domain used.

Fig. 16 shows how single run outputs can be combined to produce integrated plots (monthly and annual) that illustrate the seasonality of the site footprints. Monthly averaged footprint plots for 2008 for Weybourne show seasonal changes in the air masses histories. The 12 monthly Weybourne footprints show how there is a subtle seasonal pattern of more Arctic air in the spring and summer months and a wider range of the footprint in winter.

For the station of interest, the domain of influence of the NAME run is split subjectively into the main geographical areas that have differing source characteristics, especially differentiating between land and marine sectors. This geographical sector map for Weybourne is shown in Fig. 17.

The NAME output represents the 10,000 inert tracer particles released during each 3 hourly period and where they are likely to have travelled on their way to the site and the output is an integration of the number of particles per grid cell over the 10 day period and represents a probability that the air passed over that area near to the ground, similar to the emission sensitivities in FLEXPART (Hirdman et al., 2010b). The particle distribution output from NAME is used to extract the information about how many particles have passed over each sector during the 10 days of travel for each 3 hour period by counting the total number of particles that pass over each grid box in each geographical area. The distribution of air particles passing over each sector can be shown as a percentage of the total number of particles in the domain from each 3 hour release period (Fig. 18). The first plot in Fig. 18 shows 4 years of distributions (2006–2009) and the second shows a smaller time period where the frequency and size of the variations of influence are clearer and can be used to visually isolate events and changes in air mass sector influences. Nearly all trajectories pass over multiple geographical areas. The monthly averaged regional distribution of air masses arriving at Weybourne illustrates the seasonality of the Arctic and North Sea air masses that are more frequent in spring as shown in Fig. 19a. Fig. 19b shows a similar plot for the Cape Verde observatory, showing how this method can pick up the seasonality of the regions that influence a site (the synoptic climatology), with much greater Saharan influence in winter (Carpenter et al., in press). These are similar to the seasonal distribution of trajectory types in the studies shown in Figs. 6 and 12.

To account for the fact that each trajectory in its ten day passage to the station will have passed over more than one geographical area, various permutations of combinations of these regions have been fitted into convenient classifications. Seven subjective air-mass classifications were defined for Weybourne and are shown in Fig. 20. They are denoted as Arctic only, Arctic and Europe, Atlantic, European, Local (UK and North Sea), Scandinavian and Greenland, America (no European).

The time integrated particle concentration or dosage (gs/m^3) in each sector shows a wide variation, with sporadic peaks. Various tests were carried out to test for the most statistically robust method for assigning a threshold for the amount of particles in a sector that would make that sector contribute significant influence on the air mass arriving at the station. The final chosen method was to select a threshold of 10% of the maximum dosage for each sector. Obviously, being closer to the station, there were more particles in the UK sector than

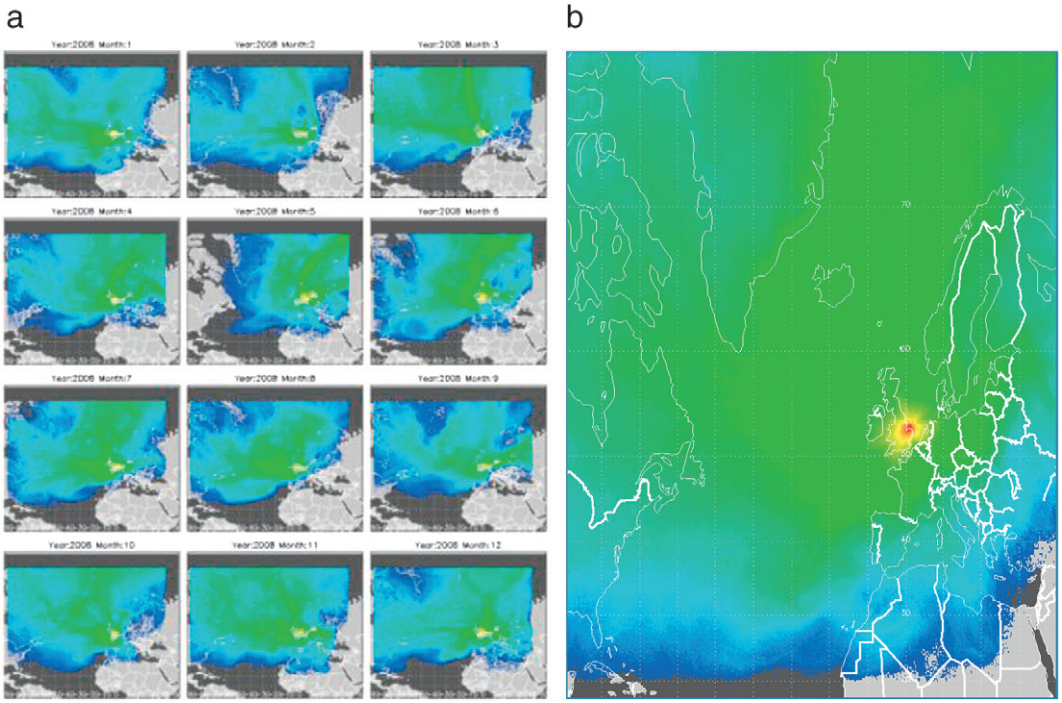

Fig. 16. Weybourne 10 day 2008 integrated footprints: a) monthly b) annual.

the American so that made the concentration threshold for assigning the American sector as significant more sensitive.

Thresholds for each sector were derived for assigning each trajectory the sectors it passed through. These thresholds were chosen by plotting dosage (in gs/m^3) against % time in that region and finding the % thresholds for a given sector at a dosage of 0.001 gs/m^3 (for the UK and North Sea the threshold dosage was chosen to be higher at 0.002 and for the American sector a lower threshold dosage of 0.0003 was chosen). These thresholds came to: Arctic (10%), Scandinavia (10%), Europe (10%), Atlantic (13%), Greenland and Iceland (5%), America (3%), UK (24%) and North Sea (17%). This form of classification flags up the more distant sectors like America as being labelled as important when the same % in other sectors would not have been significant. The local UK sector was only flagged as particularly important when the trajectory spent over 24% of the time over the UK.

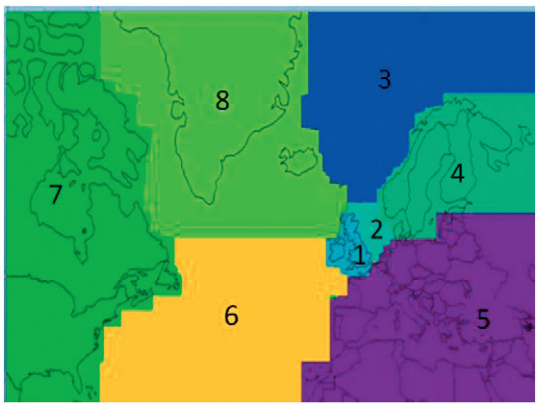

Fig. 17. Regional grid divisions for Weybourne for 10 day domain.

Permutations of the seven sectors were combined to derive a trajectory classification (as shown in Fig. 20) for the entire time series, in this case as three hourly time series. The final result is a time series with flags assigned to trajectory types numbered 1 to 7.

The station composition data time series was averaged into 3 hourly intervals so as to be comparable with the 3 hourly air mass trajectory type time series from the NAME-based classification. Fig. 21 shows how the chemistry can vary when the regions that the air masses pass over before reaching the station change, showing how ozone and NO_2 increase during the period of a high European and UK influence (e.g. 25th–28th July and 30th July–-1st August). The average values for each of the air mass trajectory type were calculated as well as the standard deviation and mean for winter (DJF), summer (JJA), spring (MAM) and autumn (SON) data for January 2006-September 2009. Fig. 22 shows the three year seasonally averaged composition in each air mass type. SO_2 was highest in the local sector and lowest in Arctic air masses and O_3 was highest in European, local and Scandinavian air masses in summer but highest in American and Arctic air masses in winter.

The methodology of using NAME to classify composition time series can be analysed on various timescales other than as shown in Fig. 22 for the 4 year average of each air mass type. It would be of interest to calculate yearly averages so as to follow annual trends in composition for each air mass type. Alternatively, this technique is useful for focussing in on particular periods of interest to

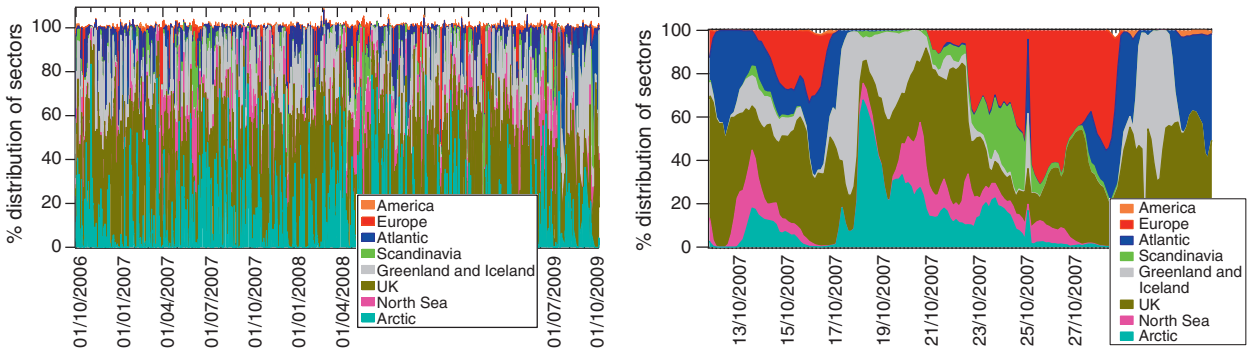

Fig. 18. Division of Weybourne sector influences for each 3 hourly period (January 2006–September 2009) and zoomed in section to show the small scale changes in regional influences.

understand sudden changes in composition or to isolate the origin of pollution spikes. This described methodology illustrates the use of dispersion models for building up a database of station footprints which can be used in a number of ways to extract the regional influences on average levels of a variety of atmospheric species at any timescale from hours to years. Trends in meteorological and synoptic scale influences on a site of interest can also be tracked with this method, to track seasonal variations in air mass origin and assess long term variations. Comparing the use of the NAME dispersion model to other techniques described in this review, it is most similar to the regional assignment of trajectories as passing through particular geographical sectors discussed in Section 4.1 and Fig. 8 (especially the FLEXPART dispersion model in Fig. 8b). However, owing to the spread of influence in the NAME dispersion model from turbulence, the distribution over multiple regions is easier to track than with a single line trajectory. Using the NAME model in this way is very similar to the FLEXPART regional assignment and relative residence times calculated by Halse et al. (2011) and particularly Hirdman et al. (2010a) (Fig. 12) and Hirdman et al. (2010b).

6. Conclusions

Interpreting air mass history and the role of transport remains an important tool for interpreting observed atmospheric composition which, owing to meteorology, is influenced by a variety of local and long range transport processes. Local wind direction and speed have been mostly replaced by a number of computational techniques (e.g. trajectory and dispersion models) that quantify the far field influence. The increase in the accuracy of trajectory and dispersion models is related to improvements in the resolution of available meteorological fields, which leads to a better resolution of the atmospheric physics and the ability to interpret the movement, mixing and transport of atmospheric constituents.

This review paper has detailed the evolution of methodologies to interpret atmospheric composition measurements according to air mass history. In situ wind measurements or the use of agglomerated meteorological fields in trajectory or dispersion models can be used to assess the influence of changing air masses history on composition with varying degrees of

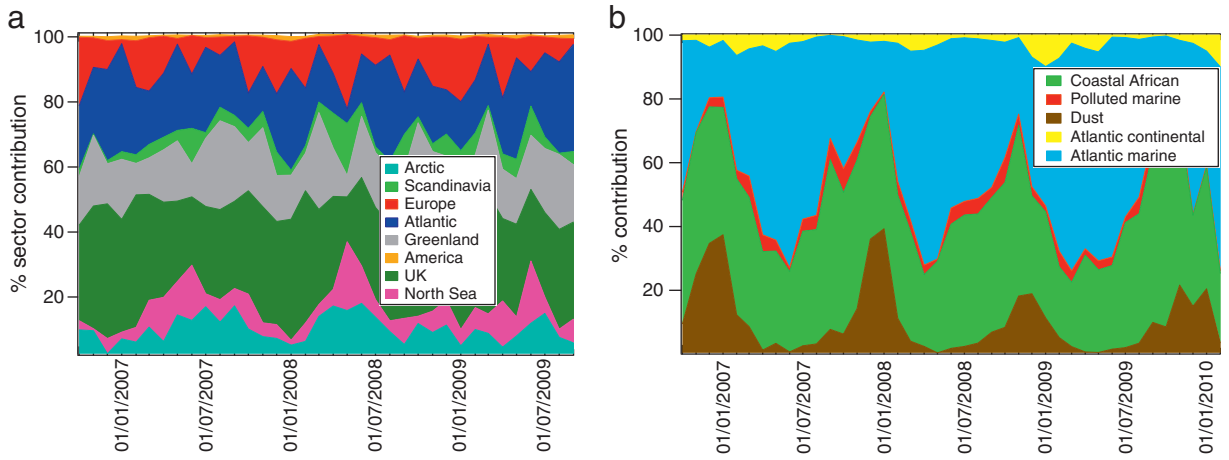

Fig. 19. Monthly averaged regional distribution at a) Weybourne and b) Cape Verde Observatory (see Carpenter et al., in press) for detailed analysis for Cape Verde).

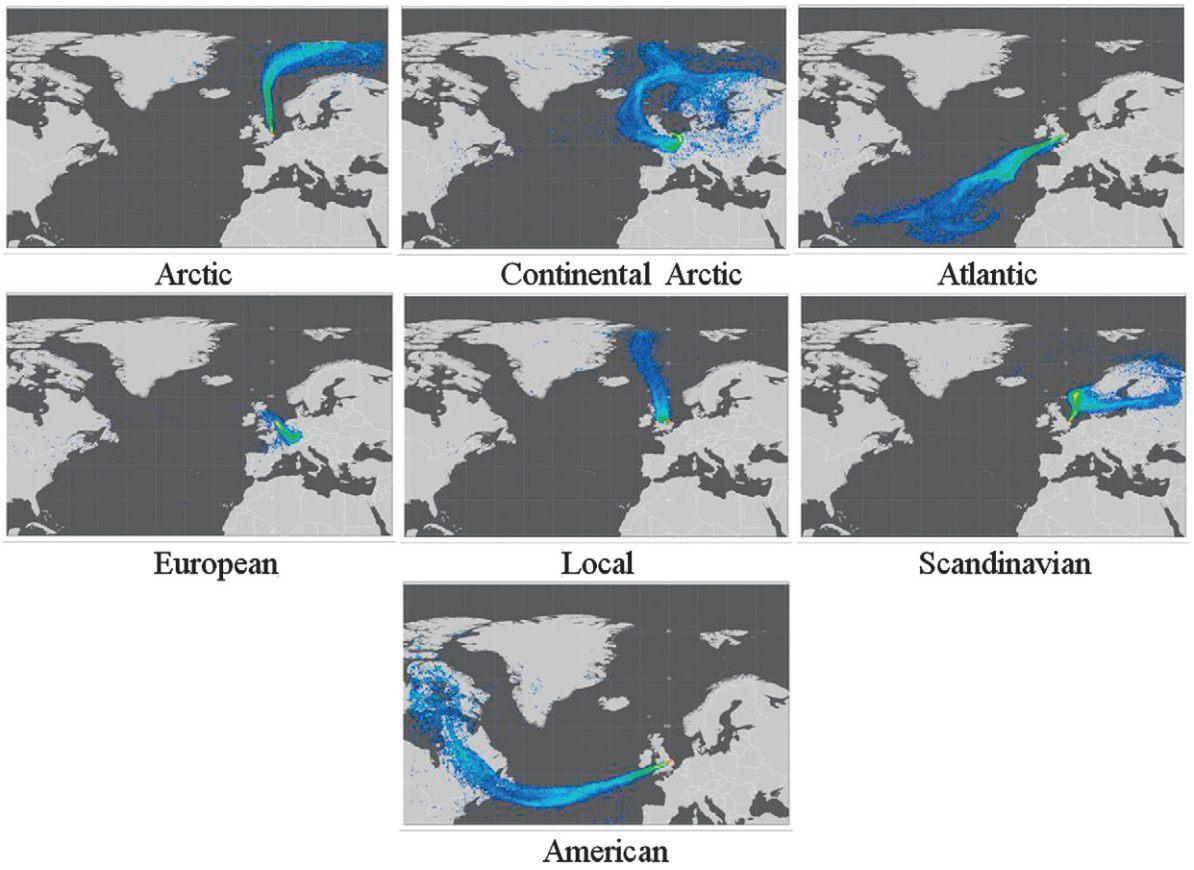

Fig. 20. Examples of the 7 Weybourne air mass classifications (3 hourly trajectories).

success. The various uses of these wind measurements and trajectory and dispersion models for the interpretation of ground based and aircraft data have been illustrated with examples.

The historical use of local wind direction or daily meteorological patterns is still of use for obtaining a short term picture of the variation of air masses but is not accurate enough to detect

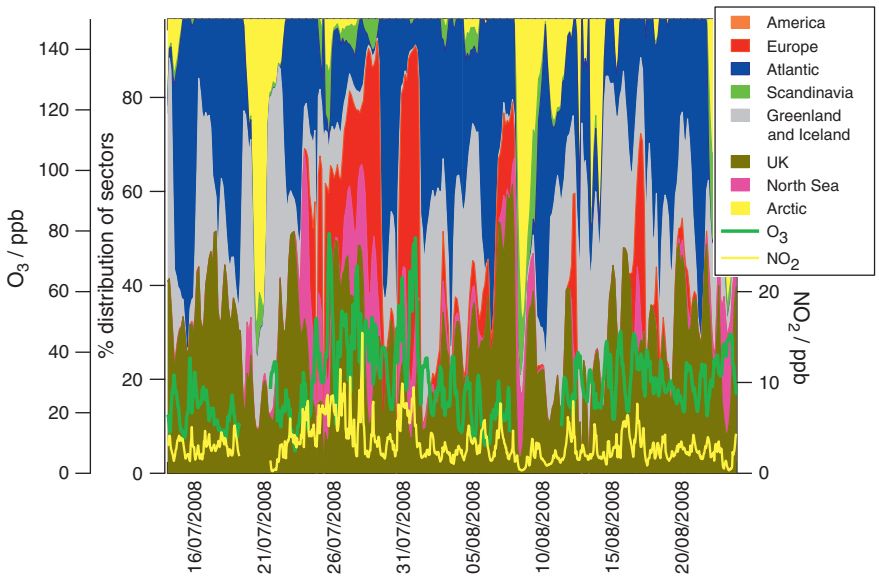

Fig. 21. Weybourne composition variations with regional distribution of air mass history. As European and UK regional influence increases, ozone increases as well as NO_2.

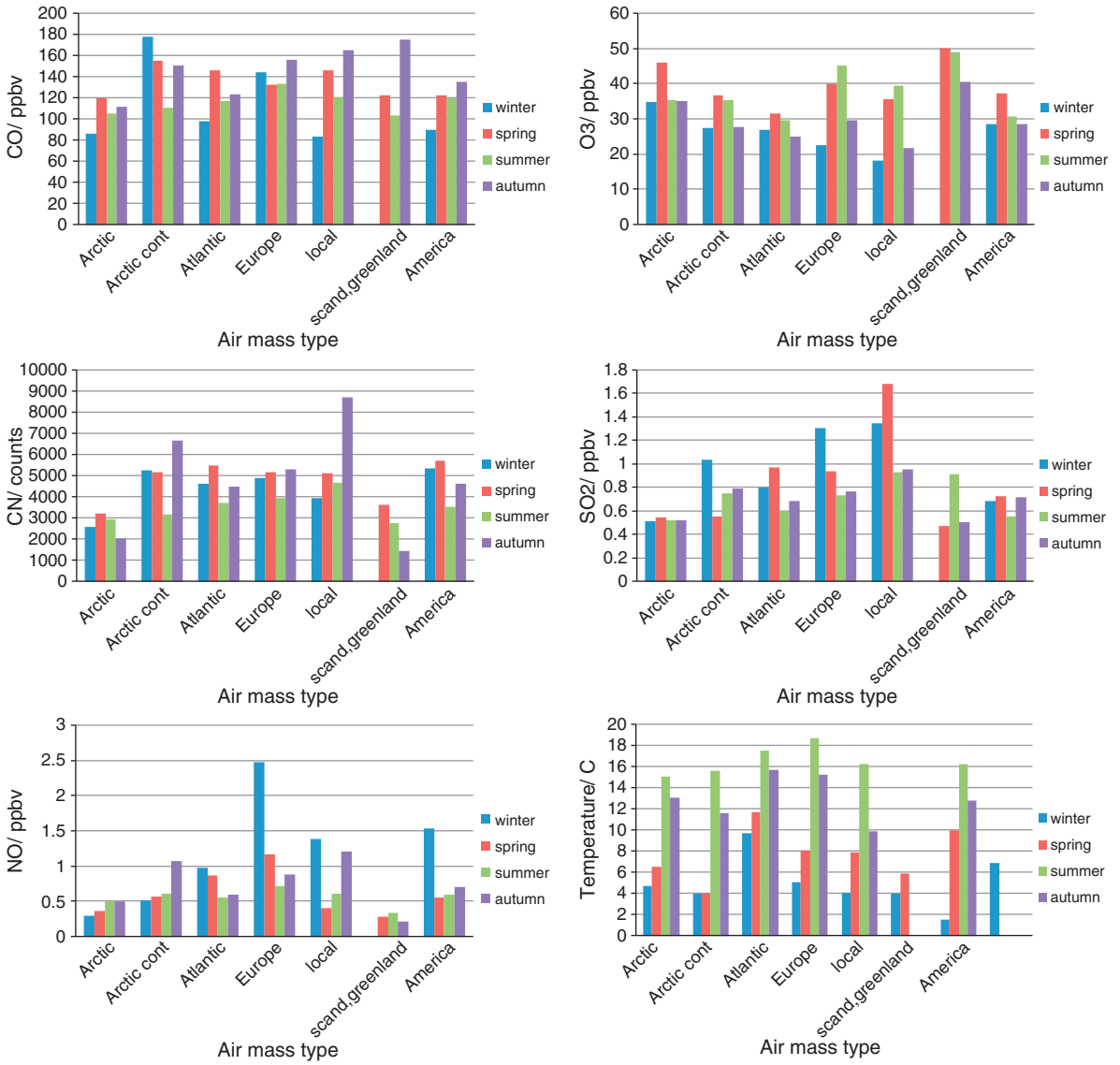

Fig. 22. Composition (CO, O_3, CN, SO_2, NO) and temperature distribution in different trajectory types 4 year average 2006–2009 winter (DJF), spring (MAM), summer (JJA) and autumn (SON).

the synoptic scale transport captured in trajectory and dispersion models. The pros and cons of trajectory and dispersion models have been explained and it is evident that the output of both types of models and indeed also of wind direction measurements is being combined with increasingly powerful computational techniques to interpret the data. Trajectory analysis remains a powerful method for tracking air mass pathways and can often be run quickly using freely available software within minutes, but it does not take into account the turbulence in the atmosphere, an important feature of the atmosphere that is incorporated into dispersion models. Dispersion models show the spread of particle movement backwards from a point, illustrating the effect of turbulence on scattering air mass particles over a large area. Dispersion models take into account the probability of air masses passing over discrete areas, by showing the varying concentrations in each area, something that is much more useful than the linear point to point air mass position of back trajectories.

Residence time analysis links trajectories and dispersion models with geographical sectors that have been specifically delineated (e.g. the European continent, the Asian high emission zone, marine areas, the Sahara or country boundaries or sectors within a fixed radius of the site). The residence time techniques that relate measured concentrations to sectors have been explained with examples and include Concentration Field (CF) and Redistributed Concentration Field (RCF) and Reverse Domain Filling (RDF). Conditional Probability Function (CPF) and Potential Source Contribution Function (PSCF) relate the ratio of high concentrations (or exceedance values) to average concentrations and Incremental Probability (IP) relates the difference between high residence times and average ones during high concentration episodes. The

technique aims to pinpoint the areas that influence the site when particular atmospheric species concentrations are high, as do Footprint Emission Sensitivity maps which use trajectories or dispersion models to extract regional or global emission information from observed concentration increases over a baseline. Many of these studies were complemented with further source apportionment studies (e.g. Positive Matrix Factorisation (PMF), Chemical Mass Balance (CMB) and Principal Component Analysis (PCA)) that extract the source of the emissions to more than just a region but to a particular type of natural or anthropogenic emissions.

Another method for classifying trajectories into groups according to similar characteristics is by using cluster analysis. The various hierarchical and non-hierarchical techniques used for clustering regional and synoptic influences on a site are explained as well as the statistical methods for comparing composition measurements between the various cluster types.

A step by step example of the procedures used to interpret composition measurements with corresponding backward trajectories or dispersion models is detailed by demonstrating a new methodology (using the NAME atmospheric dispersion model) for assessing the effect of air mass origin on the atmospheric composition at a long term measurement station. The case study for four years for the Weybourne Atmospheric Observatory illustrates how atmospheric composition measurements can be exploited and how this is an especially effective technique for stations with a variety of composition influences (e.g. a combination of marine and continental influences or clean and polluted sectors). The NAME model has been used to calculate a multi-year time series of trajectory types and sectors of influence. This methodology can provide a long term view of the type of air masses and conditions that affect each site and the inter-annual and intra-annual variability of the air mass types reaching the site as well as the composition of those air masses.

With the vast amount of studies that have been carried out since the 1980s in this area of combining climatology studies with composition measurements, there have still not been many studies that have looked at more than 2 or 3 years of measurements, not enough to discern whether composition changes are due to long term synoptic changes (changing regional influences) or changes in the composition within each region. Most studies segregate air masses by regional influence, but few look at the time lag between leaving that region and reaching the site and combine that with species lifetimes. With the exception of aircraft experiments or for high altitude measurement stations studies, there have been few studies to segregate the vertical influences as well as the surface regional influences.

Understanding and being unable to unpick the history of air-masses in the atmosphere remains an important tool for assessing influence of emissions and change in the atmosphere.

Acknowledgements

We would like to thank the authors of the work referenced herein for allowing reproduction of their various figures. We would like to thank the UK Met Office for supplying the Unified Model Meteorological data and the use of the NAME model. We would also like to thank University of Leicester's High Performance Computing services for supplying the necessary computing power for running, plotting and storing the model output and Roland Leigh at Leicester University for developing the code to plot and interpret the NAME model output on the global grid scale. We would like to thank Brian Bandy, William Sturges and Stuart Penkett and many past and present scientists at UEA for providing the Weybourne dataset and the BADC for hosting it, as well as NCAS and UEA for funding the observatory. I would like to thank the reviewer for providing invaluable insights and advice for improving this paper. Finally, we would like to thank the National Centre for Atmospheric Science (NCAS) and NERC for funding.

References

Aalto, T., Hatakka, J., Paatero, J., Tuovinen, J.P., Aurela, M., Laurila, T., Holmen, K., Trivett, N., Viisanen, Y., 2002. Tropospheric carbon dioxide concentrations at a northern boreal site in Finland: basic variations and source areas. Tellus Ser. B: Chem. Phys. Meteorol. 54, 110–126.

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.

Alleman, L.Y., Lamaison, L., Perdrix, E., Robache, A., Galloo, J.C., 2010. PM_{10} metal concentrations and source identification using positive matrix factorization and wind sectoring in a French industrial zone. Atmos. Res. 96, 612–625.

Altshuller, A.P., 1986. Relationships between direction of wind flow and ozone inglow concentrations at rural locations outside of St Louis, MO. Atmos. Environ. 20, 2175–2184.

Amodio, M., Bruno, P., Caselli, M., de Gennaro, G., Dambruoso, P.R., Daresta, B.E., Ielpo, P., Gungolo, F., Placentino, C.M., Paolillo, V., Tutino, M., 2008. Chemical characterization of fine particulate matter during peak PM_{10} episodes in Apulia (South Italy). Atmos. Res. 90, 313–325.

Apadula, F., Gotti, A., Pigini, A., Longhetto, A., Rocchetti, F., Cassardo, C., Ferrarese, S., Forza, R., 2003. Localization of source and sink regions of carbon dioxide through the method of the synoptic air trajectory statistics. Atmos. Environ. 37, 3757–3770.

Ashbaugh, L.L., 1983. A statistical trajectory technique for determining air-pollution source regions. J. Air Pollut. Control Assoc. 33, 1096–1098.

Ashbaugh, L.L., Malm, W.C., Sadeh, W.Z., 1985. A residence time probability analysis of sulphur concentrations at Grand-Canyon National Park. Atmos. Environ. 19, 1263–1270.

Baker, J., 2010. A cluster analysis of long range air transport pathways and associated pollutant concentrations within the UK. Atmos. Environ. 44.

Baur, F., Hess, P., Nagel, H., 1944. Kalendar der Grosswetterlagen Europas 1881–1939. DWD, Bad Homburg.

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.

Begum, B.A., Kim, E., Jeong, C.H., Lee, D.W., Hopke, P.K., 2005. Evaluation of the potential source contribution function using the 2002 Quebec forest fire episode. Atmos. Environ. 39, 3719–3724.

Bhugwant, C., Bessafi, M., Riviere, E., Leveau, J., 2001. Diurnal and seasonal variation of carbonaceous aerosols at a remote MBL site of La Reunion Island. Atmos. Res. 57, 105–121.

Biegalski, S.R., Hopke, P.K., 2004. Total potential source contribution function analysis of trace elements determined in aerosol samples collected near Lake Huron. Environ. Sci. Technol. 38, 4276–4284.

Blake, N.J., Penkett, S.A., Clemitshaw, K.C., Anwyl, P., Lightman, P., Marsh, A.R.W., Butcher, G., 1993. Estimates of atmospheric hydroxyl radical concentrations from the observed decay of many reactive hydrocarbons in well-defined urban plumes. J. Geophys. Res. 98, 2851–2864.

Borge, R., Lumbreras, J., Vardoulakis, S., Kassomenos, P., Rodriguez, E., 2007. Analysis of long-range transport influences on urban PM_{10} using two-stage atmospheric trajectory clusters. Atmos. Environ. 41, 4434–4450.

Brankov, E., Rao, S.T., Porter, P.S., 1998. A trajectory-clustering-correlation methodology for examining the long-range transport of air pollutants. Atmos. Environ. 32, 1525–1534.

Brenninkmeijer, C.A.M., Crutzen, P., Boumard, F., Dauer, T., Dix, B., Ebinghaus, R., Filippi, D., Fischer, H., Franke, H., Friess, U., Heintzenberg, J., Helleis, F., Hermann, M., Kock, H.H., Koeppel, C., Lelieveld, J., Leuenberger, M., Martinsson, B.G., Miemczyk, S., Moret, H.P., Nguyen, H.N., Nyfeler, P., Oram, D., O'Sullivan, D., Penkett, S., Platt, U., Pupek, M., Ramonet, M., Randa, B., Reichelt, M., Rhee, T.S., Rohwer, J., Rosenfeld, K., Scharffe, D., Schlager, H., Schumann, U., Slemr, F., Sprung, D., Stock, P., Thaler, R., Valentino, F., van Velthoven, P., Waibel, A., Wandel, A., Waschitschek, K., Wiedensohler, A., Xueref-Remy, I., Zahn, A., Zech, U., Ziereis, H., 2007. Civil aircraft for the regular investigation of the atmosphere based on an instrumented container: the new CARIBIC system. Atmos. Chem. Phys. 7, 4953–4976.

Burley, J.D., Bytnerowicz, A., 2011. Surface ozone in the White Mountains of California. Atmos. Environ. 45, 4591–4602.
Burley, J.D., Ray, J.D., 2007. Surface ozone in Yosemite National Park. Atmos. Environ. 41, 6048–6062.
Cabello, M., Orza, J.A.G., Galiano, V., Ruiz, G., 2008. Influence of meteorological input data on backtrajectory cluster analysis — a seven-year study for southeastern Spain. Adv. Sci. Res. 2, 65–70.
Cammas, J.P., Brioude, J., Chaboureau, J.P., Duron, J., Mari, C., Mascart, P., Nedelec, P., Smit, H., Patz, H.W., Volz-Thomas, A., Stohl, A., Fromm, M., 2009. Injection in the lower stratosphere of biomass fire emissions followed by long-range transport: a MOZAIC case study. Atmos. Chem. Phys. 9, 5829–5846.
Cape, J.N., Methven, J., Hudson, L.E., 2000. The use of trajectory cluster analysis to interpret trace gas measurements at Mace Head, Ireland. Atmos. Environ. 34, 3651–3663.
Cardenas, L.M., Austin, J.F., Burgess, R.A., Clemitshaw, K.C., Dorling, S., Penkett, S.A., Harrison, R.M., 1998. Correlations between CO, NOy, O-3, and non-methane hydrocarbons and their relationships with meteorology during winter 1993 on the North Norfolk Coast, UK. Atmos. Environ. 32, 3339–3351.
Carpenter, L.J. Fleming, Z.L. Read, K.A. Lee, J.D. Moller, S.J. Hopkins, J. Purvis, R. Lewis, A.C. Müller, K. Heinold, B. Herrmann, H. Wadinga Fomba, K. van Pinxteren, D. Müller, C. Tegen, I. Wiedensohler, A. Müller, T. Niedermeier, N. Achterberg, E.P. Patey, M.D. Kozlova, E.A. Heimann, M. Heard, D.E. Plane, J.M.C. Mahajan, A.S. Oetjen, H. Vaughan, S. Ingham, T. Arnold, S.R., Stone, D. Whalley, L. Evans, M. Pilling, M.J. Leigh, R.J. Monks, P.S. Karunaharan, A. Tschritter, J. Pöhler, D. Frieß, U. Holla, R. Mendes, L. Lopez, H. Faria, B. ManningA.J. and Wallace, D.W.R. in press. Seasonal characteristics of tropical marine boundary layer air measured at the Cape Verde Atmospheric Observatory. J. Atmos. Chem.
Carvalho, A., Monteiro, A., Ribeiro, I., Tchepel, O., Miranda, A.I., Borrego, C., Saavedra, S., Souto, J.A., Casares, J.J., 2010. High ozone levels in the northeast of Portugal: Analysis and characterization. Atmos. Environ. 44, 1020–1031.
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.
Chang, R.Y.W., Leck, C., Graus, M., Müller, M., Paatero, J., Burkhart, J.F., Stohl, A., Orr, L.H., Hayden, K., Li, S.M., Hansel, A., Tjernström, M., Leaitch, W.R., Abbatt, J.P.D., 2011. Aerosol composition and sources in the Central Arctic Ocean during ASCOS. Atmos. Chem. Phys. Discuss. 11, 14837–14881.
Cheng, M.-D., Lin, C.-J., 2001. Receptor modeling for smoke of 1998 biomass burning in Central America. J. Geophys. Res. 106, 22871–22886.
Choi, E., Heo, J.B., Yi, S.M., 2010. Apportioning and locating nonmethane hydrocarbon sources to a background site in Korea. Environ. Sci. Technol. 44, 5849–5854.
Cohen, M., Artz, R., Draxler, R., Miller, P., Poissant, L., Niemi, D., Ratte, D., Deslauriers, M., Duval, R., Laurin, R., Slotnick, J., Nettesheim, T., McDonald, J., 2004. Modeling the atmospheric transport and deposition of mercury to the Great Lakes. Environ. Res. 95, 247–265.
Cohen, D.D., Crawford, J., Stelcer, E., Bac, V.T., 2010. Long range transport of fine particle windblown soils and coal fired power station emissions into Hanoi between 2001 to 2008. Atmos. Environ. 44, 3761–3769.
Colette, A., Ancellet, G., Borchi, F., 2005. Impact of vertical transport processes on the tropospheric ozone layering above Europe. Part I: study of air mass origin using multivariate analysis, clustering and trajectories. Atmos. Environ. 39, 5409–5422.
Cosemans, G., Kretzschmar, J., Mensink, C., 2008. Pollutant roses for daily averaged ambient air pollutant concentrations. Atmos. Environ. 42, 6982–6991.
Crawford, J., Charnbers, S., David, D.C., Dyer, L., Wang, T., Zahorowski, W., 2007. Receptor modelling using positive matrix factorisation, back trajectories and radon-222. Atmos. Environ. 41, 6823–6837.
Crawford, J., Zahorowski, W., Cohen, D.D., 2009. A new metric space incorporating radon-222 for generation of back trajectory clusters in atmospheric pollution studies. Atmos. Environ. 43, 371–381.
Crutcher, H.L., Rhodes, R.C., Graves, M.E., Fairbairn, B., Nelson, A.C., 1986. Application of cluster-analysis to aerometric data. J. Air Pollut. Control Assoc. 36, 1116–1122.
Cui, J., Pandey Deolal, S., Sprenger, M., Henne, S., Staehelin, J., Steinbacher, M., Nédélec, P., 2011. Free tropospheric ozone changes over Europe as observed at Jungfraujoch (1990–2008): an analysis based on backward trajectories. J. Geophys. Res. 116, D10304.
Cullen, M.J.P., 1993. The unified forecast/climate model. Meteorol. Mag. 1449, 81–94.
Danielsen, E.F., Bleck, R., 1967. Research in four-dimensional diagnosis of cyclonic storm cloud system. AFCRL Rep. 67–0617, Bedford, Mass., 1967 (Available as NTIS AD 670 847, from Natl. Tech. Inf. Serv., Springfield, Va.).
Darby, L.S., 2005. Cluster analysis of surface winds in Houston, Texas, and the impact of wind patterns on ozone. J. Appl. Meteorol. 44, 1788–1806.
Davis, J.M., Eder, B.K., Nychka, D., Yang, Q., 1998. Modeling the effects of meteorology on ozone in Houston using cluster analysis and generalized additive models. Atmos. Environ. 32, 2505–2520.
Davis, R.E., Normile, C.P., Sitka, L., Hondula, D.M., Knight, D.B., Gawtry, S.P., Stenger, P.J., 2010. A comparison of trajectory and air mass approaches to examine ozone variability. Atmos. Environ. 44, 64–74.
de Foy, B., Fast, J.D., Paech, S.J., Phillips, D., Walters, J.T., Coulter, R.L., Martin, T.J., Pekour, M.S., Shaw, W.J., Kastendeuch, P.P., Marley, N.A., Retama, A., Molina, L.T., 2008. Basin-scale wind transport during the MILAGRO field campaign and comparison to climatology using cluster analysis. Atmos. Chem. Phys. 8, 1209–1224.
de Foy, B., Zavala, M., Bei, N., Molina, L.T., 2009. Evaluation of WRF mesoscale simulations and particle trajectory analysis for the MILAGRO field campaign. Atmos. Chem. Phys. 9, 4419–4438.
Delcloo, A.W., De Backer, H., 2008. Five day 3D back trajectory clusters and trends analysis of the Uccle ozone sounding time series in the lower troposphere (1969–2001). Atmos. Environ. 42, 4419–4432.
Demuzere, M., van Lipzig, N.P.M., 2010. A new method to estimate air-quality levels using a synoptic-regression approach. Part I: present-day O_3 and PM_{10} analysis. Atmos. Environ. 44, 1341–1355.
Derwent, R.G., Witham, C.S., Utembe, S.R., Jenkin, M.E., Passant, N.R., 2010. Ozone in Central England: the impact of 20 years of precursor emission controls in Europe. Environ. Sci. Policy 13, 195–204.
Doddridge, B.G., Dirmeyer, P.A., Merrill, J.T., Oltmans, S.J., Dickerson, R.R., 1994. Interannual variability over the Eastern North-Atlantic Ocean — chemical and meteorological evidence for tropical influence on regional-scale transport in the extratropics. J. Geophys. Res.-Atmos. 99, 22923–22935.
Dogan, G., Gullu, G., Tuncel, G., 2008. Sources and source regions effecting the aerosol composition of the Eastern Mediterranean. Microchem. J. 88, 142–149.
Dorling, S.R., Davies, T.D., 1995. Extending cluster-analysis- Synoptic meteorology links to characterize chemical climates at 6 northwest European monitoring stations. Atmos. Environ. 29, 145–167.
Dorling, S.R., Davies, T.D., Pierce, C.E., 1992. Cluster-analysis — a technique for estimating the synoptic meteorological controls on air and precipitation chemistry — method and applications. Atmos. Environ. 26, 2575–2581.
Draxler, R.R., Rolph, G.D., 2003. HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) Model. NOAA Air Resources Laboratory, Silver Spring, MD. Model access via NOAA ARL READY Website http://ready.arl.noaa.gov/HYSPLIT.php.
Draxler, R.R., Taylor, A.D., 1982. Horizontal parameters for long-range transport modelling. J. Appl. Meteorol. 21, 367–372.
Droppo, J.G., Napier, B.A., 2008. Wind direction bias in generating wind roses and conducting sector-based air dispersion modeling. J. Air Waste Manage. Assoc. 58, 913–918.
Du, S.Y., Rodenburg, L.A., 2007. Source identification of atmospheric PCBs in Philadelphia/Camden using positive matrix factorization followed by the potential source contribution function. Atmos. Environ. 41, 8596–8608.
Dueñas, C., Orza, J.A.G., Cabello, M., Fernández, M.C., Cañete, S., Pérez, M., Gordo, E., 2011. Air mass origin and its influence on radionuclide activities (^{7}Be and ^{210}Pb) in aerosol particles at a coastal site in the western Mediterranean. Atmos. Res. 101, 205–214.
Dvorska, A., Lammel, G., Holoubek, I., 2009. Recent trends of persistent organic pollutants in air in central Europe — air monitoring in combination with air mass trajectory statistics as a tool to study the effectivity of regional chemical policy. Atmos. Environ. 43, 1280–1287.
Ebinghaus, R., Jennings, S.G., Kock, H.H., Derwent, R.G., Manning, A.J., Spain, T.G., 2011. Decreasing trends in total gaseous mercury observations in baseline air at Mace Head, Ireland from 1996 to 2009. Atmos. Environ. 45, 3475–3480.
Eder, B.K., Davis, J.M., Bloomfield, P., 1994. An automated classification scheme designed to better elucidate the dependence of ozone on meteorology. J. Appl. Meteorol. 33, 1182–1199.
Eneroth, K., Kjellstrom, E., Holmen, K., 2003. A trajectory climatology for Svalbard; investigating how atmospheric flow patterns influence observed tracer concentrations. Phys. Chem. Earth. 28, 1191–1203.
Eneroth, K., Holmen, K., Berg, T., Schmidbauer, N., Solberg, S., 2007. Springtime depletion of tropospheric ozone, gaseous elemental mercury and non-methane hydrocarbons in the European Arctic, and its relation to atmospheric transport. Atmos. Environ. 41, 8511–8526.
Escudero, M., Stein, A.F., Draxler, R.R., Querol, X., Alastuey, A., Castillo, S., Avila, A., 2011. Source apportionment for African dust outbreaks over the Western Mediterranean using the HYSPLIT model. Atmos. Res. 99, 518–527.
Evans, M.J., Shallcross, D.E., Law, K.S., Wild, J.O.F., Simmonds, P.G., Spain, T.G., Berrisford, P., Methven, J., Lewis, A.C., McQuaid, J.B., Pilling, M.J., Bandy, B.J., Penkett, S.A., Pyle, J.A., 2000. Evaluation of a Lagrangian

box model using field measurements from EASE (Eastern Atlantic Summer Experiment) 1996. Atmos. Environ. 34, 3843–3863.
Fenneteaux, I., Colin, P., Etienne, A., Boudries, H., Dutot, A.L., Perros, P.E., Toupance, G., 1999. Influence of continental sources on oceanic air composition at the eastern edge of the North Atlantic Ocean, TOR 1992–1995. J. Atmos. Chem. 32, 233–280.
Fiore, A.M., Dentener, F.J., Wild, O., Cuvelier, C., Schultz, M.G., Hess, P., Textor, C., Schulz, M., Doherty, R.M., Horowitz, L.W., MacKenzie, I.A., Sanderson, M.G., Shindell, D.T., Stevenson, D.S., Szopa, S., Van Dingenen, R., Zeng, G., Atherton, C., Bergmann, D., Bey, I., Carmichael, G., Collins, W.J., Duncan, B.N., Faluvegi, G., Folberth, G., Gauss, M., Gong, S., Hauglustaine, D., Holloway, T., Isaksen, I.S.A., Jacob, D.J., Jonson, J.E., Kaminski, J.W., Keating, T.J., Lupu, A., Marmer, E., Montanaro, V., Park, R.J., Pitari, G., Pringle, K.J., Pyle, J.A., Schroeder, S., Vivanco, M.G., Wind, P., Wojcik, G., Wu, S., Zuber, A., 2009. Multimodel estimates of intercontinental source–receptor relationships for ozone pollution. J. Geophys. Res.-Atmos. 114.
Flesch, T.K., Wilson, J.D., Yee, E., 1995. Backward-time Lagrangian stochastic dispersion models and their application to estimate gaseous emissions. J. Appl. Meteorol. 34, 1320–1332.
Font, A., Morgui, J.A., Rodo, X., 2011. Assessing the regional surface influence through backward lagrangian dispersion models for aircraft CO_2 vertical profiles observations in NE Spain. Atmos. Chem. Phys. 11, 1659–1670.
Forster, C., Wandinger, U., Wotawa, G., James, P., Mattis, I., Althausen, D., Simmonds, P., O'Doherty, S., Jennings, S.G., Kleefeld, C., Schneider, J., Trickl, T., Kreipl, S., Jager, H., Stohl, A., 2001. Transport of boreal forest fire emissions from Canada to Europe. J. Geophys. Res.-Atmos. 106, 22887–22906.
Fox, T.D., Ludwick, J.D., 1976. Lead (Pb) concentrations associated with 1000-MB geostrophic back trajectories at Quillayute, Washington. Atmos. Environ. 10, 799–803.
Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron) — a review of applications in the atmospheric sciences. Atmos. Environ. 32, 2627–2636.
Gardner, M.W., Dorling, S.R., 1999. Neural network modelling and prediction of hourly NO_x and NO_2 concentrations in urban air in London. Atmos. Environ. 33, 709–719.
Gardner, M.W., Dorling, S.R., 2000. Meteorologically adjusted trends in UK daily maximum surface ozone concentrations. Atmos. Environ. 34, 171–176.
Gardner, M., Dorling, S., 2001. Artificial neural network-derived trends in daily maximum surface ozone concentrations. J. Air Waste Manage. Assoc. 51, 1202–1210.
Gebhart, K.A., Schichtel, B.A., Barna, M.G., 2005. Directional biases in back trajectories caused by model and input data. J. Air Waste Manage. Assoc. 55, 1649–1662.
Gebhart, K.A., Schichtel, B.A., Malm, W.C., Barna, M.G., Rodriguez, M.A., Collett Jr., J.L., 2011. Back-trajectory-based source apportionment of airborne sulfur and nitrogen concentrations at Rocky Mountain National Park, Colorado, USA. Atmos. Environ. 45, 621–633.
Gerstengabe, F.-W., Werner, P., Rüge, U., 1999. Katalog der Grosswetterlagen Europas 1881–1998 nach P. Hess und H. Brezowsky, 5. Aufl. – Potsdam-Inst. F. Klimafolgenforschung, Potsdam, Germany. 138 pp.
Gheusi, F., Ravetta, F., Delbarre, H., Tsamalis, C., Chevalier-Rosso, A., Leroy, C., Augustin, P., Delmas, R., Ancellet, G., Athier, G., Bouchou, P., Campistron, B., Cousin, J.M., Fourmentin, M., Meyerfeld, Y., 2011. Pic 2005, a field campaign to investigate low-tropospheric ozone variability in the Pyrenees. Atmos. Res. 101, 640–665.
Gilliam, R.C., Hogrefe, C., Rao, S.T., 2006. New methods for evaluating meteorological models used in air quality applications. Atmos. Environ. 40.
Gogoi, M.M., Moorthy, K.K., Babu, S.S., Bhuyan, P.K., 2009. Climatology of columnar aerosol properties and the influence of synoptic conditions: first-time results from the northeastern region of India. J. Geophys. Res.-Atmos. 114.
Gregory, G.L., Bachmeier, A.S., Blake, D.R., Heikes, B.G., Thornton, D.C., Bandy, A.R., Bradshaw, J.D., Kondo, Y., 1996. Chemical signatures of aged Pacific marine air: mixed layer and free troposphere as measured during PEM-West A. J. Geophys. Res.-Atmos. 101, 1727–1742.
Gros, V., Williams, J., Lawrence, M.G., von Kuhlmann, R., van Aardenne, J., Atlas, E., Chuck, A., Edwards, D.P., Stroud, V., Krol, M., 2004. Tracing the origin and ages of interlaced atmospheric pollution events over the tropical Atlantic Ocean with in situ measurements, satellites, trajectories, emission inventories, and global models. J. Geophys. Res. 109.
Güllü, G., Dogan, G., Tuncel, G., 2005. Atmospheric trace element and major ion concentrations over the eastern Mediterranean Sea: Identification of anthropogenic source regions. Atmos. Environ. 39, 6376–6387.
Hains, J.C., Taubman, B.F., Thompson, A.M., Stehr, J.W., Marufu, L.T., Doddridge, B.G., Dickerson, R.R., 2008. Origins of chemical pollution derived from Mid-Atlantic aircraft profiles using a clustering technique. Atmos. Environ. 42, 1727–1741.
Halse, A.K., Schlabach, M., Eckhardt, S., Sweetman, A., Jones, K.C., Breivik, K., 2011. Spatial variability of POPs in European background air. Atmos. Chem. Phys. 11, 1549–1564.
Han, Y.J., Holsen, T.M., Hopke, P.K., Yi, S.M., 2005. Comparison between back-trajectory based modeling and Lagrangian backward dispersion modeling for locating sources of reactive gaseous mercury. Environ. Sci. Technol. 39, 1715–1723.
Han, Y.M., Cao, J.J., Jin, Z.D., An, Z.S., 2009. Elemental composition of aerosols in Daihai, a rural area in the front boundary of the summer Asian monsoon. Atmos. Res. 92, 229–235.
Harrison, R.M., Grenfell, J.L., Peak, J.D., Clemitshaw, K.C., Penkett, S.A., Cape, J.N., McFadyen, G.G., 2000. Influence of airmass back trajectory upon nitrogen compound composition. Atmos. Environ. 34, 1519–1527.
Harrison, R.M., Yin, J., Tilling, R.M., Cai, X., Seakins, P.W., Hopkins, J.R., Lansley, D.L., Lewis, A.C., Hunter, M.C., Heard, D.E., Carpenter, L.J., Creasy, D.J., Lee, J.D., Pilling, M.J., Carslaw, N., Emmerson, K.M., Redington, A., Derwent, R.G., Ryall, D., Mills, G., Penkett, S.A., 2006. Measurement and modelling of air pollution and atmospheric chemistry in the UK West Midlands conurbation: overview of the PUMA Consortium project. Sci. Total. Environ. 360, 5–25.
Hart, M., De Dear, R., Hyde, R., 2006. A synoptic climatology of tropospheric ozone episodes in Sydney, Australia. Int. J. Climatol. 26, 1635–1649.
Helmis, C.G., Moussiopoulos, N., Flocas, H.A., Sahm, P., Assimakopoulos, V.D., Naneris, C., Maheras, P., 2003. Estimation of transboundary air pollution on the basis of synoptic-scale weather types. Int. J. Climatol. 23, 405–416.
A.J. Hewitt, 2010. Investigating land-air carbon fluxes using a Lagrangian model and satellite retrieved carbon dioxide, Thesis, University of Leicester.
Hirdman, D., Aspmo, K., Burkhart, J.F., Eckhardt, S., Sodemann, H., Stohl, A., 2009. Transport of mercury in the Arctic atmosphere: evidence for a springtime net sink and summer-time source. Geophys. Res. Lett. 36.
Hirdman, D., Burkhart, J.F., Sodemann, H., Eckhardt, S., Jefferson, A., Quinn, P.K., Sharma, S., Strom, J., Stohl, A., 2010a. Long-term trends of black carbon and sulphate aerosol in the Arctic: changes in atmospheric transport and source region emissions. Atmos. Chem. Phys. 10, 9351–9368.
Hirdman, D., Sodemann, H., Eckhardt, S., Burkhart, J.F., Jefferson, A., Mefford, T., Quinn, P.K., Sharma, S., Strom, J., Stohl, A., 2010b. Source identification of short-lived air pollutants in the Arctic using statistical analysis of measurement data and particle dispersion model output. Atmos. Chem. Phys. 10, 669–693.
Hopke, P.K., Allan, H.L., 2009. Chapter 1 theory and application of atmospheric source apportionment. Developments in Environmental Sciences. Elsevier, pp. 1–33.
Hopke, P.K., Ito, K., Mar, T., Christensen, W.F., Eatough, D.J., Henry, R.C., Kim, E., Laden, F., Lall, R., Larson, T.V., Liu, H., Neas, L., Pinto, J., Stolzel, M., Suh, H., Paatero, P., Thurston, G.D., 2005. PM source apportionment and health effects: 1. Intercomparison of source apportionment results. J. Expo. Sci. Environ. Epidemiol. 16, 275–286.
HTAP, 2007. Task force on hemispheric transport of air pollution. In: Keating, T.J., Zuber, A. (Eds.), Hemispheric Transport of Air Pollution 2007 Interim Report. : Air Pollut. Stud., 16. U.N. Econ. Comm. for Europe, New York.
Huang, L., Gong, S.L., Sharma, S., Lavoue, D., Jia, C.Q., 2010. A trajectory analysis of atmospheric transport of black carbon aerosols to Canadian high Arctic in winter and spring (1990–2005). Atmos. Chem. Phys. 10, 5065–5073.
Hwang, I., Hopke, P.K., 2007. Estimation of source apportionment and potential source locations of PM2.5 at a west coastal IMPROVE site. Atmos. Environ. 41, 506–518.
Im, U., Tayanç, M., Yenigün, O., 2008. Interaction patterns of major photochemical pollutants in Istanbul, Turkey. Atmos. Res. 89, 382–390.
Jackson, D.R., Methven, J., Pope, V.D., 2001. Transport in the low-latitude tropopause zone diagnosed using particle trajectories. J. Atmos. Sci. 58, 173–192.
Jenkinson, A.F., Collison, F.P., 1977. An initial climatology of gales over the North Sea, Synoptic Climatology Branch Memorandum No. 62, Meteorological Office, Bracknell.
Jones, A., Thomson, D., Hort, M., Devenish, B., 2007. The UK Met Office's next-generation atmospheric dispersion model, NAME III. Air Pollut. Model. Appl. XVII 17, 580–589.
Jorba, O., Perez, C., Rocadenbosch, F., Baldasano, J.M., 2004. Cluster analysis of 4-day back trajectories arriving in the Barcelona area, Spain, from 1997 to 2002. J. Appl. Meteorol. 43, 887–901.
Junker, C., Sheahan, J.N., Jennings, S.G., O'Brien, P., Hinds, B.D., Martinez-Twary, E., Hansen, A.D.A., White, C., Garvey, D.M., Pinnick, R.G., 2004. Measurement and analysis of aerosol and black carbon in the southwestern United States and Panama and their dependence on air mass origin. J. Geophys. Res.-Atmos. 109.
Junker, C., Wang, J.L., Lee, C.T., 2009. Evaluation of the effect of long-range transport of air pollutants on coastal atmospheric monitoring sites in and around Taiwan. Atmos. Environ. 43, 3374–3384.

Kahl, J.D., 1993. A cautinary note on the use of air trajectories in interpreting atmospheric chemistry measurements. Atmos. Environ. 27, 3037–3038.
Kaiser, A., Schelfinger, H., Spangl, W., Weiss, A., Gilge, S., Fricke, W., Ries, L., Cemas, D., Jesenovec, B., 2007. Transport of nitrogen oxides, carbon monoxide and ozone to the Alpine Global Atmosphere Watch stations Jungfraujoch (Switzerland), Zugspitze and Hohenpeissenberg (Germany), Sonnblick (Austria) and Mt. Krvavec (Slovenia). Atmos. Environ. 41, 9273–9287.
Kalkstein, L.S., Tan, G.R., Skindlov, J.A., 1987. An evaluation of 3 clustering procedures for use in synoptic climatological classification. J. Clim. Appl. Meteorol. 26, 717–730.
Kang, C.M., Kang, B.W., Lee, H.S., 2006. Source identification and trends in concentrations of gaseous and fine particulate principal species in Seoul, South Korea. J. Air Waste Manage. Assoc. 56, 911–921.
Karaca, F., Camci, F., 2010. Distant source contributions to PM_{10} profile evaluated by SOM based cluster analysis of air mass trajectory sets. Atmos. Environ. 44, 892–899.
Karaca, F., Anil, I., Alagha, O., 2009. Long-range potential source contributions of episodic aerosol events to PM_{10} profile of a megacity. Atmos. Environ. 43, 5713–5722.
Kassomenos, P., Vardoulakis, S., Borge, R., Lumbreras, J., Papaloukas, C., Karakitsios, S., 2010. Comparison of statistical clustering techniques for the classification of modelled atmospheric trajectories. Theor. Appl. Climatol. 102, 1–12.
Kasumba, J., Hopke, P.K., Chalupa, D.C., Utell, M.J., 2009. Comparison of sources of submicron particle number concentrations measured at two sites in Rochester, NY. Sci. Total. Environ. 407, 5071–5084.
Kim, N.K., Kim, Y.P., 2008. Major factors affecting the ambient particulate nitrate level at Gosan, Korea. Atmos. Res. 90, 104–114.
Kocak, M., Mihalopoulos, N., Kubilay, N., 2009. Origin and source regions of PM_{10} in the Eastern Mediterranean atmosphere. Atmos. Res. 92, 464–474.
Kuhn, T., Damoah, R., Bacak, A., Sloan, J.J., 2010. Characterising aerosol transport into the Canadian High Arctic using aerosol mass spectrometry and Lagrangian modelling. Atmos. Chem. Phys. 10, 10489–10502.
Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T., Foxall, R., Cawley, G., 2003. Extensive evaluation of neural network models for the prediction of NO_2 and PM_{10} concentrations, compared with a deterministic modelling system and measurements in central Helsinki. Atmos. Environ. 37, 4539–4550.
Laj, P., Bilde, M., Pappalardo, G., Baltensperger, U., Klausen, J., Pla -Duelmer, C., Hjorth, J., Simpson, D., Reimann, S., Clerbaux, C., Coheur, P., Richter, A., Mazière, M.d., Rudich, Y., McFiggans, G., Torseth, K., Wiedensohler, A., Morin, S., Schulz, M., 2009. Measuring atmospheric composition change. Atmos. Environ. 43 (33 Special Issue), 5351–5414. doi:10.1016/j.atmosenv.2009.08.020.
Lamb, H.H., 1972. British Isles weather types and a register of daily sequence of circulation patterns, 1861–1971. Geophysical Memoir, 116. HMSO, London. 85 pp.
Law, K.S., Fierli, F., Cairo, F., Schlager, H., Borrmann, S., Streibel, M., Real, E., Kunkel, D., Schiller, C., Ravegnani, F., Ulanovsky, A., D'Amato, F., Viciani, S., Volk, C.M., 2010. Air mass origins influencing TTL chemical composition over West Africa during 2006 summer monsoon. Atmos. Chem. Phys. 10, 10753–10770.
Lawler, M.J., Finley, B.D., Keene, W.C., Pszenny, A.A.P., Read, K.A., von Glasow, R., Saltzman, E.S., 2009. Pollution-enhanced reactive chlorine chemistry in the eastern tropical Atlantic boundary layer. Geophys. Res. Lett. 36. doi:10.1029/2008GL036666.
Lee, S., Ashbaugh, L., 2007a. Comparison of multi-receptor and single-receptor trajectory source apportionment (TSA) methods using artificial sources. Atmos. Environ. 41, 1119–1127.
Lee, S., Ashbaugh, L., 2007b. Comparison of the MURA and an improved single-receptor (SIRA) trajectory source apportionment (TSA) method using artificial sources. Atmos. Environ. 41, 4466–4481.
Lee, S., Ashbaugh, L., 2007c. The impact of trajectory starting heights on the MURA trajectory source apportionment (TSA) method. Atmos. Environ. 41, 7022–7036.
Lee, J.D., Moller, S.J., Read, K.A., Lewis, A.C., Mendes, L., Carpenter, L.J., 2009. Year-round measurements of nitrogen oxides and ozone in the tropical North Atlantic marine boundary layer. J. Geophys. Res. 114. doi:10.1029/2009jd011878.
Lefohn, A.S., Wernli, H., Shadwick, D., Limbach, S., Oltmans, S.J., Shapiro, M., 2011. The importance of stratospheric-tropospheric transport in affecting surface ozone concentrations in the western and northern tier of the United States. Atmos. Environ. 45, 4845–4857.
Leuchner, M., Rappengluck, B., 2010. VOC source–receptor relationships in Houston during TexAQS-II. Atmos. Environ. 44, 4056–4067.
Lewis, A.C., Evans, M.J., Methven, J., Watson, N., Lee, J.D., Hopkins, J.R., Purvis, R.M., Arnold, S.R., McQuaid, J.B., Whalley, L.K., Pilling, M.J., Heard, D.E., Monks, P.S., Parker, A.E., Reeves, C.E., Oram, D.E., Mills, G., Bandy, B.J., Stewart, D., Coe, H., Williams, P., Crosier, J., 2007. Chemical composition observed over the mid-Atlantic and the detection of pollution signatures far from source regions. J. Geophys. Res. 112.
Lim, J.M., Lee, J.H., Moon, J.H., Chung, Y.S., Kim, K.H., 2010. Source apportionment of PM_{10} at a small industrial area using positive matrix factorization. Atmos. Res. 95, 88–100.
Lin, C.H., Wu, Y.L., Chang, K.H., Lai, C.H., 2004. A method for locating influential pollution sources and estimating their contributions. Environ. Model. Assess. 9, 129–136.
Lupu, A., Maenhaut, W., 2002. Application and comparison of two statistical trajectory techniques for identification of source regions of atmospheric aerosol species. Atmos. Environ. 36, 5607–5618.
Makra, L., Mika, J., Bartzokas, A., Beczi, R., Borsos, E., Sumeghy, Z., 2006. An objective classification system of air mass types for Szeged, Hungary, with special interest in air pollution levels. Meteorol. Atmos. Phys. 92, 115–137.
Makra, L., Mika, J., Bartzokas, A., Beczi, R., Sumeghy, Z., 2009. Comparison of objective air-mass types and the Peczely weather types and their ability to classify levels of air pollutants in Szeged, Hungary. Int. J. Environ. Pollut. 36, 81–98.
Makra, L., Sánta, T., Matyasovszky, I., Damialis, A., Karatzas, K., Bergmann, K.-C., Vokou, D., 2010. Airborne pollen in three European cities: detection of atmospheric circulation pathways by applying three-dimensional clustering of backward trajectories. J. Geophys. Res. 115, D24220.
Makra, L., Matyasovszky, I., Guba, Z., Karatzas, K., Anttila, P., 2011. Monitoring the long-range transport effects on urban PM_{10} levels using 3D clusters of backward trajectories. Atmos. Environ. 45, 2630–2641.
Malcolm, A.L., Manning, A.J., 2001. Testing the skill of a Lagrangian dispersion model at estimating primary and secondary particulates. Atmos. Environ. 35, 1677–1685.
Malcolm, A.L., Derwent, R.G., Maryon, R.H., 2000. Modelling the long-range transport of secondary PM_{10} to the UK. Atmos. Environ. 34, 881–894.
Manning, A.J., Ryall, D.B., Derwent, R.G., Simmonds, P.G., O'Doherty, S., 2003. Estimating European emissions of ozone-depleting and greenhouse gases using observations and a modeling back-attribution technique. J. Geophys. Res.-Atmos. 108.
Manning, A.J., O'Doherty, S., Jones, A.R., Simmonds, P.G., Derwent, R.G., 2011. Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach. J. Geophys. Res. 116, D02305.
Markou, M.T., Kassomenos, P., 2010. Cluster analysis of five years of back trajectories arriving in Athens, Greece. Atmos. Res. 98, 438–457.
Martin, C.L., Allan, J.D., Crosier, J., Choularton, T.W., Coe, H., Gallagher, M.W., 2011. Seasonal variation of fine particulate composition in the centre of a UK City. Atmos. Environ. 45, 4379–4389.
Maryon, R.H., Smith, F.B., Conway, B.J., Goddard, D.M., 1991. The U.K. nuclear accident model. Prog. Nucl. Energy 26 (2), 85Ð104.
McConnell, C.L., Highwood, E.J., Coe, H., Formenti, P., Anderson, B., Osborne, S., Nava, S., Desboeufs, K., Chen, G., Harrison, M.A.J., 2008. Seasonal variations of the physical and optical characteristics of Saharan dust: results from the Dust Outflow and Deposition to the Ocean (DODO) experiment. J. Geophys. Res.-Atmos. 113.
Merrill, J.T., 1994. Isentropic air-flow probability analysis. J. Geophys. Res.-Atmos. 99, 25881–25889.
Merrill, J.T., 1996. Trajectory results and interpretation for PEM-West A. J. Geophys. Res.-Atmos. 101, 1679–1690.
Merrill, J.T., Moody, J.L., 1996. Synoptic meteorology and transport during the North Atlantic Regional Experiment (NARE) intensive: overview. J. Geophys. Res. 101, 28903–28921.
Merrill, J.T., Moody, J.L., Oltmans, S.J., Levy, H., 1996. Meteorological analysis of tropospheric ozone profiles at Bermuda. J. Geophys. Res.-Atmos. 101, 29201–29211.
Methven, J., Evans, M., Simmonds, P., Spain, G., 2001. Estimating relationships between air mass origin and chemical composition. J. Geophys. Res.-Atmos. 106, 5005–5019.
Methven, J., Arnold, S.R., O'Connor, F.M., Barjat, H., Dewey, K., Kent, J., Brough, N., 2003. Estimating photochemically produced ozone throughout a domain using flight data and a Lagrangian model. J. Geophys. Res.-Atmos. 108.
Methven, J., Arnold, S.R., Stohl, A., Evans, M.J., Avery, M., et al., 2006. Establishing Lagrangian connections between observations within air masses crossing the Atlantic during the International Consortium for Atmospheric Research on Transport and Transformation experiment. J. Geophys. Res. 111, D23S62.
Miller Jr., R.G., 1981. Simultaneous Statistical Inference, 2nd ed. Springer, New York. 299 pp.
Moody, J.L., Galusky, J.A., Galloway, J.N., 1989. The use of atmospheric transport pattern recognition techniques in understanding variation in precipitation chemistry. Atmospheric Deposition (Proceedings of the Baltimore Symposium, May 1989). IAHS Publ. No. 179.
Moy, L.A., Dickerson, R.R., Ryan, W.F., 1994. Relationship between back trajectories and tropospheric trace gas concentrations in rural Virginia. Atmos. Environ. 28, 2789–2800.
Müller, K., Lehmann, S., van Pinxteren, D., Gnauk, T., Niedermeier, N., Wiedensohler, A., Herrmann, H., 2010. Particle characterization at the

Cape Verde atmospheric observatory during the 2007 RHaMBLe intensive. Atmos. Chem. Phys. 10, 2709–2721.
Nyanganyura, D., Makarau, A., Mathuthu, M., Meixner, F.X., 2008. A five-day back trajectory climatology for Rukomechi research station (northern Zimbabwe) and the impact of large-scale atmospheric flows on concentrations of airborne coarse and fine particulate mass. S. Afr. J. Sci. 104, 43–52.
Oanh, N.T.K., Chutimon, P., Ekbordin, W., Supat, W., 2005. Meteorological pattern classification and application for forecasting air pollution episode potential in a mountain-valley area. Atmos. Environ. 39, 1211–1225.
Occhipinti, C., Aneja, V.P., Showers, W., Niyogi, D., 2008. Back-trajectory analysis and source–receptor relationships: particulate matter and nitrogen isotopic composition in rainwater. J. Air Waste Manage. Assoc. 58, 1215–1222.
Ohare, G.P., Wilby, R., 1995. A review of ozone pollution in the United-Kingdom and Ireland with an analysis using Lamb weather types. Geogr. J. 161, 1–20.
Olivier, J., Bouwman, A., Maas, C.v.d., Berdowski, J., Veldt, C., Bloos, J., Visschedijk, A., Zandveld, P., Haverlag, J., 1996. Description of EDGAR Version 2.0: a set of global emission inventories of greenhouse gases and ozone-depleting substances for all anthropogenic and most natural sources on a per country basis and on 1 degree × 1 degree grid. Natl. Inst. of Public Health and the Environ., Bilthoven, Netherlands.
Paris, J.D., Stohl, A., Ciais, P., Nedelec, P., Belan, B.D., Arshinov, M.Y., Ramonet, M., 2010. Source–receptor relationships for airborne measurements of CO_2, CO and O_3 above Siberia: a cluster-based approach. Atmos. Chem. Phys. 10, 1671–1687.
Park, S.S., Lee, K.-H., Kim, Y.J., Kim, T.Y., Cho, S.Y., Kim, S.J., 2008. High time-resolution measurements of carbonaceous species in PM2.5 at an urban site of Korea. Atmos. Res. 89, 48–61.
Pekney, N.J., Davidson, C.I., Zhou, L.M., Hopke, P.K., 2006. Application of PSCF and CPF to PMF-modeled sources of PM2.5 in Pittsburgh. Aerosol Sci. Technol. 40, 952–961.
Penkett, S.A., Burgess, R.A., Coe, H., Coll, I., Hov, O., Lindskog, A., Schmidbauer, N., Solberg, S., Roemer, M., Thijsse, T., Beck, J., Reeves, C.E., 2007. Evidence for large average concentrations of the nitrate radical (NO_3) in Western Europe from the HANSA hydrocarbon database. Atmos. Environ. 41, 3465–3478.
Petterssen, S., 1940. Weather Analysis and Forecasting, pp. 221–223.
Pochanart, P., Akimoto, H., Maksyutov, S., Staehelin, J., 2001. Surface ozone at the Swiss Alpine site Arosa: the hemispheric background and the influence of large-scale anthropogenic emissions. Atmos. Environ. 35, 5553–5566.
Pochanart, P., Akimoto, H., Kajii, Y., Potemkin, V.M., Khodzher, T.V., 2003. Regional background ozone and carbon monoxide variations in remote Siberia/East Asia. J. Geophys. Res. 108, 18.
Poirot, R.L., Allan, H.L., 2009. Chapter 2 use of trace metals as source fingerprints. Developments in Environmental Sciences. Elsevier, pp. 35–60.
Poirot, R.L., Wishinski, P.R., 1986. Visibility, sulfate and air-mass history associated with the summertime aerosol in Northern Vermont. Atmos. Environ. 20, 1457–1469.
Poirot, R.L., Wishinski, P.R., Hopke, P.K., Polissar, A.V., 2001. Comparative application of multiple receptor methods to identify aerosol sources in northern Vermont. Environ. Sci. Technol. 35, 4622–4636.
Poissant, L., 1999. Potential sources of atmospheric total gaseous mercury in the St. Lawrence River valley. Atmos. Environ. 33, 2537–2547.
Polissar, A.V., Hopke, P.K., Paatero, P., Kaufmann, Y.J., Hall, D.K., Bodhaine, B.A., Dutton, E.G., Harris, J.M., 1999. The aerosol at Barrow, Alaska: long-term trends and source locations. Atmos. Environ. 33, 2441–2458.
Polissar, A.V., Hopke, P.K., Harris, J.M., 2001a. Source regions for atmospheric aerosol measured at Barrow, Alaska. Environ. Sci. Technol. 35, 4214–4226.
Polissar, A.V., Hopke, P.K., Poirot, R.L., 2001b. Atmospheric aerosol over Vermont: chemical composition and sources. Environ. Sci. Technol. 35, 4604–4621.
Polson, D., Fowler, D., Nemitz, E., Skiba, U., McDonald, A., Famulari, D., Di Marco, C., Simmons, I., Weston, K., Purvis, R., Coe, H., Manning, A.J., Webster, H., Harrison, M., O'Sullivan, D., Reeves, C., Oram, D., 2011. Estimation of spatial apportionment of greenhouse gas emissions for the UK using boundary layer measurements and inverse modelling technique. Atmos. Environ. 45, 1042–1049.
Pongkiatkul, P., Kim Oanh, N.T., 2007. Assessment of potential long-range transport of particulate air pollution using trajectory modeling and monitoring data. Atmos. Res. 85, 3–17.
Real, E., Law, K.S., Weinzierl, B., Fiebig, M., Petzold, A., Wild, O., Methven, J., Arnold, S., Stohl, A., Huntrieser, H., Roiger, A., Schlager, H., Stewart, D., Avery, M., Sachse, G., Browell, E., Ferrare, R., Blake, D., 2007. Processes influencing ozone levels in Alaskan forest fire plumes during long-range transport over the North Atlantic. J. Geophys. Res. 112, D10S41.
Real, E., Law, K.S., Schlager, H., Roiger, A., Huntrieser, H., Methven, J., Cain, M., Holloway, J., Neuman, J.A., Ryerson, T., Flocke, F., de Gouw, J., Atlas, E., Donnelly, S., Parrish, D., 2008. Lagrangian analysis of low altitude anthropogenic plume processing across the North Atlantic. Atmos. Chem. Phys. 8, 7737–7754.
Real, E., Orlandi, E., Law, K.S., Fierli, F., Josset, D., Cairo, F., Schlager, H., Borrmann, S., Kunkel, D., Volk, C.M., McQuaid, J.B., Stewart, D.J., Lee, J., Lewis, A.C., Hopkins, J.R., Ravegnani, F., Ulanovski, A., Liousse, C., 2010. Cross-hemispheric transport of central African biomass burning pollutants: implications for downwind ozone production. Atmos. Chem. Phys. 10, 3027–3046.
Reddy, B.S.K., Kumar, K.R., Balakrishnaiah, G., Gopal, K.R., Reddy, R.R., Ahammed, Y.N., Narasimhulu, K., Reddy, L.S.S., Lal, S., 2010. Observational studies on the variations in surface ozone concentration at Anantapur in southern India. Atmos. Res. 98, 125–139.
Reidmiller, D.R., Jaffe, D.A., Chand, D., Strode, S., Swartzendruber, P., Wolfe, G.M., Thornton, J.A., 2009. Interannual variability of long-range transport as seen at the Mt. Bachelor observatory. Atmos. Chem. Phys. 9, 557–572.
Riccio, A., Giunta, G., Chianese, E., 2007. The application of a trajectory classification procedure to interpret air pollution measurements in the urban area of Naples (Southern Italy). Sci. Total. Environ. 376, 198–214.
Robinson, N.H., Newton, H.M., Allan, J.D., Irwin, M., Hamilton, J.F., Flynn, M., Bower, K.N., Williams, P.I., Mills, G., Reeves, C.E., McFiggans, G., Coe, H., 2011. Source attribution of Bornean air masses by back trajectory analysis during the OP3 project. Atmos. Chem. Phys. 11, 9605–9630.
Rodriguez, S., Alastuey, A., Alonso-Pérez, S., Querol, X., Cuevas, E., Abreu-Afonso, J., Viana, M., Pandolfi, M., de la Rosa, J., 2011. Transport of desert dust mixed with North African industrial pollutants in the subtropical Saharan Air Layer. Atmos. Chem. Phys. Discuss. 11, 8841–8892.
Rozwadowska, A., Zielinski, T., Petelski, T., Sobolewski, P., 2010. Cluster analysis of the impact of air back-trajectories on aerosol optical properties at Hornsund, Spitsbergen. Atmos. Chem. Phys. 10, 877–893.
Ryall, D.B., Maryon, R.H., 1998. Validation of the UK Met. Office's name model against the ETEX dataset. Atmos. Environ. 32, 4265–4276.
Ryall, D.B., Derwent, R.G., Manning, A.J., Simmonds, P.G., O'Doherty, S., 2001. Estimating source regions of European emissions of trace gases from observations at Mace Head. Atmos. Environ. 35, 2507–2523.
Ryall, D.B., Derwent, R.G., Manning, A.J., Redington, A.L., Corden, J., Millington, W., Simmonds, P.G., O'Doherty, S., Carslaw, N., Fuller, G.W., 2002. The origin of high particulate concentrations over the United Kingdom, March 2000. Atmos. Environ. 36, 1363–1378.
Salisbury, G., Monks, P.S., Bauguitte, S., Bandy, B.J., Penkett, S.A., 2002. A seasonal comparison of the ozone photochemistry in clean and polluted air masses at Mace Head, Ireland. J. Atmos. Chem. 41, 163–187.
Salvador, P., Artinano, B., Querol, X., Alastuey, A., 2008. A combined analysis of backward trajectories and aerosol chemistry to characterise long-range transport episodes of particulate matter: the Madrid air basin, a case study. Sci. Total. Environ. 390, 495–506.
Salvador, P., Artinano, B., Pio, C., Afonso, J., Legrand, M., Puxbaum, H., Hammer, S., 2010. Evaluation of aerosol sources at European high altitude background sites with trajectory statistical methods. Atmos. Environ. 44, 2316–2329.
Scheifinger, H., Kaiser, A., 2007. Validation of trajectory statistical methods. Atmos. Environ. 41, 8846–8856.
Schelfinger, H., Kaiser, A., 2007. Validation of trajectory statistical methods. Atmos. Environ. 41, 8846–8856.
Schichtel, B.A., Gebhart, K.A., Barna, M.G., Malm, W.C., 2006. Association of airmass transport patterns and particulate sulfur concentrations at Big Bend National Park, Texas. Atmos. Environ. 40, 992–1006.
Schmale, J., Schneider, J., Ancellet, G., Quennehen, B., Stohl, A., Sodemann, H., Burkhart, J., Hamburger, T., Arnold, S.R., Schwarzenboeck, A., Borrmann, S., Law, K.S., 2011. Source identification and airborne chemical characterisation of aerosol pollution from long-range transport over Greenland during POLARCAT summer campaign 2008. Atmos. Chem. Phys. 11, 10097–10123.
Schwarz, J., Chi, X., Maenhaut, W., Civis, M., Hovorka, J., Smolík, J., 2008. Elemental and organic carbon in atmospheric aerosols at downtown and suburban sites in Prague. Atmos. Res. 90, 287–302.
Seibert, P., 1993. Convergence and accuracy of numerical methods for trajectory calculations. J. Appl. Meteorol. 32, 558–566.
Seibert, P., Frank, A., 2004. Source–receptor matrix calculation with a Lagrangian particle dispersion model in backward mode. Atmos. Chem. Phys. 4, 51–63.
Seibert, P., Kromp-Kolb, H., Baltensperger, U., Jost, D.T., Schwikowski, M., Kasper, A., Puxbaum, H., 1994. Trajectory analysis of aerosol measurements at high alpine sites. In: Borrell, P.M., Borrell, P., Cvitai, T., Seiler, W. (Eds.), Transport and Transformation of Pollutants in the Troposphere. Academic Publishing, Den Haag, pp. 689–693.
Shadbolt, R.P., Waller, E.A., Messina, J.P., Winkler, J.A., 2006. Source regions of lower-tropospheric airflow trajectories for the lower peninsula of Michigan: a 40-year air mass climatology. J. Geophys. Res.-Atmos. 111.
Shan, W., Yin, Y., Lu, H., Liang, S., 2009. A meteorological analysis of ozone episodes using HYSPLIT model and surface data. Atmos. Res. 93, 767–776.

Sharma, S., Lavoue, D., Cachier, H., Barrie, L.A., Gong, S.L., 2004. Long-term trends of the black carbon concentrations in the Canadian Arctic. J. Geophys. Res.-Atmos. 109.

Sharma, S., Andrews, E., Barrie, L.A., Ogren, J.A., Lavoue, D., 2006. Variations and sources of the equivalent black carbon in the high Arctic revealed by long-term observations at Alert and Barrow: 1989–2003. J. Geophys. Res.-Atmos. 111.

Simmonds, P.G., Seuring, S., Nickless, G., Derwent, R.G., 1997. Segregation and interpretation of ozone and carbon monoxide measurements by air mass origin at the TOR Station Mace Head, Ireland from 1987 to 1995. J. Atmos. Chem. 28, 45–59.

Simmonds, P.G., Derwent, R.G., Manning, A.L., Spain, G., 2004. Significant growth in surface ozone at Mace Head, Ireland, 1987–2003. Atmos. Environ. 38, 4769–4778.

Sindosi, O.A., Katsoulis, B.D., Bartzokas, A., 2003. An objective definition of air mass types affecting Athens, Greece; the corresponding atmospheric pressure patterns and air pollution levels. Environ. Technol. 24, 947–962.

Solberg, S., Schmidbauer, N., Semb, A., Stordal, F., Hov, O., 1996. Boundary-layer ozone depletion as seen in the Norwegian Arctic in Spring. J. Atmos. Chem. 23, 301–332.

Solberg, S., Stordal, F., Hov, O., 1997. Tropospheric ozone at high latitudes in clean and polluted air masses, a climatological study. J. Atmos. Chem. 28, 111–123.

Solberg, S., Hov, O., Sovde, A., Isaksen, I.S.A., Coddeville, P., De Backer, H., Forster, C., Orsolini, Y., Uhse, K., 2008. European surface ozone in the extreme summer 2003. J. Geophys. Res.-Atmos. 113.

Song, Y., Dai, W., Shao, M., Liu, Y., Lu, S.H., Kuster, W., Goldan, P., 2008. Comparison of receptor models for source apportionment of volatile organic compounds in Beijing, China. Environ. Pollut. 156, 174–183.

Spichtinger, N., Winterhalter, M., Fabian, P., 1996. Ozone and grosswetterlagen analysis for the Munich metropolitan area. Environ. Sci. Pollut. Res. 3, 145–152.

Stohl, A., 1996. Trajectory statistics — a new method to establish source–receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe. Atmos. Environ. 30, 579–587.

Stohl, A., 1998. Computation, accuracy and applications of trajectories — a review and bibliography. Atmos. Environ. 32, 947–966.

Stohl, A., Wotawa, G., Seibert, P., Krompkolb, H., 1995. Interpolation errors in wind fields as a function of spatial and temporal resoloution and their impact on different types of kinematic trajectories. J. Appl. Meteorol. 34, 2149–2165.

Stohl, A., Hittenberger, M., Wotawa, G., 1998. Validation of the Lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data. Atmos. Environ. 32, 4245–4264.

Stohl, A., Spichtinger-Rakowsky, N., Bonasoni, P., Feldmann, H., Memmesheimer, M., Scheel, H.E., Trickl, T., Hübener, S.H., Ringer, W., Mandl, M., 2000. The influence of stratospheric intrusions on alpine ozone concentrations. Atmos. Environ. 34, 1323–1354.

Stohl, A., James, P., Forster, C., Spichtinger, N., Marenco, A., Thouret, V., Smit, H.G.J., 2001. An extension of measurement of ozone and water vapour by airbus in-service aircraft (MOZAIC) ozone climatologies using trajectory statistics. J. Geophys. Res. 106, 27757–27768.

Stohl, A., Eckhardt, S., Forster, C., James, P., Spichtinger, N., Seibert, P., 2002. A replacement for simple back trajectory calculations in the interpretation of atmospheric trace substance measurements. Atmos. Environ. 36, 4635–4648.

Stohl, A., Cooper, O.R., Damoah, R., Fehsenfeld, F.C., Forster, C., Hsie, E.Y., Hubler, G., Parrish, D.D., Trainer, M., 2004. Forecasting for a Lagrangian aircraft campaign. Atmos. Chem. Phys. 4, 1113–1124.

Stohl, A., Forster, C., Frank, A., Seibert, P., Wotawa, G., 2005. Technical note: the Lagrangian particle dispersion model FLEXPART version 6.2. Atmos. Chem. Phys. 5, 2461–2474.

Stohl, A., Andrews, E., Burkhart, J.F., Forster, C., Herber, A., Hoch, S.W., Kowal, D., Lunder, C., Mefford, T., Ogren, J.A., Sharma, S., Spichtinger, N., Stebel, K., Stone, R., Ström, J., Torseth, K., Wehrli, C., Yttri, K., 2006. Pan-Arctic enhancements of light absorbing aerosol concentrations due to North American boreal forest fires during summer 2004. J. Geophys. Res. 111, D22214.

Stohl, A., Berg, T., Burkhart, J.F., Fjaeraa, A.M., Forster, C., Herber, A., Hov, Ø., Lunder, C., McMillan, W.W., Oltmans, S., Shiobara, M., Simpson, D., Solberg, S., Stebel, K., Ström, J., Tørseth, K., Treffeisen, R., Virkkunen, K., Yttri, K., 2007. Arctic smoke — record high air pollution levels in the European Arctic due to agricultural fires in Eastern Europe in spring 2006. Atmos. Chem. Phys. 7, 511–534.

Stohl, A., Seibert, P., Arduini, J., Eckhardt, S., Fraser, P., Greally, B.R., Lunder, C., Maione, M., Muhle, J., O'Doherty, S., Prinn, R.G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P.G., Vollmer, M.K., Weiss, R.F., Yokouchi, Y., 2009. An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons. Atmos. Chem. Phys. 9, 1597–1620.

Strong, J., Whyatt, J.D., Hewitt, C.N., 2006. Application of multiple wind-roses to improve the modelling of ground-level ozone in the UK. Atmos. Environ. 40, 7480–7493.

Strong, J., Whyatt, J.D., Hewitt, C.N., Derwent, R.G., 2010. Development and application of a Lagrangian model to determine the origins of ozone episodes in the UK. Atmos. Environ. 44, 631–641.

Sutton, R.T., Maclean, H., Swinbank, R., Oneill, A., Taylor, F.W., 1994. High-resolution stratospheric tracer fields estimated from satellite-observations using lagrangian trajectory calculations. J. Atmos. Sci. 51, 2995–3005.

Tarasick, D.W., Jin, J.J., Fioletov, V.E., Liu, G., Thompson, A.M., Oltmans, S.J., Liu, J., Sioris, C.E., Liu, X., Cooper, O.R., Dann, T., Thouret, V., 2010. High-resolution tropospheric ozone fields for INTEX and ARCTAS from IONS ozonesondes. J. Geophys. Res. 115, D20301.

Tarasova, O.A., Senik, I.A., Sosonkin, M.G., Cui, J., Staehelin, J., Prevot, A.S.H., 2009. Surface ozone at the Caucasian site Kislovodsk High Mountain Station and the Swiss Alpine site Jungfraujoch: data analysis and trends (1990–2006). Atmos. Chem. Phys. 9, 4157–4175.

Taubman, B.F., Hains, J.C., Thompson, A.M., Marufu, L.T., Doddridge, B.G., Stehr, J.W., Piety, C.A., Dickerson, R.R., 2006. Aircraft vertical profiles of trace gas and aerosol pollution over the mid-Atlantic United States: statistics and meteorological cluster analysis. J. Geophys. Res.-Atmos. 111.

Toledano, C., Cachorro, V.E., de Frutos, A.M., Torres, B., Berjon, A., Sorribas, M., Stone, R.S., 2009. Airmass classification and analysis of aerosol types at El Arenosillo (Spain). J. Appl. Meteorol. Climatol. 48, 962–981.

Traub, M., Fischer, H., de Reus, M., Kormann, R., Heland, J., Ziereis, H., Schlager, H., Holzinger, R., Williams, J., Warneke, C., de Gouw, J., Lelieveld, J., 2003. Chemical characteristics assigned to trajectory clusters during the MINOS campaign. Atmos. Chem. Phys. 3, 459–468.

Tscherwenka, W., Seibert, P., Kasper, A., Puxbaum, H., 1998. On-line measurements of sulfur dioxide at the 3 km level over central Europe (Sonnblick Observatory, Austria) and statistical trajectory source analysis. Atmos. Environ. 32, 3941–3952.

Tsuang, B.J., 2003. Quantification on the source/receptor relationship of primary pollutants and secondary aerosols by a Gaussian plume trajectory model: part I theory. Atmos. Environ. 37, 3981–3991.

Tsuang, B.J., Chen, C.L., Lin, C.H., Cheng, M.T., Tsai, Y.I., Chio, C.P., Pan, R.C., Kuo, P.H., 2003a. Quantification on the source/receptor relationship of primary pollutants and secondary aerosols by a Gaussian plume trajectory model: part II. Case study. Atmos. Environ. 37, 3993–4006.

Tsuang, B.J., Lee, C.T., Cheng, M.T., Lin, N.H., Lin, Y.C., Chen, C.L., Peng, C.M., Kuo, P.H., 2003b. Quantification on the source/receptor relationship of primary pollutants and secondary aerosols by a Gaussian plume trajectory model: part III—Asian dust-storm periods. Atmos. Environ. 37, 4007–4017.

Turias, I.J., Gonzalez, F.J., Martin, M.L., Galindo, P.L., 2006. A competitive neural network approach for meteorological situation clustering. Atmos. Environ. 40, 532–541.

Tuzson, B., Henne, S., Brunner, D., Steinbacher, M., Mohn, J., Buchmann, B., Emmenegger, L., 2011. Continuous isotopic composition measurements of tropospheric CO_2 at Jungfraujoch (3580 m a.s.l.), Switzerland: real-time observation of regional pollution events. Atmos. Chem. Phys. 11, 1685–1696.

Varotsos, C.A., Efstathiou, M.N., Kondratyev, K.Y., 2003. Long-term variation in surface ozone and its precursors in Athens, Greece — a forecasting tool. Environ. Sci. Pollut. Res. 10, 19–23.

Vasconcelos, L.A.D., Kahl, J.D.W., Liu, D.S., Macias, E.S., White, W.H., 1996. A tracer calibration of back trajectory analysis at the Grand Canyon. J. Geophys. Res.-Atmos. 101, 19329–19335.

Venkatachari, P., Zhou, L.M., Hopke, P.K., Felton, D., Rattigan, O.V., Schwab, J.J., Demerjian, K.L., 2006. Spatial and temporal variability of black carbon in New York City. J. Geophys. Res.-Atmos. 111.

Viana, M., Kuhlbusch, T.A.J., Querol, X., Alastuey, A., Harrison, R.M., Hopke, P.K., Winiwarter, W., Vallius, A., Szidat, S., Prevot, A.S.H., Hueglin, C., Bloemen, H., Wahlin, P., Vecchi, R., Miranda, A.I., Kasper-Giebl, A., Maenhaut, W., Hitzenberger, R., 2008. Source apportionment of particulate matter in Europe: a review of methods and results. J. Aerosol Sci. 39, 827–849.

Virkkula, A., Aurela, M., Hillamo, R., Makela, T., Pakkanen, T., Kerminen, V.M., Maenhaut, W., Francois, F., Cafmeyer, J., 1999. Chemical composition of atmospheric aerosol in the European subarctic: contribution of the Kola Peninsula smelter areas, central Europe, and the Arctic Ocean. J. Geophys. Res.-Atmos. 104, 23681–23696.

Walker, H.L., Derwent, R.G., Donovan, R., Baker, J., 2009. Photochemical trajectory modelling of ozone during the summer PUMA campaign in the UK West Midlands. Sci. Total. Environ. 407, 2012–2023.

Wang, T.J., Lam, K.S., Tsang, C.W., Kot, S.C., 2004. On the variability and correlation of surface ozone and carbon monoxide observed in Hong Kong using trajectory and regression analyses. Adv. Atmos. Sci. 21, 141–152.

Wang, Y.Q., Zhang, X.Y., Draxler, R.R., 2009. TrajStat: GIS-based software that uses various trajectory statistical analysis methods to identify potential sources from long-term air pollution measurement data. Environ. Model. Softw. 24, 938–939.

Wang, Y.X., Hao, J.M., McElroy, M.B., Munger, J.W., Ma, H., Nielsen, C.P., Zhang, Y.Q., 2010. Year round measurements of O_3 and CO at a rural site near Beijing: variations in their correlations. Tellus Ser. B: Chem. Phys. Meteorol. 62, 228–241.

Watson, J.G., Chen, L.W.A., Chow, J.C., Doraiswamy, P., Lowenthal, D.H., 2008. Source apportionment: findings from the U.S. supersites program. J. Air. Waste Manage. Assoc. 58, 265–288.

Weiss-Penzias, P., Jaffe, D.A., Jaegle, L., Liang, Q., 2004. Influence of long-range-transported pollution on the annual and diurnal cycles of carbon monoxide and ozone at Cheeka Peak Observatory. J. Geophys. Res.-Atmos. 109.

Weiss-Penzias, P., Jaffe, D.A., Swartzendruber, P., Dennison, J.B., Chand, D., Hafner, W., Prestbo, E., 2006. Observations of Asian air pollution in the free troposphere at Mount Bachelor Observatory during the spring of 2004. J. Geophys. Res.-Atmos. 111.

Wimolwattanapun, W., Hopke, P.K., Pongkiatkul, P., 2011. Source apportionment and potential source locations of $PM_{2.5}$ and $PM_{2.5-10}$ at residential sites in metropolitan Bangkok, *Atmospheric*. Pollut. Res. 2, 172–181.

Wolfe, G.M., Thornton, J.A., McNeill, V.F., Jaffe, D.A., Reidmiller, D., Chand, D., Smith, J., Swartzendruber, P., Flocke, F., Zheng, W., 2007. Influence of trans-Pacific pollution transport on acyl peroxy nitrate abundances and speciation at Mount Bachelor Observatory during INTEX-B. Atmos. Chem. Phys. 20, 5309–5325.

Wotawa, G., Kroger, H., 1999. Testing the ability of trajectory statistics to reproduce emission inventories of air pollutants in cases of negligible measurement and transport errors. Atmos. Environ. 33, 3037–3043.

Wotawa, G., Kroger, H., Stohl, A., 2000. Transport of ozone towards the Alps — results from trajectory analyses and photochemical model studies. Atmos. Environ. 34, 1367–1377.

Wu, Z.J., Hu, M., Shao, K.S., Slanina, J., 2009. Acidic gases, NH_3 and secondary inorganic ions in PM_{10} during summertime in Beijing, China and their relation to air mass history. Chemosphere 76, 1028–1035.

Xia, X., Chen, H., Zhang, W., 2007. Analysis of the dependence of column-integrated aerosol properties on long-range transport of air masses in Beijing. Atmos. Environ. 41, 7739–7750.

Xiao, H., Kang, S.C., Zhang, Q.G., Han, W.W., Loewen, M., Wong, F., Hung, H., Lei, Y.D., Wania, F., 2010. Transport of semivolatile organic compounds to the Tibetan Plateau: monthly resolved air concentrations at Nam Co. J. Geophys. Res.-Atmos. 115.

Xie, Y., Berkowitz, C.M., 2007. The use of conditional probability functions and potential source contribution functions to identify source regions and advection pathways of hydrocarbon emissions in Houston, Texas. Atmos. Environ. 41, 5831–5847.

Xu, J., DuBois, D., Pitchford, M., Green, M., Etyemezian, V., 2006. Attribution of sulfate aerosols in Federal Class I areas of the western United States based on trajectory regression analysis. Atmos. Environ. 40, 3433–3447.

Xue, L.K., Wang, T., Zhang, J.M., Zhang, X.C., Deliger, Poon, C.N., Ding, A.J., Zhou, X.H., Wu, W.S., Tang, J., Zhang, Q.Z., Wang, W.X., 2011. Source of surface ozone and reactive nitrogen speciation at Mount Waliguan in western China: new insights from the 2006 summer study. J. Geophys. Res. 116, 12.

Yan, P., Tang, J., Huang, J., Mao, J.T., Zhou, X.J., Liu, Q., Wang, Z.F., Zhou, H.G., 2008. The measurement of aerosol optical properties at a rural site in Northern China. Atmos. Chem. Phys. 8, 2229–2242.

Yue, W., Stolzel, M., Cyrys, J., Pitz, M., Heinrich, J., Kreyling, W.G., Wichmann, H.E., Peters, A., Wang, S., Hopke, P.K., 2008. Source apportionment of ambient fine particle size distribution using positive matrix factorization in Erfurt, Germany. Sci. Total. Environ. 398, 133–144.

Zhu, L., Huang, X., Shi, H., Cai, X., Song, Y., 2011. Transport pathways and potential sources of PM_{10} in Beijing. Atmos. Environ. 45, 594–604.