Combining dispersion modelling with synoptic patterns to understand the wind-borne transport into the United Kingdom of the Bluetongue disease vector

Laura Burgin[1], Marie Ekström[2], Suraje Dessai[3]

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

[2]CSIRO Land and Water, Black Mountain, GPO Box 1666, Canberra ACT 2601, Australia 6600

[3] School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK

Corresponding author: Marie Ekström, CSIRO Land and Water, Black Mountain, GPO Box 1666, Canberra ACT 2601, Australia

Abstract

Bluetongue, an economically important animal disease, can be spread over long distances by carriage of insect vectors (*Culicoides* biting midges) on the wind. The weather conditions which influence the midge's flight are controlled by synoptic scale atmospheric circulations. A method is proposed that links wind-borne dispersion of the insects to synoptic circulation through the use of a dispersion model in combination with principal component analysis (PCA) and cluster analysis. We illustrate how to identify the main synoptic situations present during times of midge incursions into the UK from the European continent. A PCA was conducted on high-pass filtered mean sea level pressure data for a domain centred over north-west Europe from 2005 to 2007. A clustering algorithm applied to the PCA scores indicated the data should be divided into 5 classes for which averages were calculated, providing a classification of the main synoptic types present. Midge incursion events were found to mainly occur in two synoptic categories; 64.8% were associated with a pattern displaying a pressure gradient over the North Atlantic leading to moderate south-westerly flow over the UK and 17.9% of the events occurred when high pressure dominated the region leading to south-easterly or easterly winds. The winds indicated by the pressure maps generally compared well against observations from a surface station and analysis charts. This technique could be used to assess frequency and timings of initiations of infection in new areas on seasonal and decadal timescales, currently not possible with other dispersion or statistical modelling methods.

Key words: Bluetongue, *Culicoides,* Wind, Synoptic pattern, Map classification, Dispersion modelling.

1 Introduction

The biting midge *Culicoides* is the principal vector for several viruses causing economically important animal diseases including Bluetongue (BT), African Horse Sickness and Epizootic Haemorrhagic Disease. Northern Europe first experienced outbreaks of the disease in 2006 with periodic outbreaks since. During a peak epidemic in 2007, losses due to death, sickness and reduced productivity of infected farm animals and movement restrictions applied to infected regions were estimated to be in the order of many hundreds of millions of pounds (Hoogendam 2007; Wilson and Mellor 2008).

The spread of vector-borne diseases is influenced directly and indirectly through a large range of environmental factors that influence the pathogen, the vector and the host. Development rates, activity levels, survival, timing of emergence, distributions, abundance levels and migrations of insect populations are determined to different but significant extents by weather and climate. The pathogen spread by the vector is itself also regulated by climate, generally replicating at a faster rate under warmer conditions (Mellor 2000). Other non-climate influences such as farm management practices and disease-limitation strategies are also relevant if attempting to understand the mechanisms of disease spread (Tabachnick 2010).

Purse *et al*. (2005) suggested that the spread of *Culicoides*-borne diseases into Europe would be likely to increase under climate change due to several factors; northwards spread of the traditional Afro-Asiatic vector species as environmental conditions become more habitable, transmission by indigenous European species becoming viable in warmer temperatures, and overall increased virus persistence as winters become shorter. If so, it is likely that a modified future climate could lead to the establishment of new serotypes of BT virus (BTV) and other related viruses in coming decades. To some extent this prediction has already been confirmed by the rapid spread of BTV type 8 (BTV-8) through Europe since 2006 and its arrival into the UK in August 2007 (Wilson and Mellor 2008) and following outbreaks of BTV serotype 1 (BTV-1) and 6 (BTV-6) in the autumn of 2008 in Brittany, France and the Netherlands respectively.

In addition to longer term climatic influences, local, highly variable weather conditions have a significant impact on midge flight and hence the spread of the viruses (Pedgley, 1982; Purse *et al.* 2005; Carpenter *et al.* 2008). Cold temperatures, high wind speeds and precipitation rates are known to reduce the number of midges becoming airborne (Blackwell 1997). Midge flight can also be terminated when they are forced out of the atmosphere by unsuitable weather, such as heavy rain associated with frontal systems (Sellers and Maafour 1991). Midges are weak fliers due to their small size of approximately 1-3mm, and will typically undergo short distance flights of less than a kilometre in order to obtain food, shelter or a breeding site (Carpenter *et al*. 2008). But despite having poor flight abilities, midges can be carried for long distances on the wind (Pedgley 1982). Therefore, *Culicoides spp.* have been implicated in the spread of the disease into formerly uninfected areas several hundreds of kilometres away (e.g., Sellers *et al*. 1978; Calistri *et al.* 2004; Alba *et al.* 2004; Gloster *et al.* 2008; Hendrickx *et al.* 2008; Agren *et al.* 2010). Large-scale pressure systems, controlling wind speed and direction, have previously been linked to patterns of midge-borne disease spread. For example, outbreaks in Israel have been attributed to carriage of infected midges from Turkey on winds caused by the Persian trough system (Braverman and Chechik 1996), and the timing of African Horse Sickness outbreaks in South Africa have been linked to the warm phase of the El Niño/Southern Oscillation (Baylis *et al.* 1999). In addition to meteorological drivers, Pioz *et al*. (2012) showed that other important drivers to the velocity of BT spread are elevation (slower in high elevation) and density of dairy cattle (negatively correlated).

In the UK, gradual overland spread of the disease from mainland Europe is prevented by a natural barrier, the English Channel. Thus incursions of midges only occur at specific times when winds from mainland Europe are favourable. This study aims to understand how large-scale synoptic conditions over north-west Europe relate to surface weather conditions during times of midge incursions, giving insight to the circulation types that are most favourable to midge transport to the UK. Subsequently this relationship could be used to assess the change in disease risk to the UK, through changes in the frequency and timing of suitable synoptic conditions, at seasonal and decadal timescales, where high-resolution wind data necessary to drive a dispersion model may be of limited availability.

In this paper, a synoptic map pattern classification and dispersion modelling outputs are combined in a case study to demonstrate the importance of synoptic circulation conditions for midge transport across the English Channel. A catalogue of typical mean sea level pressure (MSLP) patterns is created to characterise daily synoptic circulation patterns for the region then related to midge incursion data derived from a dispersion model modified to estimate dispersion of 'midge particles' across the English Channel. Due to the computational expense of the dispersion model, this case study is limited to a three year period (2005-2007 inclusive). This period covers the first epidemic of the disease in northern Europe and the first incursion of the disease to the UK. The scope of the paper is limited to the analysis of wind-borne transport of midges that could potentially act as BT vectors, across the English Channel into the UK. The dispersion model cannot capture non-meteorological factors such as the complex short-distance movements of midges across land or human influences on disease dynamics through animal movements and vaccination programmes. Although some results will be specific for the study site, the overall methodology could be applied to any region should the appropriate datasets be available.

2 Data

2.1 Midge days

The UK Met Office's numerical atmospheric-dispersion modelling environment (NAME) (Jones *et al.* 2004) has been adapted to simulate the dispersion of wind-borne midges and used to identify days when midges are likely to be transported from coastal areas of the near-continent to the UK i.e. a 'midge day'. NAME utilises meteorological data from the UK Met Office's operational numerical weather prediction model, the Unified Model (Davies *et al.* 2005). The midge dispersion model was developed as part of a web-based early warning system to predict likely incursion events of BTV-infected midges for the UK government's Department for the Environment, Food and Rural Affairs. The aim being to identify areas of the UK most at risk of BT outbreaks for use in planning decisions such as movement restrictions, vaccination schemes and communication programmes to stakeholders. The service warned of the risk of an incursion into Suffolk overnight of 4/5 August 2007, which is believed to have resulted in the first UK outbreak at Baylham Farm, Ipswich, UK on 22 September 2007 (Gloster *et al.* 2008). This event, along with further testing of the model against outbreaks in Sweden (Agren *et al*. 2010), validates that the model can accurately simulate midge incursion events. The mass of 'midge particles' released into the model atmosphere is based on several meteorological thresholds. These thresholds were derived from the results of experiments carried out at the Institute for Animal Health, Pirbright, UK (Sanders, C. pers. comm. 2008). No particles are released when either rainfall exceeds 1mmhr^{-1}, when wind speed is greater than 3ms^{-1} or when temperatures are colder than 3°C at the release location, as midges will seek shelter when the weather is wet, windy or cold. Under suitable conditions particles are released, then advected following the mean wind with a turbulence component supplied by a random walk scheme. The 'flight' of a midge particle is terminated when rainfall exceeds 1mmhr^{-1} (washout of the midge), and also after twelve hours (the estimated maximum flight survival time of a midge (Sanders, C. pers. comm. 2008).

The main midge season is from April to November in northern Europe (Mellor, P. pers. comm. 2008), thus the model was run on all days during these months for the study period (2005-2007). The release sites for the midge particles were two locations on the coast of the near-continent, representing possible BT outbreak sites with the potential to cause risk to the UK (Fig. 1). Midge particles were released during the peak take-off period around sunset (1800-2100Z) and their positions were tracked for a maximum of 12 hours as they were advected downwind. Days with midge particles successfully reaching the UK across the English Channel, when the weather was suitable for midge flight and winds were suitably directed, were classed as a midge day.

2.2 Meteorological data

MSLP data from the UK Met Office numerical weather prediction model the Unified Model (Davies *et al.* 2005) at a resolution of 0.11° was used to produce the main patterns of circulation in a domain from 46°N to 56°N and 13°W to 15°E. This region of Northern Europe was chosen to include the influence of anticyclonic systems located over Eastern Europe and frontal systems moving in from the Atlantic on the occurrence of midge days. MSLP grids were extracted for all days in 2005-2007 at 00 UTC to represent the conditions during each overnight midge incursion event.

The surface wind climate associated with the main circulation patterns and the midge days was obtained from weather observations recorded at Langdon Bay, UK (51.13N, 1.35E) (Fig. 1). Hourly measurements of wind speed and direction at 10m were extracted from the UK Met Office observation database for the three year period of the study.

3 Methodology

To assess if midges require particular synoptic situations to cross the English Channel it was necessary to elucidate the circulation types that are typical for the region. There are a number of methods for extracting different modes of variation in the atmosphere in order to relate it to the surface environment. Manual classifications include the Lamb weather catalogue (Lamb 1972), a classification of winds over the British Isles into seven basic types and the European Grosswetterlagen (Hess and Brezovsky 1977), where surface and upper air charts are used to classify periods of several days into one of three main types of flow; zonal, mixed or meridional. Automated approaches include correlation based map-pattern classifications, where the similarity between pressure maps is calculated mathematically to objectively place them in discrete categories (Lund 1963), and eigenvector based classifications.

Here an automated methodology outlined by Yarnal (1993) was chosen to ensure a time efficient, largely objective, and reproducible map pattern classification. This methodology is known as the circulation-to-environment approach, in which the synoptic classification is produced first and then related to the environmental variable in question. This approach has been widely used for a variety of applications; daily precipitation (Serra *et al.* 1998), agricultural wind erosion (Ekström *et al.* 2002), dust storm frequency (Ekström *et al.* 2004), heavy snowstorms (Esteban *et al.* 2005) and tropospheric ozone episodes (Hart *et al.* 2006). The map pattern catalogue was then related to midge days to identify circulation types that are more associated with high risk of trans-channel transport of midges from the European continent to the UK. The local wind conditions during such events were then detailed using wind observations from a surface station on the coast of Kent, UK.

3.1 Map pattern classification

The automated method used here is an eigenvector-based map pattern classification based on principal component analysis (PCA) of standardised daily 00UTC MSLP patterns in combination with a clustering technique, to identify the significant modes of atmospheric circulation across the study area. The classification procedure involves several steps of analysis, each of which is detailed in the sections below.

3.1.1. High-pass data filtering

Prior to the PCA, the MSLP data was subjected to a high pass filter to remove variability on time scales longer than the typical duration of regional weather systems, as otherwise the PCA would be dominated by the strong seasonal variability in the pressure data. Unwanted temporal variability in the data was removed following a method outlined in Hewitson and Crane (1992) whereby variability on timescale longer than typical weather events are removed through the use of a moving average filter, preserving variability occurring on timescales less than the first significant harmonic. Here, to retain the spatial pattern within each daily pressure grid, a time series of average grid values was created and the moving average filter was applied to these (where the length of the moving average is the length of the significant harmonic -1 day). The moving average filter was set to 8 day as identified using the tool REDFIT (Schultz and Mudelsee 2002). The difference between the original grid values and the filtered average time series values was then calculated. These standardised pressure grids were then used in all the subsequent analyses.

3.2.1. Principal component analysis

PCA is a technique to reduce the dimensionality of a dataset by retaining those components that contribute most to its variance, whilst minimising any loss of information (e.g., Jolliffe 2002). PCA is often used in atmospheric science as a tool to find spatial or temporal variability in physical fields by condensing a data set into its underlying fundamental modes of variation (e.g., Preisendorfer 1988).

The PCA was carried out using the correlation matrix in S-mode decomposition giving spatially distributed loadings and temporally distributed scores for the selected number of PCs (e.g. Yarnal 1993). The PCA, although strictly a mathematical algorithm, involves elements of subjectivity in the selection of the number of PCs to retain and whether to rotate the selected PCs or not. The choice of optimal number of PCs to retain can be aided by a number of different methods. North *et al.* (1982) provided a rule-of-thumb which proposes that the cut-off should occur where the sampling error of a particular eigenvalue (λ) is comparable to or larger than the spacing between λ and a neighbouring value. The sampling error is given as $\delta\lambda \sim \lambda(2/N)1/2$, where N is the number of variables over which the PCA is carried out on. Two graphical aids, the scree test (Cattell 1966) and the log scree test (Davis and Kalkstein 1990) have also been used. In the former, the point where the slope of the plot levels off is assumed to represent the point at which little is added to the explained variance by adding further PCs, while in the latter a dip in the log-transformed eigenvalue is used as the indicator.

The second element of subjectivity involves whether or not to rotate the retained PCs. Buell (1975) demonstrated that in S-mode analysis, unrotated PCs give resulting loading maps with regular characteristic patterns which are statistical artefacts and nearly independent of the spatial variation in the data. A visual inspection of the unrotated PC loading patterns showed evidence of Buell patterns, suggesting the need for rotation of the selected PCs. For this application the orthogonal Varimax rotation (Kaiser 1958) was used. The Varimax transformation changes the relationship between the components but retains the orthogonality constraint. For a full discussion of the advantages of rotation see Richmann (1986).

3.1.3. Cluster analysis

To group the days with similar characteristics, based on their similarity to the different loading patterns, the PC scores were submitted to a cluster analysis. Two fundamentally different approaches can be taken when clustering data depending on the underlying structure of the data; hierarchical and non-hierarchical cluster analysis (e.g., Wilks 1995). In the former, the analysis merges subsequent pairs of observations which are most similar in k-dimensional space to build a hierarchy of sets of groups which tends to work well when there is a natural hierarchical structure to the data, for example in taxonomy or genetic sequencing. For the map-classification study there was no reason to assume that the MSLP data had an underlying hierarchical structure hence a non-hierarchical method was used.

Clustering relies on a distance measure to assess the degree of similarity in the PC scores. Here the method of k-means was used, where *k* initial cluster centres are chosen randomly and each observation is assigned to a cluster based on its Euclidean distance from the cluster centroid. In this non-hierarchical method the observations are re-assigned to globally optimise the within-cluster sum of squared distances. To reduce the risk of finding a local minimum the procedure was repeated 100 times with different initial seeds. As the number of clusters is pre-defined the degree of objectivity in the analysis may be reduced. It was therefore repeated with several different numbers of clusters (*k* was increased from 2 to 20) and the number of clusters chosen was based on two optimization criteria; the smallest total within-cluster sum-of-squared errors for all clusters and where the highest number of midge days fitted into one cluster. The spatial characteristics of each cluster were then represented by a composite of all de-seasonalized grids included in each separate cluster (e.g. Yarnal 1993).

The resulting map-pattern classification was then used to determine if local weather conditions suitable for midge take-off and subsequent carriage downwind into the UK can be related to larger scale pressure patterns. This was done by taking the individual pressure pattern for each day that is classified as a midge-day and determining which cluster it falls within. The summed totals of midge-day occurrences within each cluster then indicates under which weather regime conditions suitable for midge take-off and dispersion typically occur.

4 Results

4.1 PCA and cluster analysis

Guided by the selection procedures outlined in section 3, 5 PCs were retained. The sampling error and spacing of the eigenvalues are seen to reach the same magnitude at about PC6 (North's rule of thumb, North *et al.* 1982), (see supplementary material for results from graphical aids). Together, the 5 retained PCs also explain over 95% of the variance in the dataset (Table 1).

The loading patterns of the rotated PCs (RPCs) describe the main modes of variation in the de-seasonalized pressure grids (Fig. 2). RPC1 displays features of the North Atlantic Oscillation, with high pressure centred over the Azores and a low pressure system centred over Iceland (Fig. 2a); RPC2 shows a trough extending from the south-west to the northeast from France through to Denmark (Fig. 2b); a centre of high pressure situated over the UK dominates the pattern of variability shown by RPC3 (Fig. 2c); RPC4 (Fig. 2d) shows a distinct high pressure region to the north-east of the domain, with a low pressure system situated over continental Europe and RPC6 shows pronounced low pressure in the southeast with a ridge of high pressure extending across the UK (Fig. 2e).

The cluster analysis resulted in 5 overall pressure patterns (PPs), subsequently referred to as PP1-PP5 (Fig. 3). The seasonal and annual relative frequency of each PP, their association with midge days and persistency are displayed in Table 1.

PP1 displays a large area of high pressure over most of the near-continent (Fig. 3a) which would be expected to generate light south-westerly winds over the English Channel. This PP was found to be most frequent in winter and to persist for up to six days. The relative frequency of midge days associated with this PP is very low; annually only 2.2% of midge days are described by this PP.

PP2 was the most common pattern annually, with a relative frequency of 34.2%, and shows a pressure gradient associated with the North Atlantic Oscillation (NAO) over the west of the domain and high pressure over eastern Europe, leading to south-westerly flow over the UK (Fig 3b). PP2 was found to occur frequently throughout the year and persist for up to seven days. The largest proportion of midge days was associated with this PP; 64.8% annually and 62.9%, 65.1% and 65.4% for spring, summer and autumn respectively.

PP3 shows north-westerly flow across the UK, generated by a centre of low pressure system situated to the north-east of the domain and high pressure in the south-west (Fig 3c). This PP was the second commonest annually during the study period, with an annual relative frequency of 30.1% and it was found to persist for up to six days. A low number of midge days are associated with this PP, only 7.3% annually.

In PP4 a very strong pressure gradient is found across the region, caused by a centre of very low pressure to the north-west and high pressure in the south-east (Fig 3d). These tight isobars indicate strong south-westerly winds would be present across the UK. This PP occurred fairly infrequently throughout the study period, with an annual relative frequency of 9.2%, although it was slightly more common in winter where it occurred 16.3% of the time. PP4 persisted for up for five days. A low number of midge days were associated with this PP, only 7.8% annually, although the relative frequency in autumn was slightly higher at 12.3%.

PP5 shows a centre of high pressure situated towards the centre of the region, with a slack gradient towards lower pressure in the south-east of the domain (Fig 3e). This PP was the third most common annually, but has the second highest annual relative frequency of midge days at 17.9%. Midge days were particularly associated with this pattern in spring and summer and are slightly less frequently in autumn.

4.2 Surface climate characteristics of the pressure patterns

To verify if the winds indicated by the PPs were representative of the surface conditions, observations from a station on the south-east coast of England at Langdon Bay have been examined. These observations also provide some indication of the meteorological characteristics of the airmass present over the study region. However, differences are expected between the geostrophic winds and observational data due to the influence of the local environment.

Wind roses for each of the PPs are given in Figure 4. In general, there is a good relationship between the observed wind speeds and directions and the atmospheric circulation expected by examination of the isobars on the PPs. The wind rose for cluster 1 shows light winds from all directions, as would be expected in a slack pressure situation described by PP1. The winds for cluster 2 are mainly light to moderate with southerlies, south-westerlies and westerlies dominating. This corresponds reasonably well with PP2 which shows south-westerly geostrophic winds over the UK. PP3 indicates north-westerly winds would be predominant over the UK during dates in this cluster. The wind rose shows some agreement, with moderate and occasionally strong winds mainly from the west through to the north. The winds on dates in cluster 4 are generally moderate to strong and the wind rose shows a large dominance in direction from the west, south-west and south. These surface observations correspond well with the pressure distribution described by PP4, with dense isobars indicating strong south-westerly winds. Winds at 00UTC during dates in cluster 5 show larger components from the south, south-east and east and are generally light. There is fairly good correspondence between these winds and the pressure distribution described by PP5, but a greater north-westerly component might be expected.

The distribution of temperatures at 00UTC for each date in the clusters have also been examined at Langdon Bay (not shown). Overall, the temperatures typically spanned by each individual cluster cover a similar range from around -2 to 17°C. Cluster 3 has a slightly larger range from -3 to 19°C, but the distribution of temperatures within the cluster indicates that it is cooler overall compared to the other clusters. Clusters 2 and 5 are slightly negatively skewed and show medians that are slightly warmer than the other clusters at 10 and 11°C indicating slightly warmer temperatures overall on dates in these clusters. The synoptic conditions represented by clusters 2 and 5 occur more frequently in spring, summer and autumn, whereas the circulation patterns represented by clusters 1 and 4 occur more often in winter (Table 1). This difference in seasonality combined with differences in the origin of the airmass may explain why variations are seen in the temperature statistics for each cluster.

4.3 Representation of the synoptic situation on midge days by the map pattern classification

Division of the midge days by the cluster analysis resulted in two PPs representing ~83%, of all midge days. To analyse if these PPs accurately demonstrate the synoptic situation present during midge incursion events the pressure distribution on midge days in each cluster have been plotted as composites and analysis synoptic charts on midge days have been examined (Fig 5).

The composite maps for midge days in cluster 2 and cluster 5 show similarities to the synoptic situation described by the map classification for PP2 and PP5. The pressure distribution in PP2 and in the midge days in cluster 2 are both dominated by a gradient associated with the NAO. However, in the midge day composite a large area of high pressure is also found in the north east of the domain over northern Germany. The synoptic situation in both PP5 and in the composite for midge days in cluster 5 is described by a blocking high. However, the location of the anticyclone differs slightly; it is located toward Denmark in the midge day composite and is found more centrally in PP5.

The representativeness of the composite maps of pressure distribution on midge days in each of the main clusters have been verified against analysis synoptic charts from the UK Met Office archive. An example chart for a midge day on 4 Sep 2005 again demonstrates the NAO gradient with high pressure towards the north east. The chart for a midge day in cluster 5 on 27 Apr 2007 shows that a large blocking high dominates the weather over the UK in a similar way to the high pressure found in PP5 and the composite for midge days in cluster 5. Overall some local differences were noted between the analysis charts and the composites but the general geostrophic flow was found to be similar in each case.

5 Discussion

To enhance understanding of the synoptic conditions conducive to midge incursions across the English Channel, results from two different analysis techniques are linked, one that categorises the daily synoptic circulation types and a second that gives the dispersion footprint of midges given a specific starting point. Because the analysis of MSLP is conducted only on three years of daily data, the classification is not necessarily robust from a climatology perspective, which would require a longer time period. Nonetheless, it identified the synoptic types occurring during the three year period for which this study was conducted. This technique was used to identify the large-scale atmospheric conditions which are present during times of wind-borne incursions of midges, acting as vectors for diseases such as BT into the UK from mainland Europe.Two PPs were found to be more commonly associated with midge days than the others. These showed different attributes in terms of pressure distribution, airmass characteristics, occurrence in each season and persistency. PP2 is the most common PP associated with midge days and also the most frequent pattern which occurs annually over the region. In combination with surface observations of wind speed, wind direction and temperature it could be deduced that the tropical maritime airmass typically associated with this pattern would be suitable for the carriage of midges to the UK. Light and moderate south, south-west and westerly winds were generally present and temperatures were found to be slightly warmer at these times. Field experiment results show that midges tend to be high in numbers and become more active during humid, warm weather (Pedgley, 1982; Carpenter *et al.* 2008). They will also not become airborne during strong winds, but wind of a certain strength is required to carry them across the English Channel. Therefore, moderate winds create the highest risk situation to coastal areas in south east England. PP5 also describes synoptic conditions suitable for midge carriage to the UK. Geostrophic winds determined from the isobars and verified by surface observations show easterly and south easterly winds would be associated with this pressure distribution. Originating from the continent, this airmass has the potential to create the hottest conditions found in northern Europe. Temperatures around 25-30°C are ideal for high midge take-off rates. The airborne midges could be carried to the UK on the gentle and moderate winds associated with the high pressure which dominates this PP. The other pressure patterns describe situations which are not particularly suitable for midge take-off and/or transport to the UK. The winds in PP1 were found to be too light and variable, associated with the slack high pressure over the near continent. In PP3 the northerly winds are described by the pressure distribution, carrying midges away from the UK. PP4 shows a strong pressure gradient associated with the NAO. However, this large difference in pressure would create wind strengths above the threshold for midges to become airborne.

The representativeness of the PPs from the map-classification was assessed by comparing these to composites of the MSLP fields on midge days and analysis charts produced by the UK Met Office. In general the map-classification was found to be a suitable representation of the synoptic situation occurring on midge days, but some local differences occurred. These PPs are therefore not suitable for detailed predictions of locations likely to be at risk in the future, but could be used as a guide to the frequency and timing of incursion events. Other local effects of, e.g. topography and land/sea breezes have not been modelled here due to the resolution of the MSLP data. The association between the PPs and precipitation rates have not been examined here either; moderate to high rainfall could prevent midges from successfully crossing the English Channel, even if winds are of a suitable strength and direction. Further, the 3-year period may be too limited to have captured all the possible synoptic types which occur over the study area and therefore not be a complete representation of the main pressure distributions in the region.

Although this study is limited with regard to general representability of the map pattern classification, the concept brings insights not demonstrated by existing work on the impact of weather on the spread of midge-borne diseases. Previous studies to determine the effect of weather and climate on the spread of insect-borne disease have generally used one of three approaches; dispersion models, statistical models or process based biological models.

Dispersion modelling has typically been carried out on past weather data to determine if wind carriage of midges from infected areas could be implicated in initiation of new cases of BT (e.g. Gloster *et al.* 2008; Hendrickx *et al.* 2008; Agren *et al.* 2009; Garcia-Lastra *et al*. 2012). Trajectory models have also been used in a predictive mode to indicate disease spread on short timescales, as demonstrated by Ducheyne *et al*. (2011) for southern France using a stochastic model drawing on information about weekly wind trajectories, terrain characteristics and epidemiological growth information. These models have not yet been used to assess the risk of disease spread in a climate change scenario.

Dispersion modelling of insect vectors can also suffer from a lack of suitable methods to directly verify their results. The NAME model used here has been indirectly validated by showing it can simulate the pattern and timing of bluetongue outbreaks in previously disease free countries (Gloster *et al*. 2008; Agren *et al*. 2009). Direct tracking of midge dispersal over tens of metres has been carried out in a small-scale “mark-recapture” field experiment (Sanders & Carpenter, 2014). This technique would be near-impossible to carry out on the larger scales involved in midge incursions into new regions. The quantification of gene flow between insect populations across regions would inform dispersal routes and therefore virus incursion and spread risks that we have only been able to describe in a qualitative manner. The imminent publication of the annotated *Culicoides* genome and second-generation sequencing techniques will allow the calculation of gene flow between populations across a range of spatial scales and landscapes and determine the frequency of the wind-borne incursion events described here.

A more common approach to study the climate change impacts on species distribution is the use of ecological niche models, linking spices occurrence (through habitat definition) to current patterns in climate to define the ‘climate envelope’ of the disease (Peters *et al*. 2014). Predictions of where and when diseases are likely to spread to in the future are then made by determining where the climate envelope currently exists and/or will be found under future climate change scenarios (Harris *et al.* 2014). Early work using this technique for BT spread used regression models to relate climate variables to the presence of the vector midge (Baylis *et al.* 2001; Wittmann *et al.* 2001). However, they assumed only one species of the midge, *Culicoides imicola*, acted as the main vector for the disease. It has since been shown that more northerly species are capable of transmitting the virus and each of these species has a different climate envelope (Purse *et al.* 2007). The current northern European epidemic was initiated in Belgium in 2006, hundreds of kilometres from the nearest outbreak, from an unknown source. Studies based on the climate envelope of *Culicoides imicola*, could not have predicted such an event would occur. In a climate change context, Zuliani *et al*. (2015), ecological niche modelling was used to study the plausible northerly extent of *Culicoides sonorensis* across USA Canada border and Samy and Peterson (2016) used a similar approach, but in a global context, predicting spatial extension of the disease in central Africa, US and western Russia under future warmer climates.

In biological models the relationships between transmission variables and climate are first established, and these relationships can then be used to predict rates of spread following an outbreak (e.g. Gubbins *et al.* 2008). Alternatively, these relationships are used with future climate scenarios to predict changes in the transmission process. A spatially explicit biological model was demonstrated by Kelso and Milne (2014) for Australia to study the dispersal of the BT vector *Culicoides brevitarsis* in Australia. The method also considers wind dispersal of the midge. The dispersion being informed by relationships derived from 10m wind data with a multiplier (representing the winds speeds at midge carrying levels), the multiplier being derived from arrival times at trapping sites for studied events. However, this approach relies on accurate knowledge of the biology of the disease and its relationship with climate variables, which can vary according to vector species, host animal and the environment. Also, biological models and ecological niche models/climate envelope techniques assess the suitability of the climate for vectors and disease spread once an initial outbreak has occurred. In contrast, the presented technique here can be used to assess the likely frequency and timing of incursions of disease with explicit representation of atmospheric dynamics, without which no further spread would occur, no matter how suitable the climate.

Finally, the map-pattern classification developed here is independent of its application. Therefore the pressure patterns can be related to the occurrence of any other phenomenon thought to be influenced by large-scale weather systems in the study region. The obvious application would be to relate the classification to other long-distance midge transport routes across seas in northern Europe to further aid the assessment of disease spread risk in the region in future seasons and decades.

6 Conclusions

This study provides new insight into the relationship between large scale synoptic circulation patterns and the likelihood for incursions of midges, potentially infected with viruses such as BT, into the UK from the northern coast of France and Belgium. This relationship assumes the activity levels and flight paths of the midges are controlled by the characteristics and movements of airmasses, which are governed by large scale atmospheric flows described by the PPs deduced in this study.

Two PPs were found to characterise the synoptic situation present during most midge incursion events. The main pattern which gives rise to ~65% of all midge days was also the most common pattern found over northern Europe by the analysis. This PP indicated south-westerly winds over the

UK, generally associated with warm, moist tropical maritime air. The other pattern which is associated with ~18% of midge incursions was found to be fairly infrequent over northern Europe during the period of the analysis. The winds associated with this pattern were directed from the continent; airmasses sourced from this region are generally warm and dry. These two patterns describe synoptic situations suitable for successful midge incursions; warm weather for high take-off rates and persistent winds towards the UK from the coast of the near-continent.

The technique described here could be used to assess changes to the frequency of PPs associated with midge days on a seasonal or decadal timescale by using simulated MSLP data from climate models, allowing predictions to be made of the timings and areas at risk in the UK of midge-borne disease outbreaks. This technique also provides benefits not offered by other modelling techniques, which do not attempt to assess the transport mechanisms necessary to carry the midge to a new area and initiate infection.

Acknowledgements

The authors gratefully acknowledge comments on the paper from John Gloster, Paul Agnew and Derrick Ryall at the Met Office, UK, and Anthony Wilson and Christopher Sanders at the Pirbright Institute, UK. The Atmospheric Dispersion and Air Quality Group at the Met Office, UK is also thanked for the use of NAME. Laura Burgin was funded by Defra contract SE4204.

References

Agren E, Burgin L, Sternberg Lewerin S, Gloster J, Elvander, M. 2010. A likely way of introduction of BTV8 to Sweden in August 2008 - comparison of results from two models for atmospheric transport of the biting midge vector. *Veterinary Record* 167: 484-488

Alba A, Casal J, Domingo M. 2004. Possible introduction of bluetongue into the Balearic Islands, Spain, in 2000, via air streams. *Veterinary Record* **155**: 460-461.

Baylis M, Mellor PS, Meiswinkel R. 1999. Horse sickness and ENSO in South Africa. *Nature* **397**: 574. DOI: 10.1038/17512

Baylis M, Mellor PS, Wittmann EJ, Rogers DJ. 2001. Prediction of areas around the Mediterranean at risk of bluetongue by modelling the distribution of its vector using satellite imaging. *Veterinary Record* **149**: 639-643

Blackwell, A. 1997. Diel flight periodicity of the biting midge Culicoides impunctatus and the effects of meteorological conditions. *Medical and Veterinary Entomology* **11:** 361-367.

Braverman Y, Chechik F. 1996. Air streams and introduction of animal disease borne on Culicoides (Diptera, Ceratopogonidae) into Israel. *Revue Scientifique et Technique de l'Office International des Epizooties* **15**: 1037-1199.

Buell CE. 1975. The topography of empirical orthogonal functions. Fourth Conference on Probability and Statistics in Atmospheric Science, American Meteorological Society: Boston, USA.

Calistri P, Giovanni A, Conte A, Nannini D, Santucci U, Patta C, Rolesu S, Caporale V. 2004. Bluetongue in Italy: part I. *Veterinaria Italiana* **40**: 243-251.

Carpenter S, Szmaragd C, Barber J, Labuschagne K, Gubbins S, Mellor PS. 2008. An assessment of *Culicoides* surveillance techniques in northern Europe: have we underestimated a potential bluetongue virus vector? *Journal of Applied Ecology* **45**: 1237- 1245. DOI: 10.1111/j.1365-2664.2008.01511.x

Cattell RB. 1966. The scree test for the number of factors. *Multivariate Behavioural Research* **1**: 245-435.

Davies T, Cullen MJP, Malcolm A, Mawson MH, Staniforth A, White AA, Wood N. 2005. A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society* **131**: 1579-1782. DOI: 10.1256/qj.04.101

Davis RE, Kalkstein LS. 1990. Development of an automated spatial synoptic climatological classification. *International Journal of Climatology* **10**: 769-794. DOI: 10.1002/joc.3370100802

Ducheyne E, Lange M, Van der Stede Y, Meroc E, Durand B, Hendrickx G. 2011. A stochastic predictive model for the natural spread of bluetongue. *Preventive Veterinary Medicine* **99**: 48-59

Garcia-Lastra R, Leginagoikoa I, Plazaola JM, Ocabo B, Aduriz G, Nunes T, Juste RA, 2012. Bluetongue Virus Serotype 1 outbreak in the Basque country (Northern Spain) 2007-2008. Data support a primary vector windborne transport. *PLoS ONE* **7**:e34431.doi10.1371/journal.pone.0034421

Gloster J, Burgin L, Witham C, Athanassiadou M, Mellor PS. 2008. Bluetongue in the United Kingdom and northern Europe in 2007 and key issues for 2008. *Veterinary Record* **162**: 298-302.

Gubbins S, Carpenter S, Baylis M, Wood JLN, Mellor PS. 2008. Assessing the risk of bluetongue to UK livestock: uncertainty and sensitivity analyses of a temperature dependent model for the basic reproduction number. *Journal of the Royal Society Interface* **5**: 363-371. DOI: 10.1098/rsif.2007.1110

Ekström M, Jönnson P, Barring L. 2002. Synoptic pressure patterns associated with major wind erosion events in southern Sweden (1973-1991). *Climate Research* **23**: 51-66.

Ekström M, McTainsh GH, Chappell A. 2004. Australian dust storms: temporal trends and relationships with synoptic pressure distributions (1960-99). *International Journal of Climatology* **24**: 1581-1599. DOI: 10.1002/joc.1072

Esteban P, Jones PD, Martín-Vide J, Mases M. 2005. Atmospheric circulation patterns related to heavy snowfall days in Andorra, Pyrenees. *International Journal of Climatology* **25**: 319-329. DOI: 10.1002/joc.1103

Harris RM, Grose MR, Lee G, Bindoff NL, Porfirio LL, Fox-Hughes P. 2014. Climate Projections for ecologists. *WIREs Clim Change* **5**: 621-637.

Hart M, De Dear R, Hyde R. 2006. A synoptic climatology of tropospheric ozone episodes in Sydney, Australia. *International Journal of Climatology* **26**: 1635-1649. DOI: 10.1002/joc.1332

Hendrickx G, Gilbert M, Staubach C, Elbers A, Mintiens K, Gerbier G, Ducheyne E. 2008. A wind density model to quantify the airborne spread of Culicoides species during north-western Europe bluetongue epidemic, 2006. *Preventive Veterinary Medicine* **87**: 162-181. DOI: 10.1016/j.prevetmed.2008.06.009

Hess P, Brezovsky H. 1977. Katalog der Grosswetterlagen Europas (1881-1976). Berichte des Deutschen Wetterdienstes, Nr. 113 Bd. 15. Selbstverlag des Deutschen Wetterdienstes, Offenbach am Main.

Hewitson BC, Crane RG. 1992. Regional climate in the GISS GCM: Surface air temperature. *Journal of Climate* **5**: 1002–1011.

Hoogendam K. 2007. *International study on the economic consequences of outbreaks of bluetongue serotype 8 in north-western Europe*. Van Hall Institute: Leeuwarden, Netherlands

Jolliffe IT. 2002. *Principal Component Analysis*. 2nd Edition. Springer-Verlag: New York, USA.

Jones A, Thomson D, Hort M, Devenish B. 2007. The U.K. Met Office's next-generation atmospheric dispersion model, NAME III. In: Borrego C, Norman AL. (Eds) Air Pollution Modeling and its Application XVII, Springer, New York, USA

Kaiser HF. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* **32**: 443-482.

Kelso JK, Milne GJ. 2014. A spatial simulation model for the dispersal of the Bluetongue vector *Culicoides brevitarsis* in Australia. *PLOS ONE* **9**(8): e104646. DOI:10.1371/journal.pone.0104646.

Lamb HH. 1972. British Isles weather types and a register of the daily sequence of circulation types. *Geophysical Memoirs 16*, HMSO: London, UK

Lund IA. 1963. Map-pattern classification by statistical methods. *Journal of Applied Meteorology* **2**: 56-65.

Mellor, P.S. 2000. Replication of arboviruses in insect vectors. Journal of Comparative Pathology **123**: 231-247.

North GR, Bell TL, Cahalan RF, Moeng FJ. 1982. Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review* **110**: 699-706.

Pedgley D. 1982. *Windborne pests and diseases*. Ellis Horwood Limited: Chichester, UK.

Preisendorfer RW. 1988. *Principal component analysis in meteorology and oceanography*. Elsevier: Amsterdam, The Netherlands

Peters J, Waegeman W, Van doninck J, Ducheyne E, Calvete C, Lucientes J, Verhoest NEC, De Baets B. 2014. Predicting spatio-temporal distributions in Spain based on environmental habitat characteristics and species dispersal. *Ecological Informatics* **22**: 69-80.

Pioz M, Guis H, Crespin L, Gay E, Calavas D, Durand B, Abrial D, Ducrot C, 2012. Why did Bluetongue spread the way it did? Factors influencing the velocity of Bluetongue Virus Serotype 8 Epizootic Wave in France? *PLos ONE* **7**: e43360. Doi:10.1371/journal.pone.0043360Purse BV, Mellor PS, Rogers DJ, Samuel AR, Mertens PC, Baylis M. 2005. Climate change and the recent emergence of bluetongue in Europe. *Nature Reviews Microbiology* **3**: 171-181. DOI: 10.1038/nrmicro1090

Purse BV, McCormick BJJ, Mellor PS, Baylis M, Boorman JPT, Borras D, Burgu I, Capela R, Caracappa S, Collantes F, De Liberato C, Delgado JA, Denison E, Georgiev G, El Harak M, De La Rocque S, Lhor Y, Lucientes J, Mangana O, Miranda MA, Nedelchev N, Nomikou K, Ozkul A, Patakakis M, Pena I, Scaramozzino P, Torina A, Rogers DJ. 2007. Incriminating bluetongue virus vectors with climate envelope models. *Journal of Applied Ecology* **44**: 1231-1242. DOI: 10.1111/j.1365-2664.2007.01342.x

Richman MB. 1986. Rotation of principal components. *Journal of Climatology* **6**: 293-335

Sanders CJ, Carpenter S. 2014. Assessment of an immunomarking technique for the study of dispersal of Culicoides biting midges. *Infection, Genetics and Evolution* **28**: 583-587

Samy AM, Peterson AT. 2016. Climate change influences on the global potential distribution of Bluetongue Virus. *PLOS ONE* **11**(3):e0150489. DOI:10.1371/journal.pone.0150489.

Schultz M, Mudelsee M. 2002. REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. *Computers and Geosciences* **28**: 421-426

Sellers RF, Pedgley DE, Tucker MR. 1978. Possible windborne spread of bluetongue to Portugal, June-July 1956. *Journal of Hygiene* **81**: 189-196.

Sellers RF, Maarouf, AR. 1991. Possible introduction of Epizootic Hemorrhagic-Disease of Deer Virus (serotype-1) and Bluetongue Virus (serotype-11) into British-Columbia in 1987 and 1988 by infected Culicoides carried on the wind. *Canadian Journal of Veterinary Research*. **55**: 367-370

Serra C, Fernandez Mills G, Periago MC, Lana X. 1998. Surface synoptic circulation and daily precipitation in Catalonia. *Theoretical and Applied Climatology* **59**: 29-49. DOI: 10.1007/s007040050011

Tabachnick WJ. 2010. Challenges in predicting climate and environmental effects on vector-borne disease episystems in a changing world. *The Journal of Experimental Biology* **213**: 946-954.

Wilks DS. 1995. *Statistical methods in the atmospheric sciences*. Academic Press: London, UK.

Wilson A, Mellor P. 2008. Bluetongue in Europe: Vectors, epidemiology and climate change. *Parasitology Research* **103**: S69-S77. DOI: 10.1007/s00436-00008-01053-x.

Wittmann EJ, Mellor PS, Baylis M. 2001. Using climate data to map the potential distribution of Culicoides imicola (Diptera: Ceratopogonidae) in Europe. *Revue Scientifique et Technique de l'Office International des Epizooties* **20**: 731-740

Yarnal B. 1993. *Synoptic climatology in environmental analysis*. Belhaven Press: London, UK

Zuliani A, Masssolo A, Lysyk T, Johnson G, Marshall S, Berger K, Cork SC. 2015. Modelling the northward expansion of *Culicoides sonorensis* (Dipteria: Ceratopogonidae) under future climate scenarios. *PLOS ONE* **10**(8): e0130294. DOI:10.1371/journal.pone.0130294

Table 1. Results of PCA and PP analysis. First section shows results of the PCA followed by the seasonal and annual relative frequency distribution of each PP and seasonal and annual relative frequency distribution of midge days associated with each PP (here *N* is the number of observations in each season and for the whole year). Last five rows gives relative frequency of PP persistency (here *N* is the total number of days in each cluster).

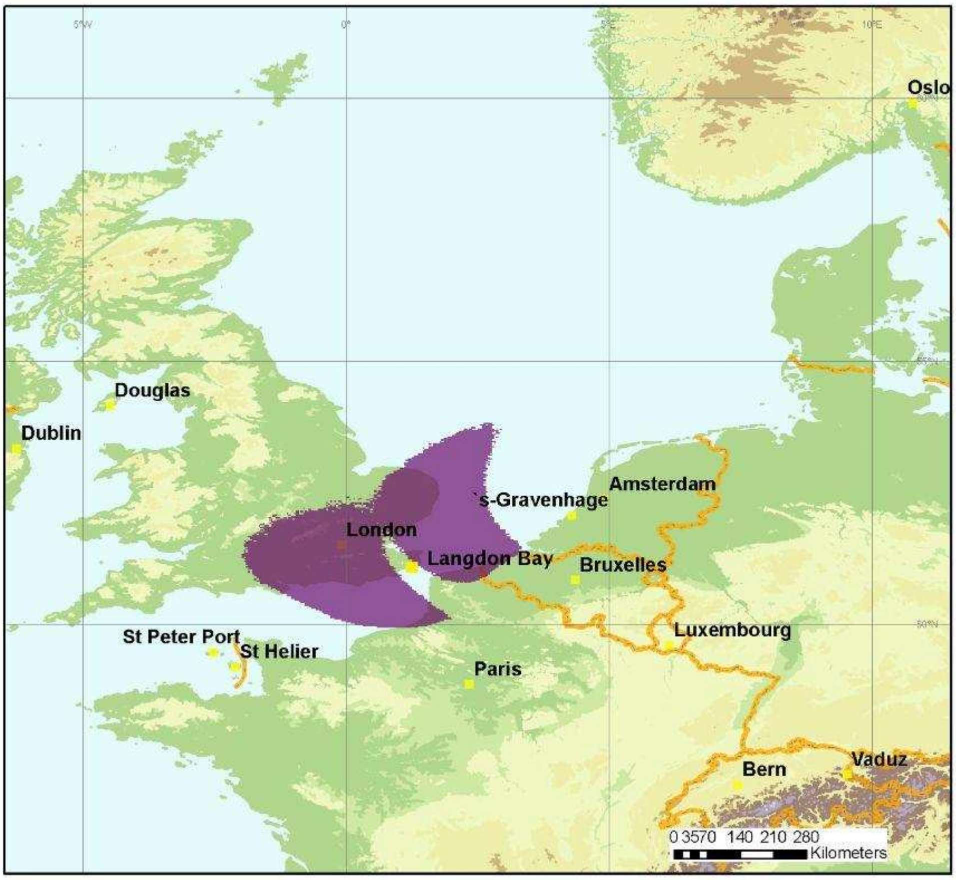
PCA results			
PC	Eigenvalue	Explained variance	Cumulative explained variance (%)
1	**1**	**1314.6**	**48.6**
2	**2**	**848.4**	**29.6**
3	**3**	**648.5**	**10.2**
4	**4**	**478.7**	**4.6**
5	**5**	**273.2**	**2.4**
6	6	117.0	1.2
7	7	102.9	0.7
8	8	73.4	0.5
9	9	51.6	0.4
10	10	17.0	0.3

Relative frequency of PPs (%)					
PP	DJF	MAM	JJA	SON	YEAR
1	14.8	6.5	2.2	5.9	7.3
2	28.5	33.3	32.6	42.5	34.2
3	25.6	33.3	39.9	21.6	30.1
4	16.3	7.2	4.3	9.2	9.2
5	14.8	19.6	21.0	20.9	19.1
N	270	276	276	273	1095

Relative frequency of midge days per PP (%)					
PP	DJF	MAM	JJA	SON	YEAR
1	0	2.9	0	3.7	2.2
2	0	62.9	65.1	65.4	64.8
3	0	8.6	11.1	3.7	7.3
4	0	5.7	3.2	12.3	7.8
5	0	20.0	20.6	14.8	17.9
N	0	35	63	81	179

Relative frequency of PP persistency (%)								
Increasing nb of days →	1	2	3	4	5	6	7	N
PP1	12.5	6.3	5.0	3.8	1.3	1.3		80
PP2	16.3	8.3	4.8	2.1	1.3	0.5	0.2	375
PP3	18.2	11.2	4.2	2.4	0.0	1.2		330
PP4	14.9	4.0	6.9	4.0	2.0			101
PP5	19.1	11.5	9.6	3.3	1.4			209

621 Figure 1. Example of output from the NAME model used to determine 'midge days'.



622

Figure 2. Loading patterns of the first five RPCs (a-e).

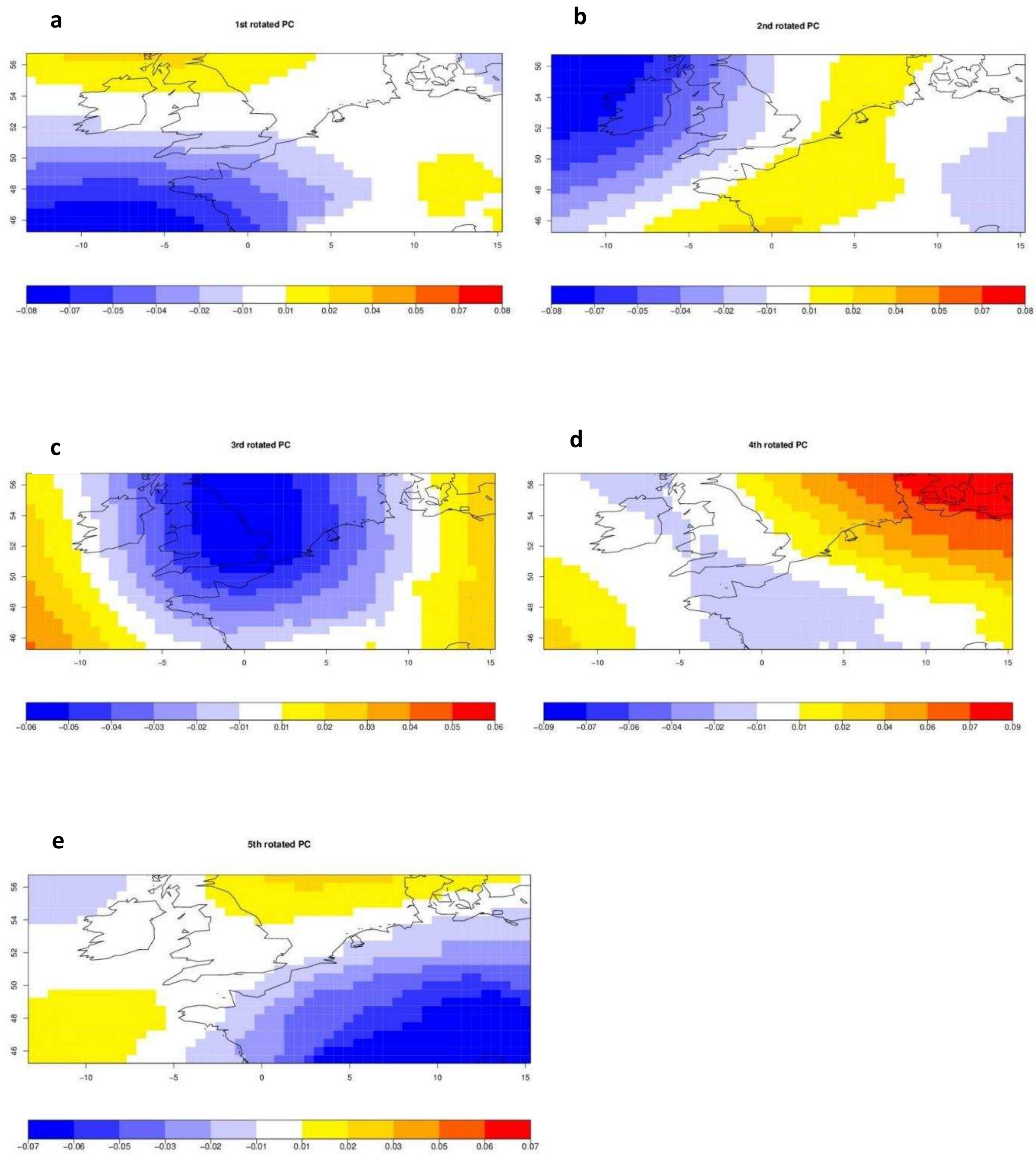


Figure 3: Average de-seasonalized pressure patterns (hPa) for each cluster (a-e).

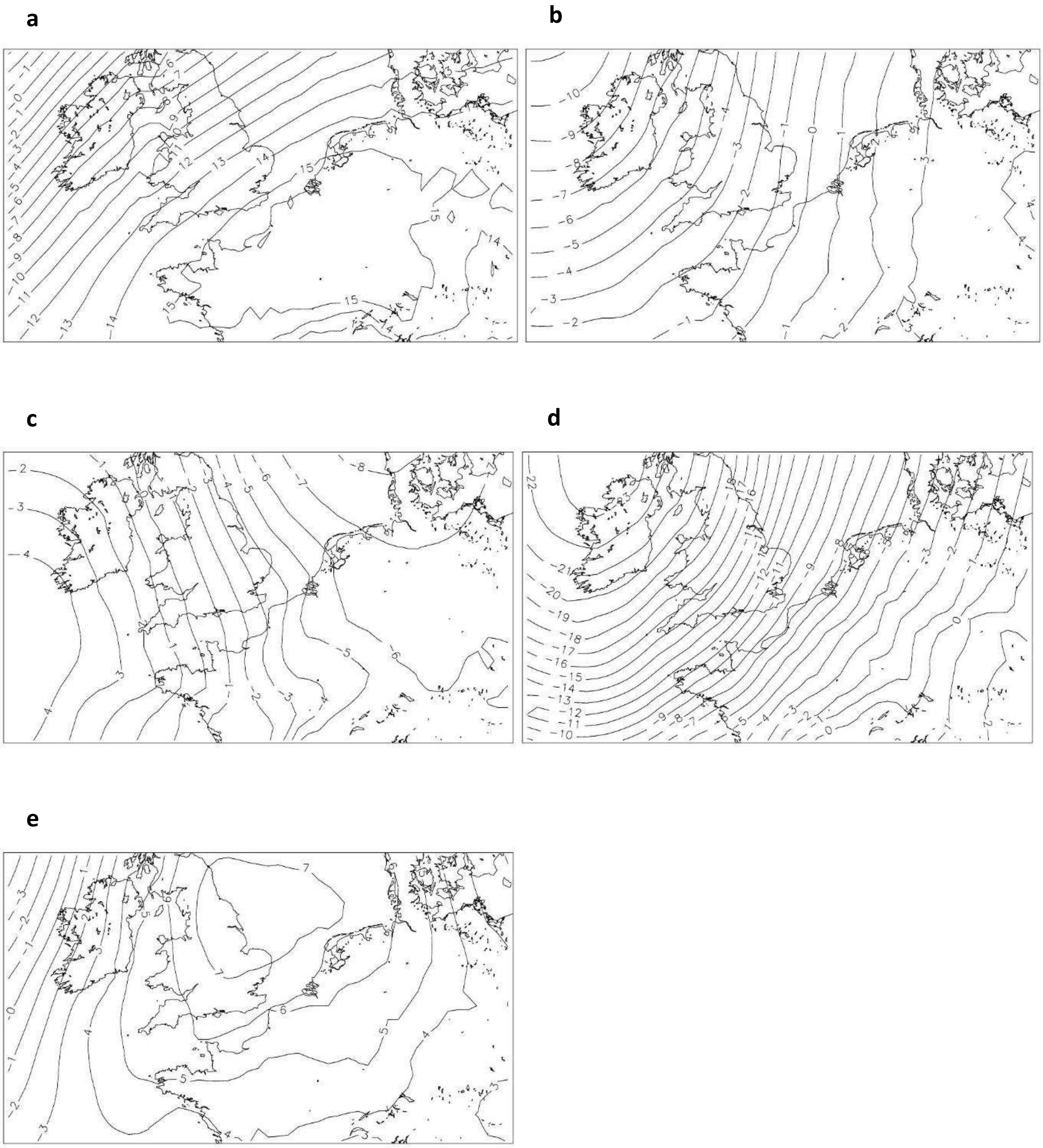


Figure 4: Windroses showing wind speed and direction at 00UTC for clusters 1-5 (a-e) at Langdon Bay. Light-grey=1-5ms^{-1}, mid-grey=5-10ms^{-1}, black=greater than 10ms^{-1}, number of calms (less than 1ms^{-1}) shown as a percentage in the centre.

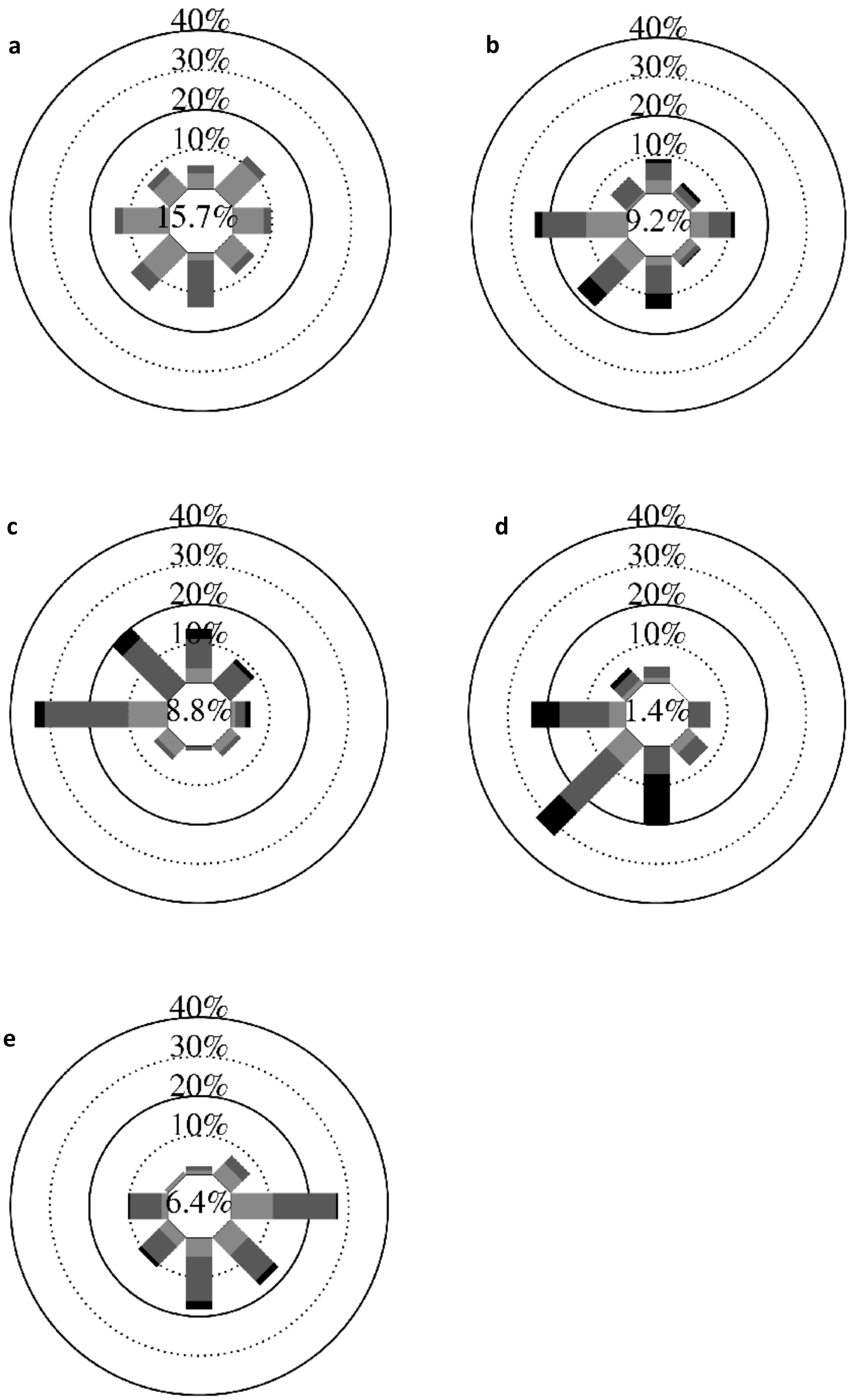


Figure 5: Composite MSLP patterns for all midge days in cluster 2 (a) and cluster 5 (c) and example synoptic charts for a midge day in cluster 2 (04/09/05) (b) and cluster 5 (27/04/07) (d).