

Response to Reviewers: Analysis of the distributions of hourly NO₂ concentrations contributing to annual average NO₂ concentrations across the European monitoring network between 2000 and 2014. Malley et al. (ACPD manuscript ID: acp-2017-910)

5 We thank the reviewers for the time they have taken to review the manuscript, and for their positives comments on our work. We respond to each reviewer's specific comments below. Reviewer comments are italicised and in blue. Unless otherwise stated, page and line numbers in our responses refer to the revised version of the manuscript.

10 **Reviewer 1:**

This was a very comprehensive review of European NO₂ measurements 2010-2014. Just some minor corrections for readers not so familiar with cluster analysis and other concepts. A few general comments may help the authors to explain some of the concepts a bit more clearly:

15

Response: We thank the reviewer for their time spent reviewing the manuscript, for their positive comments on our work, and for the minor suggestions, which have improved the paper. We respond to their general and specific suggestions below.

20 *Cluster analysis using hourly and monthly averages: Section 2.1 line 27: mention the hourly and monthly average contribution to the annual average NO₂.*

Response: We agree that specifically referring to the statistics for which increasing variance was explained is useful, and have therefore made the following amendment.

25

Original Text P7 L5: 'The aim was to maximise the explained variance between sites with as few clusters as possible, and therefore, for each 10 µg m⁻³ 2010-2014 NO_{2AA} bin, the point at which increasing the number of clusters resulted in relatively small increases in the proportion of variance explained was identified.'

30

Amended Text P7 L5: 'The aim in selecting the number of clusters of sites was to maximise the explained variance in the percentage contribution of each month, hour of day, and hourly NO₂ concentrations in 5 µg m⁻³ bins to NO_{2AA} between sites with as few clusters as possible. Therefore, for each 10 µg m⁻³ 2010-2014 NO_{2AA} bin, the point at which increasing the number of clusters resulted in relatively small increases in the proportion of variance explained was identified.'

35

This is how the cluster analysis is calculated. Could you explain how sometimes an hourly contribution might be more or less important? Does it pull out the sites that have large diurnal variability and those that stay the same.

5 **Response:** In the cluster analysis, the grouping of sites is based on the similarity across 77
variables, the percentage contribution from 12 months of the year, 24 hours of the day, and 41 x
10 5 $\mu\text{g m}^{-3}$ hourly NO_2 concentration bins. At the start of the clustering, all sites are contained in
individual clusters (each containing 1 site). At each step, the clusters are merged that result in
the smallest increase in within cluster variance, i.e. for this first step, the two sites for which the
15 77 variables are most similar are grouped into one cluster. Hence the clustering algorithm is
identifying groups of sites for which the pattern of percent contributions to $\text{NO}_{2\text{AA}}$ across the
day, month and across the hourly NO_2 concentration distribution are similar. As the reviewer
suggests, this can result in sites with large diurnal variability being grouped in one cluster, and
those with less variability grouped in another cluster (Figure 5 shows this clearly for sites with
2010-2014 $\text{NO}_{2\text{AA}}$ between 50 and 60 $\mu\text{g m}^{-3}$, with sites in Cluster 4 having large diurnal
variation (shown by the larger difference in contribution to $\text{NO}_{2\text{AA}}$ from the top and bottom 6
hours), and Clusters 1 and 5 having a smaller difference). In addition to grouping sites with
relatively large and small differences in diurnal variation, it also highlights sites with different
patterns in diurnal contribution, as shown in Figure 6, which shows a differences in the
20 proportion of hours with large contribution to $\text{NO}_{2\text{AA}}$ occurring in the morning/evening for
different types of sites.

However, the key advantage of the cluster analysis undertaken in this work is the grouping of
sites based on the similarity of monthly, hour of day, and hourly NO_2 contribution to $\text{NO}_{2\text{AA}}$
25 together, rather than evaluating these contributions separately. This facilitated the identification
of sites where, e.g. the contribution to $\text{NO}_{2\text{AA}}$ from hours of the day was similar, but monthly
contributions, or contributions from 5 $\mu\text{g m}^{-3}$ hourly NO_2 concentration bins differed. The major
distinctions and similarities between clusters are highlighted in Section 3.2 (e.g. P10 L17
highlights the similarity in monthly contribution to $\text{NO}_{2\text{AA}}$ for the two major clusters for sites
30 with 2010-2014 $\text{NO}_{2\text{AA}}$ between 60 and 70 $\mu\text{g m}^{-3}$, but differences in diurnal contribution, and in
contribution from across the diurnal NO_2 concentration distribution). The identification of
similarity across clusters in some contributions to $\text{NO}_{2\text{AA}}$ (e.g. across the year), but differences
in others (e.g. across the day or hourly NO_2 concentration distribution) underpins aspects of the
discussion, e.g. P13 L26: ‘However, the extent of the reduction from reducing emissions during
35 peak hours of the day, and from reducing the short-term peak hourly NO_2 concentrations, varies
between sites, which may reflect variability between sites in meteorological conditions or in
traffic flows (i.e. strength of the emission source in close proximity to the monitoring sites).’

We feel it is important for the reader to understand the user-defined decisions made in the clustering algorithm. These user choices include the clustering algorithm (hierarchical or non-hierarchical methods), the choice of distance measure (i.e. how the similarity across sites is quantified), and, for hierarchical methods, the method by which sites are grouped (Ward's method was used here, alternatives include Average, Minimum or Maximum linkage (Kaufman and Rousseeuw, 1990)). We have therefore acknowledged that there are various types of cluster algorithms, and that the clustering result may be affected by these choices. We have referred to a previous study that analysed different clustering algorithms, and found that Ward's method hierarchical cluster analysis performed effectively compared to alternative methods (Mangiameli et al., 1996). We have also added more text to highlight the key outputs from the cluster analysis, and directed the reader to where these are discussed.

Original Text P6 L12: 'The data capture for each hour of the day, and each month of the year was also calculated for each year at each site. For those sites in each 2010-2014 NO_{2AA} bins with >75% data capture in each month and each hour, the contributions of each month, hour of day, and hourly NO₂ concentrations in 5 µg m⁻³ bins to NO_{2AA} were calculated. These sites were further analysed to identify the major differences in the 'state' statistics across sites with similar 2010-2014 NO_{2AA} values. The relatively low number of sites in the 70-80 µg m⁻³ and >80 µg m⁻³ bins meant that variation in these statistics was examined across the sites individually. However, for the lower NO_{2AA} concentration bins, which contained up to several hundred sites, Ward's hierarchical cluster analysis was used to group sites according to similarity in the contribution of different months, hours, and hourly NO₂ concentration bins to NO_{2AA}. In this process, each monitoring site initially constituted its own cluster. Clusters were then merged based on the similarity of the 2010-2014 average values of the percent contribution to NO_{2AA} from January-December, from 00:00 to 23:00, and from each 5 µg m⁻³ hourly NO₂ concentration bin. At each step, the two clusters were merged which results in the smallest increase in within-cluster variance, and this was repeated until all sites are contained within a single cluster (Kaufman and Rousseeuw, 1990; Ward, 1963). This process is represented by a dendrogram, which was then 'cut' at an appropriate level to determine the number of clusters, and the grouping of sites. The cut-off point was determined from the visual identification of 'elbows' in the plots of proportion of variance explained against the number of clusters into which sites were divided. The aim was to maximise the explained variance between sites with as few clusters as possible, and therefore, for each 10 µg m⁻³ 2010-2014 NO_{2AA} bin, the point at which increasing the number of clusters resulted in relatively small increases in the proportion of variance explained was identified. The cut-off points selected grouped sites into between 5 and 8 clusters, which explained between 57 and 75% of total between-site variance, as shown in Figures S1-S7.'

Amended text P6 L12: ‘The data capture for each hour of the day, and each month of the year was also calculated for each year at each site. For those sites in each 2010-2014 NO_{2AA} bins with >75% data capture in each month and each hour, the percent contribution of hourly NO₂ concentrations in each of the 12 months, in each of the 24 hours of the day, and from hourly NO₂ concentrations across the year in 5 µg m⁻³ bins (i.e. from hourly NO₂ concentrations between 0-5 µg m⁻³, 5-10 µg m⁻³ ... 195-200 µg m⁻³, >200 µg m⁻³) to NO_{2AA} were calculated. These sites were further analysed to identify the major differences in the percentage contribution to NO_{2AA} from each month of the year, hour of the day, and hourly NO₂ concentrations in 5 µg m⁻³ bins across sites with similar 2010-2014 NO_{2AA} values. The relatively low number of sites in the 70-80 µg m⁻³ and >80 µg m⁻³ bins meant that variation in these statistics was examined across the sites individually. However, for the lower NO_{2AA} concentration bins, which contained up to several hundred sites, Ward’s hierarchical cluster analysis was used to group sites according to similarity in 77 variables, i.e. the percent contribution to 2010-2014 NO_{2AA} of different months (12 values), hours of the day (24 values), and 5 µg m⁻³ hourly NO₂ concentration bins (41 values, percent contribution from hourly NO₂ concentrations 0-5 µg m⁻³, 5-10 µg m⁻³ ... 195-200 µg m⁻³, >200 µg m⁻³). In this process, each monitoring site initially constituted its own cluster. Clusters were then merged based on the similarity of the 2010-2014 average values of the percent contribution to NO_{2AA} from January-December, from 00:00 to 23:00, and from each 5 µg m⁻³ hourly NO₂ concentration bin (i.e. the two most ‘similar’ clusters were merged). Euclidean distance was used as the measure of the similarity between these variables between clusters. At each step, the two clusters were merged which results in the smallest increase in within-cluster variance, and this was repeated until all sites are contained within a single cluster (Kaufman and Rousseeuw, 1990; Ward, 1963). Other hierarchical and non-hierarchical clustering algorithms and distance measures could have been used in place of the Ward’s hierarchical algorithm, and the Euclidean distance, respectively, that were applied in this work. Selection of alternative methods could influence the clustering result (Kaufman and Rousseeuw, 1990), but the Ward’s hierarchical cluster algorithm applied here has been shown here to perform effectively compared with alternative approaches, including with respect to the dispersion of observations, the presence of outliers, and irrelevant variables (Mangiameli et al., 1996).

The clustering of monitoring sites in each 10 µg m⁻³ 2010-2014 NO_{2AA} bin process is represented by a dendrogram, which was then ‘cut’ at an appropriate level to determine the number of clusters, and the grouping of sites. The cut-off point was determined from the visual identification of ‘elbows’ in the plots of the proportion of variance explained against the number of clusters into which sites were divided. The aim in selecting the number of clusters of sites was to maximise the explained variance in the percentage contribution of each month, hour of day, and hourly NO₂ concentrations in 5 µg m⁻³ bins to NO_{2AA} between sites with as few clusters as possible. Therefore, for each 10 µg m⁻³ 2010-2014 NO_{2AA} bin, the point at which increasing

the number of clusters resulted in relatively small increases in the proportion of variance explained was identified. The cut-off points selected grouped sites into between 5 and 8 clusters. A key feature of the clustering result is the proportion of variance across sites in the percent contribution to NO_{2AA} from months, hours of the day and hourly NO₂ concentration bins that is explained by dividing sites into a certain number of clusters. The clusters produced in this work explained between 57 and 75% of the variance between sites in the 10 µg m⁻³ 2010-2014 NO_{2AA} bin ranges, as shown in Figure S2. This means that there remains between 43% and 25% variability in these variables between sites with similar 2010-2014 NO_{2AA} that is not explained by dividing sites into the clusters produced in this analysis. Other key aspects of the clustering are the location of sites included in each cluster, and the similarities and differences in the value of each variable (i.e. the percent contribution from each month, hour of day and 5 µg m⁻³ hourly NO₂ concentration bins) between clusters. These aspects are described in Section 3.2.’

What factor is fed into the cluster analysis? Difference between hourly concentration and annual average and difference between monthly and annual?

Response: As outlined above, the 77 variables used to group sites encapsulated variation in the contribution to NO_{2AA} across the year, day and hourly NO₂ concentration distribution. These are the percentage contribution to the annual average from hourly NO₂ concentrations occurring in each month of the year, hour of the day, and in 5 µg m⁻³ bins. This has been made clearer in the text (see additions in response to the comment directly above).

Section 2.1 Line 12 – make clearer how the 5 ugm-3 bins are different to the 10 ugm-3 bins. I will come back to this in Fig 6.

Response: We have amended this text in response to a comment below to make clear that the 5 µg m⁻³ bins relate to *hourly* NO₂ concentrations (and the percent contribution to the annual average from hourly NO₂ concentrations in these ranges), as opposed to the 10 µg m⁻³ bins, that disaggregate sites based on annual NO₂ concentrations.

Section 2.1 Line 14-can you explain what “state” statistics is?

Response: This reference to a state statistic has been replaced with an explicit explanation of the statistics that are being referenced here.

Original Text P6 L16: ‘These sites were further analysed to identify the major differences in the ‘state’ statistics across sites with similar 2010-2014 NO_{2AA} values.’

Amended Text P6 L16: ‘These sites were further analysed to identify the major differences in the percentage contribution to NO_{2AA} from each month of the year, hour of the day, and hourly NO₂ concentrations in 5 µg m⁻³ bins across sites with similar 2010-2014 NO_{2AA} values.’

5

Would a Figure in the appendix showing what a cluster analysis looks like, with the branches shown and the regions in S8 put in blocks along the x axis and coloured even if we cant see the individual station names be useful?

10

Response: The reviewer is correct that one of the advantages of the hierarchical clustering approaches (as opposed to non-hierarchical methods such as k-means) is that observations (in this case sites) can be summarised as a dendrogram. This allows the linkages between observations and clusters to be viewed, as well as the ‘level’ at which the dendrogram is ‘cut’ to produce the set of clusters. However, for this application there are a few factors that we think reduce the utility of this visualisation. Most importantly, as identified by the reviewer, there are too many sites to be individually identified on a normal sized figure, and hence it would not be possible to use this type of plot to view the linkage between individual sites. Secondly, a key conclusion of our paper is that sites with similar variation in contributions to NO_{2AA} were largely not grouped into specific regions. Hence colouring the x-axis according to region would have multiple colours across each cluster, which repeats the information already included in Figure 3.

15

20

25

With the limitations of summarising the cluster results using a dendrogram, it was our intention in the original manuscript to use a series of figures to extract and describe the most important aspects of the cluster analysis. Firstly, it is key to understand the extent to which the number of clusters selected explains the variation between sites. We show this clearly in Figures S1-S7 for sites with 2010-2014 NO_{2AA} in each 10 µg m⁻³ concentration bin (now Figure S2), with the number of cluster selected highlighted in red. Next, it is important that the reader understands the spatial distribution of sites assigned to each cluster, and this is shown in Figure 3 in the main manuscript, and then in supplemental maps in more detail. Finally, it is then important to understand how the variables used to group sites (i.e. percentage contribution NO_{2AA} in each month, hour of day and hourly NO₂ distribution) vary between clusters, which are summarised in Figures 4-7.

30

35

We do not think that including dendrograms would increase the explanation of the clustering results that is already included in the current figures in the main manuscript and supplement. We do agree that it is important to clearly explain the concepts of cluster analysis, and its limitations. To this end we have expanded the description in the text of the cluster analysis methodology, the assumptions made, what is produced from the cluster analysis, and

explanation as to what the different figures in the paper show in terms of the clustering results (see amended text in response to a previous comment (starting at P6 L16 in the main manuscript)).

5 *Referring to Figure 1 doesn't really help the reader so much in explaining what a cluster analysis does and looks like?*

Response: The cluster analysis methodology is described on P6 (L16-P7 L23). We do not refer to Figure 1 in this description.

10

In Fig 4, 5 and 6 with the explanations of what type of sites generally each cluster is made of, it is not easy to work out what the difference is. Maybe you justify in the end that there is no difference between them in most of these figures, so the cluster analysis is not specific enough to understand the characteristics of any one station?

15

Response: As outlined above in response to a previous comment, the key advantage of the methodology adopted here is that the cluster analysis identified similarity across sites in terms of monthly, hour of day and hourly NO₂ concentration bin contributions to NO_{2AA}. This means that it was possible to determine some clusters of sites for which, e.g. the contribution of different months was similar, but for which the contribution from across the day and hourly NO₂ concentration distribution, varied. We synthesise the key differences between clusters (informed by Figures 4, 5, 6 and 7) within the results text of the main manuscript. Key results from the interpretation of these figures, and the clustering result in general are the larger monthly variation in contribution to NO_{2AA} at sites in northern Italy, the similar contribution from months of the year at sites in other regions with high NO_{2AA}, the increasing winter contribution as NO_{2AA} decreases, and the distinction between sites with similar monthly contribution, but differences in hourly contribution (larger/smaller diurnal variation), and hourly NO₂ concentration distribution contribution (broader/narrower set of hourly NO₂ concentrations determining NO_{2AA}).

20

25

30

In terms of the relevance of these results to any one particular site, it is not the aim of this paper to exhaustively describe the differences between individual sites, but to explore the main differences across the European monitoring network. However, for sites with 2010-2014 NO_{2AA} <70 µg m⁻³, where the cluster analysis was used to group sites, we do quantify the extent to which the main differences between groups of sites explain the variation between all sites in the analysis. Figures S1-S7 of the original submission (now Figure S2) show the proportion of the variation between all sites in a 10 µg m⁻³ 2010-2014 NO_{2AA} range that is explained by the number of clusters into which sites are grouped. This gives an indication of the remaining variability between sites within clusters that is not explored within this analysis. We have now

35

added additional text acknowledging that there is variability between individual sites within clusters that is not considered in this analysis (see amended text in response to a comment above (P7 L18).

5

% contribution to NO_{2AA} 2.1, p.5 Line 26: This is used throughout and is not entirely clear. % contribution to NO_{2AA} from each month of the year and hour of the day. Somehow I am left not entirely understanding how this was done – could it be explained better?

10

Response: We have added additional explanation as to how these statistics were calculated on first usage of this term (P6 L17).

15

Original Text P6 L12: ‘The data capture for each hour of the day, and each month of the year was also calculated for each year at each site. For those sites in each 2010-2014 NO_{2AA} bins with >75% data capture in each month and each hour, the contributions of each month, hour of day, and hourly NO₂ concentrations in 5 µg m⁻³ bins to NO_{2AA} were calculated.’

20

Amended Text P6 L12: ‘The data capture for each hour of the day, and each month of the year was also calculated for each year at each site. For those sites in each 2010-2014 NO_{2AA} bins with >75% data capture in each month and each hour, the percent contribution of hourly NO₂ concentrations in each of the 12 months, in each of the 24 hours of the day, and from hourly NO₂ concentrations across the year in 5 µg m⁻³ bins (i.e. from hourly NO₂ concentrations between 0-5 µg m⁻³, 5-10 µg m⁻³ ... 195-200 µg m⁻³, >200 µg m⁻³) to NO_{2AA} were calculated.’

25

Five , ten and fifteen 5 ugm-3 hourly NO2 Please explain this on p.8 line 19 and p. 9 line24 and in Figure 6. Has this analysis added a lot to the analysis. If you keep in, please explain how this is done a bit more clearly.

30

Response: We have added additional explanation as to how these statistics were calculated, and have kept them in the analysis (see amended text in response to a comment above (P6 L19)).

35

Their importance is in highlighting groups of sites where a relatively large, or relatively narrow range of hourly NO₂ concentrations contribute to annual NO₂. The EU, as well as promulgating an annual limit value, also has an hourly NO₂ limit value. In calculating the contribution from hourly NO₂ concentrations in 5 µg m⁻³ bins, it is identified in this analysis that there is a contrast between sites with highest NO_{2AA} at which reducing short-term peak concentrations in excess of the hourly EU limit value would also have an appreciable reduction in annual average concentrations, and those sites where short-term peak hourly NO₂ concentrations are less

extreme, but where annual NO₂ concentrations are nevertheless substantially above the EU limit value.

5 *Trend analysis You look at 2000-2014 trends and compare how most recent (2010-2014) levels can see different trends from the past to the present. Have you explained how inferring a trend from the past against a level at the end of the trend may not be directly correlated to a trend forwards from now? It seems like it may be more useful to look at the 2000-2004 NO₂ levels and then show whether they increased or decreased. We don't want you to do that again but just justify why the end period is*
10 *useful in understanding the trend up to that point.*

Response: We agree that discussion of the relevance for assessing trends at sites grouped by their annual NO₂ concentration at the end of the period should be added. The reason for grouping site based on the NO_{2AA} level in 2010-2014 is to provide a link to the previous section,
15 which explores the contributions to 2010-2014 NO_{2AA}. The trend analysis aims to show whether these contributions, and resulting annual NO₂ concentrations have changed since 2000, or whether they have remained the same. This is important because over the 2000-2014 period, there have been substantial NO_x emission reductions across the EU-28 Member States (EEA, 2016). The categorisation of sites based on 2010-2014 NO_{2AA} shows whether NO_{2AA} at sites that
20 exceed the EU annual NO₂ limit value during the most recent period has decreased, or whether the changes in NO_x emissions across the EU have not resulted in a decrease at these sites that currently exceed the EU standard.

25 We have added text to highlight these points, and to acknowledge, as the reviewer suggests, that by quantifying 2000-2014 trends at these sites, we are not making any prediction as to how trends are likely to progress into the future.

Additional text P8 L4: ‘Trends between 2000 and 2014 were assessed by grouping all qualifying sites based on the 2010-2014 NO_{2AA} value at each site. These groupings were used to
30 investigate whether changes in drivers of NO₂ variability (e.g. European-wide NO_x emission reductions) had had different effects at sites where the magnitude of annual NO₂ concentrations during the most recent period (2010-2014) differed.’

Additional text P18 L7: ‘Finally, assessment of trends between 2000 and 2014 does not
35 indicate how trends at each site are likely to progress into the future, which will be determined by future changes in NO_x emissions, and meteorology.’

Figures 8 and 12 are very different. Are they at odds with each other? Have you compared them? Fig

12 shows a strong linkage between NO₂ level and increases or decreases and Fig 8 seems to say that there are no significant trends

5 **Response:** Figures 8 and 12 are not at odds with one another. Figure 8 shows trends in *annual* NO₂ concentrations, while Figure 12 focusses on changes in the contribution of *hourly* NO₂ concentration to NO_{2AA}. Figure 8 does not show ‘that there are no significant trends’, it shows that at a substantial number of sites there has been a significant decreasing trend in annual NO₂ concentrations between 2000 and 2014. Figure 12 shows that at a substantial number of sites, the contribution to NO_{2AA} from relatively high hourly NO₂ concentrations has decreased, with a corresponding increase in the contribution from relatively moderate and low hourly NO₂ concentrations. This general pattern of change in the distribution of hourly NO₂ concentrations at sites across Europe is consistent with decreasing trends in annual NO₂ concentrations.

Specifics: 1. Intro p.4 Line 20: New sentence: This work

15 **Response:** ‘This work’ is the start of the new sentence.

p.4 Line 22: I would have liked more of an explanation for Fig 1 but maybe not all readers would

20 **Response:** We have added at this point reference to Section 2.1, which describes Figure 1, and the statistics that underpin this analysis in more detail.

p.6 Line 12. . . . in each of the 201-2014 NO_{2aa} bins. . .

25 **Response:** This sentence has been revised in line with the recommendation.

p. 9 Line 28. . . between 60 and 70 ugm-3 (gap)

30 **Response:** This sentence has been revised in line with the recommendation.

4 Discussion: p.12, line 5 and 6. Cant this be summarised into “there is a not a regional correlation in the clustering”?

35 **Response:** Yes, this is broadly what we are saying with these sentences. However, we have retained the original wording rather than that suggested by the reviewer because in some instances there were regional groupings (e.g. in northern Italy).

Figures

Fig 2 - > 80 is clearer than 80+

Response: We have revised all figures in line with this recommendation.

5 *Fig 4 a – >80: would be nice to label the site name in the legend/ caption?!*

Response: The site id for those sites in the 70-80 and >80 $\mu\text{g m}^{-3}$ is shown in the figure. For sites contained within the clusters, there are simply too many to list them in the paper.

10 *Fig 4b line 5 caption: . . .the sites shown as..*

Response: This caption has been revised in line with the recommendation.

Fig 6- comments earlier

15

Response: We have revised this figure to change 80+ to >80.

20 *Figure 9- Zoom in (as there is nothing in N Scandinavia , N Africa, Atlantic, Eastern Europe). >80 is better than 80+. Either there are too many colour bins or it is too difficult to see the difference between a green colour and blue. It looks like green is inside but there is no green in the legend. Very difficult to see beyond the orange or green outer circle*

25 **Response:** We have revised this figure to change 80+ to >80. We have zoomed in, and have changed how the significance of the trend is signified. Rather than the outer circle colour showing the significance of the trend, it is now the shape of each point (circle, $p < 0.05$, triangle, $p > 0.05$). Zooming in on the area of interest in the map, and having only one colour scale on the plot, shows more clearly the magnitude and significance of the trend at each site.

Figure 10,11 and 12- a,b,c- h is a bit hidden on Figure move above

30

Response: We have revised these figures in line with this recommendation.

Appendix Figs S1-S7 could be smaller and put into one figure a- g?

35 **Response:** We have combined these figures in line with this recommendation to form a new Figure S1.

Reviewer 2

5 *General: Well-written summary of NO₂ monitoring data and trends across Europe.
Ready for final publication as is. A few comments:*

Response: We thank the reviewer for the time they spent reviewing the manuscript, and are pleased that they consider the paper ready for publication. Below we respond to each of the reviewer's specific comments in turn.

10 *1) In terms of the drivers, it would have been interesting to assess correlations between actual met measurements (assuming some are nearby, if not onsite) to see if more robust assessments about the contribution drivers could be determined.*

15 **Response:** We agree, it would be an interesting analysis to assess variability in NO₂ concentrations in relation to variation in measurements of wind speed, wind direction, temperature and other meteorological parameters. However, a number of fundamental steps are required before the meteorological measurements collected across Europe, and the network of atmospheric composition measurements are effectively able to be used to
20 undertake the analysis the reviewer outlines.

Specifically, at the majority of atmospheric composition monitoring sites there is not co-located measurement of meteorological parameters. This has recently been highlighted in the UK Air Quality Expert Group's recent analysis of the utility of the UK air quality compliance monitoring network, and the need for collocated met and atmospheric
25 composition data was one of the key recommendations from this report: 'For some scientific and research applications the evidential value of compliance data would be greatly enhanced through the co-measurement and reporting of meteorological parameters' (AQEG, 2015).

30 In the absence of sufficient co-located meteorological measurements, an assessment such as that outlined by the reviewer would then rely on meteorological measurements made at nearby stations. While at rural locations nearby met and atmospheric composition sites may have a relatively high degree of representativeness, a more careful assessment of the
35 representativeness of the urban meteorological station location, in relation to urban background and urban traffic monitoring sites would be required before the met data could be considered applicable to comparison with NO₂ measurements. Previous studies have highlighted i) the large gradients in local scale meteorological conditions within urban areas (Kanda, 2007), and ii) the importance of local-scale meteorology in

determining NO₂ concentrations (Carslaw and Carslaw, 2007). We therefore consider the incorporation of meteorological measurement analysis a very substantial additional piece of analysis beyond the scope of this work. Within the existing text, we have referred to previous studies that assess the relationship between NO₂ concentrations and meteorological conditions at different types of sites (e.g. P14 L24).

However, the reviewer raises an important point about the importance of meteorological conditions in determining NO₂ concentrations, and we have therefore referenced the above AQEG report in the discussion, and reiterated the conclusion of that report that widespread co-location of met data alongside atmospheric composition monitors could increase the ability to assess the drivers of NO₂ (and other air pollutant) variability, especially at urban locations.

Additional Text P18 L23: ‘Co-located meteorological measurements with atmospheric composition measurements, as recommended by AQEG (2015), would also allow meteorological drivers of annual NO₂ concentrations (and other impacts) to be assessed in more detail.’

2) Wonder if a companion meteorological clustering analyses would have proven valuable, in both the trends analyses and the composition/contribution analyses (i.e., during stagnation events does one see different trends, or greater contribution from rush hours)?

Response: As outlined above, the availability of suitable meteorological data across Europe is not sufficient to undertake a meteorological cluster analysis that would be comparable with the cluster analysis of NO₂ measurement data. Within the manuscript we refer to previous work that has assessed the importance of meteorological conditions in determining NO₂ concentrations, which gives some insight into the reviewers question. Schafer et al. (2006) assessed the relationship between boundary layer height and NO₂ concentrations in Hannover, Germany, and found substantially less correlation with boundary layer height and NO₂ concentrations at roadside sites than at urban background sites. Hence whether stagnation events produces different trends, such as greater contribution from rush hour periods may depend on the location of the sites (next to a road or removed from a road).

3) In the US, we often don't have as much confidence in the metadata associated w/ AQ monitoring sites (e.g., site environs change over time and metadata is not updated). Appears that isn't the case in Europe, but might be interesting to "doublecheck" urban/rural & traffic/background against emissions data, or satellite landuse, or population / other surrogates for emissions.

Response: In line with the reviewer's request, we have 'double checked' the relationship between site classification and population and gridded NO_x emission estimates. The population density from the GEOSTAT 2011 dataset

(<http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography/geostat>) in the 1 km grid in which the site was located was determined. Similarly, the total NO_x emission in the 0.1° grid containing each site was determined from gridded NO_x emissions developed by the European Monitoring and Evaluation Programme (EMEP) Centre for Emission Inventories and Projections (CEIP, http://www.ceip.at/new_emep-grid/01_grid_data_2014).

The resulting range of population density and NO_x emissions in the vicinity of different classifications of monitoring sites clearly show distinctions between rural and urban/suburban sites (Figures 1 and 2 below). There is less difference between population density and NO_x emissions between suburban and urban sites, and little distinction for background and traffic sites, although there is a distinction between rural industrial and other classifications of rural sites.

The lack of distinction across urban sites is likely due to the spatial resolution of the gridded population and NO_x emission datasets that are available with full European coverage. One km grids, and 0.1° grids in particular will contain both roadside and background locations within a city.

We do not think that it is necessary, or advisable, to reclassify sites to differ from those used in the AirBase data repository. This is especially the case as these official sites classifications have been used in previous analyses (e.g. Joly and Peuch, 2012), and will likely be used in future work given that these classifications are underpinned by EU legislation (referenced on P5 L13 of the main paper). Creating a discrepancy with how sites are classified in this work and other studies would be unhelpful to those attempting to consider this work in the context of previous studies. More useful are studies which specifically aim to assess the classification of sites based on different variables and methodologies, of which there are multiple examples for Europe, which are referred in the discussion of this work, and to which we have now directed readers at the point in which site classification is discussed in the Methods.

Based on the reviewer's suggestion, we have now referenced previous studies that have looked at monitoring site classification across Europe in the Methods. We have not reproduced Figures 1 and 2 contained in this response document in the main text or supplement of the manuscript because i) the key conclusions of the analysis would not be

5 affected by a small number of sites having characteristics of a different site
classification, ii) these are only two proxy variables for site classification, and there may
have been changes that have occurred in the site environment that affect the
classification that are not reflected either in a change in population of NO_x emissions at
10 0.1° grid size that would require site-by-site analyses to identify robustly, and iii) to keep
focus on the key aspects of our paper, which relate to the contribution of hourly NO₂
variability to annual NO₂ concentrations across Europe. Specifically, for the NO_x
emission figure (Figure 2 below), we have not included it because of the relatively
coarser scale of the gridded NO_x emissions (0.1° grids). We want to avoid any confusion
15 or suggestion that these gridded emissions represent the variability in NO_x emissions
within the immediate environment of the site. This is unlikely to be the case given the
spatial heterogeneity of NO_x emissions within a city at scales below 0.1°. We also note
that this response document will be publicly available for those readers interested in
understanding variation in these proxy variables in relation to the site classifications used
in this work.

Additional text P5 L19: ‘European monitoring site classification has been evaluated
previously (Flemming et al., 2005; Joly and Peuch, 2012; Spangl et al., 2007).’

20 However, we have not included the NO_x emission figure (Figure 2 below) in the
supplement because of the coarser scale of the gridded NO_x emissions (0.1° grids). We
want to avoid any confusion or suggestion that these gridded emissions represent the
variability in NO_x emissions within the immediate environment of the site. This is
25 unlikely to be the case given the spatial heterogeneity of NO_x emissions within a city at
scales below 0.1°.

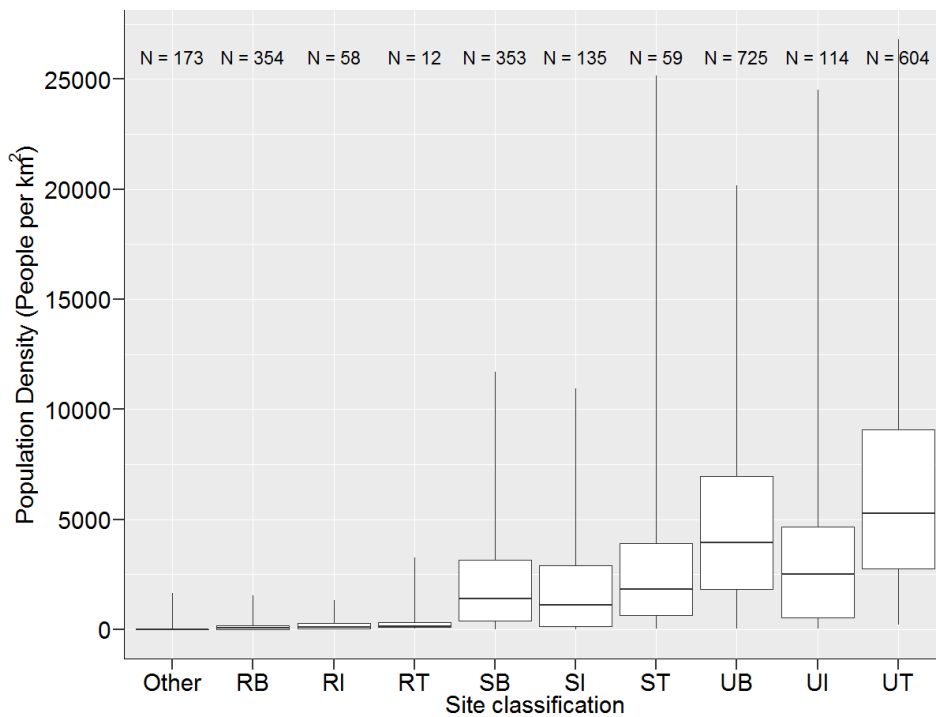


Figure 1: Range (median, 25th and 75th percentiles at top and bottom of box, 2.5th and 97.5th percentile at top and bottom of whiskers) of estimated population density in 2011 in the 1km grids in which NO₂ monitoring sites of different classifications are located. Classification abbreviations denote the site area (R = Rural, S = Suburban, U = Urban), and site type (B = Background, I = Industrial, T = Traffic).

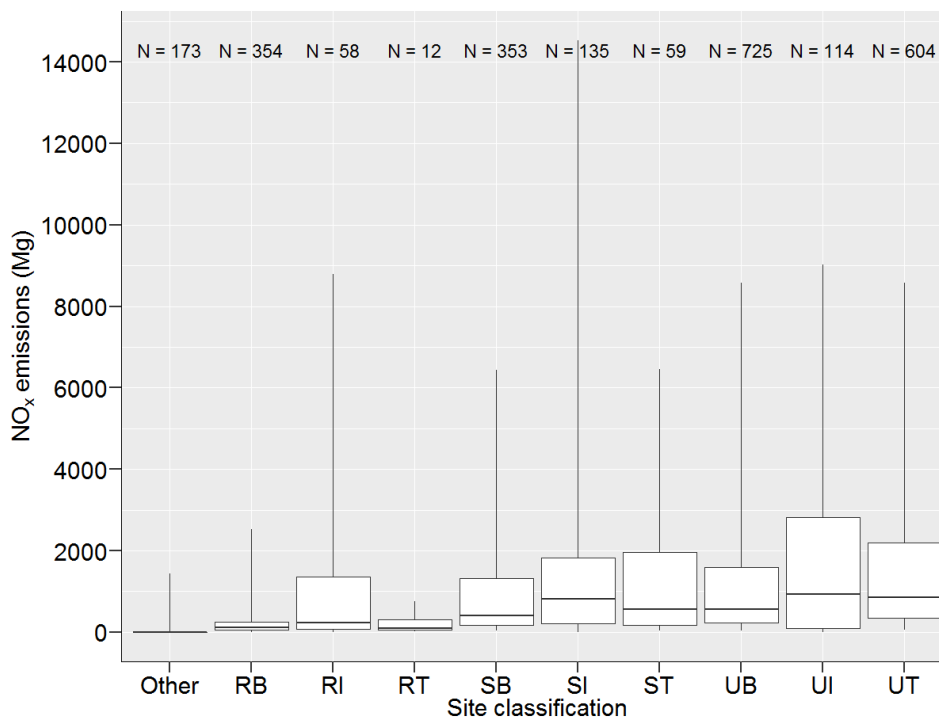


Figure 2: Range (median, 25th and 75th percentiles at top and bottom of box, 2.5th and 97.5th percentile at top and bottom of whiskers) of estimated NO_x emissions in 2014 in the 0.1° grids in which NO₂ monitoring sites of different classifications are located. Classification abbreviations denote the site area (R = Rural, S = Suburban, U = Urban), and site type (B = Background, I = Industrial, T = Traffic).

5 *Specific:*

Page 7: May want to combine the first two full paragraphs on page 7 w/ the last paragraph from page 6.

10 **Response:** Paragraphs have been combined in line with the suggestion.

Pag 12: Am struggling to understand sentence that starts on line 31. If those sites in N. Italy have annual averages dominated by winter months, why would photochemical drivers "be a more important" factor. Is this supposed to read "less important"?

15 **Response:** The sentence does not state that photochemistry is an important driver in producing *high* NO₂ concentrations. Rather it is stating that photochemical conversion of NO₂ to NO and O₃ may be a more important driver of what the concentration of NO₂ is at sites in northern Italy compared to other regions. In the context of a low summertime contributions to annual average NO₂ at northern Italian sites, this means producing low summer NO₂ concentrations (and increasing O₃ concentrations). We have rephrased to make this clearer.

20 **Original text P13 L21:** ‘This indicates that in this region the photochemical conversion of NO₂ to NO and O₃ in summer may be a more important factor in determining NO₂ concentrations at these sites than at other sites across Europe with similar NO_{2AA} concentrations.’

25 **Amended text P13 L21:** ‘This indicates that in this region the photochemical conversion of NO₂ to NO and O₃ in summer may be a more important factor in determining NO₂ concentrations (i.e. lowering NO₂ concentration during summer) at these sites than at other sites across Europe with similar NO_{2AA} concentrations.’

Page 13: add "a" before "heavily industrialised".

30 **Response:** We have amended the sentence in line with the suggestion.

35

References

40 AQEG: Evidential Value of Defra Air Quality Compliance Monitoring. Air Quality Expert Group,

- Defra Publications, available at: http://uk-air.defra.gov.uk/assets/documents/reports/cat11/1509290925_DEF-PB14312_Evidential_value_of_Defra_air_quality_compliance_moni, 2015.
- 5 Carslaw, D. C. and Carslaw, N.: Detecting and characterising small changes in urban nitrogen dioxide concentrations, *Atmos. Environ.*, 41(22), 4723–4733, doi:10.1016/j.atmosenv.2007.03.034, 2007.
- EEA: European Union emission inventory report 1990–2014 under the UNECE Convention on Long-range Transboundary Air Pollution (LRTAP). European Environment Agency Report No. 16/2016., 2016.
- 10 Flemming, J., Stern, R. and Yamartino, R. J.: A new air quality regime classification scheme for O₃, NO₂, SO₂ and PM₁₀ observations sites, *Atmos. Environ.*, 39(33), 6121–6129 [online] Available from: %3CGo, 2005.
- Joly, M. and Peuch, V. H.: Objective classification of air quality monitoring sites over Europe, *Atmos. Environ.*, 47, 111–123 [online] Available from: %3CGo, 2012.
- 15 Kanda, M.: Progress in Urban Meteorology :A Review, *J. Meteorol. Soc. Japan*, 85B, 363–383, doi:10.2151/jmsj.85B.363, 2007.
- Kaufman, L. and Rousseeuw, P. J.: *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York., 1990.
- Mangiameli, P., Chen, S. K. and West, D.: A comparison of SOM neural network and hierarchical clustering methods, *Eur. J. Oper. Res.*, 93(2), 402–417 [online] Available from: %3CGo, 1996.
- 20 Schäfer, K., Emeis, S., Hoffmann, H. and Jahn, C.: Influence of mixing layer height upon air pollution in urban and sub-urban areas, *Meteorol. Zeitschrift*, 15(6), 647–658, doi:10.1127/0941-2948/2006/0164, 2006.
- Spangl, W., Schneider, J., Moosmann, L. and Nagi, C.: *Representativeness and Classification of Air Quality Monitoring Stations*, Umweltbundesamt Report.
- 25 <http://www.umweltbundesamt.at/fileadmin/site/publikationen/REP0121.pdf>, 2007.
- Ward, J.: Hierarchical Grouping to Optimize an Objective Function, *J. Am. Stat. Assoc.*, 58, 236–244 [online] Available from: <http://www.jstor.org/stable/2282967>, 1963.

Analysis of the distributions of hourly NO₂ concentrations contributing to annual average NO₂ concentrations across the European monitoring network between 2000 and 2014

Christopher S Malley¹, Erika von Schneidemesser², Sarah Moller³, Christine F. Braban⁴, W. Kevin Hicks¹,
5 Mathew R. Heal⁵

¹Stockholm Environment Institute, Environment Department, University of York, York, UK

²Institute for Advanced Sustainability Studies, Potsdam, Germany

³National Centre for Atmospheric Science (NCAS), Department of Chemistry, University of York, York, UK

⁴NERC Centre for Ecology & Hydrology, Penicuik, UK

10 ⁵School of Chemistry, University of Edinburgh, Edinburgh, UK

Correspondence to: Christopher S. Malley (chris.malley@york.ac.uk)

Abstract. Exposure to nitrogen dioxide (NO₂) is associated with negative human health effects, both for short-term ‘peak’ concentrations and from long-term exposure to a wider range of NO₂ concentrations. For the latter, the European Union has established an air quality limit value of 40 µg m⁻³ as an annual average. However, factors such as proximity and strength of
15 local emissions, atmospheric chemistry and meteorological conditions means that there is substantial variation in the hourly NO₂ concentrations contributing to an annual average concentration. The aim of this analysis was to quantify the nature of this variation at thousands of monitoring sites across Europe through the calculation of a standard set of ‘chemical climatology’ statistics. Specifically, at each monitoring site that satisfied data capture criteria for inclusion in this analysis, annual NO₂ concentrations, as well as the percentage contribution from each month, hour of the day, and hourly NO₂ concentrations divided
20 into 5 µg m⁻³ bins were calculated.

Across Europe, 2010-2014 average annual NO₂ concentrations (NO_{2AA}) exceeded the annual NO₂ limit value at 8% of >2500 monitoring sites. The application of this ‘chemical climatology’ approach showed that sites with distinct monthly, hour of day, and hourly NO₂ concentration bin contributions to NO_{2AA} were not grouped in specific regions of Europe, and within relatively
25 small geographic regions there were sites with similar NO_{2AA}, but with differences in these contributions. Specifically, at sites with highest NO_{2AA}, there were generally similar contributions from across the year, but there were also differences in the contribution of peak vs moderate hourly NO₂ concentrations to NO_{2AA}, and from different hours across the day. Trends between 2000 and 2014 for 259 sites indicate that, in general, the contribution to NO_{2AA} from winter months has increased, as has the contribution from the rush-hour periods of the day, while the contribution from peak hourly NO₂ concentrations has decreased.
30 The variety of monthly, hour of day and hourly NO₂ contribution bin contributions to NO_{2AA}, across cities, countries and regions of Europe indicate that within relatively small geographic areas different interactions between emissions, atmospheric chemistry and meteorology produce variation in NO_{2AA} and the conditions that produce it. Therefore, measures implemented to reduce NO_{2AA} in one location may not be as effective in others. The development of strategies to reduce NO_{2AA} for an area

should consider i) the variation in monthly, hour of day and hourly NO₂ concentration bin contributions to NO_{2AA} within that area, and ii) how specific mitigation actions will affect variability in hourly NO₂ concentrations.

1 Introduction

Nitrogen dioxide (NO₂) is an air pollutant that is directly emitted and together with nitric oxide (NO) constitute primary emitted nitrogen oxides (commonly termed NO_x). NO₂ is emitted from a range of sources, most importantly from road transport (with different vehicle types emitting different proportions of NO_x as NO₂), and is also formed in the atmosphere through the reaction between primary emitted NO and ozone (O₃). Exposure to NO₂ has been linked to negative human health effects, as summarised in the World Health Organization (WHO) Review of the Evidence on the Health Aspects of Air Pollution (REVIHAAP, 2013). Alongside their conclusion of ‘consistent’ epidemiological evidence for association between short-term NO₂ exposure and negative respiratory effects, REVIHAAP (2013) also conclude that studies that have used multi-pollutant models to separate the association between long-term NO₂ exposure and respiratory effects from those of other pollutants ‘provide support for an effect of NO₂ independent from particle mass metrics’, and that this evidence was ‘suggestive of a causal relationship.’ The United States Environmental Protection Agency (US EPA) similarly concluded that epidemiological and experimental evidence indicates that there is ‘likely to be a causal relationship between long-term NO₂ exposure and respiratory effects’ (US EPA, 2016). The United Kingdom Committee on the Medical Effects of Air Pollution (COMEAP) concluded that ‘although it is possible that, to some extent, NO₂ acts as a marker of the effects of other traffic-related pollutants, the epidemiological and mechanistic evidence now suggests that it would be sensible to regard NO₂ as causing some of the health impact found to be associated with it in epidemiological studies’ (COMEAP, 2015). Recent analysis of large prospective cohorts in the United States and Canada also indicate significant associations between long-term NO₂ exposure and mortality after controlling for exposure to PM_{2.5} and O₃ (Crouse et al., 2015; Turner et al., 2016).

In the European Union (EU), the evidence of NO₂ health effects has led to the establishment of air quality standards for the protection of human health. Limit values for NO₂ are set at 200 µg m⁻³ for 1-h average concentrations (with 18 exceedances permitted per year), and 40 µg m⁻³ for annual average concentrations (European Council Directive 2008/50/EC, 2008). These limit values are in line with the WHO air quality guidelines (WHO, 2006). Hence a common basis for assessment of NO₂ levels relevant to the long-term health impact is well established across EU Member States. However, despite the consistent ‘impact’ metric, a common approach to assess the conditions producing this impact metric is currently lacking. For the protection of human health and vegetation, understanding these conditions is important in addition to assessment of the extent of the impact metric itself, as it informs how the magnitude of the impact metric arises, and therefore the options available for its mitigation.

Annual average NO₂ concentrations in a location are determined by a combination of factors, including the following.

5 i) the emissions environment, such as the proximity and magnitude of major sources of NO_x ($\text{NO} + \text{NO}_2$) and the proportion of NO_x emissions that are directly emitted as NO_2 , (AQEG, 2007; Beevers et al., 2012; Carslaw et al., 2011; Carslaw and Beevers, 2004; Grice et al., 2009; Vestreng et al., 2009). Across Europe, the major source of NO_x emissions is road transport (39% of NO_x emissions in the EU28 Member States in 2014, (EEA, 2016b)), and recent analysis has suggested that road transport NO_x emissions could be even higher when accounting for real-world emission factors (Anenberg et al., 2017). Other sources include energy production and distribution (20%), commercial, institutional and households (14%), and energy use in industry (13%) (EEA, 2016b).

10 ii) atmospheric chemistry, including the balance between secondary NO_2 formation through the reaction between primary emitted NO and O_3 , and its photolysis back to NO and O_3 as emissions are transported further from their source (Jenkin, 2014; Jenkin and Clemitshaw, 2000; Mavroidis and Chaloulakou, 2011). Other chemical processes which determine levels of NO_2 at a particular location include the reaction of NO_2 with the hydroxyl (OH) radical to form nitric acid (HNO_3), which is a major chemical removal process for NO_2 during daytime (Jenkin and Clemitshaw, 2000), and the combination of NO_2 and peroxy radicals to form peroxy acetyl nitrates (PANs) which also removes NO_2 close to emission sources. PAN species can also be transported long distances and act as a source of NO_2 remote from major emission sources (Fischer et al., 2014).

20 iii) atmospheric dispersion between source and receptor (AQEG, 2004; Donnelly et al., 2012). Meteorological conditions determine NO_2 levels by controlling the extent to which emissions can disperse (AQEG, 2004). For example, lower boundary layer heights and temperature inversions during winter can prevent pollution, including NO_2 , from ventilating from the boundary layer, leading to higher wintertime NO_2 concentrations in European cities. In contrast, in summer, daytime boundary layer heights are generally higher, and there is greater solar activity which reduces NO_2 concentrations in some locations through NO_2 photolysis (Henschel et al., 2016). Additionally, the built environment also determines the extent to which road transport NO_x emissions (or from other sources) can escape the street canyon, or immediate environment within which they are emitted. At background monitoring sites in rural locations, longer range transport can also determine NO_2 concentrations (Donnelly et al., 2012).

30 The combination of emission, chemistry and meteorology can drive substantial variation in the distributions of hourly NO_2 concentrations across the year and across the day that underpin any given annual average NO_2 concentration. The aim of this work was to investigate the spatial and temporal variation in the patterns of hourly NO_2 concentrations producing annual average NO_2 concentrations across Europe, utilising data from thousands of European NO_2 monitoring sites. At each site, a standard set of statistics were calculated to quantify the magnitude of an ‘impact’ metric, i.e. annual average NO_2 concentrations, as well as the contribution to annual average NO_2 concentrations from each month of the year, hour of the day, and from hourly NO_2 concentrations (hourly concentrations were binned to concentration ranges in this study to investigate

the influence of relatively low, moderate, or high hourly NO₂ concentrations in determining annual NO₂ concentrations). This method of integrating analysis of an impact metric with the variation in atmospheric composition producing it has been outlined previously as a ‘chemical climatology’ framework (Malley et al., 2014a), and applied to assess the conditions producing long-term health relevant PM impacts, and ozone health and vegetation impacts, at the two UK European Monitoring and Evaluation Programme (EMEP) monitoring sites (Malley et al., 2015, 2016). The application here of the chemical climatology approach to assessment of long-term health-relevant NO₂ is the first application of this methodology to a regional monitoring network encompassing thousands of monitoring sites.

Europe-wide NO₂ concentrations have been examined previously. For example, Cyrus et al. (2012) investigated variation in annual NO₂ concentrations across 36 European cities/regions to assess spatial differences in estimated NO₂ exposure for use in epidemiological analysis of air pollution health impacts. Henschel et al. (2016) used monitoring data to assess trends (1999–2010) in NO₂ concentrations in relation to the implementation of the Euro NO_x emission standards in 9 cities in 7 countries. In addition, studies focussing on one city, country, or region of Europe have also been undertaken (AQEG, 2004; Braniš, 2009; Carslaw et al., 2011; Carugno et al., 2016; Chaloulakou et al., 2008; Clapp and Jenkin, 2001; Cuevas et al., 2014; Donnelly et al., 2012, 2011; Gilge et al., 2010; Jenkin, 2014; Lozano et al., 2009; Mavroidis and Chaloulakou, 2011). These studies differ in the methods for analysis of NO₂ measurements, for example Henschel et al. (2016) select one traffic and urban background site to represent all traffic and urban background conditions in that city, while Cyrus et al. (2012) identify greater variation in annual NO₂ concentrations within each city than between cities but did not assess variation in short-term NO₂ concentration contributions to annual concentrations due to the two-week resolution of the measurements. This work builds on these previous studies by presenting a framework for the common characterisation of not only the magnitude of annual average NO₂ across monitoring sites, but also the variation in hourly NO₂ concentrations that give rise to it (Figure 1, [see Section 2.1](#)), and by applying this framework across all sites with available data. This common method of data interpretation allows the relevance of the results from previous studies assessing NO₂ variability in one location to be evaluated for other locations. Similarly, the identification of the variability in hourly NO₂ concentration contributions to annual average NO₂ highlights differences in the likely effectiveness of reducing annual NO₂ concentrations through the implementation of mitigation strategies which may differently affect hourly NO₂ concentrations during particular times across the year, and across the day, or which focus on infrequent peak or frequent moderate hourly NO₂ concentrations. Finally, the application of this chemical climatology approach to monitoring data (or modelled data) in other regions would also allow for a consistent comparison of the variability in hourly NO₂ contributions to annual average NO₂ between regions.

30 **2 Methods**

Data for this analysis was obtained from the AirBase European public data repository (<https://www.eea.europa.eu/data-and-maps/data/aqreporting-2>). Under the EU Air Quality Directive, NO₂ must be measured according to the European Standard,

EN 14211, 2005 (European Council Directive 2008/50/EC, 2008). This is a chemiluminescence technique involving the reduction of NO₂ to NO using heated (300–350 °C) molybdenum surfaces followed by the gas-phase reaction between NO and O₃. Other measurement methods, including those used in Local Authority air quality monitoring and European Monitoring and Evaluation Programme (EMEP) monitoring, include low time resolution NO₂ diffusion tube measurement and higher temporal resolution photolytic converter-chemiluminescence NO₂ measurements, respectively. Both diffusion tubes and molybdenum converter-chemiluminescence analysers have positive biases in the measurements due to non-specific NO₂ chemical measurement (Villena et al., 2012). The photolytic method has fewer artefacts but at very low NO₂ concentration may have interferences from PAN (Reed et al., 2016). In total, data from 3162 sites monitoring between 2000 and 2014 was assessed against the data capture criteria defined for this analysis (details below). Of the sites which met the data capture criteria, 79% measured NO₂ by chemiluminescence analysers (with molybdenum-converters), 1% by Differential Optical Absorption Spectroscopy (DOAS), and for the remainder the measurement method was listed as unknown. Data from all available sites in the AirBase repository were used in this analysis, as these measurements are used to assess compliance with the EU air quality limit values. The metadata associated with each monitoring site includes its location, and its classification in terms of general area representativeness (urban, suburban or rural) and station type (traffic, industrial or background) (European Commission, 2013; JRC-AQUILA, 2013). The combination of area and station type produce the classifications used in this paper (e.g. rural background, urban traffic). [European monitoring site classification has been evaluated previously](#) (Flemming et al., 2005; Joly and Peuch, 2012; Spangl et al., 2007).

2.1 Chemical climate statistics

The rationale and framework of a ‘chemical climatology’ analysis were outlined in Malley et al. (2014a). The aim is to link a specific impact of atmospheric composition to the conditions that produce it. To this end a chemical climate consists of three elements for which standard sets of statistics are calculated for each site (Figure 1). First, an *impact* of atmospheric composition is selected as the focus of the chemical climate. Here the impact is the risk of health effects from long-term exposure to NO₂, as quantified through annual average NO₂ concentration (henceforth abbreviated to NO_{2AA}), in line with the EU regulatory guidelines. The *state* of atmospheric composition describes the underlying distributions of hourly NO₂ concentrations that contribute to NO_{2AA}. For this analysis, the percent contribution to NO_{2AA} from each month of the year and hour of the day was calculated. In addition, hourly NO₂ concentrations for each year at each site were divided into 5 µg m⁻³ bins (i.e. 0-5, 5-10, 10-15... µg m⁻³) and the percent contribution to NO_{2AA} from each of these bins was calculated to investigate the relative contribution of low, moderate and high NO₂ concentrations to NO_{2AA}. These state statistics help to indicate likely causal *drivers* producing the impact metric, which, for NO_{2AA}, include emissions, meteorology and atmospheric chemistry (see Section 4). To characterise present-day health-relevant NO₂ ‘chemical climates’ across Europe, but to average out the influence of any inter-annual variation in annual NO₂ concentrations, the statistics shown in Figure 1 were averaged between 2010 and 2014. Long-term trends in the statistics were evaluated for the period 2000 and 2014, for sites with sufficient data capture during this 15-year period.

At each site $\text{NO}_{2\text{AA}}$ was calculated if there were >75% hourly observations in that year. The 2010-2014 average was calculated if more than three of the five years had valid $\text{NO}_{2\text{AA}}$ values. These sites were then grouped into $10 \mu\text{g m}^{-3}$ bins of average 2010-2014 $\text{NO}_{2\text{AA}}$ concentration, in order separately to evaluate the conditions associated with very low $\text{NO}_{2\text{AA}}$ concentrations (i.e. between 0 and $10 \mu\text{g m}^{-3}$), through intermediate $\text{NO}_{2\text{AA}}$ concentrations (i.e. $10\text{-}20 \mu\text{g m}^{-3}$, $20\text{-}30 \mu\text{g m}^{-3}$...), to the highest observed $\text{NO}_{2\text{AA}}$ concentrations ($>80 \mu\text{g m}^{-3}$).

The data capture for each hour of the day, and each month of the year was also calculated for each year at each site. For those sites in each 2010-2014 $\text{NO}_{2\text{AA}}$ bins with >75% data capture in each month and each hour, the percent contributions of each hourly NO_2 concentrations in each of the 12 months, in each of the 24 hours hour-of the day, and from hourly NO_2 concentrations across the year in $5 \mu\text{g m}^{-3}$ bins (i.e. from hourly NO_2 concentrations between $0\text{-}5 \mu\text{g m}^{-3}$, $5\text{-}10 \mu\text{g m}^{-3}$... $195\text{-}200 \mu\text{g m}^{-3}$, $>200 \mu\text{g m}^{-3}$) to $\text{NO}_{2\text{AA}}$ were calculated. These sites were further analysed to identify the major differences in the percentage contribution to $\text{NO}_{2\text{AA}}$ from each month of the year, hour of the day, and hourly NO_2 concentrations in $5 \mu\text{g m}^{-3}$ bins-state-statistics across sites with similar 2010-2014 $\text{NO}_{2\text{AA}}$ values. The relatively low number of sites in the $70\text{-}80 \mu\text{g m}^{-3}$ and $>80 \mu\text{g m}^{-3}$ bins meant that variation in these statistics was examined across the sites individually. However, for the lower $\text{NO}_{2\text{AA}}$ concentration bins, which contained up to several hundred sites, Ward's hierarchical cluster analysis was used to group sites according to similarity in 77 variables, i.e. the percent contribution to 2010-2014 $\text{NO}_{2\text{AA}}$ of different months (12 values), hours of the day (24 values), and $5 \mu\text{g m}^{-3}$ hourly NO_2 concentration bins to $\text{NO}_{2\text{AA}}$ (41 values, percent contribution from hourly NO_2 concentrations $0\text{-}5 \mu\text{g m}^{-3}$, $5\text{-}10 \mu\text{g m}^{-3}$... $195\text{-}200 \mu\text{g m}^{-3}$, $>200 \mu\text{g m}^{-3}$). In this process, each monitoring site initially constituted its own cluster. Clusters were then merged based on the similarity of the 2010-2014 average values of the percent contribution to $\text{NO}_{2\text{AA}}$ from January-December, from 00:00 to 23:00, and from each $5 \mu\text{g m}^{-3}$ hourly NO_2 concentration bin (i.e. the two most 'similar' clusters were merged). Euclidean distance was used as the measure of the similarity between these variables between clusters. At each step, the two clusters were merged which results in the smallest increase in within-cluster variance, and this was repeated until all sites are contained within a single cluster (Kaufman and Rousseeuw, 1990; Ward, 1963). Other hierarchical and non-hierarchical clustering algorithms and distance measures could have been used in place of the Ward's hierarchical algorithm, and the Euclidean distance, respectively, that were applied in this work. Selection of alternative methods could influence the clustering result (Kaufman and Rousseeuw, 1990), but the Ward's hierarchical cluster algorithm applied here has been shown here to perform effectively compared with alternative approaches, including with respect to the dispersion of observations, the presence of outliers, and irrelevant variables (Mangiameli et al., 1996).

The clustering of monitoring sites in each $10 \mu\text{g m}^{-3}$ 2010-2014 $\text{NO}_{2\text{AA}}$ bin is process is represented by a dendrogram, which was then 'cut' at an appropriate level to determine the number of clusters, and the grouping of sites. The cut-off point was determined from the visual identification of 'elbows' in the plots of the proportion of variance explained against the number

of clusters into which sites were divided. The aim in selecting the number of clusters of sites was to maximise the explained variance in the percentage contribution of each month, hour of day, and hourly NO₂ concentrations in 5 µg m⁻³ bins to NO_{2AA} between sites with as few clusters as possible, ~~and therefore~~, for each 10 µg m⁻³ 2010-2014 NO_{2AA} bin, the point at which increasing the number of clusters resulted in relatively small increases in the proportion of variance explained was identified.

5 The cut-off points selected grouped sites into between 5 and 8 clusters, ~~which explained between 57 and 75% of total between-site variance.~~ A key feature of the clustering result is the proportion of variance across sites in the percent contribution to NO_{2AA} from months, hours of the day and hourly NO₂ concentration bins that is explained by dividing sites into a certain number of clusters. The clusters produced in this work explained between 57 and 75% of the variance between sites in the 10 µg m⁻³ 2010-2014 NO_{2AA} bin ranges, as shown in Figures ~~S1S2-S7~~. This means that there remains between 43% and 25% variability in these variables between sites with similar 2010-2014 NO_{2AA} that is not explained by dividing sites into the clusters produced in this analysis. Other key aspects of the clustering are the location of sites included in each cluster, and the similarities and differences in the value of each variable (i.e. the percent contribution from each month, hour of day and 5 µg m⁻³ hourly NO₂ concentration bins) between clusters. These aspects are described in Section 3.2.

15 Finally, sites in each of these groups which had long-running NO₂ measurements between 2000 and 2014 were identified. Across the 15-year period, a valid statistic (i.e. that met the data capture criteria outlined above) was required in at least 12 of the 15 years for the trend in that statistic to be estimated. The magnitude, direction and significance of trends in NO_{2AA}, and in month, hour of day and hourly concentration contributions to NO_{2AA} were calculated using two methods that accounted for autocorrelation in each statistic, but differed in assumptions about the type of underlying distribution of each statistic at each
20 site. A number of parametric and non-parametric trend analysis methods were used to assess consistency of the results when different methods typically found in the literature for environmental trend analysis were applied.

Firstly, the magnitude and direction of the trend was estimated using the non-parametric Theil-Sen statistic, i.e. the median of the trends between all pairs of data points (Theil, 1950a, 1950b, 1950c). The significance of the trend was assessed using block
25 bootstrap resampling to account for autocorrelation in each statistic at each site, using algorithms implemented in the Openair package within the R statistical software (Carslaw and Ropkins, 2012; R Core Team, 2016). The block length used was $n^{1/3}$ (n = length of time series), the default in the Openair package based on the analysis in Kunsch et al. (1989). Secondly, a parametric first-order auto-regressive (AR(1)) model was fitted to each statistic at each site, which also accounts for autocorrelation in the data (Box and Jenkins, 1976).

30 Trend estimations that do not account for autocorrelation in the data were also used. The non-parametric Mann-Kendall statistic was used to estimate the significance of the trend in each statistic at each site (Mann, 1945), in combination with the Theil-Sen statistic for the magnitude and direction of the trend. Finally, trends were also estimated using parametric ordinary least-squares linear regression. These two methods did not account for autocorrelation in the trend estimation, and were therefore

used to compare the consistency of results obtained with the non-parametric (Theil Sen/black bootstrap) and parametric (AR(1)) methods that did account for auto-correlation. All methods used in this work only estimated linear trends, and trends with $p \leq 0.05$ were designated as statistically significant. Trends between 2000 and 2014 were assessed by grouping all qualifying sites based on the 2010-2014 NO_{2AA} value at each site. These groupings were used to investigate whether changes in drivers of NO₂ variability (e.g. European-wide NO_x emission reductions) had had different effects at sites where the magnitude of annual NO₂ concentrations during the most recent period (2010-2014) differed.

3 Results

3.1 Impact metric: Annual average NO₂ concentration

Across Europe, 2587 sites met the data availability criteria to determine the 2010-2014 NO_{2AA} concentration (Table 1), of which 8% exceeded the 40 µg m⁻³ EU annual NO₂ limit value (Figure 2). These sites were spread across Europe with 22 out of the 34 countries represented having at least one site where 2010-2014 NO_{2AA} exceeded 40 µg m⁻³. Most of these exceedances were at urban traffic sites (77%), consistent with previous analysis (EEA, 2016a). The 2010-2014 NO_{2AA} concentration exceeded 40 µg m⁻³ at over 50% of urban traffic sites in Belgium, Germany, Italy, Norway, Serbia, Switzerland, and at over 40% of such sites in France, Greece, Portugal and the UK. There were also 20 urban background sites spread across cities in the UK (London, Manchester), Germany (Frankfurt), France (Paris), Spain (Madrid, Barcelona) and Italy (Rome, Monza) where 2010-2014 NO_{2AA} was > 40 µg m⁻³. However, most urban background sites had 2010-2014 NO_{2AA} concentrations between 10-20 (34% of all urban background sites), 20-30 (44%) and 30-40 µg m⁻³ (18%). In contrast, most sites with the lowest concentrations were rural background sites; 61% of sites with 2010-2014 NO_{2AA} concentrations in the 0-10 µg m⁻³ bin were rural background, and 93% of all rural background sites had NO_{2AA} concentrations < 20 µg m⁻³.

3.2 State: Monthly, hour of day and hourly concentration bin contribution to annual average NO₂

Of the 2587 sites that had sufficient data capture to calculate 2010-2014 NO_{2AA} concentrations, 2198 (85%) had sufficient data capture (>75%) in each month of the year and each hour of the day to assess hourly NO₂ contributions to NO_{2AA} (Table 2). This included seven of the eight sites with 2010-2014 NO_{2AA} concentrations > 80 µg m⁻³, and five of the seven sites with NO_{2AA} between 70-80 µg m⁻³. As described in Section 2, the large number of sites with 2010-2014 NO_{2AA} in each of the lower concentration bins were clustered according to similarity in monthly, hour of day, and hourly NO₂ concentration bin contributions to NO_{2AA} (summarised in Table 2). The proportion of sites located in different regions of Europe assigned to each cluster is shown in Figure 3 (regions are defined in Figure S38), and the locations of sites in each cluster for each NO_{2AA} bin are shown in Figures S49-S1045. For the majority of clusters across the 2010-2014 NO_{2AA} bins, the sites were geographically widespread across Europe. This indicates that the distinct variations in monthly, hour of day, and hourly NO₂ concentration bin contributions to NO_{2AA} identified by the cluster analysis are not confined to specific European regions, e.g. for the 60-70 µg m⁻³ 2010-2014 NO_{2AA} bin, sites in Central Europe made up a substantial proportion of sites in four of the five

clusters (Figure 3). However, there were also clusters that did not contain sites in particular Europe regions, and clusters that were dominated by sites from one region. Most notably there were clusters in each $\text{NO}_{2\text{AA}}$ bin between $20\text{-}30\ \mu\text{g m}^{-3}$ and $60\text{-}70\ \mu\text{g m}^{-3}$ that predominantly contained sites from northern Italy and neighbouring areas. The distinct variations in the contribution of months, hour of day and hourly NO_2 concentration bins that lead to cluster separation have been summarised as (i) the differences in seasonal contribution to $\text{NO}_{2\text{AA}}$, as shown in Figures 4a and 4b, (ii) the contribution from the 6 hours with the largest contribution to $\text{NO}_{2\text{AA}}$, and the 6 hours with the lowest contribution to $\text{NO}_{2\text{AA}}$, as a proxy for the amplitude of diurnal variation in hour of day contributions to $\text{NO}_{2\text{AA}}$ (Figure 5), and (iii) the contribution from the five, ten, and fifteen $5\ \mu\text{g m}^{-3}$ hourly NO_2 concentration bins with the largest contribution to $\text{NO}_{2\text{AA}}$, as a proxy for the breadth of hourly NO_2 concentrations that make a substantial contribution to $\text{NO}_{2\text{AA}}$ (Figure 6).

10

The majority of sites where 2010-2014 $\text{NO}_{2\text{AA}}$ exceeded $40\ \mu\text{g m}^{-3}$ were between 40 and $70\ \mu\text{g m}^{-3}$. However, analysis of the conditions producing 2010-2014 $\text{NO}_{2\text{AA}}$ at the sites with highest 2010-2014 $\text{NO}_{2\text{AA}}$ (i.e. $> 70\ \mu\text{g m}^{-3}$) illustrates patterns that also occurred across the majority of sites that exceeded the EU limit value. All sites with highest 2010-2014 $\text{NO}_{2\text{AA}}$ concentration ($> 80\ \mu\text{g m}^{-3}$) lacked seasonal variation in the contribution of hourly NO_2 concentrations to $\text{NO}_{2\text{AA}}$ (Figure 4a) and had largest contributions to $\text{NO}_{2\text{AA}}$ from daytime hours, with morning and evening peaks, reflecting their urban traffic classification. Whilst the 30-35% contribution to $\text{NO}_{2\text{AA}}$ from peak hours (Figure 5), defined as the six hours with largest contribution to $\text{NO}_{2\text{AA}}$, was similar across these sites, the timing of the peak periods differed, e.g. due to differences in the magnitude of the morning and evening peaks in NO_2 concentration. For example, the top 6 hours occurred exclusively during the afternoon/evening at two of the highest $\text{NO}_{2\text{AA}}$ sites, 5 of the 6 peak hours occurred in the evening at two sites, 4 out of 6 peak hours occurred in the evening at one site, and two sites had an equal number of top 6 hours during the morning and evening (Figure 7).

In addition, the contribution from different hourly NO_2 concentration bins to $\text{NO}_{2\text{AA}}$ varied across these sites with highest $\text{NO}_{2\text{AA}}$. At the three French sites, a narrower range of $5\ \mu\text{g m}^{-3}$ concentration bins made substantial contributions to $\text{NO}_{2\text{AA}}$ compared to the three German and one UK sites. This is illustrated by the larger contribution of the top five contributing $5\ \mu\text{g m}^{-3}$ hourly NO_2 bins (27-34% of $\text{NO}_{2\text{AA}}$) at the three French sites, compared to the German and UK sites (21-22%) (Figure 6). Consequently, while $\text{NO}_{2\text{AA}}$ was similar across these sites, the frequency of short-term peak NO_2 concentrations was not. Hourly NO_2 concentrations above the $200\ \mu\text{g m}^{-3}$ hourly NO_2 EU limit value contributed 6% to $\text{NO}_{2\text{AA}}$ at GB0682A, and 2-4% at the three German sites, compared with only 0.3-1% at the French sites. Controlling hourly EU limit value exceedances would therefore have a greater impact on reducing $\text{NO}_{2\text{AA}}$ at the UK and German sites compared with the French sites. Conversely, reduction in the relatively moderate hourly NO_2 concentrations ($60\text{-}120\ \mu\text{g m}^{-3}$) would yield greater reductions in $\text{NO}_{2\text{AA}}$ at these French sites (where contribution from $60\text{-}120\ \mu\text{g m}^{-3}$ to $\text{NO}_{2\text{AA}}$ varied between 60 and 68%) compared to the German and UK sites (45-50%). The patterns of variation in contribution from months, hour of day, and hourly NO_2

concentrations at the sites with 2010-2014 $\text{NO}_{2\text{AA}}$ between 70 and 80 $\mu\text{g m}^{-3}$ were broadly similar to those identified for the $>80 \mu\text{g m}^{-3}$ sites (Figures 4-6).

As shown in Section 3.1 the majority of sites where 2010-2014 $\text{NO}_{2\text{AA}}$ exceeded 40 $\mu\text{g m}^{-3}$ were between 40 and 70 $\mu\text{g m}^{-3}$.

5 For sites with 2010-2014 $\text{NO}_{2\text{AA}}$ in the range 60-70 $\mu\text{g m}^{-3}$, most were urban traffic sites (Figure 2), and 61% were grouped into two of the five clusters identified (Cluster 2 and Cluster 3). These sites were geographically widespread (Figure 3, Figure S49), and had similar patterns of hourly NO_2 contribution to $\text{NO}_{2\text{AA}}$ as at sites with higher 2010-2014 $\text{NO}_{2\text{AA}}$. There was little variation in seasonal contribution to $\text{NO}_{2\text{AA}}$ (Figure 4a), a 30-35% contribution in the top 6 hours of the day to $\text{NO}_{2\text{AA}}$, and similar variation in contribution of hourly NO_2 concentration bins to $\text{NO}_{2\text{AA}}$. Cluster 2 sites had lower diurnal variation in the contribution to $\text{NO}_{2\text{AA}}$, as shown by a relatively lower contribution from the peak 6 hours of the day to $\text{NO}_{2\text{AA}}$ (31% compared with 34% for Cluster 3), and a higher contribution from the minimum 6 hours of the day compared with Cluster 3 (17% vs 13%). In addition, sites in Cluster 2 were determined by a relatively narrower set of hourly NO_2 concentrations, with the top five contributing 5 $\mu\text{g m}^{-3}$ hourly NO_2 bins contributing 34% of $\text{NO}_{2\text{AA}}$, compared with only 28% of $\text{NO}_{2\text{AA}}$ for Cluster 3 sites (Figure 6).

15

Other clusters contained sites with 2010-2014 $\text{NO}_{2\text{AA}}$ between 60 and 70 $\mu\text{g m}^{-3}$ where the conditions producing $\text{NO}_{2\text{AA}}$ were distinct to those sites with higher $\text{NO}_{2\text{AA}}$ (and sites in Clusters 2 & 3). For example, the three sites in Cluster 1 were in northern Italy and western Austria, and had substantial seasonal variation in contribution to $\text{NO}_{2\text{AA}}$ (Figure 4a), with 33% of $\text{NO}_{2\text{AA}}$ derived during winter, and only 18% during summer. There was also less diurnal variation in contribution to $\text{NO}_{2\text{AA}}$ than for sites in Cluster 2 and 3, with the peak 6 hours contributing 30%, and the minimum 6 hours 20% (Figure 5). In contrast, the distinction between the three sites in Cluster 4, and others was the larger diurnal variation in contribution to $\text{NO}_{2\text{AA}}$, with the peak 6 hours contributing 34%, and the minimum 6 hours only 11%. The Cluster 4 sites, in southern Germany and northern Italy, also had the broadest range of concentrations contributing to $\text{NO}_{2\text{AA}}$, with only 24% of $\text{NO}_{2\text{AA}}$ accounted for by the top five 5 $\mu\text{g m}^{-3}$ hourly NO_2 bins.

25

Sites with $\text{NO}_{2\text{AA}}$ in the range 50-60 $\mu\text{g m}^{-3}$ were grouped into 8 clusters, with the majority (85%) grouped into four clusters (Table 2). As for the 60-70 $\mu\text{g m}^{-3}$ sites, one cluster (Cluster 2) included sites predominantly in northern Italy (as well as Switzerland, southern France, and Hungary), and which had substantial seasonal variation in contribution to $\text{NO}_{2\text{AA}}$ (Figure 4a). In contrast the sites in the other 3 major clusters (Clusters 1, 3, and 4, containing 66% of sites) were more similar to those sites with higher $\text{NO}_{2\text{AA}}$ ($>70 \mu\text{g m}^{-3}$), with little seasonal variation in contribution to $\text{NO}_{2\text{AA}}$. The distinctions between sites in these three clusters were the extent of variation in diurnal contribution to $\text{NO}_{2\text{AA}}$ (Cluster 1 sites had the smallest difference in contribution from the peak and minimum 6 hours, Cluster 3 had the largest (Figure 5)), and range of hourly NO_2 concentrations that contributed to $\text{NO}_{2\text{AA}}$ (Cluster 1 sites had the narrowest range, Cluster 3 sites the widest). Sites with 2010-2014 $\text{NO}_{2\text{AA}}$

30

between 50 and 60 $\mu\text{g m}^{-3}$ in each of these major clusters were quite geographically widespread, and located in cities throughout Europe.

For sites with 2010-2014 $\text{NO}_{2\text{AA}}$ between 40 and 50 $\mu\text{g m}^{-3}$, there was also a cluster (Cluster 2) with sites in northern Italy and Austria and which had the largest seasonal variation in contribution to $\text{NO}_{2\text{AA}}$ (Figure 4a). Sites in other clusters (Clusters 3 and 5) also had relatively large winter contribution to $\text{NO}_{2\text{AA}}$ compared with most sites with higher $\text{NO}_{2\text{AA}}$, and these sites were located across central and southern Europe, rather than confined to northern Italy (Figure 3, Figure S611). However, many sites with $\text{NO}_{2\text{AA}}$ in this range, spread across a large number of countries, also showed little seasonal variation in contribution to $\text{NO}_{2\text{AA}}$. These sites grouped in Clusters 1, 4 and 6, and were distinguished by relatively low diurnal variation and narrower range of hourly NO_2 concentrations determining $\text{NO}_{2\text{AA}}$ for sites in Clusters 1 and 6 compared to sites in Cluster 4.

Although at the majority of sites the 2010-2014 $\text{NO}_{2\text{AA}}$ concentration did not exceed the 40 $\mu\text{g m}^{-3}$ EU limit value, negative health outcomes may be associated with lower concentrations (REVIHAAP, 2013). For $\text{NO}_{2\text{AA}} < 40 \mu\text{g m}^{-3}$, sites across all clusters tended to have larger contributions from winter to $\text{NO}_{2\text{AA}}$ and lower contributions from summer. As expected, as $\text{NO}_{2\text{AA}}$ decreased there was also a narrower range of hourly NO_2 concentrations that contributed to $\text{NO}_{2\text{AA}}$, and therefore the top 5, 10 and 15 largest contributing concentration bins contributed a larger proportion of $\text{NO}_{2\text{AA}}$ than at sites with higher $\text{NO}_{2\text{AA}}$ (Figure 6). There were also clusters of sites in specific locations with distinct hourly NO_2 concentration contributions to $\text{NO}_{2\text{AA}}$. For example, in the 30-40 $\mu\text{g m}^{-3}$ and 20-30 $\mu\text{g m}^{-3}$ 2010-2014 $\text{NO}_{2\text{AA}}$ bins, there were clusters containing sites in northern Italy and neighbouring areas (Cluster 1 and Cluster 6, respectively) that had the largest seasonal difference in contribution to $\text{NO}_{2\text{AA}}$ across all clusters in these 2010-2014 $\text{NO}_{2\text{AA}}$ bins. Similar to the sites with higher $\text{NO}_{2\text{AA}}$ concentrations, the clustering of sites with 2010-2014 $\text{NO}_{2\text{AA}} < 40 \mu\text{g m}^{-3}$ grouped sites from across wide geographic areas based on (i) a relatively narrow distribution of hourly NO_2 concentrations contributing to $\text{NO}_{2\text{AA}}$ (e.g. 30-40 $\mu\text{g m}^{-3}$, Cluster 2; 20-30 $\mu\text{g m}^{-3}$, Cluster 4, Figure 6), or (ii) a relatively large diurnal variation in contribution to $\text{NO}_{2\text{AA}}$ (e.g. 30-40 $\mu\text{g m}^{-3}$, Cluster 4 and Cluster 7; 20-30 $\mu\text{g m}^{-3}$, Cluster 4, Figure 5). In addition, the contribution from morning and evening hours to those hours of the day with peak contribution also varied, as was the case for the sites with higher $\text{NO}_{2\text{AA}}$ (Figure 7).

3.3 Long-term trends

Trends between 2000 and 2014 were calculated at 259 sites. The estimated trends using parametric and non-parametric approaches, and methods which did and did not account for autocorrelation produced generally consistent results for the changes in $\text{NO}_{2\text{AA}}$ and monthly, hour of day, and hourly NO_2 concentration bin contributions to $\text{NO}_{2\text{AA}}$. Figure 8 shows the proportion of sites with significant decreasing, increasing and non-significant trends in $\text{NO}_{2\text{AA}}$ between 2000 and 2014. The proportion of sites in each 2010-2014 $\text{NO}_{2\text{AA}}$ bin with significant decreasing trends was similar for the Theil-Sen/Mann-Kendall (Figure 8a), linear regression (Figure 8b) and first-order autoregressive (AR(1)) (Figure 8d) methods, and 20-30%

lower for the Theil-Sen/block bootstrap method (Figure 8c). However, the magnitude and direction of the trends in $\text{NO}_{2\text{AA}}$ across all sites were very similar for the Theil-Sen/block bootstrap or AR(1) methods (Figure S116). The spatial distribution of 2000-2014 changes in $\text{NO}_{2\text{AA}}$ is shown in Figure 9, and the proportion of sites in each 2010-2014 $\text{NO}_{2\text{AA}}$ bins with significant decreasing, increasing ($p \leq 0.05$), and non-significant trends in monthly, hour of day, and hourly NO_2 concentration bin contribution to $\text{NO}_{2\text{AA}}$ are shown in Figures 10, 11 and 12, respectively. These figures show the parametric AR(1)-estimated trends because, in contrast to linear regression, autocorrelation was adjusted for using this method. The corresponding summaries of the non-parametric Theil-Sen/block bootstrap-derived trend estimates are provided in supplemental information for comparison, and show broadly similar patterns (Figures S127-S1529). The statistics referred to below are from the AR(1) trend estimation.

The estimated trend in $\text{NO}_{2\text{AA}}$ was negative at the majority of sites. For sites where 2010-2014 $\text{NO}_{2\text{AA}}$ exceeded $40 \mu\text{g m}^{-3}$, the trend was not statistically significant at a larger proportion of sites (54%, i.e. 14 out of 27 sites) than those with lower 2010-2014 $\text{NO}_{2\text{AA}}$ (Figure 8d). Sites with 2010-2014 $\text{NO}_{2\text{AA}}$ concentrations in the ranges 30-40, 20-30, 10-20, and 0-10 $\mu\text{g m}^{-3}$ had significant decreasing trends in $\text{NO}_{2\text{AA}}$ at 68%, 74%, 69% and 68% of sites, respectively (Figure 8). Although most sites had a non-significant change in monthly contribution to $\text{NO}_{2\text{AA}}$ between 2000 and 2014, there was a larger proportion of sites with a significant decreasing trend in the contribution to $\text{NO}_{2\text{AA}}$ from summer months, in particular from June, and significant increasing contribution from winter months (Figure 10). The predominant pattern of change in hour of day contribution to $\text{NO}_{2\text{AA}}$ was an increasing contribution from 'peak' morning and evening hours, and a decreasing contribution to $\text{NO}_{2\text{AA}}$ from night-time hours, and hours in the middle of the day (Figure 11). For sites in the lowest 2010-2014 $\text{NO}_{2\text{AA}}$ category (0-10 $\mu\text{g m}^{-3}$), a smaller proportion of sites had significant increasing or decreasing trends in the contribution across the day compared to sites with highest 2010-2014 $\text{NO}_{2\text{AA}}$. Finally, across all 2010-2014 $\text{NO}_{2\text{AA}}$ bins, the predominant pattern of change in the contribution of hourly NO_2 concentrations to $\text{NO}_{2\text{AA}}$ between 2000 and 2014 was an increasing contribution from relatively low hourly NO_2 concentrations, and a decreasing contribution from relatively high hourly NO_2 concentrations (Figure 12).

4 Discussion

The calculation of a standard set of statistics for thousands of monitoring sites showed that sites with similar $\text{NO}_{2\text{AA}}$ differed in the contribution from hourly NO_2 concentrations during different months, hours of the day, and across the hourly NO_2 concentration distribution. Clustering of sites based on these distinct contributions to $\text{NO}_{2\text{AA}}$ showed that sites with similar $\text{NO}_{2\text{AA}}$ located relatively close to one another were often grouped into different clusters, and that sites relatively far from one another were often grouped within the same clusters (Figure 3). The observations that multiple distinct patterns contributed to $\text{NO}_{2\text{AA}}$ in a particular European region builds on analysis by Cyrus et al. (2012) showing that variation in $\text{NO}_{2\text{AA}}$ within a particular city was greater than variation in $\text{NO}_{2\text{AA}}$ across the 36 cities analysed. The present study shows that there can also

be variation in the contributions from months, hours of the day, and hourly NO₂ concentration bins to NO_{2AA} within relatively small geographic areas across Europe

4.1 Drivers: Links to meteorology, emissions, and chemistry

The specific location of a monitoring site determines the relative contributions of meteorological conditions, atmospheric chemistry and NO_x emission sources to the NO_{2AA} concentration. In this work, sites with 2010-2014 NO_{2AA} above 40 μg m⁻³ were almost all urban traffic sites (JRC-AQUILA, 2013). At these sites, the lack of a seasonal pattern in contribution to NO_{2AA} indicates that the major driver for NO_{2AA} is the NO_x emission sources in close proximity to the site. In addition to primary NO₂ emissions, the presence of O₃ at these urban traffic sites will also lead to secondary formation of NO₂ (from primary NO emissions). This reaction can proceed in a few seconds at polluted locations (AQEG, 2004), within the timescale of dispersion from source to roadside monitor, and suggests that long-range transport does not play the major role in driving exceedance of the NO_{2AA} EU limit value. These conditions occur in cities across Europe. Schaefer et al. (2006) also showed in Hanover, Germany, that pollutant concentrations closest to major roads in street canyons were substantially less correlated with boundary layer mixing height compared with urban background stations. Hence NO_{2AA} at these sites may also be less influenced than sites with lower NO_{2AA} by seasonal differences in boundary layer height, which are generally lower during the day in winter (Seidel et al., 2012). Reductions in hourly NO₂ concentrations across the whole year, e.g. through local NO_x emission reductions, would therefore yield similar benefits in reducing NO_{2AA} for the majority of sites exceeding the EU NO₂ annual limit value, rather than just in winter months, owing to the relatively equal contribution from hourly NO₂ concentrations during all seasons (Figure 4a). However, the extent of the reduction from reducing emissions during peak hours of the day, and from reducing the short-term peak hourly NO₂ concentrations, varies between sites, which may reflect variability between sites in meteorological conditions or in traffic flows (i.e. strength of the emission source in close proximity to the monitoring sites).

Some sites, generally located in northern Italy and neighbouring areas, with 2010-2014 NO_{2AA} between 40 and 70 μg m⁻³ had substantially larger contributions from winter months to NO_{2AA}. This indicates that in this region the photochemical conversion of NO₂ to NO and O₃ in summer may be a more important factor in determining NO₂ concentrations (i.e. lowering NO₂ concentration during summer) at these sites than at other sites across Europe with similar NO_{2AA} concentrations. Northern Italy is a heavily industrialised and relatively polluted region of Europe (Carugno et al., 2016; Cyrus et al., 2012; Hazenkamp-Von Arx et al., 2004). Large NO_x and VOC emissions in this region, combined with relatively high levels of solar radiation in summer have previously been factors attributed to high levels of summertime photochemical ozone production (Masiol et al., 2017). Hence reducing hourly NO₂ concentrations during the winter would be more effective in reducing NO_{2AA}, but reducing NO_x emissions during the summer could reduce photochemical O₃ production in northern Italy and neighbouring areas. In addition to changes in atmospheric photochemistry between seasons, higher wintertime NO₂ concentrations have also been attributed to higher emissions from residential heating, and seasonal changes in meteorological conditions (Henschel et al., 2016). For example, Palarz et al. (2017) calculated the frequency of temperature inversions at 00:00, 06:00, 12:00, and 18:00

UTC across Europe during different seasons between 1981 and 2015. In winter, southern Europe, including northern Italy, had the highest frequency of temperature inversions at 00:00, 06:00, and 18:00 across mainland Europe. However, in summer, temperature inversions were less common in northern Italy at 06:00 compared to south-western Europe (Iberian peninsula), and at 18:00 compared to south-eastern Europe, i.e. in summer meteorological conditions were such that temperature
5 inversions tended to dissipate earlier in northern Italy compared to the Iberian peninsula, and form later than in south-eastern Europe. This indicates that, compared to the rest of Europe, in northern Italy there may be a larger difference between winter and summer in the frequency of temperature inversions and the trapping of pollution close to the surface that may contribute to the large difference in seasonal contribution to $\text{NO}_{2\text{AA}}$ at sites in this region.

10 There was also a greater contribution from winter compared to summer months to $\text{NO}_{2\text{AA}}$ for the majority of sites across Europe with $\text{NO}_{2\text{AA}}$ below the EU limit. This likely reflects the greater distance of these sites from direct NO_x emission sources, given that most of them were not urban traffic sites. The longer timescales of dispersion becomes more important in controlling NO_2 concentrations at these sites, and during summer NO_2 can be photolysed to NO and O_3 during transport from source to receptor. Pollutant concentrations at urban sites further from emissions sources were also shown to be more dependent on boundary
15 layer height compared to traffic sites (Schäfer et al., 2006), which during the day are higher in summer compared to winter (Seidel et al., 2012), increasing the volume into which NO_x emissions can be dispersed.

The conditions producing $\text{NO}_{2\text{AA}}$ have changed since 2000, even though at some sites this has not resulted in a statistically significant change in $\text{NO}_{2\text{AA}}$. Between 2000 and 2014, NO_x emissions from all sources were estimated to have decreased by
20 39%, and by 46% for road transport specifically (EEA, 2016b). However, Carslaw et al. (2011, 2016) showed that road transport NO_x emission decreases may not have been as large as expected from the implementation of vehicle emission standards. Carslaw et al. (2016) showed at roadside sites in London that reductions in NO_2 concentrations had been smaller than decreases in NO_x concentrations between 1995 and 2015, and the NO_2/NO_x emission ratio increased from less than 5% in 1997, to 25% in 2010, and 15% in 2014. The lower ratio in 2014 was attributed to a reduction in NO_2/NO_x ratios in NO_x
25 emissions from heavy duty vehicles as a result of Euro IV and Euro V vehicle emissions standards. A similar increase in the NO_2/NO_x emission ratio was also shown across the EU by Grice et al. (2009). The lower than estimated reduction in NO_x emissions from road transport, and the increase in the fraction of NO_x emitted as NO_2 , may be factors contributing to non-significant trends in $\text{NO}_{2\text{AA}}$ estimated at some sites, especially at the urban traffic sites with highest 2010-2014 $\text{NO}_{2\text{AA}}$.

30 The major changes in the hourly NO_2 contributions to $\text{NO}_{2\text{AA}}$ between 2000 and 2014 were a decreasing contribution from summer months, and an increasing contribution from winter months, an increasing contribution from hours during peak road traffic hours, and decreasing contributions from the highest hourly NO_2 concentrations. Henschel et al. (2016) analysed changes in seasonal NO_2 concentrations in 9 European cities and reported a decrease in summer NO_2 concentrations, and an increase in winter concentrations, which is consistent with the seasonal contribution trends calculated here. The present

analysis indicates that the factors contributing to enhanced wintertime NO₂ concentrations have become more important in determining NO_{2AA} between 2000 and 2014 at the majority of sites. This includes larger NO_x emissions from building heating during winter (Henschel et al., 2016), and meteorological conditions, with shallower daytime boundary layers (Seidel et al., 2012) and more frequent, deeper and stronger daytime temperature inversions in winter compared with summer (Palarz et al., 2017). Palarz et al. (2017) estimated that between 1981 and 2015, the frequency of night-time temperature inversions across most of mainland Europe had not changed significantly in winter or summer, and during summer, there was little change in the depth and strength of these inversions. In contrast, during winter, significant decreasing trends in the depth and strength of temperature inversions were calculated across most of mainland Europe. Reductions in the depth into which emissions can disperse during wintertime temperature inversions may contribute to the increasing contribution of hourly NO₂ concentrations during winter to annual NO₂ concentrations. However, Palarz et al. (2017) did not estimate trends in the frequency, depth and strength temperature inversions during the daytime, when the majority of NO_x emissions occur, and did not specifically assess changes between 2000 and 2014 for which trends were calculated in this work.

The increasing contribution to NO_{2AA} from rush-hour periods between 2000 and 2014 across sites with a wide range of 2010-2014 NO_{2AA} underlines the importance of local road transport NO_x emission sources in determining NO_{2AA} at monitoring sites across Europe. The increase in contribution to NO_{2AA} from rush-hour periods may reflect smaller decreases in NO₂ concentrations during these hours compared to other hours, lower than expected decreases in total road transport NO_x emissions during the period 2000-2014, as well as increases in the fraction of road transport NO_x emissions emitted as NO₂ (Carslaw et al., 2016).

Finally, the 2000-2014 trends indicate that conditions producing the highest hourly NO₂ concentrations have decreased in frequency. This may reflect decreases in NO_x emissions overall, but may also reflect changes in traffic patterns in particular cities (Baptista et al., 2014; CEBR, 2014; INRIX, 2016; Thomas, 2016), or meteorological conditions that concentrate NO₂ close to the monitoring sites. For example, Wagner et al. (2015) showed that in an urban street canyon there was a substantially stronger relationship between decreasing peak NO₂ concentrations and increasing mixing layer height than for mean NO₂ concentrations, for which there was no significant relationship. Therefore any increase in mixing layer height may be expected to disproportionately reduce highest hourly NO₂ concentrations as opposed to more moderate levels. Zhang et al. (2013) concluded that daytime (12:00 UTC) boundary layer height had increased between 1973 and 2010 based on daily radiosonde observations at 25 stations across Europe (average 73 m decade⁻¹), with no indication of seasonal differences.

Some studies have used satellite data to estimate trends in NO_{2AA} across Europe (Blond et al., 2007; Geddes et al., 2016; Konovalov et al., 2010; Schneider et al., 2015). Across Europe, Geddes et al. (2016) estimate an average 1996-2012 trend in annual NO₂ of -2.2 % y⁻¹. The current analysis decomposes the seasonal, hour of day, and hourly NO₂ concentration changes that have produced this average trend, but also highlights variation in the magnitude, and statistical significance of NO_{2AA}

trends across Europe. Many sites with 2010-2014 $\text{NO}_{2\text{AA}}$ exceeding the annual EU limit value had a non-significant trend in $\text{NO}_{2\text{AA}}$, which may not be reflected in the satellite derived trend estimates due to the spatial resolution of the satellite NO_2 estimates (e.g. 0.1×0.1 degrees in Geddes et al. (2016)).

4.2 Implications for policies aimed at reducing annual NO_2 concentrations across Europe

5 Previous studies have shown that specific mitigation actions aimed at reducing NO_2 concentrations can have varying effects in achieving this goal. For example, Jeanjean et al. (2017) modelled the effect of seven mitigation actions on mean NO_2 concentrations for a street in London, including additional trees lining the street, physical barriers between road and pedestrian areas, and paint applied to buildings. The individual implementation of these measures changed mean NO_2 concentrations beside roads between a 0.1% increase and a 7.4% decrease. Additionally, the effect of some mitigation measures on reducing
10 NO_2 concentrations during a particular season, hour of day, or on peak vs moderate concentrations can vary. For example, NO_x emission reduction from reducing indoor building temperatures disproportionately occur in winter when heating is required (Chiesa et al., 2014).

The analysis presented here does not specifically identify the likely effect of individual mitigation actions aimed at reducing
15 annual NO_2 concentrations on $\text{NO}_{2\text{AA}}$ at the different groups of sites identified. However, a greater understanding of the variability in hourly NO_2 concentration driving $\text{NO}_{2\text{AA}}$ at different sites, for which a methodology and practical application is the focus of this paper, increases the ability to assess how changes in hourly NO_2 variation at a site would affect $\text{NO}_{2\text{AA}}$, e.g. due to a particular mitigation action. For example, total traffic flows in cities in the Netherlands (Almelo), and in Greece (Athens) are lower in July and August compared with winter months, and the distribution of peak traffic flow also differ, with
20 these differences attributed to fewer commutes during the holiday period (Mavroidis and Ilia, 2012; Stathopoulos and Karlaftis, 2001; Thomas et al., 2008; Weijermars, 2007). Measures to reduce congestion in these cities may therefore disproportionately reduce traffic flows during the winter, when congestion is worse.

This work, focusing on linking hourly NO_2 variation to annual NO_2 concentrations, shows that there are sites across Europe
25 with similar $\text{NO}_{2\text{AA}}$, but with distinct variation in seasonal contributions. With this knowledge, the effect of such a policy which disproportionately reduces winter NO_x emissions may therefore be more effective at those sites within an $\text{NO}_{2\text{AA}}$ bin that have larger winter contributions to $\text{NO}_{2\text{AA}}$, and at those sites with relatively lower $\text{NO}_{2\text{AA}}$, as winter contributions to $\text{NO}_{2\text{AA}}$ tended to increase as $\text{NO}_{2\text{AA}}$ decreased. More generally, measures implemented to reduce $\text{NO}_{2\text{AA}}$ in one location may be less effective in others, and the development of strategies to reduce $\text{NO}_{2\text{AA}}$ in a particular area should consider the variation in
30 monthly, hour of day and hourly NO_2 concentrations contributions to $\text{NO}_{2\text{AA}}$ within that area. Similarly, consideration should also be given to how different mitigation actions will affect hourly NO_2 concentrations across the year, hour of day and peak vs moderate hourly NO_2 concentrations, as this will determine the extent to which that measure will be effective at different locations. In summary, if how $\text{NO}_{2\text{AA}}$ arises at sites across a region is known, then the ability of a monitoring network is

enhanced to inform on how changes in hourly NO₂ concentrations, resulting from the implementation of a particular mitigation action, will affect an impact or regulatory metric such as annual NO₂ concentrations.

4.3 Uncertainties and Limitations

5 A key consideration is whether the sites used in this analysis are representative of the range of contributions from hourly NO₂ concentrations producing NO_{2AA} across Europe. As outlined above, even within relatively small geographic areas NO_{2AA}, and contributions from months, hours and hourly NO₂ concentration bins to NO_{2AA}, can vary. Hence even though the NO₂ monitoring network across EU Member States is among the densest monitoring networks in the world, there may be locations where the combination of proximity to emission source, and factors affecting dispersion of NO_x emissions mean that the contribution of months, hours of the day and hourly NO₂ concentrations is distinct from the monitored locations within that
10 city. Multiple studies have previously assessed monitoring site representativeness across Europe, using a wide variety of methods, including classifying sites based on population or other non-pollutant proxy variables (Joly and Peuch, 2012; Spangl et al., 2007), and classifying sites based on the similarity of air pollutant concentrations (Austin et al., 2013; Malley et al., 2014b; Tarasova et al., 2007). This analysis shows that, for NO₂, sites in close geographic proximity to one another can have different hourly NO₂ contributions to annual average NO₂, while there can be substantial similarity in those conditions at sites
15 distant from one another. This is in contrast to assessments of site representativeness for other pollutants, such as ozone, where distinct geographic groupings were identified from analysis of ozone datasets (Malley et al., 2014b; Tarasova et al., 2007). This work therefore emphasises that while there have been multiple methodologies outlined for classifying monitoring sites, one consideration is that a classification based on pollutant measurements is likely to be pollutant-specific, and therefore may differ depending on the pollutant selected to undertake the classification.

20

The trend analysis undertaken in this work estimated the linear trend in NO_{2AA}, and monthly, hour of day, and hourly NO₂ concentration bins contributions. Broadly consistent results were obtained using parametric and non-parametric methods, and with and without adjustment for autocorrelation. The proportion of sites with different trend assignments (i.e. significant increasing, decreasing or non-significant) across the four trend estimates calculated for each statistic differed by at most 20-
25 30%. This indicates that assumptions about the underlying distribution of the data, and the level of autocorrelation did not introduce bias into the trend estimates that might lead to different conclusions on the general patterns of change. However, changes in the drivers of NO₂ concentrations have not occurred linearly (Carslaw and Carslaw, 2007), and the trend analysis does not account for these step changes and non-linearities. NO₂ concentrations have been shown to be affected by particular events; e.g. the 2008-2009 economic recession (Castellanos and Boersma, 2012; Cuevas et al., 2014), increases in urban buses
30 and diesel vehicles (Carslaw and Carslaw, 2007), or changes in the NO₂/NO_x emission ratio (Carslaw et al., 2016). Future work to investigate non-linear changes in the statistics calculated in this work could provide additional insight into how hourly NO₂ concentrations determining NO_{2AA} change in response to the implementation of specific policies or mitigation strategies. Additionally, the trend analysis was conducted at sites with sufficient data capture between 2000 and 2014. The 259 sites

which met these criteria were not evenly distributed across Europe, but were disproportionately grouped in central Europe, specifically in Germany, where 53% of the sites were located, followed by Austria (15%) and France (14%). Hence the trend results presented here are determined predominantly by the specific changes that have occurred at sites in central Europe, rather than reflecting changes across the whole of Europe. A geographically broader comparison of the consistency of changes in $\text{NO}_{2\text{AA}}$, and the conditions producing it across Europe, would be enhanced by the calculation of trends in the statistics derived here at sites in southern, eastern and northern Europe as the length of time series increase at sites in these regions, and by assessment of spatial autocorrelation across sites included in the trend analysis. Finally, assessment of trends between 2000 and 2014 does not indicate how trends at each site are likely to progress into the future, which will be determined by future changes in NO_x emissions, and meteorology.

10

The aim of this analysis was to demonstrate that calculating a standard set of ‘chemical climate’ statistics at a large number of monitoring sites could increase the information gained from the monitoring network. The calculation of a standard set of statistics has shown that, in addition to the variety of $\text{NO}_{2\text{AA}}$ across European sites shown previously (Cyrus et al., 2012), sites with similar $\text{NO}_{2\text{AA}}$ across Europe differ in how $\text{NO}_{2\text{AA}}$ is derived. Hence this study provides information as to how reducing NO_2 during specific months of the year, hours of the day, or focusing on peak vs moderate hourly NO_2 concentrations is likely to impact $\text{NO}_{2\text{AA}}$ at different types of sites across Europe. However, once emitted, NO_2 has multiple other impacts. For example, NO_2 is a key precursor for the formation of surface O_3 (Karlsson et al., 2017; Monks et al., 2015), PAN (Fischer et al., 2014; Jenkin and Clemitshaw, 2000), secondary inorganic nitrate (Fuzzi et al., 2015; Putaud et al., 2010), and contributes to nitrogen deposition (Fowler et al., 2015; Galloway et al., 2008). To explore the conditions which result in the contribution of NO_2 to these other impacts, an alternative set of chemical climate statistics would be required. Future work to derive and apply these statistics to the European monitoring network data could allow for similarity or differences in the conditions producing annual NO_2 compared to these other impacts, and therefore facilitate a more comprehensive assessment of the effects of air pollution mitigation strategies. Co-located meteorological measurements with atmospheric composition measurements, as recommended by AQEG (2015), would also allow meteorological drivers of annual NO_2 concentrations (and other impacts) to be assessed in more detail. Finally, the chemical climate statistics derived in this work have shown the variety of hourly NO_2 contributions producing $\text{NO}_{2\text{AA}}$ across Europe. Application of these statistics to NO_2 measurements from other regions could also facilitate comparison of the variation in monthly, hour of day, and hourly NO_2 distributions that contribute to $\text{NO}_{2\text{AA}}$ in other regions.

15

20

25

5 Conclusions

30

An annual average limit value for NO_2 concentrations for longer-term exposure of $40 \mu\text{g m}^{-3}$ has been established by the European Union for the protection of human health. Analysis of monitoring network measurements is key to assessing compliance with this limit value across Europe. The aim of this analysis was to link the magnitude of this impact metric to the

monthly, hour of day, and hourly NO₂ concentrations variation that produces it. Hence these contributions were calculated for 2010-2014 average annual NO₂ concentrations for more than 2500 sites monitoring NO₂ across Europe, and their trends estimated at 259 sites between 2000 and 2014.

- 5 This study shows that, in general, sites with distinct monthly, hour of day, and hourly NO₂ concentration bin contributions to annual NO₂ were not grouped into distinct geographic areas, and there were a range of patterns of hourly NO₂ concentration variability producing similar annual NO₂ concentrations with a particular area. This is a consequence of different interactions between emissions, atmospheric chemistry and meteorology at different types of site. Hence within a city, country, or European region, there is variation in the extent to which annual NO₂ concentrations could be reduced by focussing on, e.g. reducing peak NO₂ concentrations compared to more moderate concentrations, winter vs summer NO₂ concentrations, and rush hour peaks compared to NO₂ levels across the day. Therefore, the assessment of actions to mitigate annual average NO₂ concentrations should evaluate their effect on hourly NO₂ concentrations across the year, across the day, and across the hourly NO₂ concentration distribution. Development of NO₂ mitigation strategies should take into account that across Europe a variety of seasonal and diurnal patterns of hourly NO₂ contribute to annual NO₂ concentrations, and therefore that specific measures implemented to reduce NO_{2AA} in one location may not be as effective in others.
- 10
- 15

Acknowledgements

The authors are grateful to E. Weatherhead for insightful discussions regarding the trend analyses applied in this work. S Moller acknowledges the Natural Environment Research Council for funding (grant number NE/N005430/1).

References

- 20 Anenberg, S. C., Miller, J., Minjares, R., Du, L., Henze, D. K., Lacey, F., Malley, C. S., Emberson, L., Franco, V., Klimont, Z. and Heyes, C.: Impacts and mitigation of excess diesel-related NO_x emissions in 11 major vehicle markets, *Nature*, 545, 467–471, 2017.
- AQEG: Nitrogen Dioxide in the United Kingdom. Air Quality Expert Group Report. Available at: <https://uk-air.defra.gov.uk/library/aqeg/publications>, 2004.
- 25 AQEG: Trends in Primary Nitrogen Dioxide in the UK. Air Quality Expert Group Report. Available at: <https://uk-air.defra.gov.uk/assets/documents/reports/aqeg/primary-no-trends.pdf>, 2007.
- AQEG: Evidential Value of Defra Air Quality Compliance Monitoring. Air Quality Expert Group, Defra Publications, available at: http://uk-air.defra.gov.uk/assets/documents/reports/cat11/1509290925_DEF-PB14312_Evidential_value_of_Defra_air_quality_compliance_moni, 2015.
- 30 Austin, E., Coull, B., Zanobetti, A. and Koutrakis, P.: A framework to spatially cluster air pollution monitoring sites in US

- based on the PM_{2.5} composition, *Environ. Int.*, 59, 244–254, 2013.
- Baptista, P., Melo, S. and Rolim, C.: Energy, environmental and mobility impacts of car-sharing systems. Empirical results from Lisbon, Portugal, *Procedia - Soc. Behav. Sci.*, 111, 28–37, doi:10.1016/j.sbspro.2014.01.035, 2014.
- Beevers, S. D., Westmoreland, E., de Jong, M. C., Williams, M. L. and Carslaw, D. C.: Trends in NO_x and NO₂ emissions from road traffic in Great Britain, *Atmos. Environ.*, 54, 107–116, doi:10.1016/j.atmosenv.2012.02.028, 2012.
- Blond, N., Boersma, K. F., Eskes, H. J., van der A, R. J., Van Roozendaal, M., De Smedt, I., Bergametti, G. and Vautard, R.: Intercomparison of SCIAMACHY nitrogen dioxide observations, in situ measurements and air quality modeling results over Western Europe, *J. Geophys. Res.*, 112(D10), D10311, doi:10.1029/2006JD007277, 2007.
- Box, G. E. and Jenkins, G. M.: *Time Series Analysis, Control, and Forecasting*. Holden Day. San Francisco, CA., 1976.
- 10 Braniš, M.: Air quality of Prague: Traffic as a main pollution source, *Environ. Monit. Assess.*, 156(1–4), 377–390, doi:10.1007/s10661-008-0491-3, 2009.
- Carslaw, D., Beevers, S., Westmoreland, E., Williams, W., Tate, J., Murrells, T., Stedman, J., Li, Y., Grice, S., Kent, A. and Tsagatakis, I.: Trends in NO_x and NO₂ emissions and ambient measurements in the UK. Available at: https://uk-air.defra.gov.uk/assets/documents/reports/cat05/1108251149_110718_AQ0724_Final_report.pdf, 2011.
- 15 Carslaw, D. C. and Beevers, S. D.: Investigating the potential importance of primary NO₂ emissions in a street canyon, *Atmos. Environ.*, 38(22), 3585–3594, doi:10.1016/j.atmosenv.2004.03.041, 2004.
- Carslaw, D. C. and Carslaw, N.: Detecting and characterising small changes in urban nitrogen dioxide concentrations, *Atmos. Environ.*, 41(22), 4723–4733, doi:10.1016/j.atmosenv.2007.03.034, 2007.
- Carslaw, D. C. and Ropkins, K.: openair - An R package for air quality data analysis, *Environ. Model. Softw.*, 27–28, 52–61, 20 doi:doi:10.1016/j.envsoft.2011.09.008, 2012.
- Carslaw, D. C., Murrells, T. P., Andersson, J. and Keenan, M.: Have vehicle emissions of primary NO₂ peaked?, *Faraday Discuss.*, 189(0), 439–454, doi:10.1039/C5FD00162E, 2016.
- Carugno, M., Consonni, D., Randi, G., Catelan, D., Grisotto, L., Bertazzi, P. A., Biggeri, A. and Baccini, M.: Air pollution exposure, cause-specific deaths and hospitalizations in a highly polluted italian region, *Environ. Res.*, 147, 415–424, 25 doi:10.1016/j.envres.2016.03.003, 2016.
- Castellanos, P. and Boersma, K. F.: Reductions in nitrogen oxides over Europe driven by environmental policy and economic recession, *Sci. Rep.*, 2, doi:10.1038/srep00265, 2012.
- CEBR: The future economic and environment costs of gridlock in 2030. Report for INRIX. Centre for Economics and Business Research. July 2014. Available at: <https://ibttta.org/sites/default/files/documents/MAF/Costs-of-Congestion-INRIX-Cebr-Report%20%283%29.pdf>, 2014.
- 30 Chaloulakou, A., Mavroidis, I. and Gavriil, I.: Compliance with the annual NO₂ air quality standard in Athens. Required NO_x levels and expected health implications, *Atmos. Environ.*, 42(3), 454–465, doi:10.1016/j.atmosenv.2007.09.067, 2008.
- Chiesa, M., Perrone, M. G., Cusumano, N., Ferrero, L., Sangiorgi, G., Bolzacchini, E., Lorenzoni, A. and Ballarin Denti, A.: An environmental, economical and socio-political analysis of a variety of urban air-pollution reduction policies for primary

- PM10 and NOx: The case study of the Province of Milan (Northern Italy), *Environ. Sci. Policy*, 44, 39–50, 2014.
- Clapp, L. J. and Jenkin, M. E.: Analysis of the relationship between ambient levels of O₃, NO₂ and NO as a function of NO_x in the UK, *Atmos. Environ.*, 35(36), 6391–6405, doi:10.1016/S1352-2310(01)00378-8, 2001.
- COMEAP: Interim Statement on Quantifying the Association of Long-Term Average Concentrations of Nitrogen Dioxide and Mortality. UK Committee on the Medical Effects of Air Pollution. Available at: <https://www.gov.uk/government/publications/nitrogen-dioxide-interim->, 2015.
- Crouse, D. L., Peters, P. A., Hystad, P., Brook, J. R., van Donkelaar, A., Martin, R. V., Villeneuve, P. J., Jerrett, M., Goldberg, M. S., Pope C. Arden, I. I. I., Brauer, M., Brook, R. D., Robichaud, A., Menard, R. and Burnett, R. T.: Ambient PM_{2.5}, O₃, and NO₂ Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC), *Environ. Health Perspect.*, 123(11), 1180–1186, doi:10.1289/ehp.1409276, 2015.
- Cuevas, C. A., Notario, A., Adame, J. A., Hilboll, A., Richter, A., Burrows, J. P. and Saiz-Lopez, A.: Evolution of NO₂ levels in Spain from 1996 to 2012, *Sci. Rep.*, 4, doi:10.1038/srep05887, 2014.
- Cyrys, J., Eeftens, M., Heinrich, J., Ampe, C., Armengaud, A., Beelen, R., Bellander, T., Beregszaszi, T., Birk, M., Cesaroni, G., Cirach, M., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dedele, A., Dimakopoulou, K., Eriksen, K., Galassi, C., Graulevičiene, R., Grivas, G., Gruzieva, O., Gustafsson, A. H., Hoffmann, B., Iakovides, M., Ineichen, A., Krämer, U., Lanki, T., Lozano, P., Madsen, C., Meliefste, K., Modig, L., Mölter, A., Mosler, G., Nieuwenhuijsen, M., Nonnemacher, M., Oldenwening, M., Peters, A., Pontet, S., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Sugiri, D., Stephanou, E. G., Taimisto, P., Tsai, M. Y., Vaskövi, É., Villani, S., Wang, M., Brunekreef, B. and Hoek, G.: Variation of NO₂ and NO_x concentrations between and within 36 European study areas: Results from the ESCAPE study, *Atmos. Environ.*, 62, 374–390, doi:10.1016/j.atmosenv.2012.07.080, 2012.
- Donnelly, A., Misstear, B. and Broderick, B.: Application of nonparametric regression methods to study the relationship between NO₂ concentrations and local wind direction and speed at background sites, *Sci. Total Environ.*, 409(6), 1134–1144, doi:10.1016/j.scitotenv.2010.12.001, 2011.
- Donnelly, A., Broderick, B. and Misstear, B.: Relating Background NO₂ Concentrations in Air to Air Mass History Using Non-Parametric Regression Methods: Application at Two Background Sites in Ireland, *Env. Model Assess*, 17, 363–373, 2012.
- EEA: Air Quality in Europe - 2016 report. European Environment Agency Report No. 28/2016. Available at: <http://www.eea.europa.eu/publications/air-quality-in-europe-2016>, 2016a.
- EEA: European Union emission inventory report 1990–2014 under the UNECE Convention on Long-range Transboundary Air Pollution (LRTAP). European Environment Agency Report No. 16/2016., 2016b.
- European Commission: Guidance on the Commission Implementing Decision laying down rules for Directives 2004/107/EC and 2008/50/EC of the European Parliament and of the Council as regards the reciprocal exchange of information and reporting on ambient air. European Commission, 2013.
- European Council Directive 2008/50/EC: On ambient air quality and cleaner air for Europe. 21st May 2008, L 152/1, 2008.
- Fischer, E. V., Jacob, D. J., Yantosca, R. M., Sulprizio, M. P., Millet, D. B., Mao, J., Paulot, F., Singh, H. B., Roiger, A., Ries,

- L., Talbot, R. W., Dzepina, K. and Deolal, S. P.: Atmospheric peroxyacetyl nitrate (PAN): a global budget and source attribution, *Atmos. Chem. Phys.*, 14(5), 2679–2698 [online] Available from: %3CGo, 2014.
- Flemming, J., Stern, R. and Yamartino, R. J.: A new air quality regime classification scheme for O-3, NO₂, SO₂ and PM₁₀ observations sites, *Atmos. Environ.*, 39(33), 6121–6129 [online] Available from: %3CGo, 2005.
- 5 Fowler, D., Steadman, C. E., Stevenson, D., Coyle, M., Rees, R. M., Skiba, U. M., Sutton, M. A., Cape, J. N., Dore, A. J., Vieno, M., Simpson, D., Zaehle, S., Stocker, B. D., Rinaldi, M., Facchini, M. C., Flechard, C. R., Nemitz, E., Twigg, M., Erisman, J. W., Butterbach-Bahl, K. and Galloway, J. N.: Effects of global change during the 21st century on the nitrogen cycle, *Atmos. Chem. Phys.*, 15(24), 13849–13893, doi:10.5194/acp-15-13849-2015, 2015.
- Fuzzi, S., Baltensperger, U., Carslaw, K., Decesari, S., Denier Van Der Gon, H., Facchini, M. C., Fowler, D., Koren, I.,
10 Langford, B., Lohmann, U., Nemitz, E., Pandis, S., Riipinen, I., Rudich, Y., Schaap, M., Slowik, J. G., Spracklen, D. V., Vignati, E., Wild, M., Williams, M. and Gilardoni, S.: Particulate matter, air quality and climate: Lessons learned and future needs, *Atmos. Chem. Phys.*, 15(14), 8217–8299, doi:10.5194/acp-15-8217-2015, 2015.
- Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. a, Seitzinger, S. P. and Sutton, M. a: Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions, *Science (80-.)*,
15 320(May), 889–892, doi:10.1126/science.1136674, 2008.
- Geddes, J. A., Martin, R. V., Boys, B. L. and van Donkelaar, A.: Long-term trends worldwide in ambient NO₂ concentrations inferred from satellite observations, *Environ. Health Perspect.*, 124(3), 281–289, doi:10.1289/ehp.1409567, 2016.
- Gilge, S., Plass-Duelmer, C., Fricke, W., Kaiser, A., Ries, L., Buchmann, B. and Steinbacher, M.: Ozone, carbon monoxide and nitrogen oxides time series at four alpine GAW mountain stations in central Europe, *Atmos. Chem. Phys.*, 10(24), 12295–
20 12316, doi:10.5194/acp-10-12295-2010, 2010.
- Grice, S., Stedman, J., Kent, A., Hobson, M., Norris, J., Abbott, J. and Cooke, S.: Recent trends and projections of primary NO₂ emissions in Europe, *Atmos. Environ.*, 43(13), 2154–2167, doi:10.1016/j.atmosenv.2009.01.019, 2009.
- Hazenkamp-Von Arx, M. E., Götschi, T., Ackermann-Liebrich, U., Bono, R., Burney, P., Cyrus, J., Jarvis, D., Lillienberg, L., Luczynska, C., Maldonado, J. A., Jaén, A., De Marco, R., Mi, Y., Modig, L., Bayer-Oglesby, L., Payo, F., Soon, A., Sunyer, J., Villani, S., Weyler, J. and Künzli, N.: PM_{2.5} and NO₂ assessment in 21 European study centres of ECRHS II: Annual
25 means and seasonal differences, *Atmos. Environ.*, 38(13), 1943–1953, doi:10.1016/j.atmosenv.2004.01.016, 2004.
- Henschel, S., Le Tertre, A., Atkinson, R. W., Querol, X., Pandolfi, M., Zeka, A., Haluza, D., Analitis, A., Katsouyanni, K., Bouland, C., Pascal, M., Medina, S. and Goodman, P. G.: Trends of nitrogen oxides in ambient air in nine European cities between 1999 and 2010, *Atmos. Environ.*, 117, 234–241, doi:10.1016/j.atmosenv.2015.07.013, 2016.
- 30 INRIX: London Congestion Trend. INRIX. Available at: <http://londonfirst.co.uk/wp-content/uploads/2016/05/London-Congestion-Trends-FINAL.pdf>, 2016.
- Jeanjean, A. P. R., Gallagher, J., Monks, P. S. and Leigh, R. J.: Ranking current and prospective NO₂ pollution mitigation strategies: An environmental and economic modelling investigation in Oxford Street, London, *Environ. Pollut.*, 225, 587–597, doi:10.1016/j.envpol.2017.03.027, 2017.

- Jenkin, M. E.: Investigation of the impact of short-timescale NO_x variability on annual mean oxidant partitioning at UK sites, *Atmos. Environ.*, 90, 43–50, doi:10.1016/j.atmosenv.2014.03.033, 2014.
- Jenkin, M. E. and Clemitshaw, K. C.: Ozone and other secondary photochemical pollutants: chemical processes governing their formation in the planetary boundary layer, *Atmos. Environ.*, 34(16), 2499–2527 [online] Available from: %3CGo, 2000.
- 5 Joly, M. and Peuch, V. H.: Objective classification of air quality monitoring sites over Europe, *Atmos. Environ.*, 47, 111–123 [online] Available from: %3CGo, 2012.
- JRC-AQUILA: JRC-AQUILA Position Paper: assessment on siting criteria, classification and representativeness of air quality monitoring stations. Available at: <http://ec.europa.eu/environment/air/pdf/SCREAM%20final.pdf>, 2013.
- Karlsson, P. E., Klingberg, J., Engardt, M., Andersson, C., Langner, J., Karlsson, G. P. and Pleijel, H.: Past, present and future
10 concentrations of ground-level ozone and potential impacts on ecosystems and human health in northern Europe, *Sci. Total Environ.*, 576, 22–35, doi:10.1016/j.scitotenv.2016.10.061, 2017.
- Kaufman, L. and Rousseeuw, P. J.: *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York., 1990.
- Konovalov, I. B., Beekmann, M., Richter, A., Burrows, J. P. and Hilboll, A.: Multi-annual changes of NO_x emissions in megacity regions: Nonlinear trend analysis of satellite measurement based estimates, *Atmos. Chem. Phys.*, 10(17), 8481–8498,
15 doi:10.5194/acp-10-8481-2010, 2010.
- Lozano, A., Usero, J., Vanderlinden, E., Raez, J., Contreras, J. and Navarrete, B.: Air quality monitoring network design to control nitrogen dioxide and ozone, applied in Malaga, Spain, *Microchem. J.*, 93(2), 164–172, doi:10.1016/j.microc.2009.06.005, 2009.
- Malley, C. S., Braban, C. F. and Heal, M. R.: New Directions: Chemical climatology and assessment of atmospheric
20 composition impacts., *Atmos. Environ.*, 87, 261–264, doi:doi: 10.1016/j.atmosenv.2014.01.027, 2014a.
- Malley, C. S., Braban, C. F. and Heal, M. R.: The application of hierarchical cluster analysis and non-negative matrix factorization to European atmospheric monitoring site classification., *Atmos. Res.*, 138, 30–40, 2014b.
- Malley, C. S., Heal, M. R., Mills, G. and Braban, C.: Trends and drivers of ozone human health and vegetation impact metrics from UK EMEP supersite measurements (1990–2013), *Atmos. Chem. Phys. Discuss.*, 15, 1869–1914, 2015.
- 25 Malley, C. S., Heal, M. R., Braban, C. F., Kentisbeer, J., Leeson, S. R., Lingard, J. J. N., Ritchie, S., Maggs, R., Beccaceci, S., Quincey, P., Brown, R. J. C. and Twigg, M. M.: The contributions to long-term health-relevant particulate matter at the UK EMEP supersites between 2010 and 2013: Quantifying the mitigation challenge, *Environ. Int.*, 95, 98–111, 2016.
- Mangiameli, P., Chen, S. K. and West, D.: A comparison of SOM neural network and hierarchical clustering methods, *Eur. J. Oper. Res.*, 93(2), 402–417 [online] Available from: %3CGo, 1996.
- 30 Mann, H. B.: Nonparametric tests against trend, *Econometrica*, 13, 245–259, 1945.
- Masiol, M., Squizzato, S., Formenton, G., Harrison, R. M. and Agostinelli, C.: Air quality across a European hotspot: Spatial gradients, seasonality, diurnal cycles and trends in the Veneto region, NE Italy, *Sci. Total Environ.*, 576, 210–224, doi:10.1016/j.scitotenv.2016.10.042, 2017.
- Mavroidis, I. and Chaloulakou, A.: Long-term trends of primary and secondary NO₂ production in the Athens area. Variation

- of the NO₂/NO_x ratio, *Atmos. Environ.*, 45(38), 6872–6879, doi:10.1016/j.atmosenv.2010.11.006, 2011.
- Mavroidis, I. and Ilia, M.: Trends of NO_x, NO₂ and O₃ concentrations at three different types of air quality monitoring stations in Athens, Greece, *Atmos. Environ.*, 63(2), 135–147, doi:10.1016/j.atmosenv.2012.09.030, 2012.
- Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., Fowler, D., Granier, C., Law, K. S., Mills, G. E., Stevenson, D. S., Tarasova, O., Thouret, V., von Schneidemesser, E., Sommariva, R., Wild, O. and Williams, M. L.: Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer, *Atmos. Chem. Phys.*, 15(15), 8889–8973, 2015.
- Palarz, A., Celinski-Myslaw, D. and Ustrnul, Z.: Temporal and spatial variability of surface-based inversions over Europe based on ERA-Interim reanalysis, *Int. J. Clim.*, 1097–88, doi:10.1002/joc.5167, 2017.
- 10 Putaud, J. P., Van Dingenen, R., Alastuey, A., Bauer, H., Birmili, W., Cyrys, J., Flentje, H., Fuzzi, S., Gehrig, R., Hansson, H. C., Harrison, R. M., Herrmann, H., Hitzenberger, R., Hüglin, C., Jones, A. M., Kasper-Giebl, A., Kiss, G., Koussa, A., Kuhlbusch, T. A. J., Löschau, G., Maenhaut, W., Molnar, A., Moreno, T., Pekkanen, J., Perrino, C., Pitz, M., Puxbaum, H., Querol, X., Rodriguez, S., Salma, I., Schwarz, J., Smolik, J., Schneider, J., Spindler, G., ten Brink, H., Tursic, J., Viana, M., Wiedensohler, A. and Raes, F.: A European aerosol phenomenology - 3: Physical and chemical characteristics of particulate matter from 60 rural, urban, and kerbside sites across Europe, *Atmos. Environ.*, 44(10), 1308–1320, doi:10.1016/j.atmosenv.2009.12.011, 2010.
- 15 R Core Team: R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>, 2016.
- Reed, C., Evans, M. J., Di Carlo, P., Lee, J. D. and Carpenter, L. J.: Interferences in photolytic NO₂ measurements: Explanation for an apparent missing oxidant?, *Atmos. Chem. Phys.*, 16(7), 4707–4724, doi:10.5194/acp-16-4707-2016, 2016.
- 20 REVIHAAP: Review of evidence on health aspects of air pollution – REVIHAAP Project technical report. World Health Organization (WHO) Regional Office for Europe. Bonn. Available: http://www.euro.who.int/__data/assets/pdf_file/0004/193108/REVIHAAP-Final-technical-rep., 2013.
- Schäfer, K., Emeis, S., Hoffmann, H. and Jahn, C.: Influence of mixing layer height upon air pollution in urban and sub-urban areas, *Meteorol. Zeitschrift*, 15(6), 647–658, doi:10.1127/0941-2948/2006/0164, 2006.
- 25 Schneider, P., Lahoz, W. A. and Van Der A, R.: Recent satellite-based trends of tropospheric nitrogen dioxide over large urban agglomerations worldwide, *Atmos. Chem. Phys.*, 15(3), 1205–1220, doi:10.5194/acp-15-1205-2015, 2015.
- Seidel, D. J., Zhang, Y., Beljaars, A., Golaz, J.-C., Jacobson, A. R. and Medeiros, B.: Climatology of the planetary boundary layer over the continental United States and Europe, *J. Geophys. Res. Atmos.*, 117(D17), n/a-n/a, doi:10.1029/2012JD018143, 30 2012.
- Spangl, W., Schneider, J., Moosmann, L. and Nagi, C.: Representativeness and Classification of Air Quality Monitoring Stations, Umweltbundesamt Report. <http://www.umweltbundesamt.at/fileadmin/site/publikationen/REP0121.pdf>, 2007.
- Stathopoulos, A. and Karlaftis, M.: Temporal and spatial variations of real-time traffic data in urban areas, *Transp. Res. Rec.*, (1768), 135–140, doi:10.3141/1768-16, 2001.

- Tarasova, O. A., Brenninkmeijer, C. A. M., Joeckel, P., Zvyagintsev, A. M. and Kuznetsov, G. I.: A climatology of surface ozone in the extra tropics: cluster analysis of observations and model results, *Atmos. Chem. Phys.*, 7(24), 6099–6117 [online] Available from: %3CGo, 2007.
- Theil, H.: A rank-invariant method of linear and polynomial regression analysis I, *Proc. Kon. Ned. Akad. V. Wetensch. A*, 53, 386–392, 1950a.
- Theil, H.: A rank-invariant method of linear and polynomial regression analysis II, *Proc. Kon. Ned. Akad. V. Wetensch. A*, 53, 521–525, 1950b.
- Theil, H.: A rank-invariant method of linear and polynomial regression analysis III, *Proc. Kon. Ned. Akad. V. Wetensch. A*, 53, 1397–1412, 1950c.
- 10 Thomas, M.: Research for TRAN committee - The World is Changing. Transport Too. Policy Department B: Structural and Cohesion Policies, European Parliament. Available at: [http://www.europarl.europa.eu/RegData/etudes/STUD/2016/563424/IPOL_STU\(2016\)563424_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2016/563424/IPOL_STU(2016)563424_EN.pdf), 2016.
- Thomas, T., Weijermars, W. and van Berkum, E.: Variations in urban traffic volumes, *Eur. J. Transp. Infrastruct. Res.*, 8(3), 251–263, 2008.
- 15 Turner, M. C., Jerrett, M., Pope III, C. A., Krewski, D., Gapstur, S. M., Diver, W. R., Beckerman, B. S., Marshall, J. D., Su, J., Crouse, D. L. and Burnett, R. T.: Long-Term Ozone Exposure and Mortality in a Large Prospective Study, *Am. J. Respir. Crit. Care Med.*, 193(10), 1134–1142, doi:10.1164/rccm.201508-1633OC, 2016.
- US EPA: Integrated Science Assessment for Oxides of Nitrogen - Health Criteria. United States Environmental Protection Agency Report Number EPA/600/R-15/068. Available at: <https://www.epa.gov/isa/integrated-science-assessment-isa-nitrogen-dioxide-health-criteria>, 2016.
- 20 Vestreng, V., Ntziachristos, L., Semb, A., Reis, S., Isaksen, I. S. A. and Tarrason, L.: Evolution of NOx emissions in Europe with focus on road transport control measures, *Atmos. Chem. Phys.*, 9(x), 1503–1520, doi:10.5194/acp-9-1503-2009, 2009.
- Villena, G., Bejan, I., Kurtenbach, R., Wiesen, P. and Kleffmann, J.: Interferences of commercial NO2 instruments in the urban atmosphere and in a smog chamber, *Atmos. Meas. Tech.*, 5(1), 149–159, doi:10.5194/amt-5-149-2012, 2012.
- 25 Wagner, P. and Schäfer, K.: Influence of mixing layer height on air pollutant concentrations in an urban street canyon, *Urban Clim.*, 1–13, doi:10.1016/j.uclim.2015.11.001, 2015.
- Ward, J.: Hierarchical Grouping to Optimize an Objective Function, *J. Am. Stat. Assoc.*, 58, 236–244 [online] Available from: <http://www.jstor.org/stable/2282967>, 1963.
- Weijermars, W.: Analysis of urban traffic patterns using clustering. TRAIL Thesis Series T2007/3, The Netherlands TRAIL Research School. PhD Thesis. University of Twente. Available at: <https://ris.utwente.nl/ws/portalfiles/portal/6069766>, 2007.
- WHO: Air quality guidelines: Global update 2005. Particulate matter, ozone, nitrogen dioxide and sulfur dioxide. World health organization regional office for europe. Available: Http://apps.who.int/iris/bitstream/10665/69477/1/who_sde_phe_oh_06.02_eng.Pdf [ac, 2006.
- Zhang, Y., Seidel, D. J. and Zhang, S.: Trends in planetary boundary layer height over europe, *J. Clim.*, 26(24), 10071–10076,

Table 1: The number and classification of sites for which 2010-2014 average annual NO₂ concentrations were calculated, separated into 10 µg m⁻³ bins.

Concentration bin	Number of sites	Rural Background	Rural Industrial	Rural Traffic	Suburban Background	Suburban Industrial	Suburban Traffic	Urban Background	Urban Industrial	Urban Traffic	Other
0-10 µg m⁻³	317	193	22	0	25	20	1	13	11	1	31
10-20 µg m⁻³	775	136	26	2	176	56	5	247	50	32	45
20-30 µg m⁻³	752	23	10	2	130	41	21	316	36	126	47
30-40 µg m⁻³	418	2	0	3	21	13	12	129	15	195	28
40-50 µg m⁻³	202	0	0	4	1	5	16	16	1	145	14
50-60 µg m⁻³	76	0	0	1	0	0	2	4	1	63	5
60-70 µg m⁻³	32	0	0	0	0	0	2	0	0	27	3
70-80 µg m⁻³	7	0	0	0	0	0	0	0	0	7	0
>80 µg m⁻³	8	0	0	0	0	0	0	0	0	8	0
Total	2587	354	58	12	353	135	59	725	114	604	173

5

10

Table 2: The number of sites in each 2010-2014 average annual NO₂ (NO_{2AA}) concentration bin with sufficient data capture to calculate monthly, hour of day, and hourly concentration contributions to 2010-2014 NO_{2AA}, split by the clusters demarcating distinct variation in monthly, hour of day, and hourly NO₂ concentration bin contributions to NO_{2AA}.

Concentration bin	Number of clusters	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Total
0-10 µg m ⁻³	5	110	38	75	9	31	-	-	-	263
10-20 µg m ⁻³	7	121	182	102	154	75	23	2	-	659
20-30 µg m ⁻³	7	104	127	164	44	147	63	3	-	652
30-40 µg m ⁻³	8	30	27	59	42	78	39	35	34	344
40-50 µg m ⁻³	8	31	12	14	35	27	34	15	7	175
50-60 µg m ⁻³	8	18	12	9	16	5	1	3	1	65
60-70 µg m ⁻³	5	3	11	7	4	3	-	-	-	28
70-80 µg m ⁻³	Sites analysed individually	-	-	-	-	-	-	-	-	5
>80 µg m ⁻³	Sites analysed individually	-	-	-	-	-	-	-	-	7

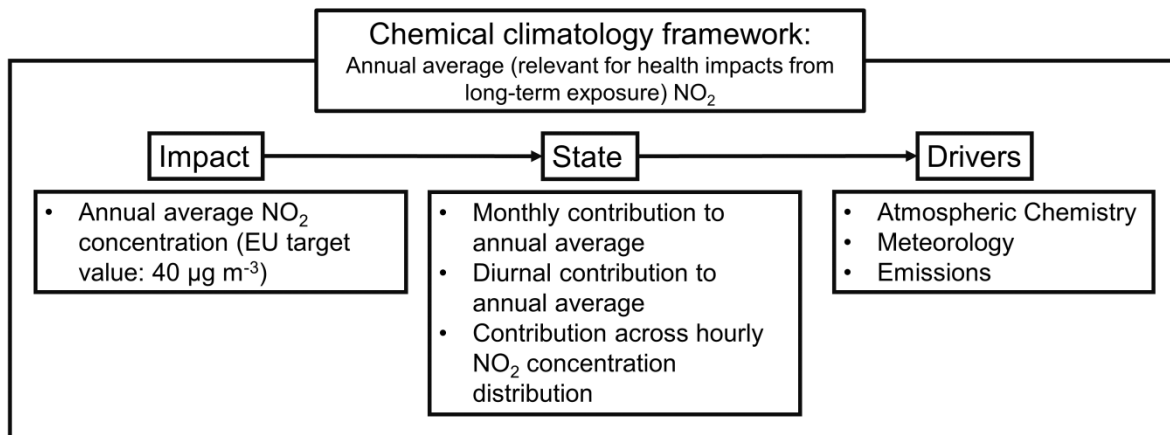


Figure 1: Outline of Chemical Climatology elements used to evaluate the conditions producing ‘long-term health-relevant’ NO₂ concentrations across Europe.

5

10

15

20

25

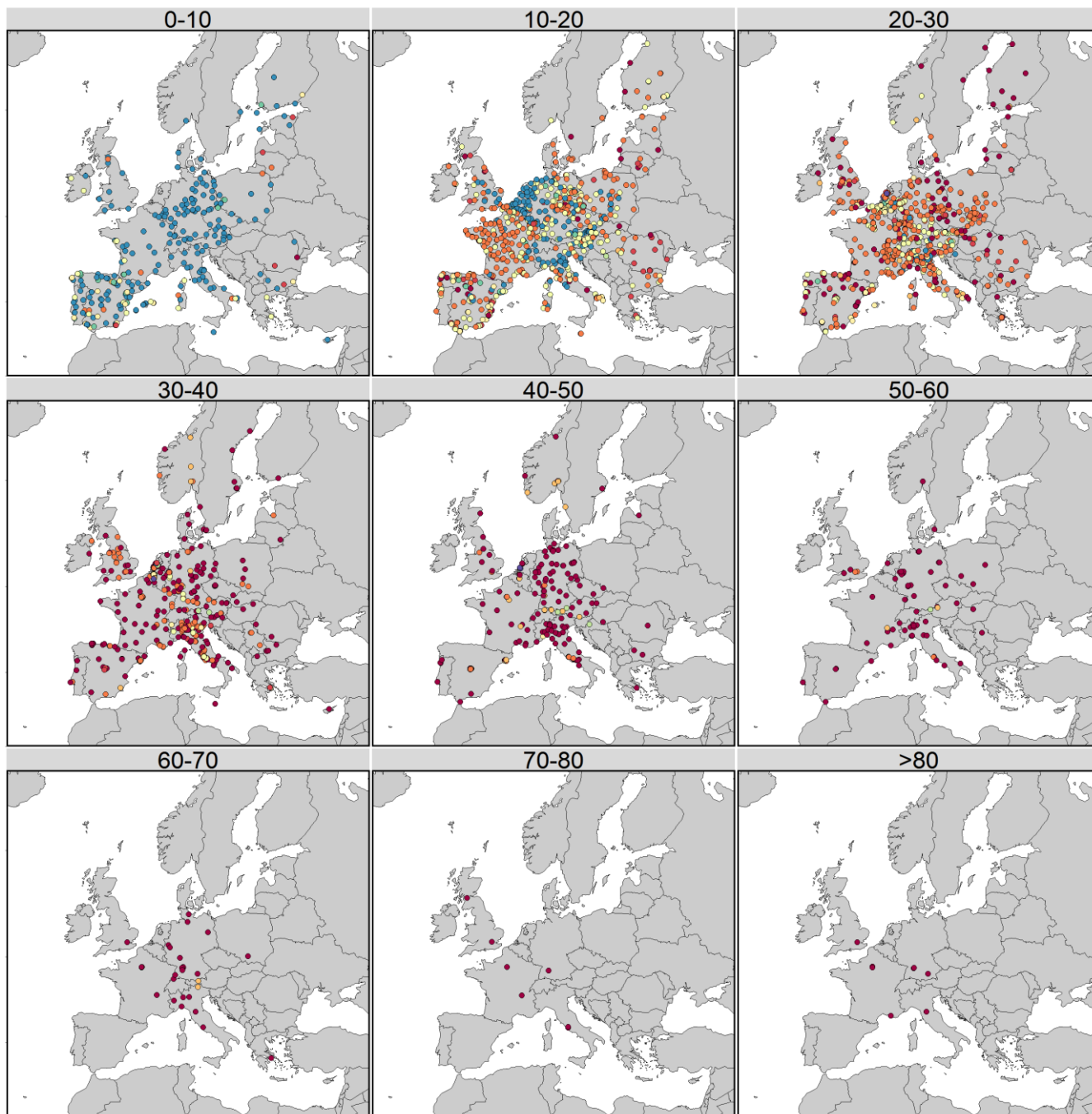


Figure 2: Locations of NO₂ monitoring sites across Europe. Sites are grouped into 10 µg m⁻³ concentration bins based on 2010-2014 average annual NO₂ concentration, and are coloured according to the EU AirBase site-type classification.

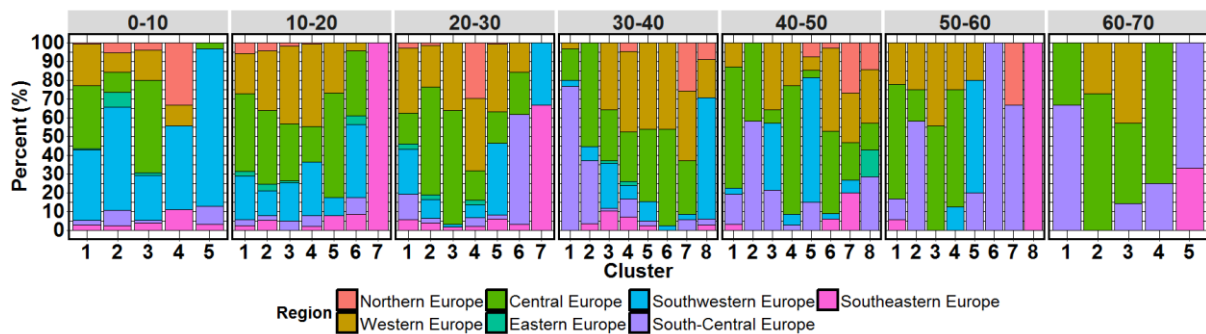


Figure 3: The proportion of sites in each 2010-2014 average annual NO₂ (NO_{2AA}) 10 µg m⁻³ concentration bin in different European regions (shown in Figure S33), separated by clusters demarcating distinct variation in monthly, hour of day, and hourly NO₂ concentration bin contributions to NO_{2AA}.

5

10

15

20

25

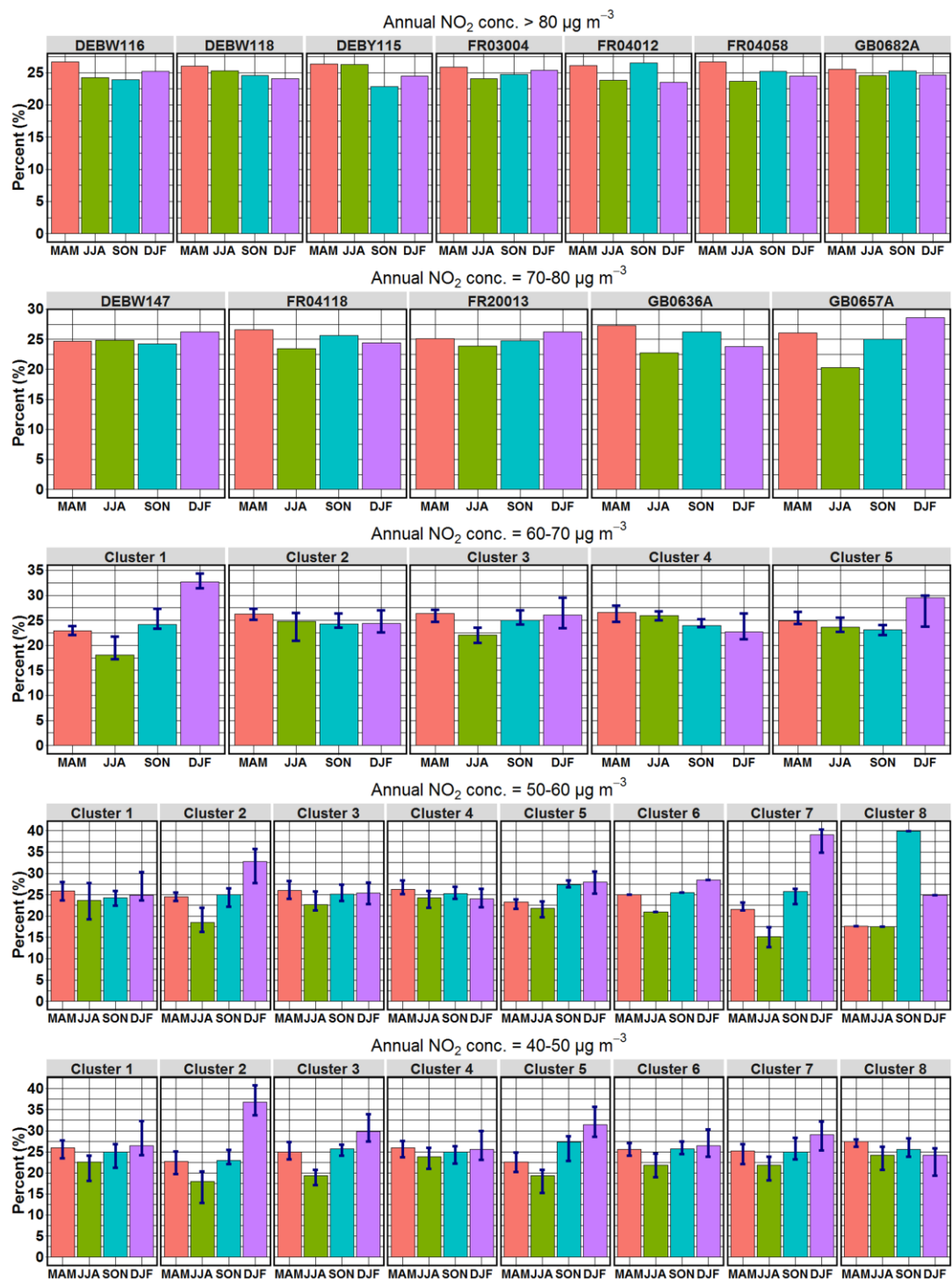


Figure 4a: The contribution to 2010-2014 average annual NO_2 concentrations (NO_{2AA}) from spring (March, April, May), summer (June, July, August), Autumn (September, October, November), and Winter (December, January, February) months for sites split

into bins based on the magnitude of 2010-2014 $\text{NO}_{2\text{AA}}$ concentration. For the >80 and $70-80 \mu\text{g m}^{-3}$ bins, these contributions are shown for sites individually (identified by Airbase ID), while for lower concentration bins sites are grouped into clusters with distinct variation in monthly, hour of day, and hourly concentration contribution to $\text{NO}_{2\text{AA}}$. For these bins, the height of the bar shows the median contribution to 2010-2014 $\text{NO}_{2\text{AA}}$ from that season across all sites, with the 5th and 95th percentile values across the sites shows as uncertainty bars.

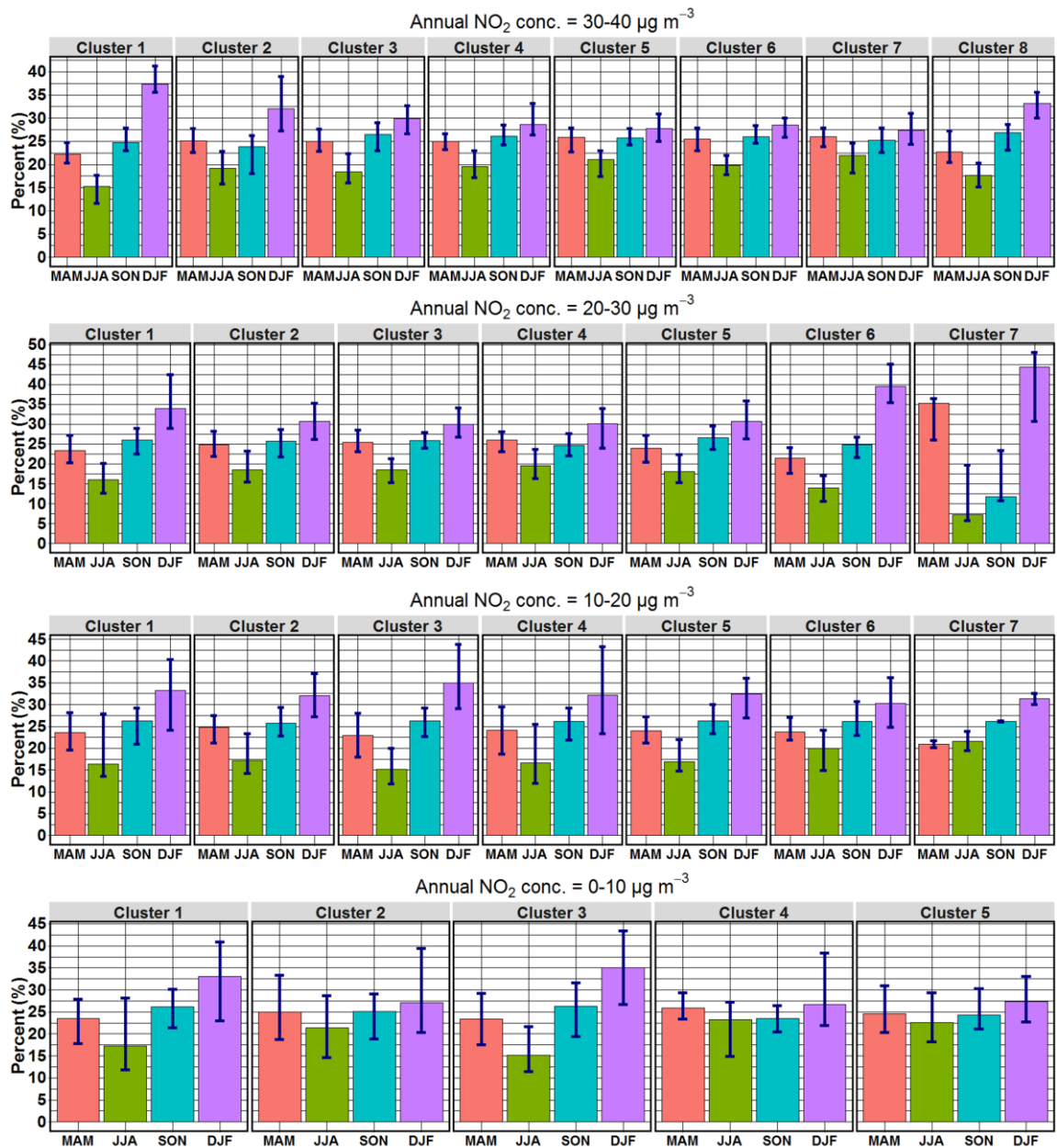


Figure 4b: The contribution to 2010-2014 average annual NO_2 concentrations ($\text{NO}_{2\text{AA}}$) from spring (March, April, May), summer (June, July, August), Autumn (September, October, November), and Winter (December, January, February) months for sites split into bins based on the magnitude of 2010-2014 $\text{NO}_{2\text{AA}}$ concentration. Sites are grouped into clusters with distinct variation in monthly, hour of day, and hourly concentration contribution to $\text{NO}_{2\text{AA}}$. For these bins, the height of the bar shows the median contribution to 2010-2014 $\text{NO}_{2\text{AA}}$ from that season across all sites, with the 5th and 95th percentile values across the sites shows as uncertainty bars.

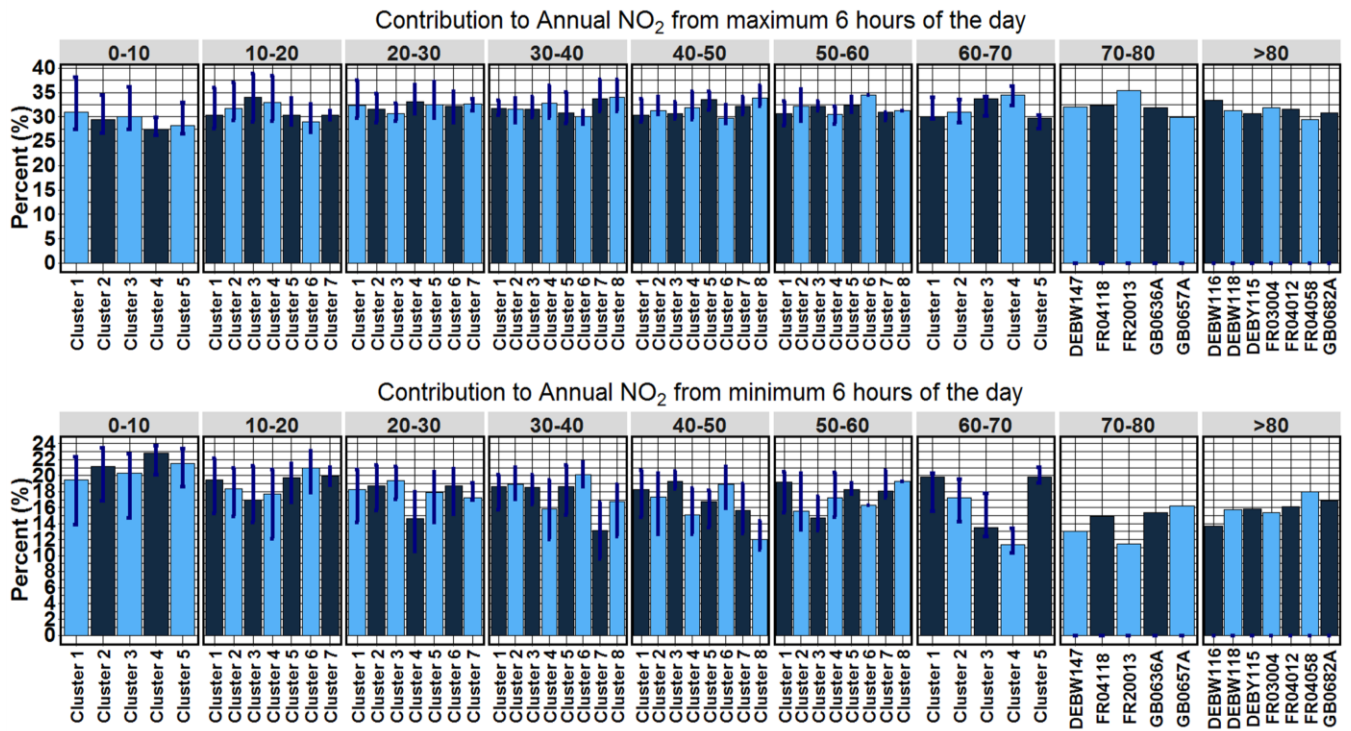


Figure 5: The contribution to 2010-2014 average annual NO₂ concentrations (NO_{2AA}) from (a) the six hours with the largest contribution to 2010-2014 NO_{2AA}, and (b) the six hours with the smallest contribution to 2010-2014 NO_{2AA}, for sites separated by 2010-2014 NO_{2AA} concentrations, and clusters demarcating distinct variation in monthly, hour of day, and hourly NO₂ concentration bin contributions to NO_{2AA}. These statistics are a proxy for variation in contribution to NO_{2AA} across the diurnal cycle.

10

15

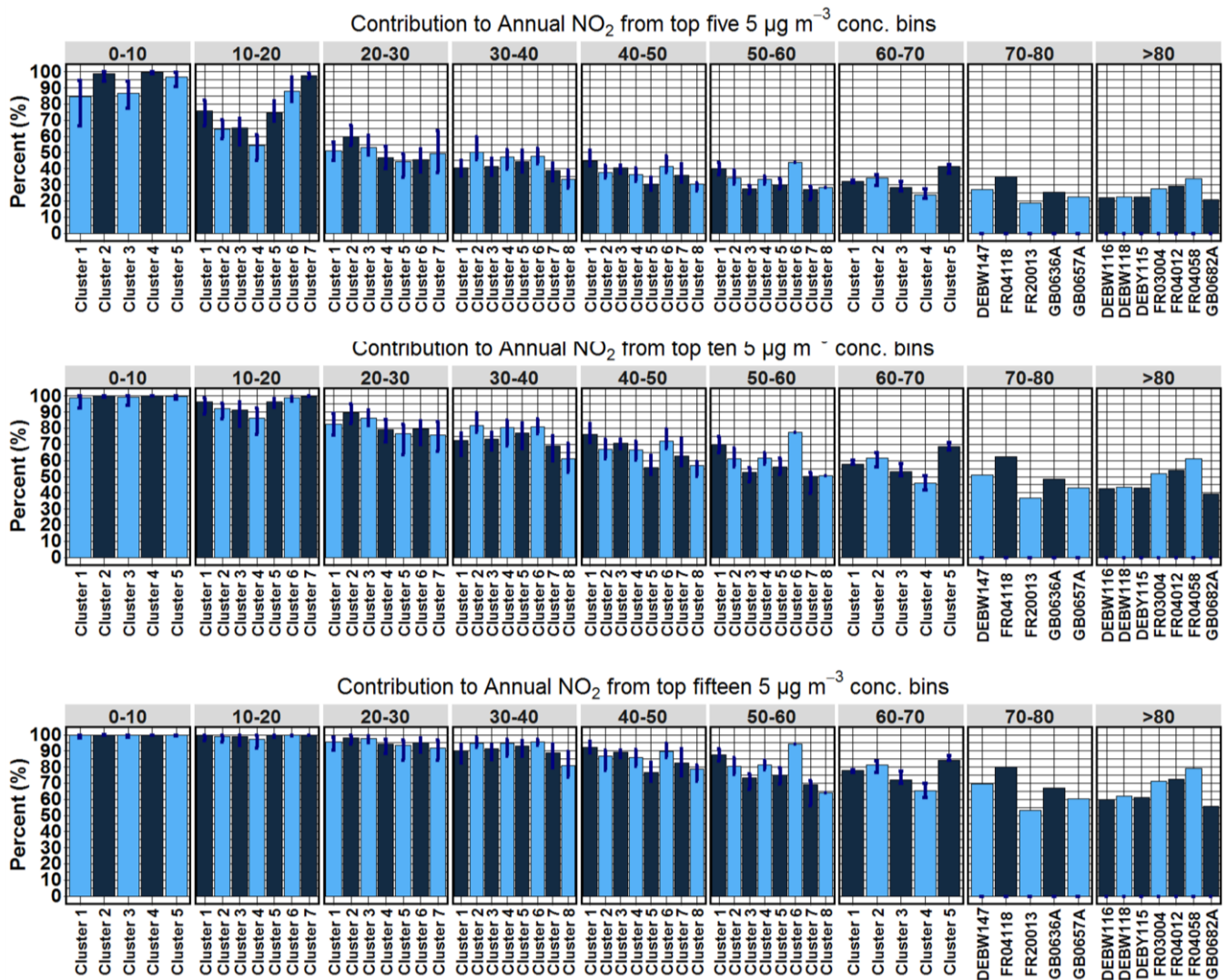


Figure 6: The contribution to 2010-2014 average annual NO₂ concentrations (NO_{2AA}) from the (a) five, (b) ten, and (c) fifteen 5 µg m⁻³ hourly NO₂ concentration bins with the largest contribution to 2010-2014 NO_{2AA}, for sites with 2010-2014 NO_{2AA} in 10 µg m⁻³ bins, and separated by clusters demarcating distinct variation in monthly, hour of day, and hourly NO₂ concentration bin contributions to NO_{2AA}. These values are a proxy to compare the range of hourly NO₂ concentrations contributing to NO_{2AA}, and the near 100% contribution for the lower 2010-2014 NO_{2AA} bins results from the fact that for NO_{2AA} of 0-10 µg m⁻³, there are far fewer high hourly NO₂ concentration values.

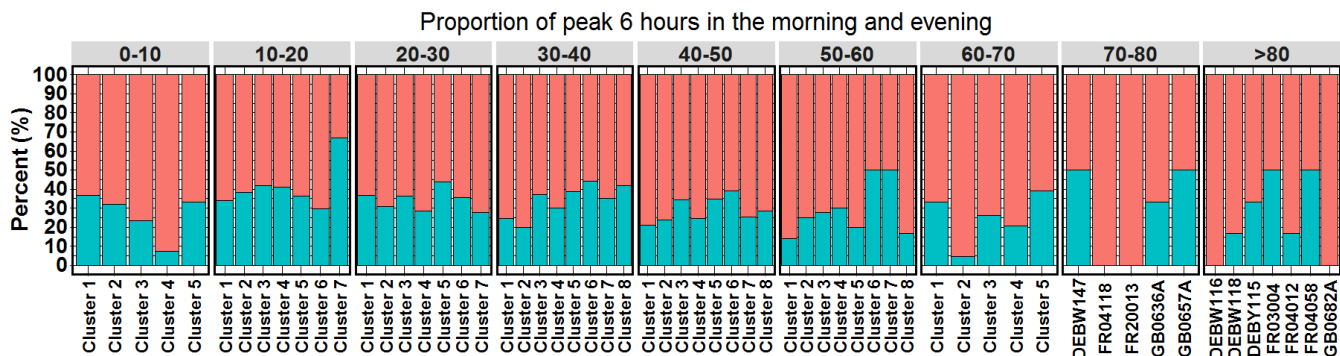


Figure 7: Proportion of the 6 hours of the day that make the largest contribution to 2010-2014 average annual NO₂ concentrations (NO_{2AA}) that occurred in the morning (before 12:00, blue), and in the afternoon/evening (after 12:00, red), for sites with 2010-2014 NO_{2AA} in 10 µg m⁻³ bins, and separated by clusters demarcating distinct variation in monthly, hour of day, and hourly NO₂ concentration bin contributions to NO_{2AA}.

10

15

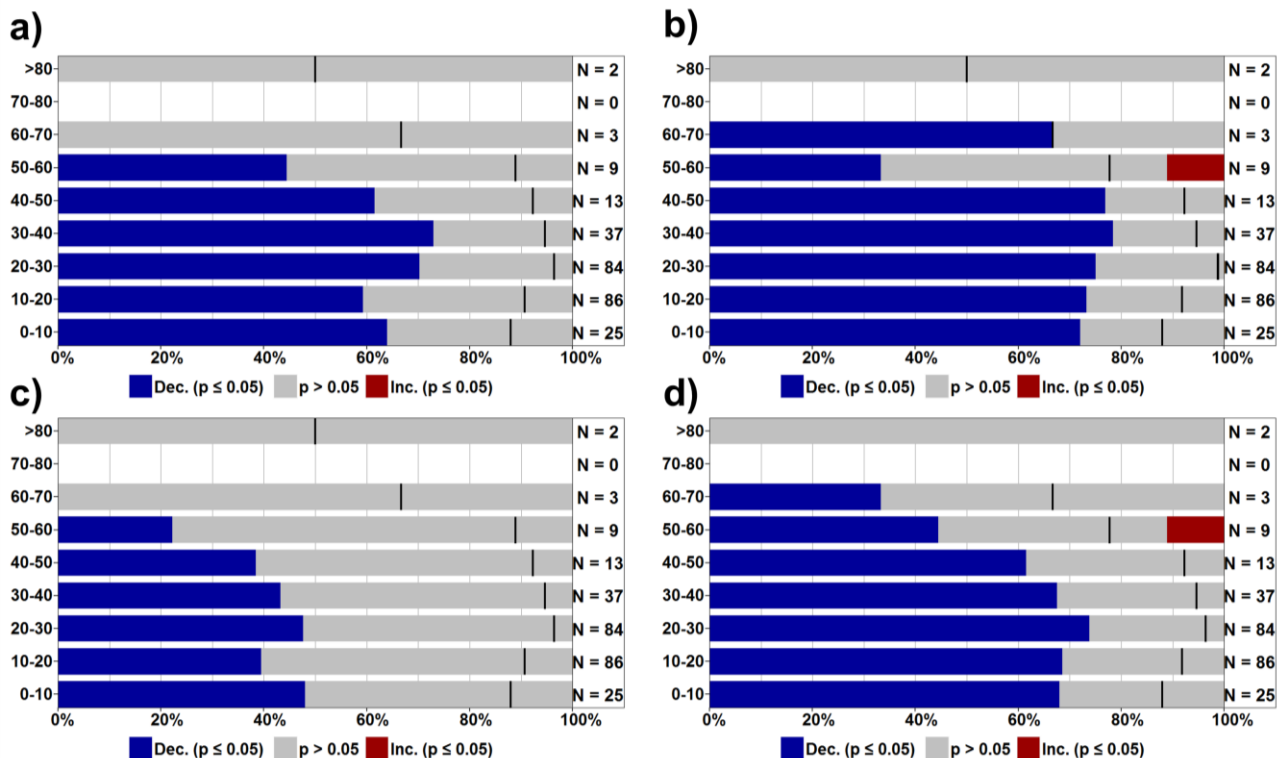


Figure 8: Proportion of monitoring sites with 2010-2014 average annual NO_2 concentrations ($\text{NO}_{2\text{AA}}$) in $10 \mu\text{g m}^{-3}$ bins that had significant decreasing (blue), significant increasing (red) ($p \leq 0.05$), and non-significant trends (grey) in $\text{NO}_{2\text{AA}}$ between 2000 and 2014, calculating using a) non-parametric Theil-Sen statistic and Mann Kendall test, b) parametric ordinary least-squares linear regression, c) non-parametric Theil-Sen statistic and block bootstrap resampling, d) parametric first-order autoregressive (AR(1)) model. The black line represents the division between decreasing and increasing trends within the non-significant bar.

5

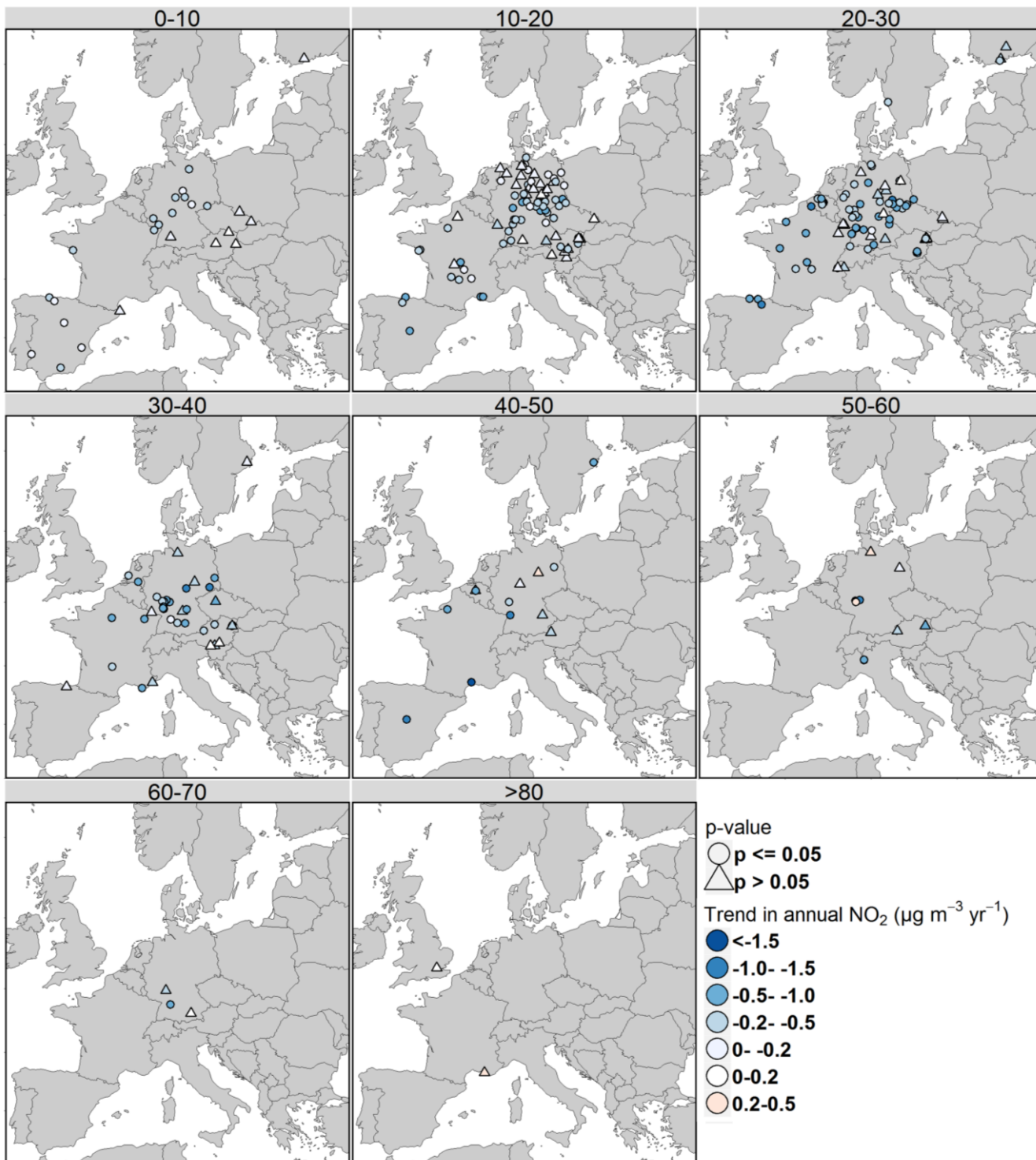


Figure 9: Magnitude and significance of trend in annual average NO_2 concentrations between 2000 and 2014, with sites separated

into panels based on 2010-2014 average annual NO₂ concentrations (NO_{2AA}). The fill colour in each point denotes the magnitude and direction of the Theil-Sen trend at a site, and the outer colour denotes whether the trend was statistically significant ($p \leq 0.05$, green), or not statistically significant ($p > 0.05$, orange). Trend estimates were calculated using a first-order autoregressive (AR(1)) model.

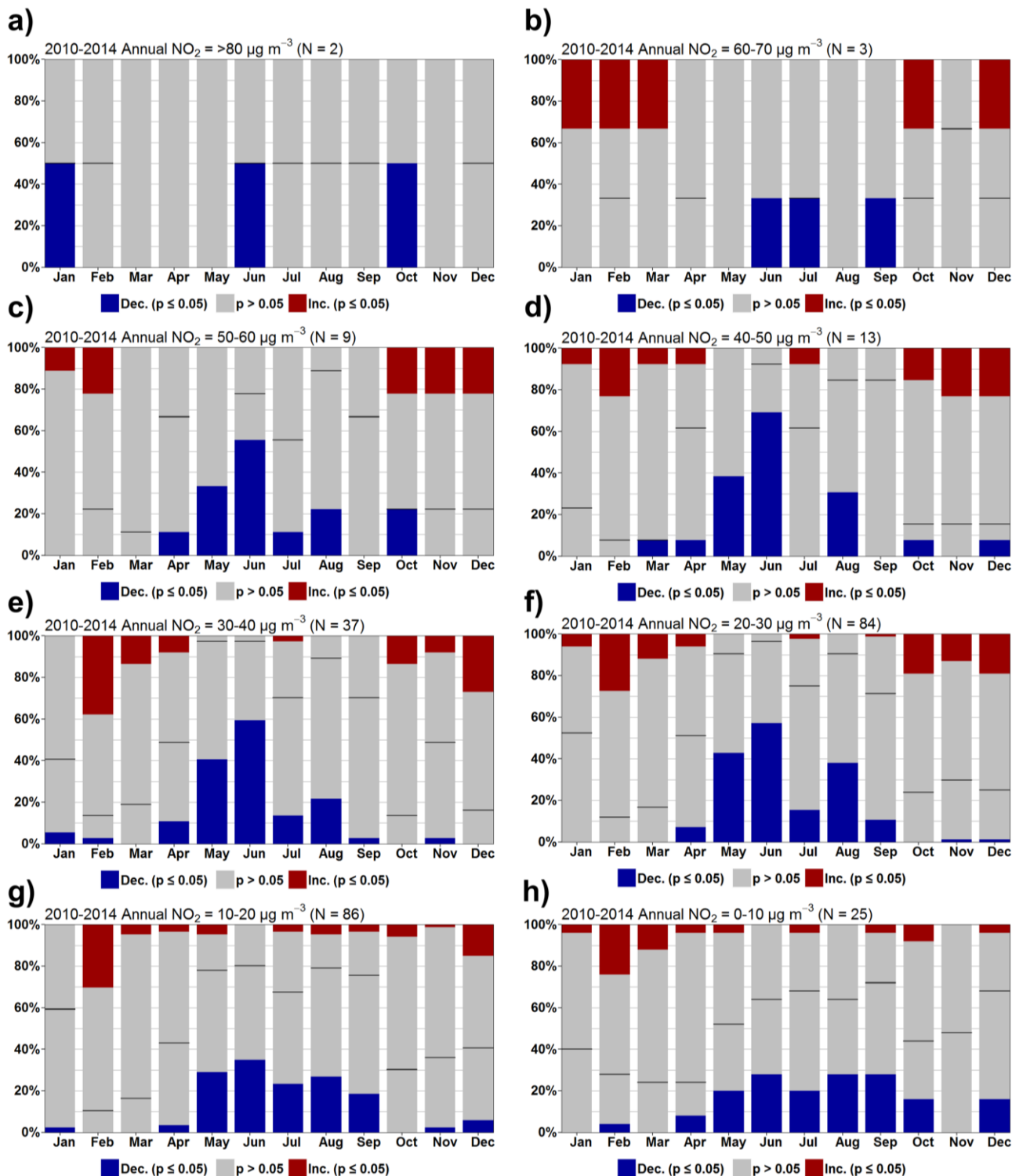


Figure 10: Proportion of sites with significant decreasing (blue), increasing (red) ($p \leq 0.05$), and non-significant (grey) trends in the

monthly percentage contribution to annual average NO₂ between 2000 and 2014, for sites with 2010-2014 average annual NO₂ concentrations (NO_{2AA}) of a) >80 µg m⁻³, b) 60-70 µg m⁻³, c) 50-60 µg m⁻³, d) 40-50 µg m⁻³, 30-40 µg m⁻³, 20-30 µg m⁻³, 10-20 µg m⁻³, 0-10 µg m⁻³. Trend estimates were calculated using a first-order autoregressive (AR(1)) model. The black line represents the division between decreasing and increasing trends within the non-significant bar.

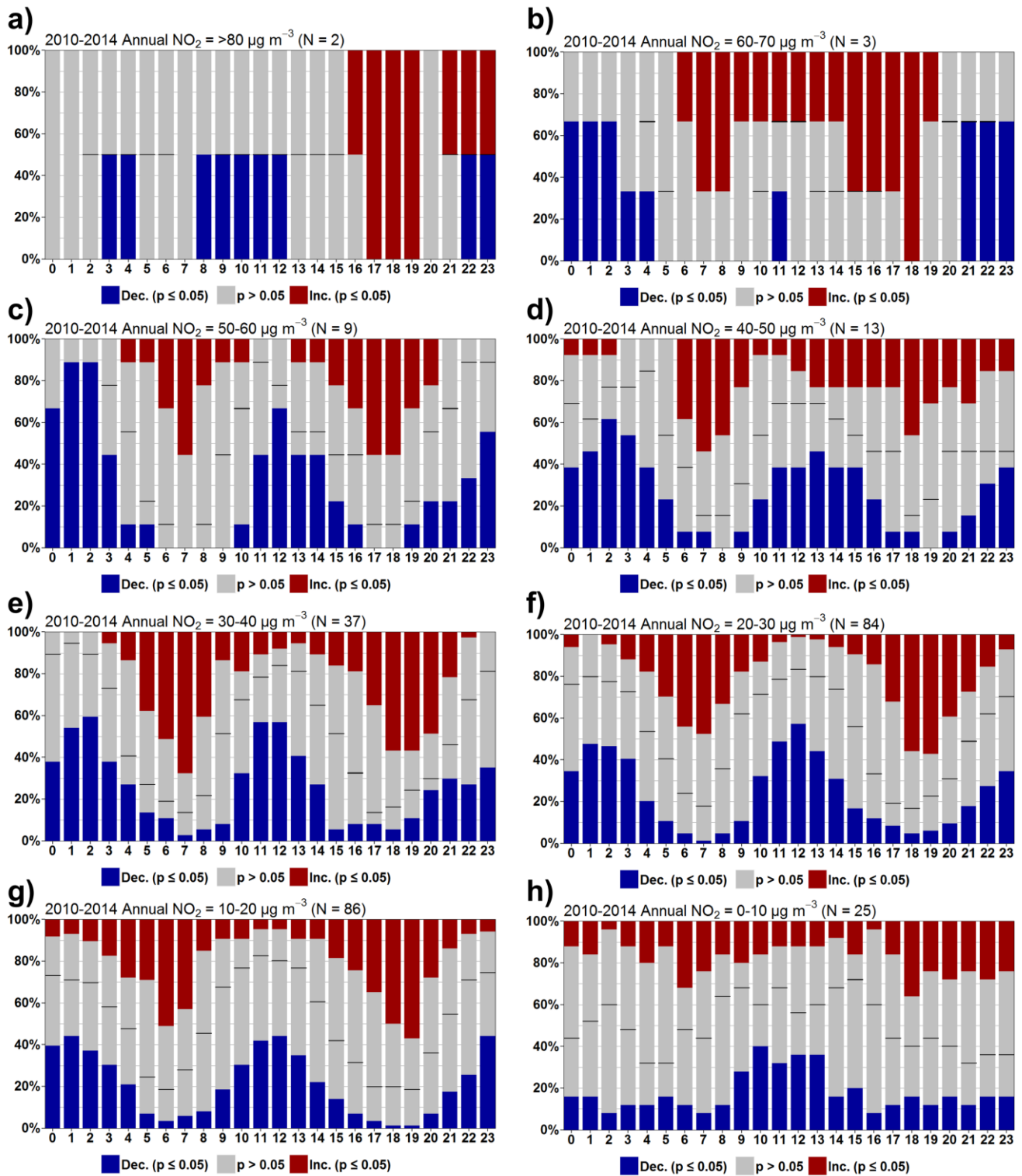


Figure 11: Proportion of sites with significant decreasing (blue), increasing (red) ($p \leq 0.05$), and non-significant (grey) trends in the

percentage contribution of each hour of day to annual average NO₂ between 2000 and 2014, for sites with 2010-2014 average annual NO₂ concentrations (NO_{2AA}) of a) >80 µg m⁻³, b) 60-70 µg m⁻³, c) 50-60 µg m⁻³, d) 40-50 µg m⁻³, 30-40 µg m⁻³, 20-30 µg m⁻³, 10-20 µg m⁻³, 0-10 µg m⁻³. Trend estimates were calculated using a first-order autoregressive (AR(1)) model. The black line represents the division between decreasing and increasing trends within the non-significant bar.

5

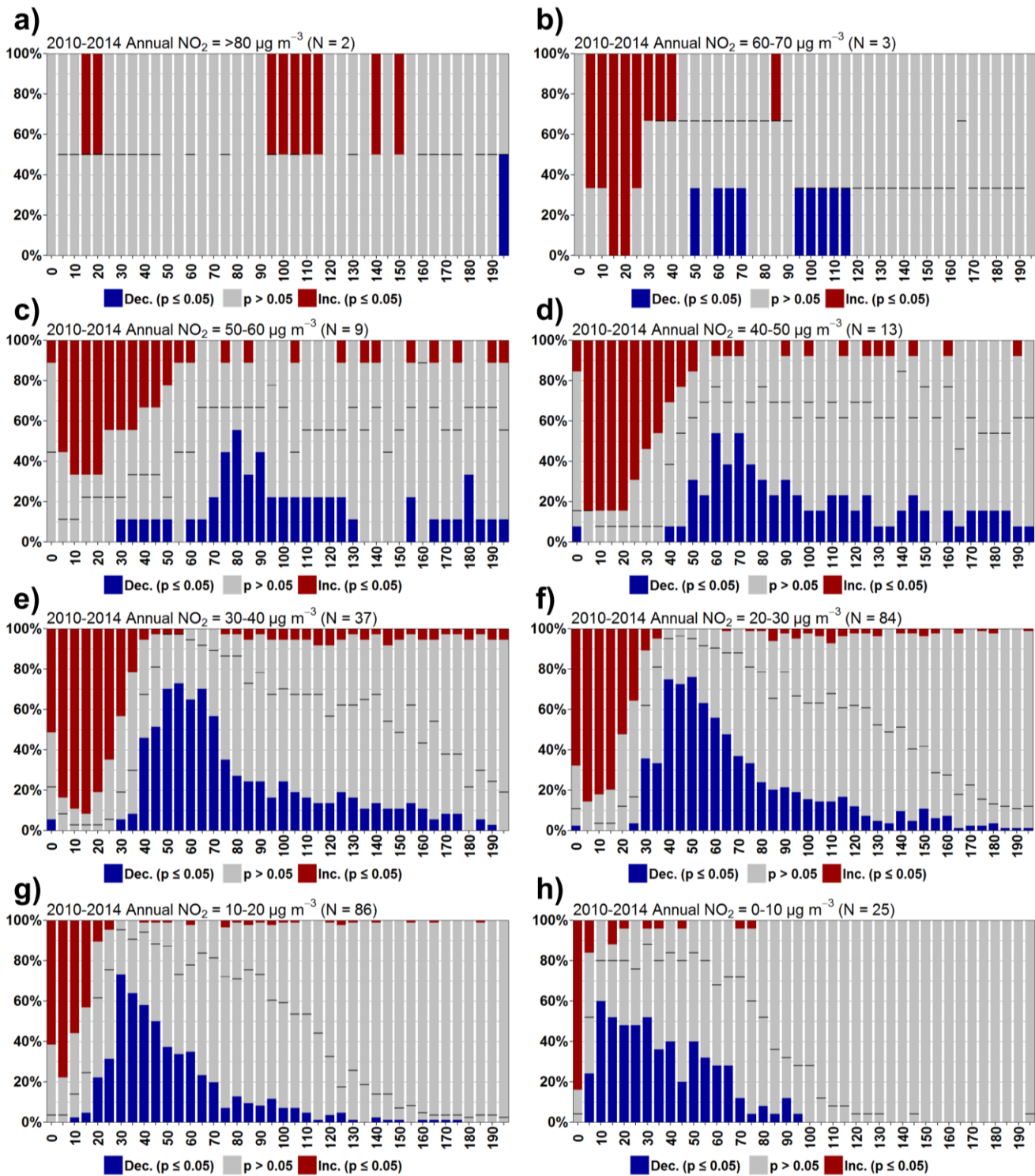


Figure 12: Proportion of sites with significant decreasing (blue), increasing (red) ($p < 0.05$), and non-significant (grey) trends in the

percentage contribution from hourly NO₂ concentrations in 5 µg m⁻³ bins to annual average NO₂ between 2000 and 2014, for sites with 2010-2014 average annual NO₂ concentrations (NO_{2AA}) of a) >80 µg m⁻³, b) 60-70 µg m⁻³, c) 50-60 µg m⁻³, d) 40-50 µg m⁻³, 30-40 µg m⁻³, 20-30 µg m⁻³, 10-20 µg m⁻³, 0-10 µg m⁻³. Trend estimates were calculated using a first-order autoregressive (AR(1)) model. The black line represents the division between decreasing and increasing trends within the non-significant bar.