atmosphere

MDPI

Article

Source Tracing of PM2.5 in a Metropolitan Area Using a Low-Cost Air Quality Monitoring Network: Case Study of Denver, Colorado, USA

Nima Afshar-Mohajer [1,*] and Mirella Shaban [2]

[1] Division of Environmental Sciences, Gradient Corporation, Boston, MA 02108, USA
[2] Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22904, USA; qwe2qh@virginia.edu
* Correspondence: nima.a-mohajer@gradientcorp.com

Abstract: Air quality assessments often require source apportioning of the air pollutants observed at the receptor site. Conventional source apportionment models are subject to high uncertainties due to the lack of accurate emission profiles of all the contributing sources and a limited number of measurements at the receptor sites. Recent advances in the development and application of low-cost $PM_{2.5}$ sensors have facilitated the formation of a more robust database with greater numbers of measurements per location and time. The main objective of this study is to combine a large database of $PM_{2.5}$ concentration records to records from low-cost sensors in Denver, Colorado, during January 2021. Using wind speed and wind direction at the receptors, we developed a visualization tool for source tracing of $PM_{2.5}$ with resulting statistical analyses and back-trajectory modeling. For this purpose, a combination of in-house and existing packages of R scripts along with National Oceanic and Atmospheric Administration (NOAA)'s trajectory model and climate and weather toolkits were used. In general, the results show that the $PM_{2.5}$ measurements obtained from such a network of $PM_{2.5}$ sensors incorporated with hourly wind field data, which are publicly available, can provide a powerful screening tool to discover the transport pathways of $PM_{2.5}$ before requiring costly source apportionment approaches. The fraction of $PM_{2.5}$ concentration detected by each sensor in regard to wind direction and speed bins were quantified using this method. The results of cluster analysis identified the area groups in respect to wind speed and wind direction bins, which shines a light on how far and in which direction polluting sources are. Finally, the back-trajectory modeling outputs illustrated the exact travel path of the $PM_{2.5}$-laden air parcels of each day to each sensor.

Keywords: HYSPLIT; source apportionment; concentration mapping; cluster analysis; urban aerosol

Citation: Afshar-Mohajer, N.; Shaban, M. Source Tracing of PM2.5 in a Metropolitan Area Using a Low-Cost Air Quality Monitoring Network: Case Study of Denver, Colorado, USA. *Atmosphere* **2024**, *15*, 797. https://doi.org/10.3390/atmos15070797

Academic Editor: Sunling Gong

Received: 1 May 2024
Revised: 14 June 2024
Accepted: 28 June 2024
Published: 2 July 2024

1. Introduction

The inhalation of airborne fine particulate matter with an aerodynamic particle diameter smaller than 2.5 μm ($PM_{2.5}$) is associated with well-known respiratory and cardiovascular diseases [1–3]. $PM_{2.5}$ is highly effective at penetrating the alveoli of the lungs. Alveoli are tiny air sacs that cover the interior of the lung. Inhaled $PM_{2.5}$ serves as a carrier of toxic organic compounds (e.g., polycyclic aromatic hydrocarbons) and inorganic chemicals (e.g., heavy metals). Since gas exchange occurs at the alveoli, the heterogeneity of the $PM_{2.5}$ composition contributes to the severity of adverse health impacts [4]. Compared to rural areas, the toxicity of $PM_{2.5}$ in urban areas affects a larger population and is more complex due to the variety of $PM_{2.5}$ polluting sources [5,6]. Statistical analysis by Santibañez et al. (2013) on 5-year $PM_{2.5}$ concentration data of Santiago, Chile showed that for every 10 $\mu g/m^3$ rise in annual mean $PM_{2.5}$ concentration, the risk of emergency hospital admissions due to stroke increased by 1.29% [7]. A similar study by Shah et al. (2015) observed that on average, the annual $PM_{2.5}$ estimated in multiple Mediterranean cities demonstrated an

increase of 10 μg/m^3 and led to an increase of 0.51% (95% CI: 0.12–0.90%) in cardiovascular admissions and 1.36% (95% CI: 0.23–2.49) in respiratory hospitalizations [8].

One of the first steps in the control and mitigation of $PM_{2.5}$ concentrations in an urban area is to identify the polluting sources. Various source apportionment methods have been developed to quantify the contribution of each of the known sources to $PM_{2.5}$ concentration in urban areas. Most of these models (e.g., chemical mass balance or positive matrix factorization) require identifying the chemical composition of the $PM_{2.5}$ samples collected at receptor sites as well as the source profiles of the contributors [9]. The collection of a sufficient mass of air pollutant samples for compositional analyses is expensive, labor-intensive, and cumbersome. $PM_{2.5}$ sampling stations that feed these source apportionment models are scarce and consequently offer limited spatio-temporal resolution.

Alternatively, there is growing attention on utilizing source apportionment techniques relying on statistical methods and back trajectory models [9–12]. Potential source contribution function (PSCF) models use the residence time of air parcel trajectories travelling from different starting points toward each receptor grid cell of a study area and estimate the probability of the air parcel arriving to the grid cells from different sources based on the meteorological trajectory pathways [13]. Although the PSCF models cannot apportion the contribution of sources to the measured receptor data, when there are adequate numbers of air parcels ending at a receptor grid cell, they obtain a reliable map of identified sources for the entire receptor site [9]. It is hypothesized that the reliability of such source identification maps improves with increasing the number of concentration measurement records in space and time.

To assess this hypothesis, data recorded by a well-established network of low-cost sensors monitoring ground-level (<10 m) $PM_{2.5}$ was selected. Although the low-cost sensors implement light scattering to estimate $PM_{2.5}$ and therefore do not collect samples for further laboratory analyses, such networks offer a screening tool for the source tracing of $PM_{2.5}$ by statistical analyses of the arriving air parcel trajectory and investigate their association with wind speed/direction. Furthermore, the application of low-cost $PM_{2.5}$ sensors (<$1500) improves the spatial resolution of the concentration maps [14–16]. With growing interest in air quality monitoring in metropolitan areas, these sensors can record concentrations via a dense network of sensors with a time-resolution as small as one minute. Thus, the resulting dataset of the ground-level $PM_{2.5}$ concentrations obtains a refined and extensive gridded concentration map of the receptor study area.

One month of continuous data measured by a low-cost $PM_{2.5}$ monitoring network in Denver was considered. In fact, as the capital and one of the fastest-growing metropolitan areas in the United States, Denver was an excellent example for the purpose of the present study. This $PM_{2.5}$ monitoring campaign by the Colorado Department of Public Health & Environment (DDPHE) in partnership with Denver Public Schools (DPS) started in 2018 and has since then launched dozens of $PM_{2.5}$ sensors at selected public schools located in regions with higher asthma rates and lower household incomes.

We propose that when there is a lack of source profile data to conduct conventional source apportionment modeling, the transport pathways of the target air pollutant can be explored using the increased number of recorded concentrations paired with the wind field data. The $PM_{2.5}$ concentrations at the receptors are then traced back using statistical schemes coupled with a back-trajectory model to identify the $PM_{2.5}$ origin for a certain period at a target receptor site. The approach to this research is as follows: (1) the concentration data were mapped and temporal variations were visualized; (2) the hourly-averaged ground-level concentration maps overlaid on the (a) corresponding wind vector fields with a spatial resolution of 3 km × 3 km and (b) cluster analysis was performed to quantify shares of each designated wind speed and wind direction group; and (3) back-trajectory modeling was implemented to further refine the results of the statistical analyses.

2. Methods

2.1. $PM_{2.5}$ Monitoring Network

The low-cost $PM_{2.5}$ monitoring network was composed of 24 Canary-S sensors developed by Lunar Outpost Inc., Arvada, CO, USA, with the temporal resolution of one minute sampling periods. The Canary-S sensors utilized Plantower PMS5003 $PM_{2.5}$ sensors utilizing laser beams to count passing airborne particles using their real-time reflectivity [17]. These sensors are programmed to convert the counts into mass-based $PM_{2.5}$ concentration in $\mu g/m^3$. Any row of measurements with unrealistic values (i.e., temperature below 30 °C or negative relative humidity and concentration) were omitted. Any sensor with more than 20% of omitted entries was entirely excluded. The raw measurements of each sensor were corrected using the calibration equation determined by deSouza et al. (2022) [17]. These calibration equations were derived from non-linear correction for environmental parameters including dew point, temperature, and relative humidity in comparison to the "true" $PM_{2.5}$ concentrations measured in tandem by co-location of the instruments complying with the federal equivalent method, meaning designated by the USEPA as a reliable method for monitoring $PM_{2.5}$ concentrations [17]. The accuracy of the measurements using Plantower5003 sensors is ± 0.1 $\mu g/m^3$ [18].

The concentration monitoring period spanned over the month of January 2021 (i.e., 1 January 2021 at 00:00:01 to 31 January 2021 at 23:59:59). One month was found to be long enough for statistical reliability yet short enough to avoid deterioration in the detection accuracy as reported by deSouza et al. (2023) [19], which is a well-known caveat of using low-cost sensors. The positioning of the low-cost sensors in regard to the counties in the immediate vicinity of the city of Denver are displayed in Figure 1. The sensors are labeled by CS followed by a number. Except for two sensors, one located in Arapahoe County (i.e., CS21) and the other one in Jefferson County (CS12), all the other sensors were either inside the city or the county of Denver. Across all the sensors, the pair of CS02 and CS03 sensors were only 236 m apart. Therefore, they were expected to record similar $PM_{2.5}$ measurements.

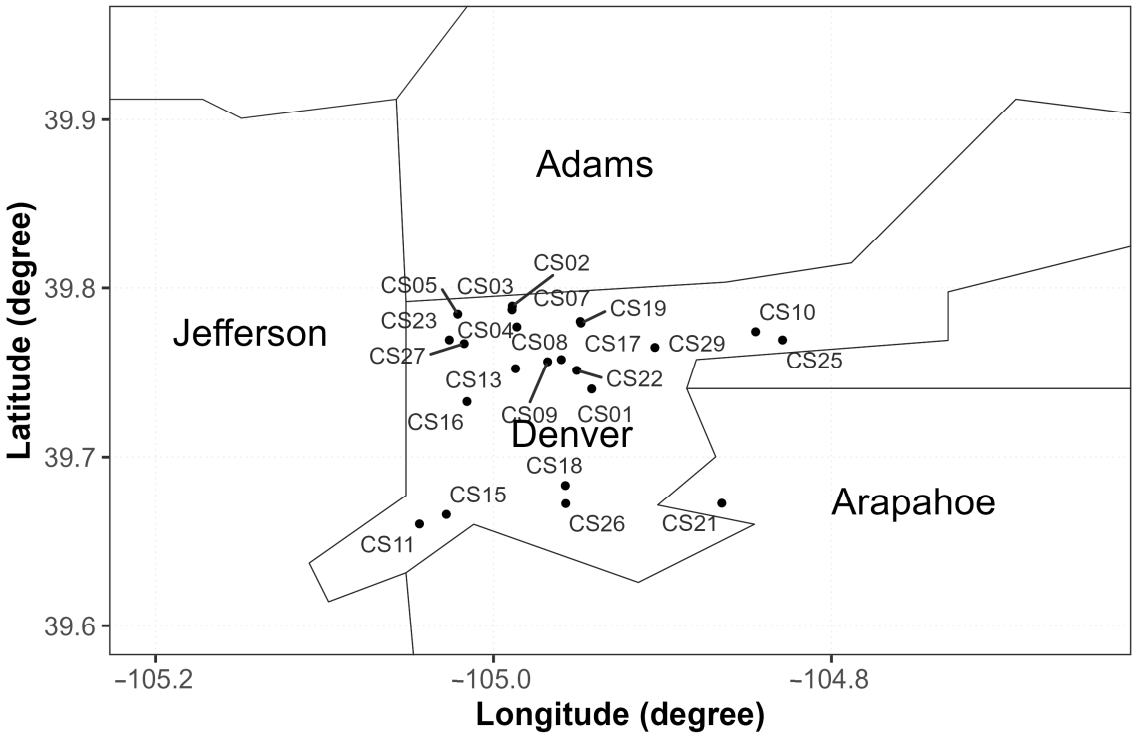

Figure 1. Low-cost $PM_{2.5}$ monitoring network used for the concentration mapping of the city of Denver.

2.2. Wind Fields

The $PM_{2.5}$ transport from the sources to the low-cost sensors is impacted by the wind field (i.e., wind direction and speed at each grid point and time). In this study, the ground-level wind speed and wind direction data were accessed from the High-Resolution Rapid Refresh (HRRR) radar-assimilated database provided by the NOAA with a 3 km grid resolution for each hour. The wind vector components of each grid cell accessed via NOAA's Weather and Climate Toolkit. The wind fields were adjusted for the map projection of Lambert conformal conic using in-house R scripts.

2.3. Statistical and Clustering Analyses

Statistical analyses were conducted using a combination of in-house R scripts and the modification of the openair package developed by Carslaw and Ropkins (2012) [20]. First, hourly variations of the $PM_{2.5}$ concentrations for the entire study area, presented as the arithmetic mean of measurements, were plotted for each day of January 2021. Then, spatial distributions of the $PM_{2.5}$ concentrations averaged by day were mapped. After recognizing the days and hours with elevated $PM_{2.5}$ concentrations, 24h average wind fields for each day were mapped.

In order to link the concentrations measured by the sensors to polluting sources, a series of statistical analyses was conducted on the $PM_{2.5}$ concentrations database and the daily wind field. Various bivariate polar plots were developed by partitioning the hourly wind speeds into 30 intervals and the wind direction into 10 degrees bins for the gridded space ruled by the HRRR dataset, then superimposing that on the $PM_{2.5}$ concentration maps developed by kriging the hourly-averaged records for the entire study area. Specifically, we evaluated the associations between the hourly directional contribution of each wind speed bin and the $PM_{2.5}$ concentration of the corresponding hour.

To identify elevated $PM_{2.5}$ concentrations, originating from the same source, cluster analysis on hourly $PM_{2.5}$ concentrations and corresponding wind speed/direction bins was conducted. The clustering model was constructed around the idea that receptors with simultaneous spikes in $PM_{2.5}$ concentrations might have the same contributing polluting sources. Therefore, wind speed, wind direction, and recorded concentrations were used to form similar clusters [21]. The clustering analysis of this study used k-means to partition M datapoints (i.e., concentrations measured by each sensor) in N dimensions (i.e., wind speed and wind direction) into K clusters. For any given number of clusters initially input to the model, this iterative algorithm minimized the Euclidean distance between the observations (i.e., set of the three hourly-averaged parameters of wind speed, wind direction, and $PM_{2.5}$ concentration records with equal weight) and centroids of each cluster on a polar coordinate. Each observation was assigned to the cluster that had the closest centroid. Finally, the positions of all the cluster centroids were re-calculated until no further cluster centroid re-positioning was feasible. The determination of the optimal number of clusters (K) was based on internal compactness measures using the total within the sum of the squared Euclidian distance and silhouette index and taking the highest index as the number of clusters [21].

2.4. Back-Trajectory Modeling

Back-trajectory analysis was used to explore differences in the air mass origins across the clusters identified in the previous analysis and track them over time. For this purpose, HYbrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT), which is a Lagrangian backward trajectory model developed by the Air Resources Laboratory of the National Oceanic and Atmospheric Administration (NOAA-ARL), was implemented. This model is widely used to trace the source points of a certain air pollutant at a regional scale and beyond [22].

R programming script was developed to attribute hourly $PM_{2.5}$ measurements by the sensor with the time-matched air mass trajectories calculated by the NOAA HYSPLIT model (version 5.2.1). Meteorological data were adopted from the Climate Forecast System

Reanalysis (CFSR) provided by the National Center for Atmospheric Research. The starting height of the trajectories was assumed to be 10 m above ground level. To link the arrival time of each air mass to the recorded $PM_{2.5}$ concentration by each receptor sensor, a four-day back trajectory analysis for each day of January 2021 was performed. The geographical center of the city of Denver was considered as the starting point of all the trajectories.

3. Results

3.1. Descriptive Analytics

Time-series of the recorded data by each sensor are shown in Figure 2. According to these plots, except CS13, all the sensors demonstrated that the peak of the 24-h $PM_{2.5}$ concentration took place on January 28th, with the maximum concentration of 50.4 $\mu g/m^3$ recorded by CS07. The average of the 24-h maximum $PM_{2.5}$ concentrations across all the sensors on 28 January was 48.6 $\mu g/m^3$. This day was the only day at which the 24-h $PM_{2.5}$ concentration exceeded the USEPA's National Ambient Air Quality Standard of 35 $\mu g/m^3$. During the 10-day period of 12 January and 21 January, the mean value of the daily $PM_{2.5}$ concentrations, averaged over all sensors, did not exceed 12 $\mu g/m^3$.

The heatmap of the hourly changes of the $PM_{2.5}$ concentrations over the city of Denver during January 2021, averaged across all the sensors, is depicted in Figure 3. 1 January had elevated hourly $PM_{2.5}$ concentrations between 4 to 9 a.m., which may be due to fireworks and the transportation of the residents celebrating the New Year events. The public schools in Denver were either closed during the first ten days of the month or running virtually due to SARS-CoV-2-related restrictions. There were sporadic raises in the hourly $PM_{2.5}$ concentration on the morning of 11 January. Starting from 22 January, the hourly $PM_{2.5}$ concentrations rose during the daytime and reached a peak mean hourly $PM_{2.5}$ concentration of 62 $\mu g/m^3$ on January 28. The mean of the hourly $PM_{2.5}$ concentration reduced to values smaller than 10 $\mu g/m^3$ in the late morning of January 30, which was a Saturday, and remained low (as denoted by the navy color) throughout the rest of the month. The average hourly $PM_{2.5}$ concentration demonstrated two diurnal period peaks between 6 and 9 a.m. (22% higher than the day mean) and between 6 and 9 pm (27% higher than the day mean).

To gain a better insight into the potential causes of the temporal variability of hourly $PM_{2.5}$ concentrations, a time-series of daily temperatures of the Denver Metropolitan Area during January 2021 was plotted by processing the hourly temperature data of NOAA's HRRR (see Figure 4). As seen, there was a dramatic decrease in the mean daily ground-level temperature from 28 January with respect to 27 January (-10 °C vs. -3 °C). This implies that the notable peak in hourly and daily $PM_{2.5}$ concentrations on January 28th may be attributed to the fossil fuel burned to compensate for the temperature drop and thus heat residential and commercial buildings. As follows, this hypothesis is appraised by comparing increasing or decreasing trends in the daily average plots of the other criteria air pollutants (excluding lead) accessed from the Air Quality System database of the USEPA for January 2021.

Spatial distributions of the 24-h $PM_{2.5}$ concentrations at each sensor location for each day of January 2021 were mapped in Figure 5a. The highest variability of the 24-h $PM_{2.5}$ concentration occurred during the last week of January. The north central part of the study area displayed the highest $PM_{2.5}$ levels, which may be due to the vehicular traffic of daily commuters to Boulder, Colorado. There was no significant difference between the daily concentrations of the western half of the study area compared to the eastern half of the study area. However, on average, there were 13% higher values of 24-h $PM_{2.5}$ levels in the northern half of the study area in respect to the southern half. The overall variability across the sensors on each day was small (3–21%). Consistent with the previous plots, all the sensors recorded a 24-h $PM_{2.5}$ concentration below 8 $\mu g/m^3$ from 14 January to 21 January. Sensor CS13, positioned in Denver downtown near state administrative buildings such as the US courthouse, was the only sensor behaving notably different from the others. This sensor indicated elevated daily $PM_{2.5}$ concentrations between 2 January and 4 January,

which may be attributed to the higher vehicle traffic of residents going back to work after the New Year holidays. However, this sensor showed the lowest concentration when compared to all other sensors for the days after 22 January, signaling erroneous measurements due to a circuit break, calibration requirement, or sensor inlet blockage. deSouza et al. (2022) also reported intermittent power surge issues with CS13 in their sensor calibration study [17]. Therefore, recorded concentrations by Sensor CS13 were unreliable and were excluded in future analyses of the study.

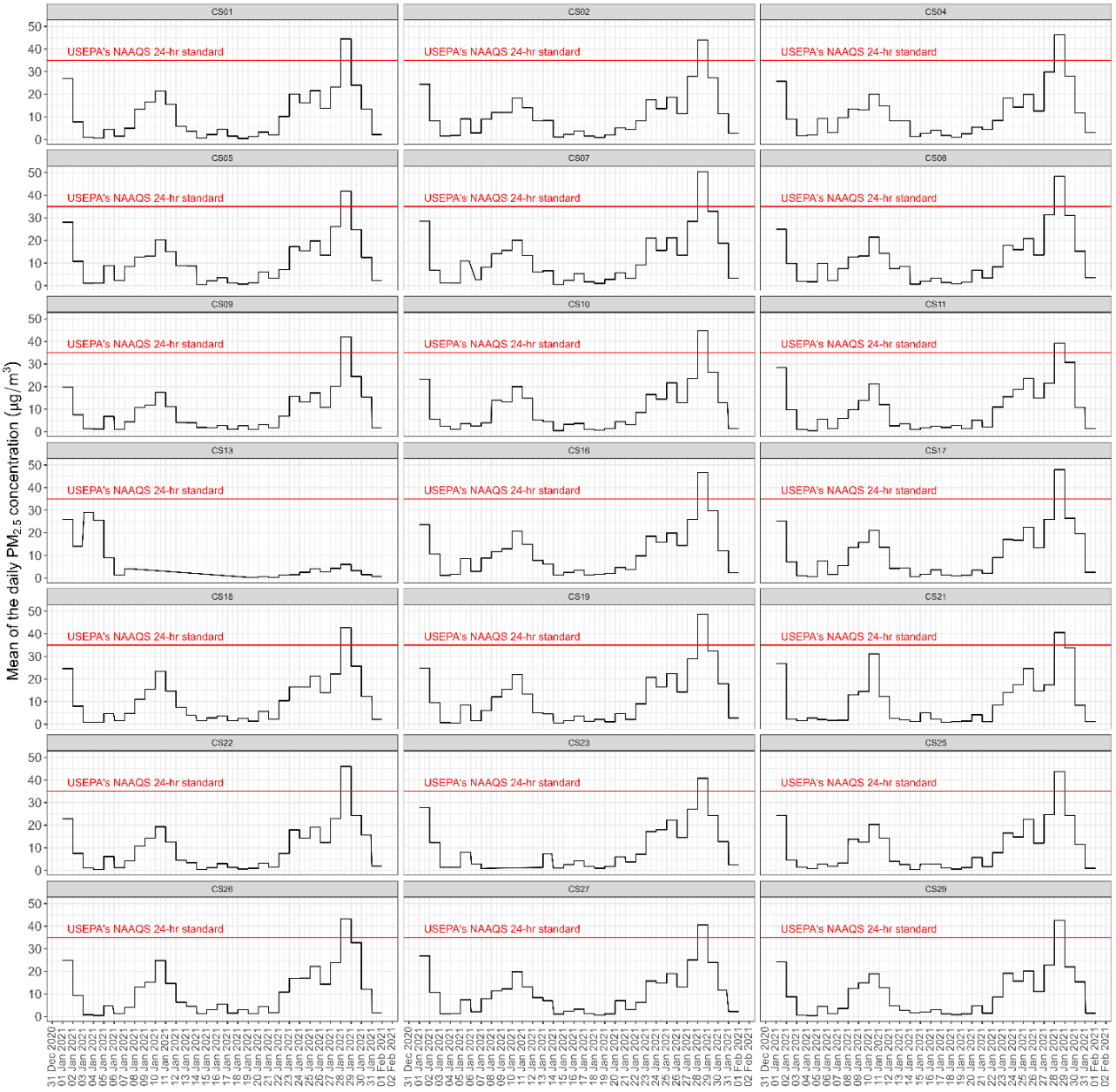

Figure 2. Time-series of the arithmetic mean of 24-h $PM_{2.5}$ concentrations recorded by each low-cost $PM_{2.5}$ sensor.

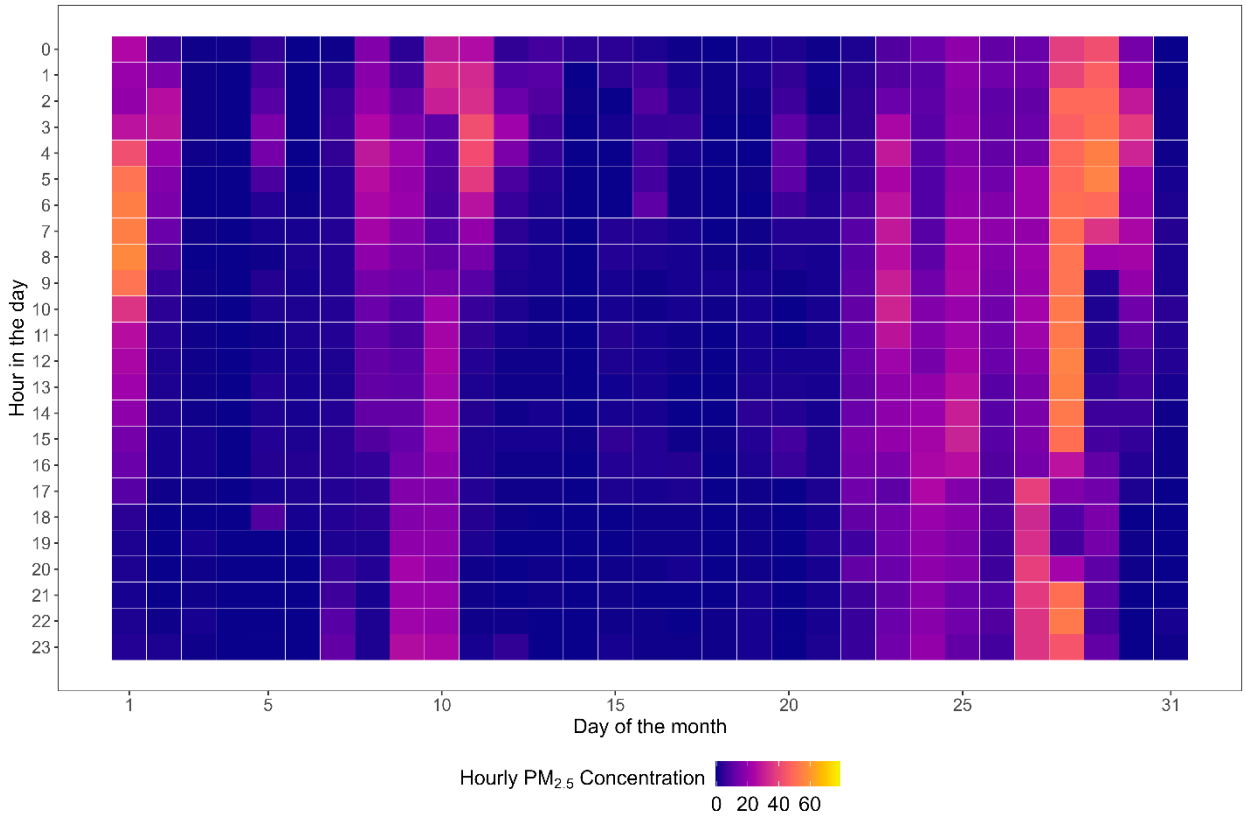

Figure 3. Hourly variability of the $PM_{2.5}$ concentration averaged over the city of Denver using data from all low-cost sensors in January 2021.

The spatial distribution plots of the 24-h $PM_{2.5}$ concentration on each day, based on recorded levels by the USEPA's air quality system (AQS), are displayed in Figure 5b. Considering the entire Denver metropolitan area, the temporal trend provided by both sets were similar. In other words, both the low-cost monitoring network and the USEPA's AQS indicated both the increasing and decreasing patterns summarized in Table 1. However, as seen, during 12–21 January the AQS network reports higher values for the average 24-h $PM_{2.5}$ concentration that is opposite for all other days of the month. A higher standard deviation of the daily $PM_{2.5}$ concentration by the low-cost monitoring network during the two key periods of 27–28 January and 29–31 January implies the high spatial variability of the data, which is disguised in the USEPA's AQS results. This highlights the significance of having more monitors per unit area in an air quality analysis.

One should note that the AQS monitors are mainly located in the north and northwest part of the area and therefore are blind to the daily spatio-temporal changes of $PM_{2.5}$ concentrations in the majority of the Denver Metropolitan area.

The daily dominant wind directions overlaid on the heatmap of the 24-h $PM_{2.5}$ concentrations of the entire study area are displayed in Figure 6. Arrows for each day on this plot denote 24-h average vectors pointing towards the wind direction and scaled by the ground-level wind speed over the city of Denver. 14 January showed the largest arrow for the mean wind speed and had the highest average wind speed of 5.5 ± 2.8 m/s maxing near the northwestern part of the area near Jefferson County. Although more than 95% of the measured ground-level wind speeds at each grid cell of the study area during January 2021 were smaller than 4.8 m/s, a significant association (i.e., p-value < 0.05) between the 24-h wind speed and $PM_{2.5}$ concentration, averaged over the entire area on the same day, was observed. In contrast to the wind intensity, wind direction varied drastically by day. During the last week of January, with an elevated daily $PM_{2.5}$ concentration, the daily prevailing winds mainly blew from north to south between 25 and 30 January, except for 29 January, on which the mean wind direction was from the southwest to the northeast. This

implies that the main contributor to the relatively higher daily $PM_{2.5}$ concentration came from the northern part of Denver. One should note that the northerly wind also prevailed on 14, 15, and 18 January, but it blew at a notably greater wind speed and (23–43% higher) resulted in relatively low daily $PM_{2.5}$ concentrations.

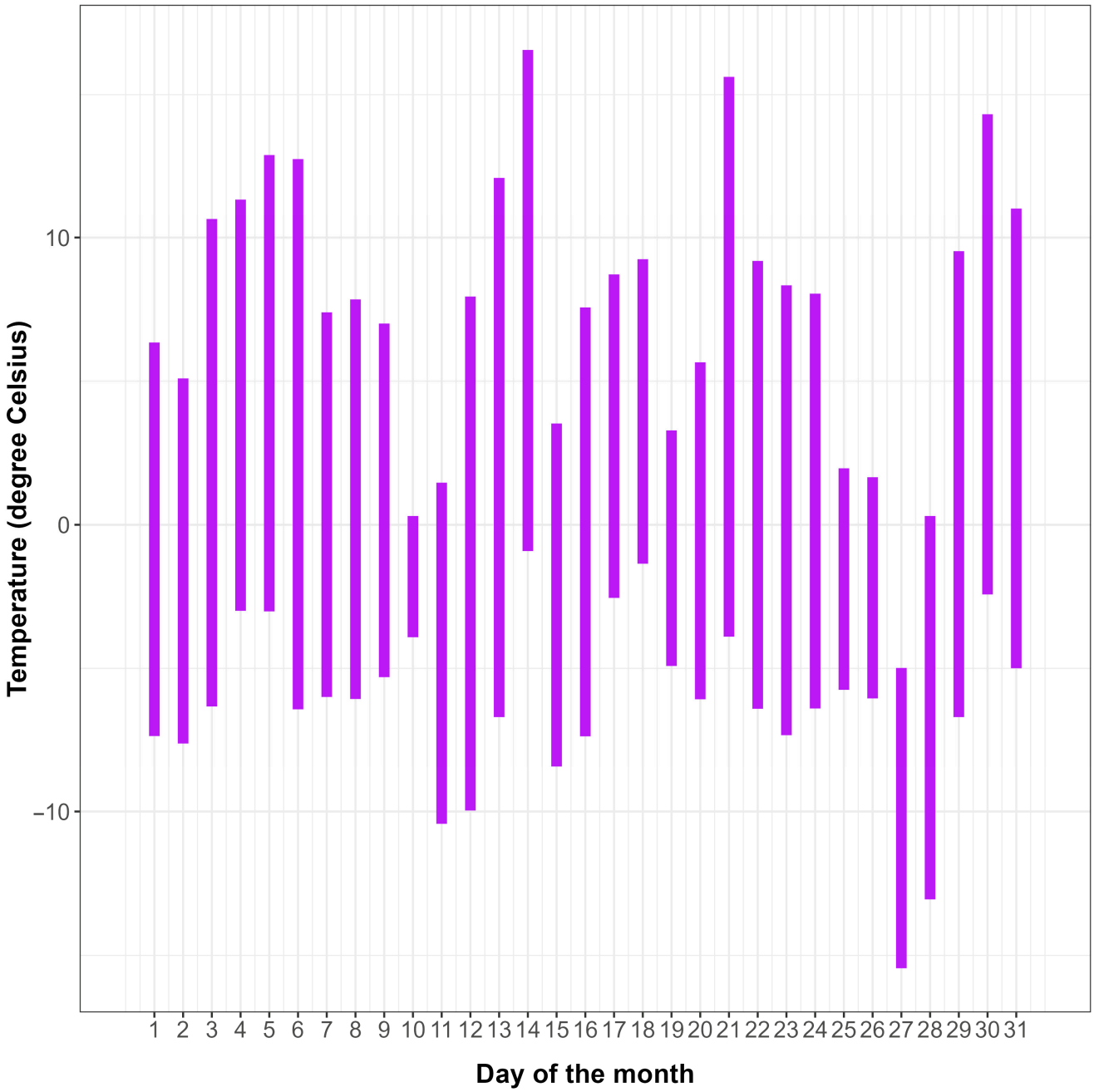

Figure 4. Variability of the hourly ground-level temperature averaged over the grid cells of NOAA's HRRR model of the study area during January 2021.

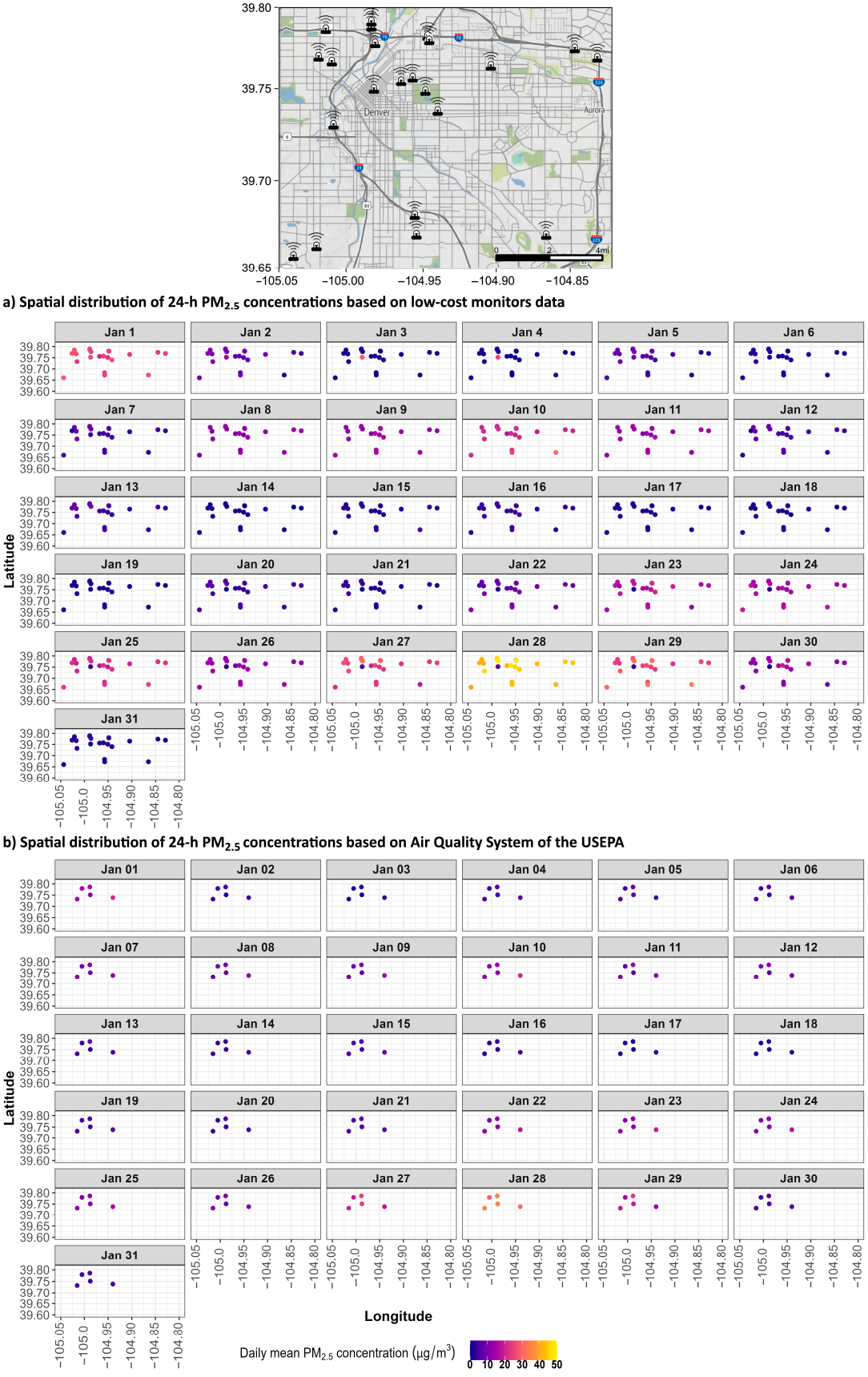

Figure 5. Spatial distribution of the 24-h $PM_{2.5}$ concentrations across the city of Denver during January 2021 obtained from (**a**) the low-cost $PM_{2.5}$ monitors managed by the Colorado DDPHE and DPS and (**b**) the air quality system established by the USEPA.

Table 1. Generic trends in 24-h $PM_{2.5}$ concentrations inferred from both monitoring networks within the Denver metropolitan area.

Period 2021	Range via Low-Cost Monitoring Network (Mean ± Standard Deviation, μg/m^3)	Range via USEPA's AQS Monitoring Network (Mean ± Standard Deviation, μg/m^3)
1 January	25.3 ± 2.0	15.2 ± 3.2
2–6 January	5.2 ± 4.9	4.4 ± 1.6
7–11 January	15.3 ± 3.9	10.6 ± 2.7
12–21 January	3.2 ± 2.1	4.3 ± 2.2
22–25 January	15.1 ± 5.4	12.3 ± 2.4
26 January	12.8 ± 2.5	9.2 ± 1.3
27–28 January	33.3 ± 11.7	28.5 ± 6.2
29–31 January	14.8 ± 10.7	9.0 ± 6.4

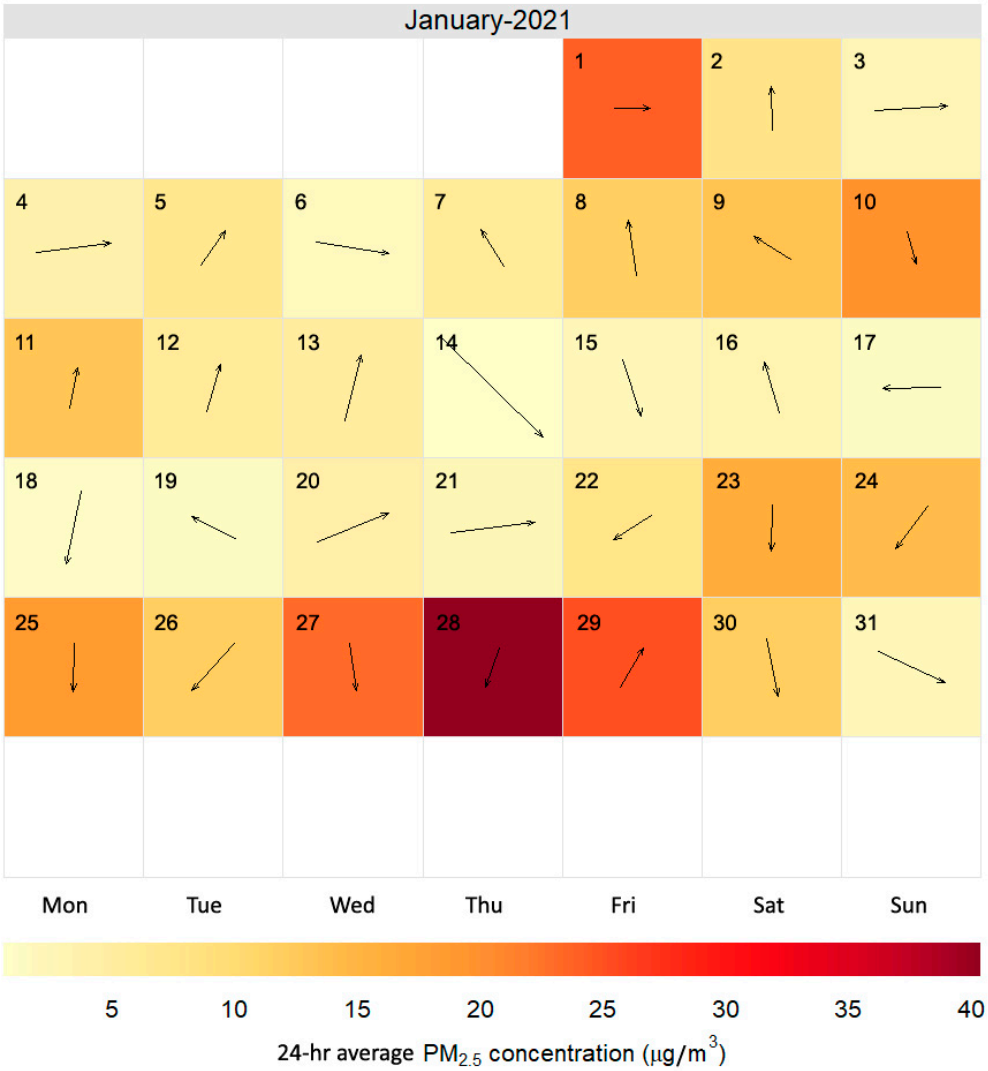

Figure 6. Daily variability of the average $PM_{2.5}$ concentrations over the city of Denver on weekly calendar of January 2021 and the corresponding prevailing wind direction (arrow direction indicates the dominant wind direction and arrow size indicates the wind intensity of the day).

3.2. Source Tracing by Statistical Analyses

The frequency of the elevated $PM_{2.5}$ concentrations measured by each sensor, averaged hourly and accounted for selected wind direction and speed bins, reveals that the majority of the elevated $PM_{2.5}$ traveled from the north/northeast. However, all five sensors located in the southern part of the city had significant levels of $PM_{2.5}$ transported from the southwest (see Figure 7).

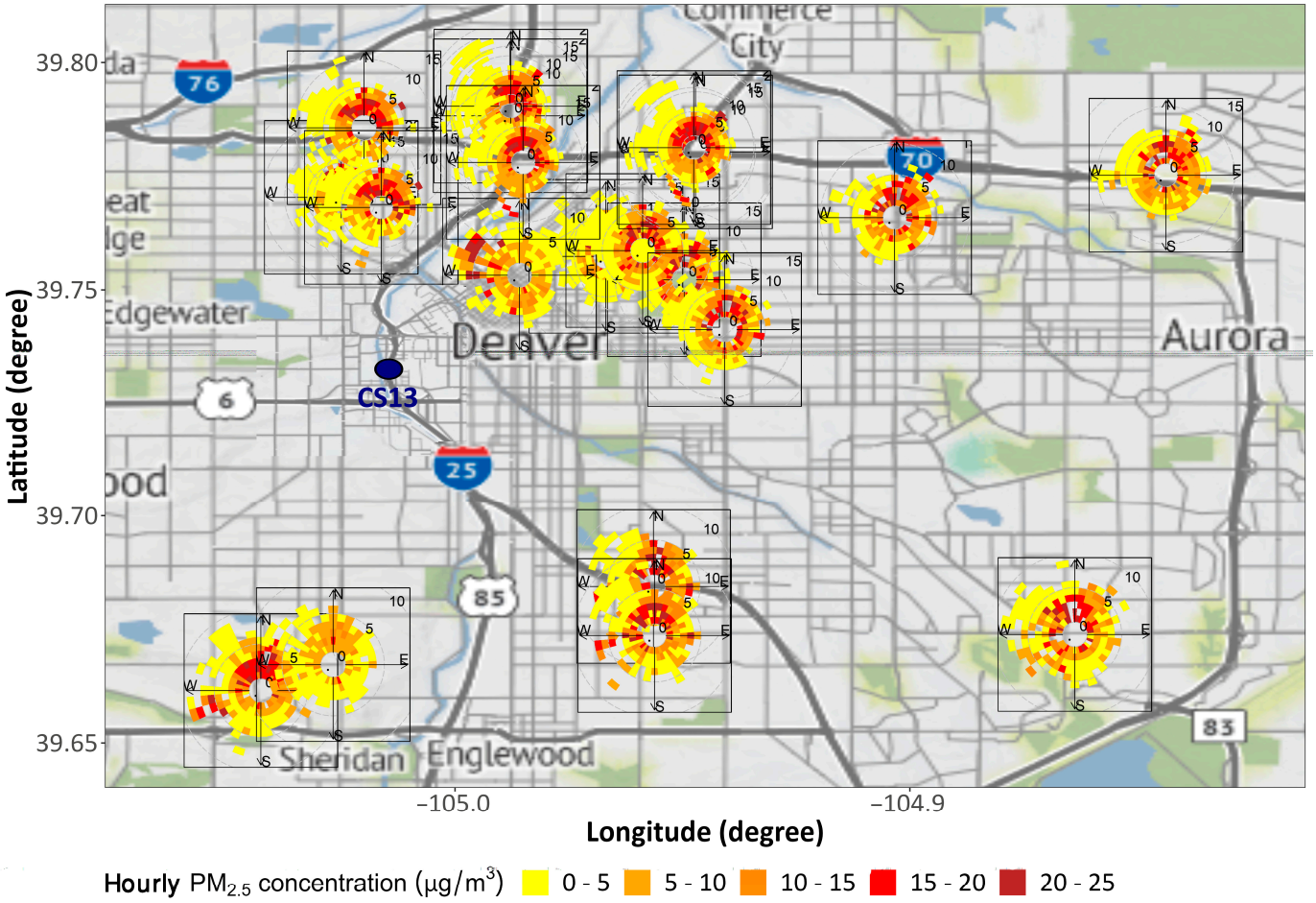

Figure 7. Contribution of each wind direction and speed combination to the recorded hourly $PM_{2.5}$ concentration by each sensor.

One benefit of developing spatial maps of the $PM_{2.5}$ clusters is to match these four clusters with the limited number of USEPA AQS monitoring criteria air pollutants. As seen in Figures 5 and 8, there were only a handful of AQS monitors in the northwest part of the city. Therefore, the spatial distribution of each air pollutant on each day cannot be assessed. Only SO_2 levels were higher in the northwest part of the area at all days. Similar to $PM_{2.5}$, other air pollutants (excluding O_3) demonstrated peak levels on 28 January. Elevated concentrations for the daily maximum SO_2 during 26–28 January signal an increase in emissions from coal fired power plants. Since both daily averages of $PM_{2.5}$ and PM_{10} concentrations peaked on 28 January, primary sources rather than secondary organic aerosols are likely responsible for the airborne pollution. The daily changes of O_3 concentrations followed an opposite trend compared to the rest of the other air pollutants. The opposite temporal trend of O_3 and NO_2 has also been observed by other researchers monitoring urban areas [23–25], and is due to the diurnal variation of vehicular traffic emissions, photochemical reactions, and meteorological parameters such as cloud cover and precipitation changing the concentrations of the oxidative radicals and therefore affecting sink and source patterns for NO_2 and O_3 [26]. The analysis of the Moderate

Resolution Imaging Spectroradiometer (MODIS) on the National Aeronautics and Space Administration's Terra satellite and its Visible Infrared Imaging Radiometer Suite (VIIRS) instrument revealed that there were two major fires in the south/southwest area of the study on 27 January that justify the elevated CO and NO_2 concentrations on 27 and 28 January based on CO and NO_2 plots (Figure 8) increasing toward the south.

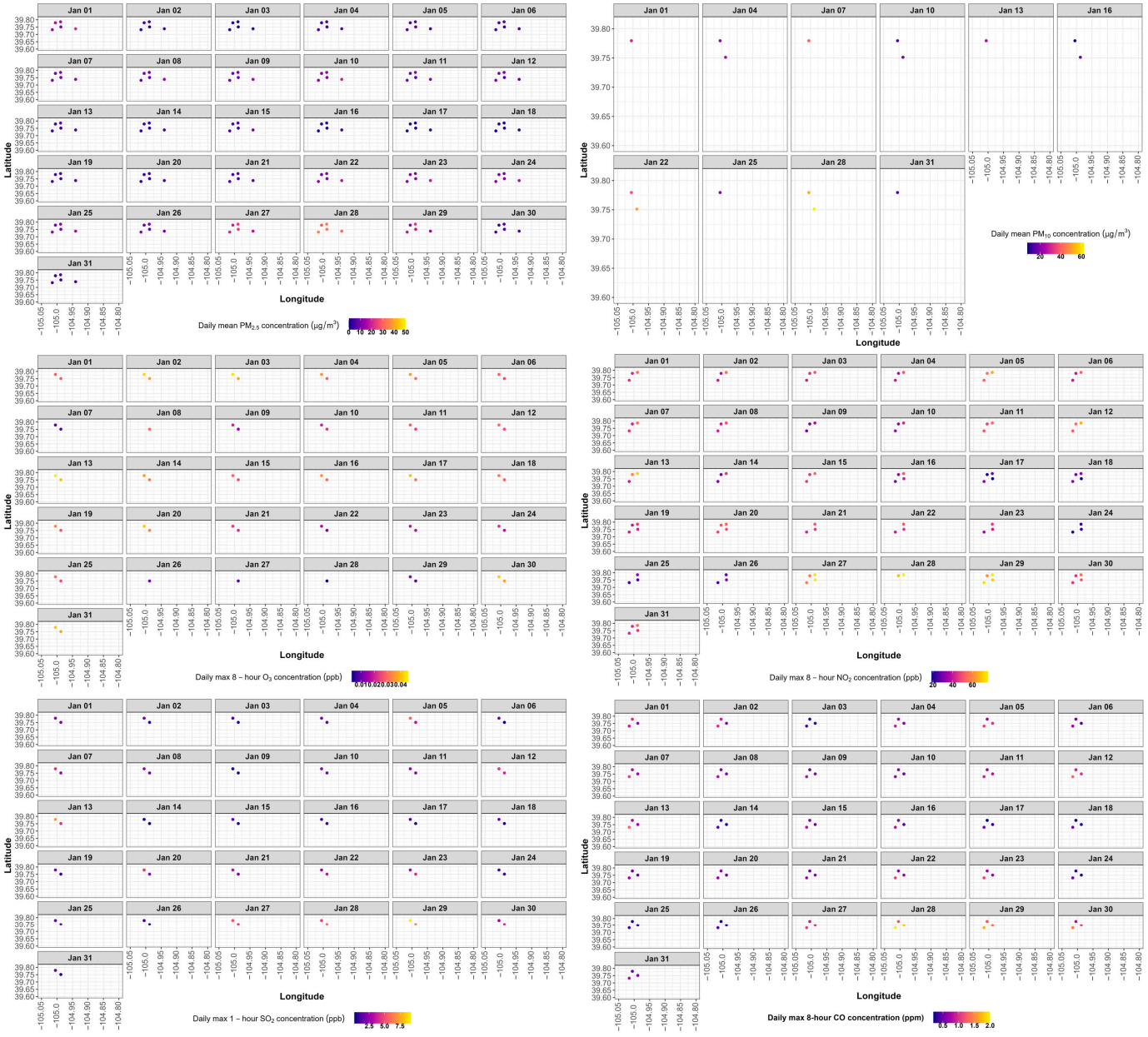

Figure 8. Spatial distribution plots of the daily maximum concentration of the criteria air pollutants on different days of January 2021 monitored by USEPA AQS.

The results of the cluster analysis are presented in Figure 9. Based on the clustering criteria, four groups (labeled by C1 through C4) with similar trends in $PM_{2.5}$ concentration levels were identified. The samples grouped in Cluster C1 encompassed the southwest to the center of the area. C1 had the highest daily average $PM_{2.5}$ concentrations attributed to the highest measurement records during the early morning hours of the weekdays, potentially due to the inflow of vehicular traffic from Lakewood to downtown Denver. On weekends, the west side of the city, covered by Clusters C1 and C2, was more $PM_{2.5}$-polluted than the other areas. The time-series plots of Cluster C3 in the northeastern part of

the study area (the furthest from downtown) correlated to those receptors with the highest concentrations in the morning time (until 6 am during workdays). Located in the northwest part of the area, Cluster C4 was composed of receptors with $PM_{2.5}$ concentrations resulting from the highest wind intensity. Compared to the other clusters, this cluster had the lowest concentrations on the workdays.

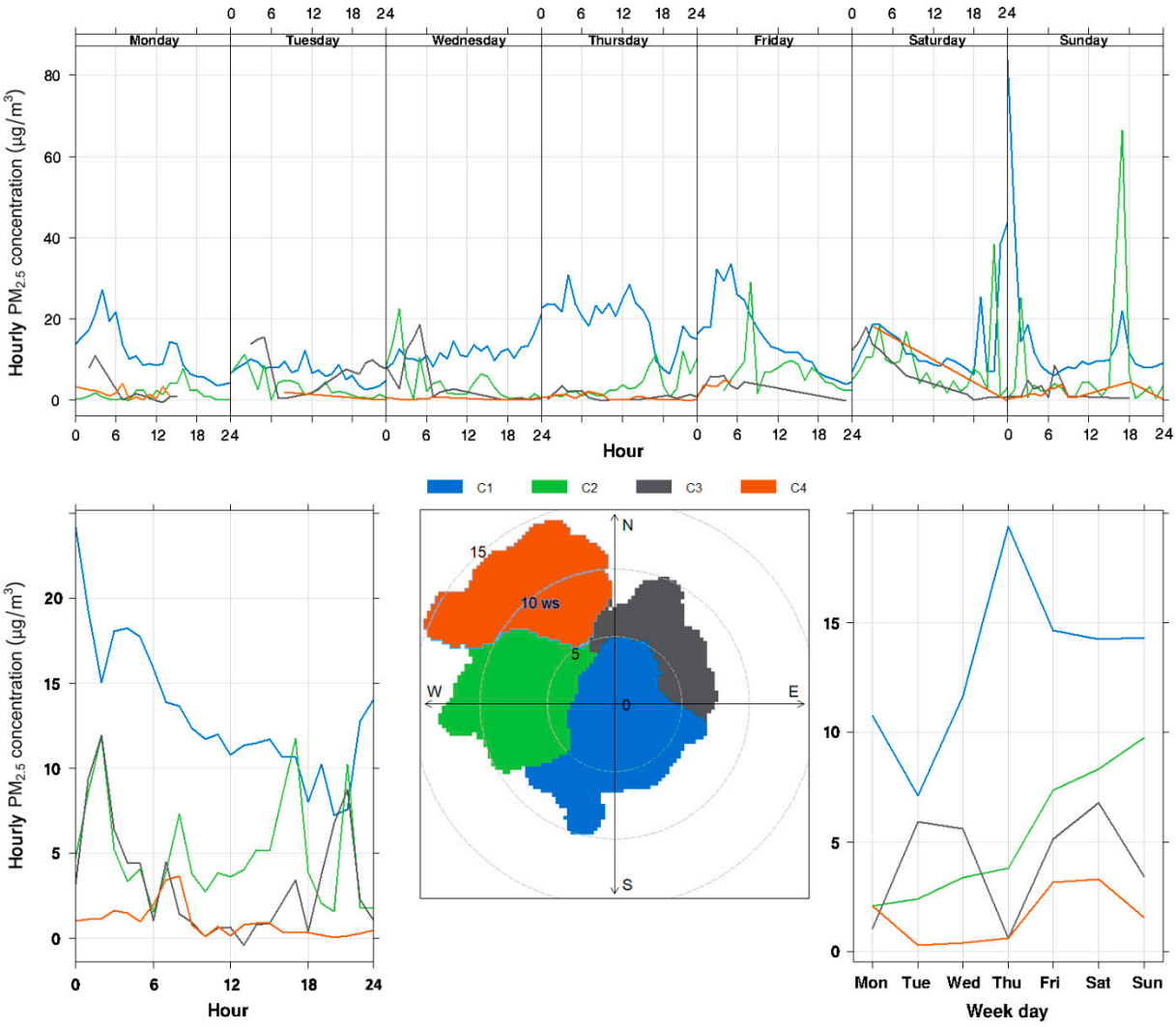

Figure 9. Identified clusters in regard to temporal trends in $PM_{2.5}$ concentration levels based on the wind field.

On days with greater daily average PM_{10} concentrations and relatively lower daily average $PM_{2.5}$ concentrations (e.g., 7 January and 22 January as seen in the first row of plots in Figure 8), the PM_{10} concentrations are relatively high while the $PM_{2.5}$ concentrations are not. Therefore, there is a considerable fraction of airborne coarse particles with aerodynamic diameters between 2.5 and 10 µm. This implies that the $PM_{2.5}$ pollution in the northwest part of the area, corresponding to Cluster C4, may be formed from tire abrasions on the pavement, resuspension of the road dust as the traffic moves, or wind-blown soil [27,28]. Valerino et al. (2017) analyzed the chemical composition of the $PM_{2.5}$ samples collected in Golden, Colorado, and in agreement with the present findings, demonstrated that organic matter comprised about 75% of $PM_{2.5}$ [29]. Similar results were drawn by Dutton et al., 2010 [30] during the Denver Aerosol Sources and Health study wherein their 5.5-year campaign of $PM_{2.5}$ samples over Denver, they speciated the samples to be 86% organic. This fraction included SOA and was relatively stable through all weekdays. The rest of

the $PM_{2.5}$ samples of their study comprised heavy metals and reduced significantly over the weekend.

3.3. Source Tracing by Back Trajectories

The four-day back trajectories of ground-level air parcels resulting from the HYSPLIT model for all low-cost sensor input as receptors and for each day of January 2021 are depicted in Figure 10. The back-trajectory lines of the first few days after the new year with relatively high $PM_{2.5}$ concentrations were stipulated to be regional. As discussed, the overall $PM_{2.5}$ pollution averaged over the city of Denver during the later part of January 2021 was reduced, which emanated from the north/northeast. The back-trajectory lines implied that the main sources of $PM_{2.5}$ in the last week of January 2021, which was the most $PM_{2.5}$ polluted period in this study, might have travelled from the northwest as a long-range transport. This observation matches with the daily average wind direction of the last week of January 2021, during which north-to-south winds were dominant except on January 29 (see Figure 6).

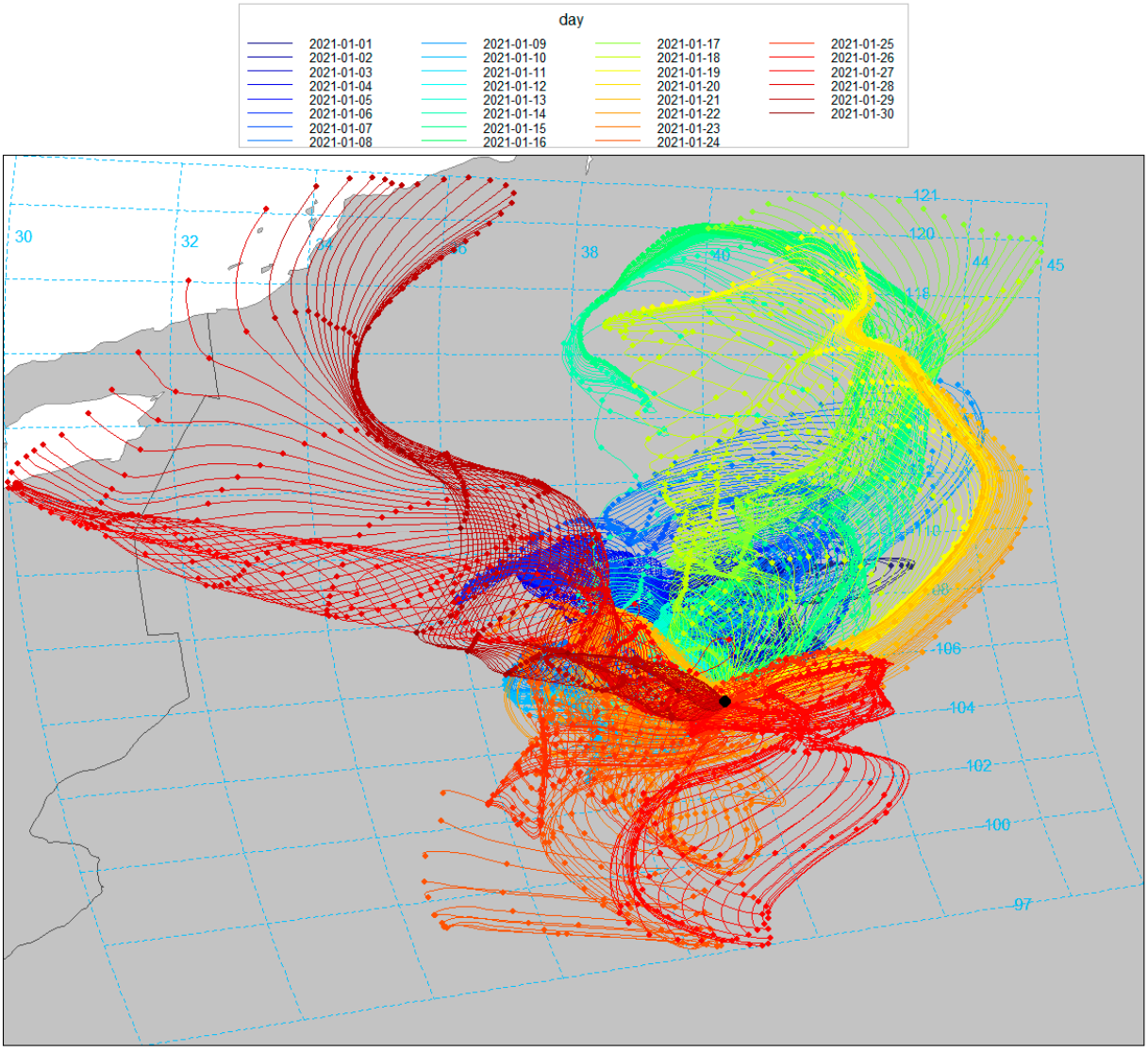

Figure 10. Four-day back trajectories of the air parcels for each day of January 2021.

4. Limitations

The meteorological data used in the modeling are based on the HRRR model. The data records presented by HRRR are recorded on an hourly basis and represent snapshots

of the calculated conditions at those times over 3 km by 3 km horizontal grids. Since all low-cost sensors are located at approximately the same height, the vertical spread of the $PM_{2.5}$ concentration cannot be determined. Consequently, vertical shear and the impact of the mixed layer that depends on the meteorological data and are integrated into NOAA's HYSPLIT model were ignored. $PM_{2.5}$ concentration records at various heights for the same location are required for leveraging that feature and reducing the uncertainties of the model.

The meteorology modeling system of HYSPLIT runs could be replaced by the Weather Research and Forecasting (WRF) that can provide temporal resolution as high as one minute (i.e., comparable to the temporal resolution of the low-cost sensors) and tailored for month-long studies. In this study, we were limited to apply CFS with a coarser time resolution of 6 h due to computational cost limitations.

5. Conclusions

The application of a dense network of low-cost sensors in air quality monitoring of megacities and metropolitan areas offers a new benefit of source tracing elevated $PM_{2.5}$ concentration within the study area. Using the large size of the dataset resulting from these sensors coupled with hourly wind field data, multiple statistical methods including cluster analysis along with back-trajectory modeling via NOAA's HYSPLIT could be utilized to identify the transport pathways of $PM_{2.5}$ in the metropolitan city of Denver, Colorado. This is performed by clustering bins of recorded $PM_{2.5}$ levels at certain receptors for each hour based on their hourly wind vector data for the receptor. This technique serves as a screening tool to identify the qualitative trend in $PM_{2.5}$ concentration variability and potential source types when compared with corresponding spatio-temporal maps of the other air pollutants. Finally, back-trajectory modeling obtains more details regarding the point of origins for $PM_{2.5}$ levels recorded and averaged for the study area. Overall, in the absence of emission profiles of suspected sources, the presented methodology serves as a screening tool to trace potential stationary sources.

Author Contributions: Conceptualization, N.A.-M.; Methodology, N.A.-M.; Formal analysis, N.A.-M.; Writing—original draft, N.A.-M.; Writing—review & editing, N.A.-M. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available in a publicly accessible repository that does not issue DOIs Publicly available datasets were analyzed in this study. This data can be found here: [https://github.com/TehyaStockman/Denver-PM-Project/tree/main/Raw_Data/2021/202101] accessed on 14 June 2024.

Acknowledgments: Support from Tehya Stockman for accessing the sensor data of the Love my Air Program is appreciated.

Conflicts of Interest: Nima Afshar-Mohajer and Mirella Shaban declare no conflict of interest.

References

1. Laden, F.; Neas, L.M.; Dockery, D.W.; Schwartz, J. Association of fine particulate matter from different sources with daily mortality in six US cities. *Environ. Health Perspect.* **2000**, *108*, 941–947. [CrossRef]
2. Pope, C.A., III; Turner, M.C.; Burnett, R.T.; Jerrett, M.; Gapstur, S.M.; Diver, W.R.; Krewski, D.; Brook, R.D. Relationships between fine particulate air pollution, cardiometabolic disorders, and cardiovascular mortality. *Circ. Res.* **2015**, *116*, 108–115. [CrossRef]
3. Xing, Y.F.; Xu, Y.H.; Shi, M.H.; Lian, Y.X. The impact of $PM_{2.5}$ on the human respiratory system. *J. Thorac. Dis.* **2016**, *8*, E69. [PubMed]
4. Kim, S.; Yang, J.; Park, J.; Song, I.; Kim, D.G.; Jeon, K.; Kim, H.; Yi, S.M. Health effects of $PM_{2.5}$ constituents and source contributions in major metropolitan cities, South Korea. *Environ. Sci. Pollut. Res.* **2022**, *29*, 82873–82887. [CrossRef]

5. Pinto, J.P.; Lefohn, A.S.; Shadwick, D.S. Spatial variability of $PM_{2.5}$ in urban areas in the United States. *J. Air Waste Manag. Assoc.* **2004**, *54*, 440–449. [CrossRef]
6. Radulescu, C.; Iordache, S.; Dunea, D.; Stihi, C.; Dulama, I.D. Risks assessment of heavy metals on public health associated with atmospheric exposure to $PM_{2.5}$ in urban area. *Rom. J. Phys.* **2015**, *60*, 1171–1182.
7. Santibañez, D.A.; Ibarra, S.; Matus, P.; Seguel, R. A five-year study of particulate matter ($PM_{2.5}$) and cerebrovascular diseases. *Environ. Pollut.* **2013**, *181*, 1–6.
8. Shah, A.S.; Lee, K.K.; McAllister, D.A.; Hunter, A.; Nair, H.; Whiteley, W.; Langrish, J.P.; Newby, D.E.; Mills, N.L. Short term exposure to air pollution and stroke: Systematic review and meta-analysis. *BMJ* **2015**, *350*, h1295. [CrossRef] [PubMed]
9. Hopke, P.K. Review of receptor modeling methods for source apportionment. *J. Air Waste Manag. Assoc.* **2016**, *66*, 237–259. [CrossRef]
10. Godoy, M.L.D.; Godoy, J.M.; Roldao, L.A.; Soluri, D.S.; Donagemma, R.A. Coarse and fine aerosol source apportionment in Rio de Janeiro, Brazil. *Atmos. Environ.* **2009**, *43*, 2366–2374. [CrossRef]
11. Contini, D.; Genga, A.; Cesari, D.; Siciliano, M.; Donateo, A.; Bove, M.C.; Guascito, M.R. Characterisation and source apportionment of PM_{10} in an urban background site in Lecce. *Atmos. Res.* **2010**, *95*, 40–54. [CrossRef]
12. Gebhart, K.A.; Schichtel, B.A.; Malm, W.C.; Barna, M.G.; Rodriguez, M.A.; Collett, J.L., Jr. Back-trajectory-based source apportionment of airborne sulfur and nitrogen concentrations at Rocky Mountain National Park, Colorado, USA. *Atmos. Environ.* **2011**, *45*, 621–633. [CrossRef]
13. Ashbaugh, L.L.; Malm, W.C.; Sadeh, W.Z. A residence time probability analysis of sulfur concentrations at Grand Canyon National Park. *Atmos. Environ.* **1985**, *19*, 1263–1270. [CrossRef]
14. Jain, S.; Presto, A.A.; Zimmerman, N. Spatial modeling of daily $PM_{2.5}$, NO_2, and CO concentrations measured by a low-cost sensor network: Comparison of linear, machine learning, and hybrid land use models. *Environ. Sci. Technol.* **2021**, *55*, 8631–8641. [CrossRef]
15. Lu, Y.; Giuliano, G.; Habre, R. Estimating hourly $PM_{2.5}$ concentrations at the neighborhood scale using a low-cost air sensor network: A Los Angeles case study. *Environ. Res.* **2021**, *195*, 110653. [CrossRef]
16. Zikova, N.; Masiol, M.; Chalupa, D.C.; Rich, D.Q.; Ferro, A.R.; Hopke, P.K. Estimating hourly concentrations of PM2. 5 across a metropolitan area using low-cost particle monitors. *Sensors* **2017**, *17*, 1922. [CrossRef]
17. deSouza, P.; Kahn, R.; Stockman, T.; Obermann, W.; Crawford, B.; Wang, A.; Crooks, J.; Li, J.; Kinney, P. Calibrating networks of low-cost air quality sensors. *Atmos. Meas. Tech.* **2022**, *15*, 6309–6328. [CrossRef]
18. Considine, E.M.; Reid, C.E.; Ogletree, M.R.; Dye, T. Improving accuracy of air pollution exposure measurements: Statistical correction of a municipal low-cost airborne particulate matter sensor network. *Environ. Pollut.* **2021**, *268*, 115833. [CrossRef]
19. deSouza, P.; Barkjohn, K.; Clements, A.; Lee, J.; Kahn, R.; Crawford, B.; Kinney, P. An analysis of degradation in low-cost particulate matter sensors. *Environ. Sci. Atmos.* **2023**, *3*, 521–536. [CrossRef]
20. Carslaw, D.C.; Ropkins, K. Openair—An R package for air quality data analysis. *Environ. Model. Softw.* **2012**, *27–28*, 52–61. [CrossRef]
21. Carslaw, D.C.; Beevers, S.D. Characterising and understanding emission sources using bivariate polar plots and k-means clustering. *Environ. Model. Softw.* **2013**, *40*, 325–329. [CrossRef]
22. Draxler, R.R.; Rolph, G.D. *HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory) Model Access via NOAA ARL READY Website (http://ready.arl.noaa.gov/HYSPLIT.php)*; NOAA Air Resources Laboratory: Silver Spring, MD, USA, 2010.
23. Han, S.; Bian, H.; Feng, Y.; Liu, A.; Li, X.; Zeng, F.; Zhang, X. Analysis of the relationship between O_3, NO and NO_2 in Tianjin, China. *Aerosol Air Qual. Res.* **2011**, *11*, 128–139. [CrossRef]
24. Agudelo–Castaneda, D.M.; Teixeira, E.C.; Pereira, F.N. Time–series analysis of surface ozone and nitrogen oxides concentrations in an urban area at Brazil. *Atmos. Pollut. Res.* **2014**, *5*, 411–420. [CrossRef]
25. Zoran, M.A.; Savastru, R.S.; Savastru, D.M.; Tautan, M.N. Assessing the relationship between ground levels of ozone (O_3) and nitrogen dioxide (NO_2) with coronavirus (COVID-19) in Milan, Italy. *Sci. Total Environ.* **2020**, *740*, 140005. [CrossRef] [PubMed]
26. Yu, H.L.; Lin, Y.C.; Kuo, Y.M. A time series analysis of multiple ambient pollutants to investigate the underlying air pollution dynamics and interactions. *Chemosphere* **2015**, *134*, 571–580. [CrossRef] [PubMed]
27. Lenschow, P.; Abraham, H.J.; Kutzner, K.; Lutz, M.; Preuß, J.D.; Reichenbächer, W. Some ideas about the sources of PM_{10}. *Atmos. Environ.* **2001**, *35*, S23–S33. [CrossRef]
28. Achilleos, S.; Evans, J.S.; Yiallouros, P.K.; Kleanthous, S.; Schwartz, J.; Koutrakis, P. PM_{10} concentration levels at an urban and background site in Cyprus: The impact of urban sources and dust storms. *J. Air Waste Manag. Assoc.* **2014**, *64*, 1352–1360. [CrossRef] [PubMed]
29. Valerino, M.J.; Johnson, J.J.; Izumi, J.; Orozco, D.; Hoff, R.M.; Delgado, R.; Hennigan, C.J. Sources and composition of $PM_{2.5}$ in the Colorado Front Range during the DISCOVER-AQ study. *J. Geophys. Res. Atmos.* **2017**, *122*, 566–582. [CrossRef]
30. Dutton, S.J.; Rajagopalan, B.; Vedal, S.; Hannigan, M.P. Temporal patterns in daily measurements of inorganic and organic speciated $PM_{2.5}$ in Denver. *Atmos. Environ.* **2010**, *44*, 987–998. [CrossRef]