

SEASONALITY DEDUCTION PLATFORM: FOR PM₁₀, PM_{2.5}, NO, NO₂ AND O₃ IN RELATIONSHIP WITH WIND SPEED AND HUMIDITY

Shubhi Harbola^{a, b}*, Volker Coors^b

^a University of Stuttgart, Stuttgart, Germany - Shubhi.Harbola@ieee.org

^b University of Applied Science, Stuttgart, Germany - (Shubhi.Harbola, Volker.Coors)@hft-stuttgart.de

Commission IV

KEY WORDS: Interactive platform, agglomerative clustering, cities planning, seasonality, unsupervised classification, dendrogram, pollution parameters, visual analytic, meteorological data, Spearman rank

ABSTRACT:

Human and ecosystem health is affected by the risk of air pollution. A comprehensive understanding of the parameters generating pollution and governing their nature in time is essential to devise functional policies focusing on minimising the concentration of the pollutants. The effect of pollution parameters on meteorological data and existing in between relationships, have been the focus of the researcher's planning of better city future. Thorough study of resources utilisation is required for contributing to framing effective, sustainable development, government policies management, and advance public services convenience. For protecting the environmental quality, renewable resources like solar and wind are more incorporated in techniques supporting better city planning. This paper considers the hourly time series Particular Matter (PM) PM_{2.5} and PM₁₀, Nitrogen Oxide (NO), and Nitrogen Dioxide (NO₂), and Ozone (O₃) along with measured wind flow and humidity. This study's objective is to assess the temporal seasonality patterns of these parameters in Stuttgart, Germany. The temporal variations over the city center in Stuttgart are analysed using unsupervised approach to perform seasonal hierarchical clustering on a series of parameters NO, NO₂, O₃, PM₁₀, and PM_{2.5}, wind speed and humidity. Furthermore, the correlations between meteorological and pollution parameters are analysed using the Spearman rank correlation method. Moreover, a dashboard is developed to provide the user desired time frame visualisation of these parameters. Proposed work would provide empirical meaning and seasonality comparison among the above mentioned parameters combined with interactive dashboard support. The analyses of the presented results clearly demonstrates the relationship between air pollutants, wind, humidity together in combine temporal activities frame. Thus, it would help city planner and policies maker with advanced knowledge of seasonality for meteorological and pollution parameters conditions.

1. INTRODUCTION

The human activities not only contributed to the lifestyle advancement and developments, meanwhile also to pollution, and change in the climate as byproducts. Very small aerial pollutants are discharged from the chimneys, industrial waste, vehicle smokes, and construction sites, that can be inhaled with the air leading to heart diseases, lung and respiratory problems all over the world. The traffic-related pollutants like Particular Matter (PM) PM_{2.5} and PM₁₀, Nitrogen Oxide (NO) and Nitrogen Dioxide (NO₂), and Ozone (O₃) remain at a high level. The air quality is affected by, NO, NO₂, O₃, PM₁₀, PM_{2.5} and their atmospheric concentrations. The lung tissue damage, cardiovascular and chronic respiratory diseases, could be hassled by coming in exposure to PM₁₀ and PM_{2.5} *i.e.*, particles with aerodynamic diameters less than 10 and 2.5 μm , respectively (Chen and Zhao, 2011). Over the urban areas, the elevated levels of pollution parameters are incorporated with both local emission sources and regional transportation (Chen and Zhao, 2011, Jasen et al., 2013). Regional transportation with diesel vehicles are the main sources of particular matters and contribute a significant portion to their levels (Wallace and Hobbs, 1977, Hardin and Kahn, 1999). Many studies have been performed to discover the seasonality of the pollution parameters along with the meteorological datasets, *e.g.*, wind speed, wind direction, temperature, humidity, precipitation, pressure. Some existing literature concluded that when the wind speeds are lower than 3.5 m/s, and the temperature higher than 21.1 °C than often high concentrations of PM₁₀, and PM_{2.5} were detected with reference to a study of PM in Ohio USA (Fraser et

al., 2003, Arthur and Owen, 2003). Moreover, some emphasised on deriving that pollution parameters are correlated to humidity and wind flow during winter (Elminir, 2005). Hien *et al.* show that wind speed, and temperature highly control the concentration of particulate matter (Hien et al., 2002). Few studies link pollutant characteristics to the meteorological parameters as, with wind effects and humidity again (Garrett and Casimiro, 2011). Several above discussed studies used smoothing and filtering techniques, ignoring the data noise and modifying the originality of temporal dataset. The comprehensive study of meteorological parameters and their contribution to PM_{10-2.5}, NO, NO₂, O₃ are poorly understood. Above research suggests that there is still a number of questions that remain to be addressed such as temporal wind nature and pollution parameters correlations, how humidity governs the PM_{10-2.5}, and NO, NO₂, O₃ relationships for user desired time frame, without modifying the authenticity of the original temporal dataset, remain to be addressed. A better insight into the system by improving human interaction with the meteorological data (Harbola and Coors, 2018) in relationship with pollution parameters. Thus this motivates for this proposed research. The problem of air pollution has caused considerable public concern in Stuttgart (Germany). Therefore, investigations into the spatio-temporal variation of concentrations of PM and gaseous pollutants across Stuttgart are necessary and essential. To keep track of the mass concentrations of PM_{10-2.5}, NO, NO₂, O₃ have been monitored in all important cities of Germany. Data from provincial and more effective center weather monitoring in Stuttgart were selected. Temporal variations of meteorological and pollution parameters were assessed and their trends of vari-

*Corresponding author.

ation between each other with respect to time for Stuttgart were investigated. Thus, unsupervised hierarchical clustering and correlation method which work on the original temporal datasets by taking into consideration the above listed gaps, are still required. Therefore the current study proposes hierarchical clustering and Spearman rank correlation method with the following contributions: (i) in depth temporal analysis of pollution and meteorological parameters using hierarchical clustering method, without applying any smoothing and noise removal technique on the collected temporal dataset, (ii) the time-frame of analysis is user-defined, (iii) dendrogram and heatmap temporal dataset visualisation to highlight the behavior of these parameters and to enhance accuracy, and (iv) comparative study of the pollution parameters and their effects with interactive dashboard view. The proposed work would provide foreknowledge of meteorological parameters nature in relationship to pollution parameters of an area, thereby helping and supporting in optimal selection of green sites with highlighting and tuning the air pollution quality. This would encourage more utilisation of renewable energy for safe and better city planning, which in turn would help for efficient management and development of the city's green resources. The increasing air pollution in big industrial cities would be alarmed and reduced for the future with this analysis. The remaining paper is organised as follows, proposed methods and datasets employed are discussed in section 2 and section 3, respectively, section 4 demonstrates the results and discussion, followed by conclusion in section 5.

$$D_{i,j} = D(A_i, A_j) = \|A_i - A_j\|^2 \quad (1)$$

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (2)$$

2. METHODOLOGY

The proposed method analysed seasonality in seven parameters using hierarchical clustering and Spearman rank correlation. Initially, values of each parameter are preprocessed before applying the clustering. The preprocessing involves normalising of the data followed by temporal filtering. The mean and standard deviation of a parameter are calculated. The values of a parameter are then subtracted by mean, followed by division with standard deviation, to get the normalised value. Further, the temporal filtering is applied on these normalised values. In the current study, the temporal filtering based on four quarters in a year is applied. First-quarter Q_1 is spring (March to May), second-quarter Q_2 is summer (June to August), third-quarter Q_3 is autumn (September to November), and fourth-quarter Q_4 is winter (December to February). These four time quarters division help in depth seasonality analysis of the considered seven parameters. Unsupervised agglomerative hierarchical clustering is applied on the temporal dataset (values) of a quarter (*i.e.* output of temporal filtering). The proximity matrix in hierarchical clustering helps in identifying the similarity of the clusters and combines most similar clusters hierarchically until the desired number of clusters are obtained. Ward's method in hierarchical clustering minimises the variance within the cluster by using the objective function of the error sum of squares (Ward, 1963). The pair of clusters that leads to a minimum increase in total within cluster variance after merging is searched. This increase is a weighted squared distance (D) between cluster centers (A_i, A_j) as shown in equation. 1 (Cormack, 1971). In order to provide more detailed comparison and seasonality trends analysis, each quarter is considered for all the parameters. This has been divided into two sets of 15 days starting and 15 days back in each quarter of a year. The sum of the squares starting from the clusters found by Ward's method is kept minimised. This gives a hint through the merging cost. The number of clusters is kept reducing until the merging

cost increases and then used the cluster number, right before the merging cost increased simultaneously (Paul and Murphy, 2009). Moreover, a dendrogram is used to obtain the final number of clusters as k . The dendrogram is a technique of agglomerative hierarchical clustering that gives a tree-like diagram that records the sequences of merges or splits. In addition Spearman rank correlation analysis between the meteorological and pollution parameters helps to derive the relationship among these parameters. Spearman rank correlation is defined in equation. 2, where d^2 represents square the difference, ρ is the correlation coefficient, n is the number of measurements, and k is the number of clusters.

Moreover, an interactive dashboard is developed to provide an in depth analytic and seasonality patterns clarity in between the meteorological and pollution parameters for user desired inputs in the four time quarters. This dashboard is called as seasonality analysis kit. The user could select the parameters over the desired time frame and compare the patterns interactively. The interactive dashboard is still in the first phase and would be more refined in future work. The developed work provides a comprehensive understanding of the relationship among the pollution parameters like NO, NO₂, O₃, PM₁₀, PM_{2.5}, and the meteorological parameters wind flow and humidity.

3. DATASET

Stuttgart pollution parameters and meteorological temporal datasets are used in this study. In the corner of Hauptstaetter Strasse 70173 Stuttgart, the historical data from 2015 to 2019 are taken from central Stuttgart station sensor¹. This dataset contains the wind speed and direction and humidity along with NO, NO₂, O₃, PM₁₀, PM_{2.5}, with temporal information attached in a 30-minute time interval. Amongst multiple values of a parameter in a single day, the mean value is considered in this study. The areas dataset is organised separately into an individual month by using time information, with past data first, followed by current data then subdivision into four considered quarters Q_1 , Q_2 , Q_3 , and Q_4 . This helps to perform pollution parameters and meteorological temporal datasets seasonality test and in depth analysis.

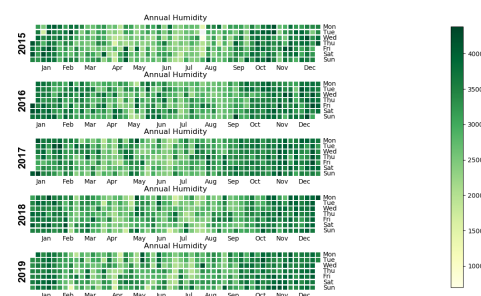


Figure 1: Annual humidity data value per day over the years (2015 to 2019).

4. RESULTS

The proposed seasonality analysis was implemented using Python and executed with four cores on Intel® Core™ i7- 4770 CPU @3.40 GHz. Stuttgart's 2015 to 2019 years of historical data with a temporal resolution of 30 min was separated by month to create monthly data over the years for both meteorological and pollution parameters. Figure. 1 and Figure. 2, show the data values recorded in a day over the 2015 to 2019 years in a heat maps representation for humidity and NO₂ respectively. In these generated heat maps, the intensity of the color was governed by the

¹<https://www.stadtklima-stuttgart.de>

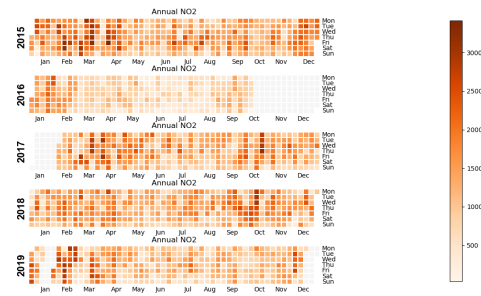


Figure 2: Annual NO₂ data value per day over the years (2015 to 2019).

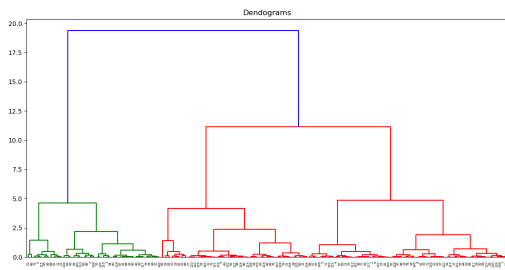


Figure 3: Dendrograms for selecting clusters in the temporal data set (here humidity as a considered parameter).

magnitude of parameter values. A similar heat map display existed for other parameters as well. Selected parameter (anyone *i.e.*, wind speed and humidity along with NO, NO₂, O₃, PM₁₀, PM_{2.5}), having higher values (range) over the time, had been assigned a darker color in the respective heat map. An unsupervised approach was used to perform comprehensive seasonal hierarchical clustering on a series of meteorological and pollution parameters. The comprehensive analysis for seasonality was studied based on four quarters (Q₁, Q₂, Q₃, Q₄) over the years. In performing the hierarchical clustering, k was taken as 6. This value of k was found empirically by performing some sensitivity tests, like, (i) if the value of k was higher (*i.e.*, number of clusters was equal to the total values in a quarter) than the clustering outcome was similar to Figure. 1 and Figure. 2, and this was not able to represent the seasonality pattern, (ii) if the value of k was lower (*i.e.*, $k = 1, 2, 3, 4$), then also there was information loss, and (iii) the dendrograms were generated as an output from unsupervised hierarchical clustering with the primary use to allocate objects to clusters in the best possible way. Figure. 3 shows the obtained dendrogram for selecting clusters (possible numbers) in the temporal data set, where in this Figure. 3, *e.g.*, the humidity was considered. Similar parameter analyses were conducted for rest of the parameters. Therefore k was taken as 6 in the present study. The unsupervised hierarchical clustering here aimed at inferring the inner structure and trends presented within the meteorological and pollution data, trying to cluster them into six classes depending on similarities among them.

In order to provide a more detailed comparison and seasonality trends analyses, quarter time frames were considered for all the parameters. Further, a quarter was divided into two parts comprising of the first fifteen days and the last fifteen days in a month. This helped in discovering all the possible changes in the quarter for each of the considered parameter. The obtained outputs of the in depth unsupervised clustering analysis performed for NO₂, where Figure. 4, and Figure. 5 give the clustering outputs for NO₂ for first and last 15 days in Q₁, and Figure. 6, and Figure. 7 in

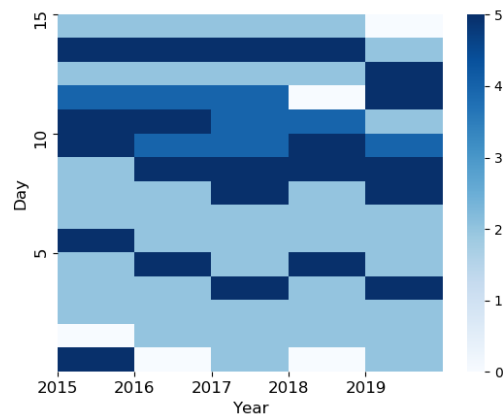


Figure 4: Clustering output for NO₂ for first 15 days in Q₁ over 2015 to 2019.

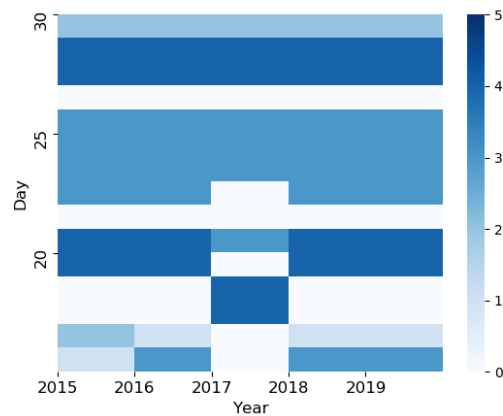


Figure 5: Clustering output for NO₂ for last 15 days in Q₁ over 2015 to 2019.

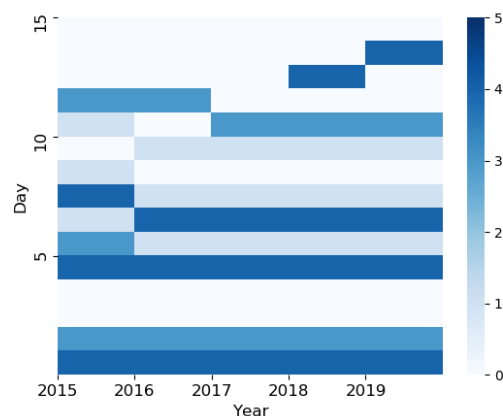


Figure 6: Clustering output for NO₂ for first 15 days in Q₂ over 2015 to 2019.

Q₂ respectively. Similarly, Figure. 8, and Figure. 9 depict the clustering outputs for NO₂ for first and last 15 days in Q₃, and Figure. 10, and Figure. 11 in Q₄ respectively. Like developed hierarchical clustering, similar outputs were generated for other

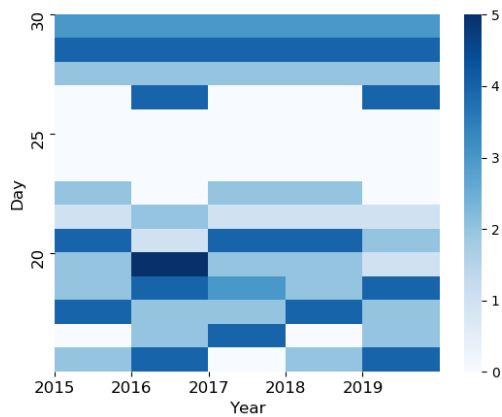


Figure 7: Clustering output for NO₂ for last 15 days in Q₂ over 2015 to 2019.

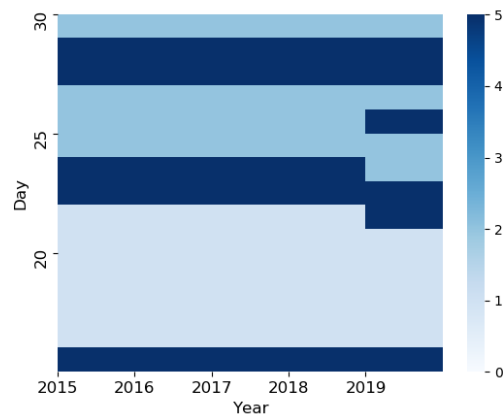


Figure 9: Clustering output for NO₂ for last 15 days in Q₃ over 2015 to 2019.

parameters in each respective quarters with the first and last fifteen days comparisons.

Further, the correlation analysis between the meteorological and pollution parameters were done to enhance the probability of deriving the relationships among these parameters. Figure. 12 helps to study the complex relationships among parameters very well. In addition, the user could select the parameters over the desired time frame and compare the patterns interactively with the help of the developed dashboard. The screenshots of the proposed dashboard are shown in Figure. 13, where wind speed (*e.g.*, case) was selected as a parameter with respect to Q₁, Q₂, Q₃, Q₄ over the years to visualise seasonality. Similarly more parameters could be selected from the seasonality analysis kit.

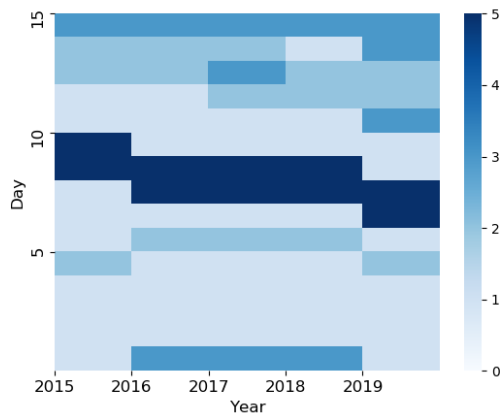


Figure 8: Clustering output for NO₂ for first 15 days in Q₃ over 2015 to 2019.

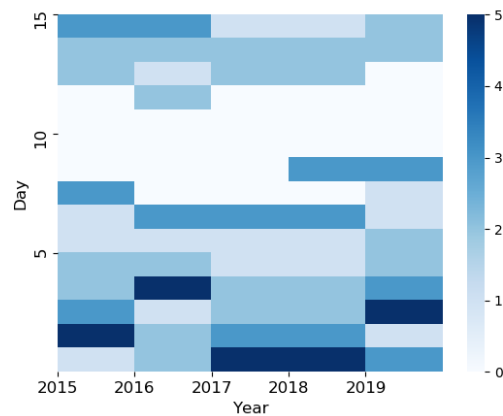


Figure 10: Clustering output for NO₂ for first 15 days in Q₄ over 2015 to 2019.

4.1 Discussion

The hierarchical cluster analyses for meteorological and pollution parameters was to highlight the trends at which any given pair of quarters (over the years) joined together in clustering diagram with each class assigned specific color code. A sequential scale of color brewer blues scale color map used for showing classes (0 to 5) with the color frequency differentiates low values class from high values class. The blended progression using, typically of a single hue, from the least to the most opaque shades, represents

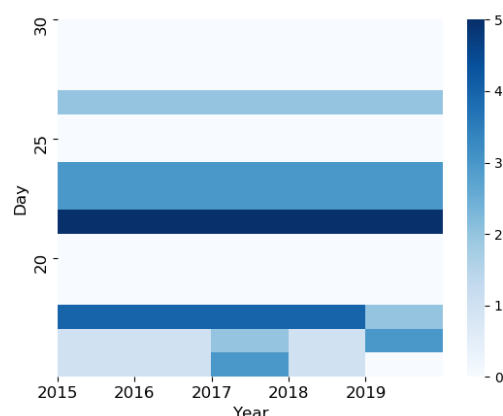


Figure 11: Clustering output for NO₂ for last 15 days in Q₄ over 2015 to 2019.

low to high values. Each year dataset for the considered parameter over the four quarters that joined together sooner (in clustering) are more similar to each other than those that are joined together later. The total within-cluster variance is minimised dur-

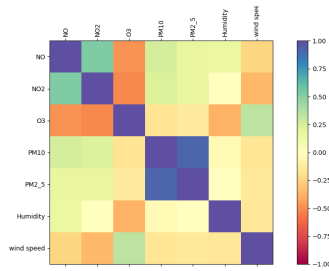


Figure 12: Correlation output between meteorological and pollution parameters.

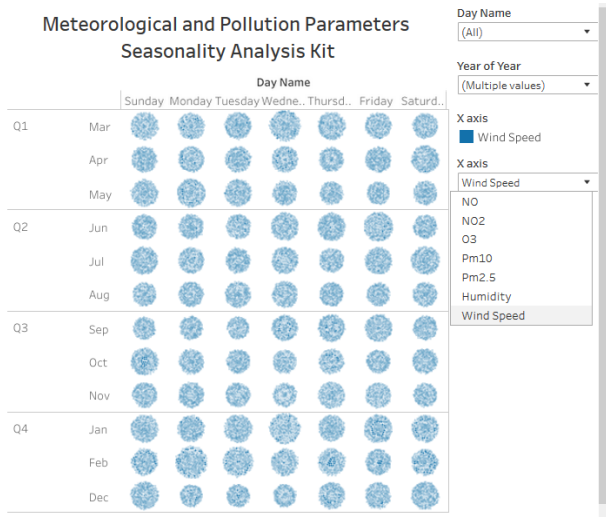


Figure 13: Interactive dashboard for meteorological and pollution parameters.

ing clustering. At each step, the paired clusters with minimum between-cluster distance are merged. As a result it is observed that NO and NO₂ concentrations are high in Q₃ autumn, and Q₄ winter over 2015 to 2019 respectively (Figure. 8, and Figure. 9, Figure. 10, and Figure. 11). Both are strongly correlated to each other with similar trends over the years, also same can be seen in the correlation graph in Figure. 12. Comparing Figure. 4 and Figure. 5 provides that in Q₁ there exists volatility in the first and last fifteen days. From 3rd to 6th day during 2015 to 2019 there exist a pattern with high NO₂ concentrations. The same pattern repeats again from 8th to 9th in 2016 to 2019 and on 14th in 2015 to 2018. However, in last fifteen days from 21st to 22nd, and 26th to 27th for 2015 to 2019, low magnitudes of NO₂ are measured for Stuttgart. As shown in Figure. 6, and Figure. 7 from 12th to 15th in Q₂ NO₂ concentrations are lowest during 2015 to 2017 and, reached highest in 2018 to 2019. For last fifteen days from 16th to 19th the concentrations reached highest during 2016 to 2019. However, from 25th to 27th NO₂ measurement was negligible in 2015 to 2018, with exceptional high concentrations during 2016 and 2019. The first and last days clustering output (in Figure. 8, and Figure. 9) for Q₃ from 8th to 10th recorded high values again in 2017 to 2019. On the other hand from 17th to 23th NO₂ concentrations are low for 2015 to 2018 but, measured highest in 2019. Figure. 10, and Figure. 11 conclude that in Q₄ from 8th to 10th in 2015 to 2017 the concentrations are lowest and high in 2018 to 2019. Moreover, from 21th to 22th in 2015 to 2019 the NO₂ concentrations approached highest again.

However, O₃ concentrations are more in Q₁ spring, Q₂ summer, and less in Q₃ autumn with exceptional increase in Q₄ winter dur-

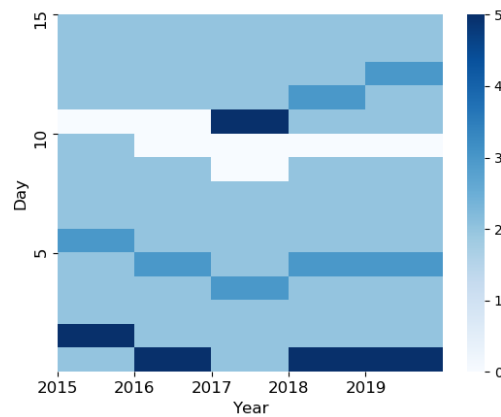


Figure 14: Clustering output for O₃ for first 15 days in Q₁ over 2015 to 2019.

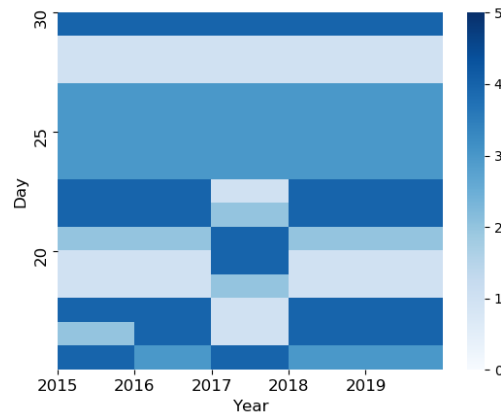


Figure 15: Clustering output for O₃ for last 15 days in Q₁ over 2015 to 2019.

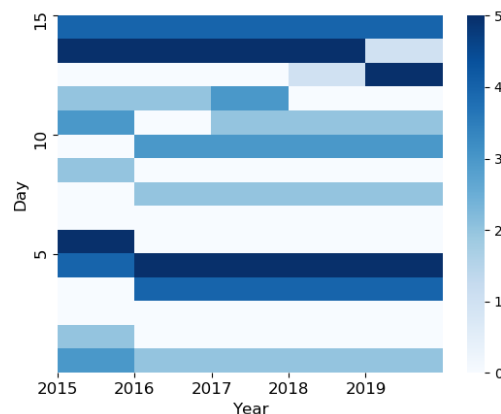


Figure 16: Clustering output for O₃ for first 15 days in Q₂ over 2015 to 2019.

ing 2015 to 2019. These (above) statements also validates that O₃ and NO₂ are negatively correlated to each other which also supports the obtained correlation in Figure. 12. Further, O₃ concentration analysis for Q₁ has been shown in Figure. 14, and Fig-

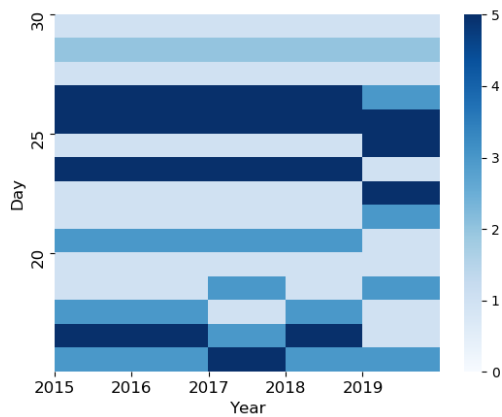


Figure 17: Clustering output for O₃ for last 15 days in Q₂ over 2015 to 2019.

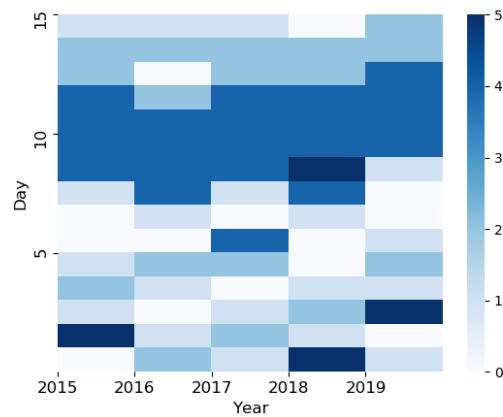


Figure 20: Clustering output for O₃ for first 15 days in Q₄ over 2015 to 2019.

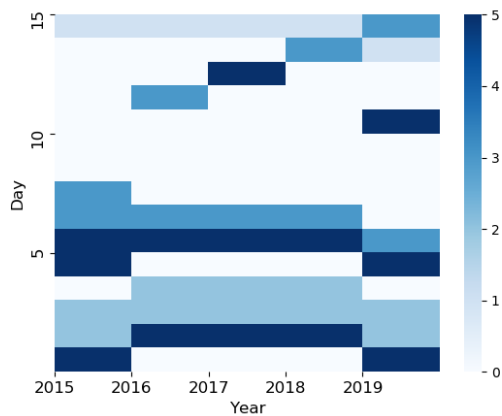


Figure 18: Clustering output for O₃ for first 15 days in Q₃ over 2015 to 2019.

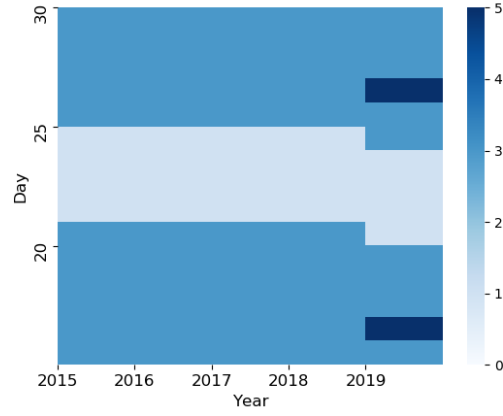


Figure 21: Clustering output for O₃ for last 15 days in Q₄ over 2015 to 2019.

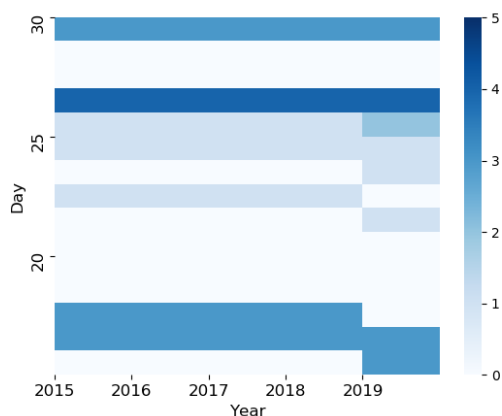


Figure 19: Clustering output for O₃ for last 15 days in Q₃ over 2015 to 2019.

ure. 15, where from from 1st to 3rd day concentrations are highest in 2016, 2018 and 2019. O₃ concentrations approach lowest from 27th to 29th in 2015 to 2019. As shown in Figure. 16, and Figure. 17, in Q₂ from 13th to 14th and 25th to 27th day O₃ con-

centrations are highest in 2015 to 2019. During Q₃ with reference to Figure. 19, and Figure. 20, from 8th to 10th O₃ concentrations are lowest during 2015 to 2018 with measured highest in 2019. From 26th to 27th the concentrations are increasing during 2015 to 2019. O₃ variation in Q₄ is shown in Figure. 20, and Figure. 21, where O₃ is high from 8th to 13th, 16th to 21st and 25th to 30th and reaches highest in 2019. Moreover, humidity magnitudes are lowest in Aug and then starts increasing from Sep *i.e.* in Q₃ autumn, to Q₄ winter, over the years 2015 to 2019 as shown in Figure. 26, and Figure. 27. This shows that humidity is negatively correlated to O₃ however, positively correlated to NO₂ which also get justified by the correlation graph in Figure. 12. Humidity clustering output delivered that in Q₁ (in Figure. 22, and Figure. 23) from 10th to 12th, and 26th to 27th the humidity measurements are highest in 2015 to 2019. As shown in Figure. 24, and Figure. 25 for Q₂ from 3rd to 4th, 9th to 11th, and 28th to 29th highest humidity measured over 2015 to 2019. Moreover from 10th to 15th humidity measured lowest in 2015 with sudden increasing spikes in 2016 to 2019 for Q₃ (in Figure. 26, and Figure. 27). Similarly, from 25th to 28th the humidity increased to highest during 2015 to 2019. With reference to Figures², in Q₄ from 10th to 14th, and 19th to 22st, respectively, measured humidity increased in 2015 to

²Due to the manuscript page limit, all the generated figures are uploaded in the GitHub https://www.github.com/shharbola/SDSC20_Images/

2019.

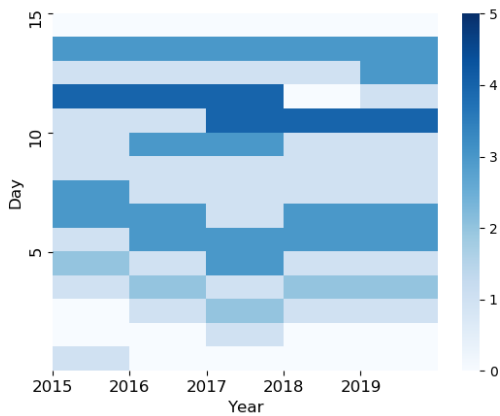


Figure 22: Clustering output for humidity for first 15 days in Q₁ over 2015 to 2019.

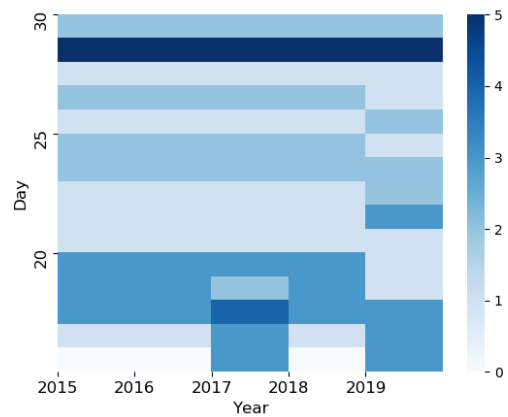


Figure 25: Clustering output for humidity for last 15 days in Q₂ over 2015 to 2019.

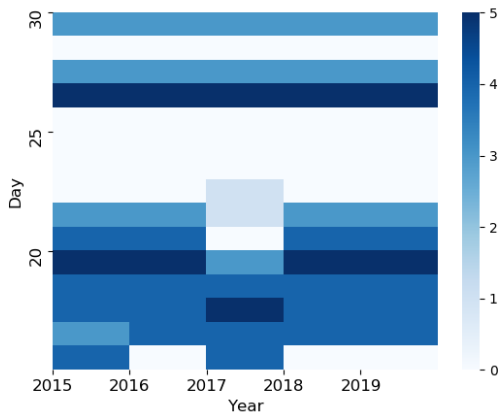


Figure 23: Clustering output for humidity for last 15 days in Q₁ over 2015 to 2019.

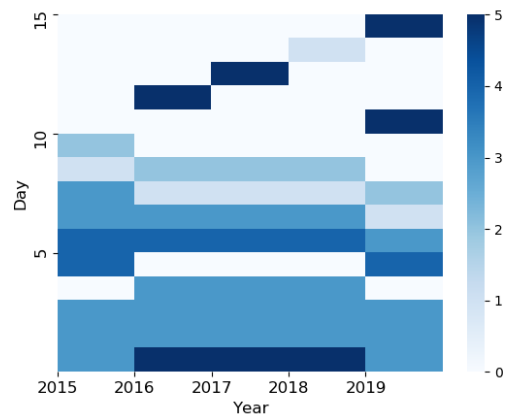


Figure 26: Clustering output for humidity for first 15 days in Q₃ over 2015 to 2019.

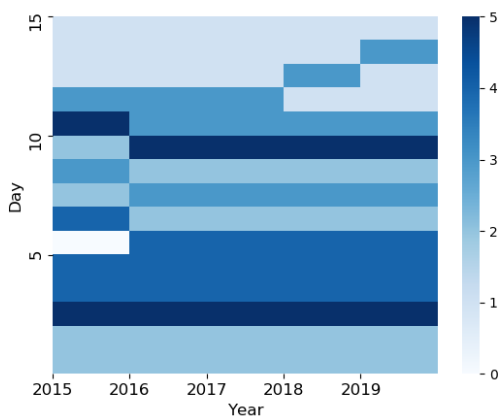


Figure 24: Clustering output for humidity for first 15 days in Q₂ over 2015 to 2019.

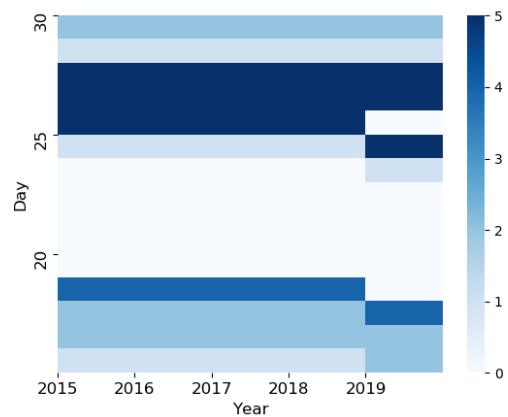


Figure 27: Clustering output for humidity for last 15 days in Q₃ over 2015 to 2019.

Moreover, wind datasets are highly volatile in nature over the years 2015 to 2019. High magnitude wind speeds are recorded

more during Q₂ summer, Q₃ autumn, to Q₄ winter, with sudden high spikes are observed during the seasonal cycle changes mostly in the months of Jan, Mar, Jul, Sep and Dec, as analysed with the help of figures². These analyses devised that wind

speed is positively correlated to NO₂ however, negatively correlated to O₃. In Q₁ from 3rd to 7th, and 26th to 29th in 2015 to 2019 there exists pattern of low speed winds. Q₂ first fifteen days from 6th to 12th, and 27th to 30th wind speed keeps increasing and reached highest with volatility in years 2015 to 2019. Moreover, in Q₃ wind speed frequently changes from mild to increasing (also reached highest) magnitudes from 19th to 20th, and 26th to 29th in 2015 to 2019. Q₄ from 8th to 14th, and 21th to 25th during 2015 to 2019 mild speed winds are measured. Furthermore, PM concentrations are more in Q₄ winter, and also in Mar, May, and Jul concluded from generated clustering outputs figures². Analysing for Q₁ from 1st to 15th, and 23th to 27th in years 2015 to 2017 represents existence of highest PM₁₀ concentrations with constantly increasing level however, with strong ban policies for diesel and old vehicles use by the German government and other regulatory movement restrictions and climate awareness, the PM concentrations are little controlled and reduced (comparison to earlier years) in 2018 and 2019. In Q₂ from 13th to 15th PM₁₀ concentration increased in 2015 to 2016 and reached highest during years 2018 to 2019. From 27th to 30th high concentration measured in 2015 to 2016 and then reduced to lowest in 2019 again during Q₂. Furthermore, in Q₃ first and last fifteen days of clustering output there exists frequently changing PM₁₀ concentration from lowest to increasing in 27th to 30th with the concentration reached highest during 2015 to 2019. In Q₄ from 2nd to 4th, and 13th to 15th in 2015 to 2019 PM₁₀ concentration measured highest. These interpretations (above analyses conclusions) provide a quick facts-crosscheck supporting the present alarming air quality situation in the Stuttgart city and requirement of probable more control measures. In addition, performed correlation analyses on pollution and meteorological datasets helped to uncover the important interrelationships, and also justified clustering analyses outcomes. Figure. 12 contributes following important points: i) NO and NO₂ are 77% positively correlated to each other, with 27% positively correlated to PM_{10-2.5}, and negatively correlated to wind speed by 53%, ii) O₃ is 50% positively correlated to wind speed, 77% negatively correlated to NO and NO₂, and 27% negatively correlated to PM_{10-2.5}, iii) Humidity is 27% positively correlated to NO and NO₂ and 50% negatively correlated to O₃, iv) wind speed is 27% negatively correlated to PM_{10-2.5}, and v) PM₁₀, and PM_{2.5} are positively correlated to each other with more than 87%. Moreover, the developed seasonality analysis kit is used to provide interactive selections of considered meteorological and pollution parameters to analyse the concurred pattern in the dataset, in a time based frame over the years. Currently, the designed dashboard is in its first phase with color based clustering display for each quarter over the years. This has helped in making the seasonality analyses tests easy, user interactive and comparable in the time domain.

5. CONCLUSION

Meteorological data have been the attention of the researchers of the smart city planning for thorough utilisation and management of resources, which help in effective government management, convenient public services, and sustainable industrial development. Using renewable energy supply would provide a healthy and amiable city, and increased welfare in more general terms. To ensure incorporation into the planning process, the renovation of the existing planning is indeed the most promising field for climate-related intervention. From a designer's perspective, the authors have stressed the need to include energy-conscious strategies to improve environmental quality. The integration of new knowledge, innovative technologies in sustainable transformation is the motive of this paper. The interpretations (above analyses conclusions) provide a quick facts-crosscheck supporting the present alarming air quality situation in the city and requirement

of probable more control measures. The interactive dashboard seasonality analysis kit of meteorological and pollution parameters would help to plan the future with more green policies. Designed dashboard in this work could be further improved with the ensemble of more parameters with more historical data set. The future focus for the authors would be to improve the analysis and utilising the outputs on interactive visual analysis dashboard web applications. Moreover, use of other methods like decision trees, neural networks and association rules, would be explored in subsequent research works for in-depth understanding of temporal relationships amongst the parameters. Meanwhile, the devised seasonality analysis of meteorological and pollution parameters over the years has the potential for selecting better government supported green policies and creating environmental awareness among humankind, and moreover, provide a foreknowledge for better city planning.

ACKNOWLEDGEMENTS

The Stuttgart wind dataset is downloaded from the History of the Section of Urban Climatology within the Office for Environmental Protection of the city of Stuttgart website.

REFERENCES

- Arthur, T., D. and Owen, M., D., 2003. Atmospheric environment. Temporal, spatial and meteorological variations in hourly Pm2.5 concentration extremes in New York City 38, pp. 1547–1558.
- Chen, C. and Zhao, B., 2011. Atmospheric environment. Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor pp. 1–9.
- Cormack, R. M., 1971. A review of classification. Journal of the Royal Statistical Society 134(3), pp. 321–367.
- Elminir, H., K., 2005. Total environment. Dependence of urban air pollutants on meteorology 350, pp. 225–237.
- Fraser, M., P., Yue, Z., W. and Buzcu, B., 2003. Atmospheric environment. source appointment of fine particulate matter in Houston TX, using organic molecular markers 37, pp. 2117–2123.
- Garrett, P. and Casimiro, E., 2011. Environment science pollution. Short-term effects of fine particulate matter (PM2.5) and ozone on daily mortality in Lisbon, Portugal 18, pp. 1585–92.
- Harbola, S. and Coors, V., 2018. Geo-visualisation and visual analytics for smart cities: A survey. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.
- Hardin, M. and Kahn, R., 1999. Academic press: New york. Aerosols and climate change, 1999.
- Hien, P., D., Bac, V., T., Tham, H., C., Nhan, D., D. and Vinh, L., D., 2002. Atmospheric environment. Influence of meteorological conditions on Pm2.5 and Pm2.5-10 concentrations during the monsoon season in Hanoi, Vietnam 36, pp. 3473–3484.
- Jasen, N., A. H., Fischer, P., Marra, Ameling, C. and Cassee, F., R., 2013. Total environment. Short-term effects of PM2.5, PM10 and PM2.5-10 on daily mortality in The Netherlands 463, pp. 20–36.
- Paul and Murphy, A., 2009. Data mining. Distances between Clustering, Hierarchical Clustering pp. 36–350.
- Wallace, J., M. and Hobbs, P., V., 1977. Academic press: New york. Atmospheric science, an introductory survey.
- Ward, J. H., 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58, pp. 236–244.