The use of a new diagram for the analysis of the daily cycles in the air-pollution data

Marija Zlata Božnar*, Boštjan Grašič and Primož Mlakar

MEIS d.o.o.
Mali Vrh pri Šmarju 78
SI-1293 Šmarje-Sap, Slovenia
E-mail: marija.zlata.boznar@meis.si
E-mail: bostjan.grasic@meis.si
E-mail: primoz.mlakar@meis.si
*Corresponding author

Dejan Gradišar

Jozef Stefan Institute,
Jamova 39
Ljubljana, Slovenia
E-mail: dejan.gradisar@ijs.si

Juš Kocijan

Jozef Stefan Institute,
Jamova 39
Ljubljana, Slovenia
and
University of Nova Gorica,
Nova Gorica, Slovenia
E-mail: jus.kocijan@ijs.si

Abstract: From a temporal viewpoint, air pollution has significant daily patterns/cycles of behaviour. These cycles are conditioned by anthropogenic and natural phenomena. In both cases, a detailed observation and an understanding of the daily cycles rules or daily patterns of air pollution can be significant and at the same time can contribute to more effective measures to reduce the harmful impact of air pollution on human health. In this paper the new sunflower diagram is presented. The key advantage of the sunflower diagram is the ease of understanding the result and the ability to present information in the form of a graphic pattern, allowing the user to quickly understand the content. Using the sunflower diagram, we will present an analysis of the meteorological parameters that are important for understanding air pollution and air-pollution data for different locations in Slovenia.

Keywords: daily cycles, analysis tool, sunflower diagram, air-pollution flower, weather flower, wind flower

Reference to this paper should be made as follows: Božnar, M. Z., Grašič, B., *Author*

Mlakar, P., Gradišar, D. and Kocijan, J. (xxxx) 'The use of advanced analytical tools for the analysis of the daily cycles in the air-pollution data', *Int. J. xxxxxxxxxxx xxxxxxxxxxx*,

Biographical notes:

Marija Zlata Božnar received her PhD degree in 1997 from the University of Ljubljana, Slovenia. She started her research work in 1990 at the Jožef Stefan Institute in Ljubljana. Since 2007 she has been a director and half-owner of the company MEIS d.o.o. She is a member of the committee of EURASAP (the European association for the science of air pollution).

Boštjan Grašič received his PhD degree in 2008 from the University of Nova Gorica, Slovenia, the Faculty for environmental sciences. He started his research work in the research-oriented company AMES d.o.o., in 1996 and continued in MEIS d.o.o., as a researcher. His main research work is focused on air-pollution modelling.

Primož Mlakar received his PhD in 1996 from the University of Ljubljana, Slovenia. He started his research work at the Jozef Stefan Institute, Ljubljana in 1983 and continued at AMES d.o.o., where he was an Assistant Director. From 2007, he has worked in the research-oriented company (SME) MEIS as a leader of several projects and a research group, where he is also the co-owner. He is the leader of the MEIS research unit.

Dejan Gradišar received his Ph.D. degree in Electrical Engineering from the University of Ljubljana in 2006. He is currently working as a researcher at the Department of Systems and Control, IJS. He is also an assistant at the Faculty of Electrical engineering, University of Ljubljana. His research interests are in the field of production optimization and dynamical systems modelling.

Juš Kocijan is currently a senior researcher at the Jozef Stefan Institute and Professor of Electrical Engineering at the University of Nova Gorica, Slovenia. His research interests include the modelling and control of dynamic systems. His other activities include: serving as editor and is on the editorial boards of research journals, serving as a member of the IFAC Technical Committee on Computational Intelligence in Control. He is a senior member of the IEEE, EEE Control Systems Society, and a member of the Slovenian Society for Simulation and Modelling (SLOSIM) and the Automatic Control Society of Slovenia.

This paper is a revised and expanded version of a paper entitled "Analysis of the daily cycles in the data on air pollution through the use of advanced analytical tools" presented at the 17th International conference on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes. Budapest, Hungary, May 9-12, 2016.

1 Introduction

Nature has a diurnal cycle. Firstly, this applies to basic meteorological processes that describe the atmosphere. These processes subsequently also affect the atmospheric pollution. When exploring the entire atmospheric pollution cycle, many emissions already have a distinctive diurnal cycle such as emissions from traffic and domestic heating sources (Mayer, 1999, Sandradewi et al, 2008). The emissions' diurnal cycle and meteorological conditions can cause diurnal cycles of air pollutants such as ozone (Garland and Derwent, 1979), when we speak about the concentrations in the lower atmosphere. In order to understand, and subsequently reduce the bad influence of pollutants on the health of people, it is necessary to recognize such diurnal cycles of concentrations in the atmosphere, and to present them in a manner which is easily understandable to the general public. An understandable presentation of atmospheric pollution is one of the preconditions for people to be able to organize their daily activities in a way to minimize their exposure to polluted air, if we cannot eliminate the pollution entirely.

Researchers of atmospheric pollution dedicate their attention to the aspect which is particularly interesting to researchers and regulatory bodies, namely the statistical analysis and the presentation of the concentration of pollutants in the atmosphere (Thunis et al., 2011). This aspect is posed by the regulatory requirements (e.g. EU directives), but we keep forgetting the aspect of presenting the statistics to the end users (the affected population). Thus, the research group at MEIS d.o.o., (Božnar et al., 2015) has defined a new way of a graphical method of analysis, which is suitable for a statistical analysis of the diurnal cycles. We called the basic graph, "sunflower". In this paper, we will firstly present this graph, and subsequently demonstrate a few cases how this graph may be useful in statistical analysing, and for the presentation of the data about the concentration of pollutants in selected locations.

The new data presentation method has already been published, including a detailed description of the graph (Božnar et al., 2015). In this paper we will focus on its applicative value in air-pollution modelling science.

2 The methodology of using the sunflower diagram on meteorological and air-pollution data

We firstly encountered the problem, i.e. seeking an appropriate presentation of the diurnal cycles, when we studied the meteorological variables – global solar radiation. We defined a new way of a graphical method potentially useful for the analysis, which we called the "sunflower". A detailed definition of the method is available in the paper (Božnar et al., 2015). The objective of this chart is an explicit graphical presentation of the characteristics of the diurnal cycle of the observed environmental parameters.

A graphical display of the characteristics allows a quick perception of the information (Cleveland and McGill, 1984). It is even more important that people are trained mainly in a way that they are able to easily find both the similarities and differences in the details between two more or less similar images. Such similarities and differences are of course noticed much faster in an image than in a numerical table chart.

The facts on human image perception (Cleveland and McGill, 1984) were utilized in the quest for a better solution to the presentation. The conventional way to present time series and model responses of atmospheric parameters for longer time range, are line graphs. The main problem in the perception of these graphs when used for the presentation of longer time spans is that the high density of the drawn lines makes it more difficult to see the daily characteristics. An even worse situation is to compare visually the differences of the lines among the days and to identify the daily similarities and differences. Furthermore, it is most difficult, if not impossible to visually compare the daily differences in a line graph for the selected atmospheric variable over a season or over various geographical locations. Such a presentation becomes completely unclear when we utilize such a graph over a multi-year time range. Nevertheless, all these disadvantages and the presentation problem can be solved with a new periodical-data presentation called the sunflower diagram.

Examples of two sunflower diagrams are presented in Figure 1. On the left, we can see the analysis of the measured half-hour data for January 2015, for the global solar radiation at the Pustice station in the continental part of Slovenia. On the right, we can see the same analysis for July 2015. The sunflower is actually a double circular histogram. Our idea was to make a histogram analysis and sort firstly in the time dimension and then follow that in the dimension of a meteorological variable. Sorting in the time dimension allows for a statistical analysis on data that occur in a particular hour of the day over a longer period, from several days to several years. A histogram analysis of the data within a particular class of the chosen hour is a very effective graphical tool for showing the basic statistical properties of the analysed data, contributing both (timeline sorting and histogram) techniques in a circular shape that mocks and extends the shape of the analogue clock is our novel idea. The course of the preparation of this chart will be presented with the case of global solar radiation measurements. Firstly, we have to sort all the measurement values into groups for each hour of the day, namely into groups from 1 to 24, whereby a measurement must be sorted in the interval (each interval is independent), when the measurements were actually taken, and the interval mark is the full hour mark at the end of the interval. Each given group is then separately presented in the form of a spike (segment) on the sunflower diagram in the direction from the centre towards the appropriate time. Measurements performed within one class for each individual hour are subsequently sorted into classes by values. The central class consists of all the values which are of no interest for the analysis of the daily cycle (darkness in this case, when there is no solar radiation). In the other classes, which are presented on a spike, we sort out the values on a scale from the lowest borderline to the highest expected value, however, the last class may be left open in the upward direction (as is the case in the figure for the values above 1,100 W/m^2), when we expect only a few very large values. Each individual value class on the figure is presented as a spike section of a clock in its own time class. Hereby, all the individual spike parts are aligned on a virtual radial line running from the centre point to the mark of the corresponding time on the outer circle of a circular graph. Each value class has its own codifying colour. In the event of potentially black/white image reproductions we also present the values with the width of the spike section. The length of the spike section is proportional to the percentage of the class in comparison to the whole. The whole is identified as a number of all (mark besides the graph: “all”) measurement intervals, which occur from the beginning of the examined statistical interval until the end. This interval for the statistical treatment is written beside the graph in the form of a date. Specifically, it also makes sense to write how many of the used measurement values were useful (“good”) and how many of them were unknown and incorrect (“unknown”).

In the final figure (the case of the left or right image in Figure 1) we may clearly see the representation of individual global solar radiation value classes according to the hours of the day. In the winter month of January, we have a relatively short period of light during the day (from 7a.m. to 5p.m.), thus the interesting value classes are only visible in these hours. We do not indicate clock spikes during the night time, but present their measurement values together for all the hours as a part of the whole, indicated in the centre of the graph. However, the bright part of the day in the July sunflower is visibly longer, according to the colour scale, we can also see the highest values. The representation of the value classes by individual hours determines the natural course of the global solar radiation deviations from the theoretically expected course. It indicates when on average a cloud or fog reduces the global solar radiation, which penetrates to the ground. On the way, as we have just described the presentation of the global solar radiation on the sunflower diagram, we can introduce other variables. Unlike the traditional wind rose (Zelenko and Lisac, 1994), with the help of this wind flower (we introduce the new name to emphasize the use of the sunflower diagram for wind speed data) we are able to give the wind speed (i.e. the categorization of the absolute speed in classes and by hours without considering the direction of the wind) or any other data on the concentration of atmospheric pollutants for the chosen location and the pollutant. It is also important that the statistical interval expressed with the sunflower is at least a day long, but it does not have any limitations regarding the duration (day, week, month, quarter as a season, year, several years etc.). The sunflower diagram and the air-pollution flowers (name for its application on air-pollution data) are very appropriate for the comparison of the daily statistics not only between various periods of time but also between various locations (e.g. for the same month between various locations across the country). Such cases will be shown below.

3 Results

In Figure 2 we present an analysis of the PM_{10} measured values for the town of Zagorje on a complex terrain in central Slovenia (coordinates: 46°07'51.9"N, 14°59'45.8"E). The analysis comprises daily cycles which were measured every 30 minutes for a period of one year. We created air-pollution flowers for each individual month. Thus, we are able to quickly determine the information on various daily pollution patterns for various months. Zagorje is located in a narrow half-closed valley, which partly expands to a basin, the winds are weak and thermal inversions are common during the winter. Particle pollution comes from the traffic, local heating sources, and the production of lime (Božnar et al., 2014, Mlakar et al., 2012, Mlakar et al., 2014). The figures show in a very clear and graphical way the following information about typical daily air-pollution regimes with PM_{10}: Firstly, the yellow and brown spikes present relatively high concentrations; green, blue and grey spikes consequentially show more and more clean air. Very clean air as a share in every interval is also presented in the centre of each air-pollution flower. Firstly, we see that during the winter months from December to February inclusively, the pollution is presented during the entire day, the highest values in December and January occur in the late afternoon and last until midnight, or an hour longer, which indicates that they are most likely due to heating (people start the heating when they return home from work). The opposite of this are the months of June, July and August, when the air is more or less clear for the major part of the day, a few higher concentrations may be noticed in the late afternoon, from 3p.m. to 8p.m., which indicates the influence of the traffic (the station is located very close to the main road through the town). In the spring and autumn, we can notice the transition from one described daily pollution pattern to another.

Celje (coordinates: 46°14'04.3"N 15°15'45.1"E) is another town in central Slovenia that has problems with pollution exceeding for PM_{10}. It lies at the southern edge of a large basin surrounded by steep high hills on the south. Cold winters are characterized by low winds and often thermal inversions. The automatic air-pollution measuring station is placed in the town centre. It measures the impact of urban traffic. Figure 3 presents the measurements for 8 individual consecutive days for PM_{10} in the town of Celje. If we use the air-pollution flower diagram for individual days, we are able to monitor very precisely when in a day and in a week, high and low concentrations occurred, and are thus more able to find the reasons for them (or confirm the hypotheses about their origin when we do not have a spatial model in a sufficiently precise spatial and time resolution). The lay public may use it for planning when to let fresh air into their houses, and outdoor activities.

Figure 4 presents the air-pollution flower diagram for ozone and NO_2 for two selected months; both graphs indicate these two pollutants in Celje, where the station is located in the centre of the town, where the impact of urban traffic is felt. The scales on the display for ozone and NO_2 are chosen in a way that we only notice increased concentrations. A daily pollution with NO_2 is clearly visible both for April and September, and the elevated values follow the expected increase in the urban traffic during rush hours both in the morning and late afternoon. However, the picture is complementary with the ozone, as is apparent from the well-known cycle of formation and decomposition of ozone due to solar radiation and ozone precursors, among which, NO and NO_2 are also very important. The two graphs for ozone and NO_2 for the same month show the complementarity between the ozone and NO_2. Exceedingly high concentrations of both are not present together at the same time. The air-pollution flower diagram for NO_2 and ozone vary from the PM_{10} air-pollution flower diagrams. This indicates that the elevated concentrations hardly occur during some hours of the day, which is clearly visible by the fact that the spikes at these hours are significantly shorter than those when the pollution is significant. If the concentrations occur in the clean air class, their presence on the graph is indicated in the numerically written share in the middle of the circle, however, they are not present as parts of the spikes.

Figures 2, 3 and 4 show different possible usages of the sunflower diagram that we can also call an air-pollution flower when applied to air-pollution data.

Figure 2 is an example of an analysis of huge amounts of measured data. In particular case charts are shown for every month, but the same can be done by comprising together several years of data and then analysing the same months on one chart. Using such an approach, air-pollution patterns representative for several years can be shown.

Figure 3 is the opposite of Figure 2. One chart represents only one day of half-hourly or hourly data of the observed variable. Such a representation allows an analysis of fine details and differences between the days. Its main advantage over a line graph is that the values of the variables are more easily paired with the corresponding hours of their occurrence, especially when used for more than one week of data.

Figure 4 presents an example of the comparative analyses of two variables (more variables are also possible). This enables a quick identification of the differences among the presented variables.

To conclude the presentation of the broad possibilities of using the sunflower graph, we present a weather-flower diagram in Figure 5. The weather flower is a slightly modified version of the sunflowers. This graph presents three main parameters of the weather forecast, namely one individual day on each separate graph. These concentric rings indicate the forecast for cloudiness, which is the nearest to the centre of the graph, the following ring indicates solar radiation, and the last ring indicates the predicted precipitation based on the same principle as the sunflower diagram. Figure 5 shows an example of the use with a pictorial legend. With such a presentation, we are able to show a much more detailed weather forecast without using numerous pictograms or line charts, which are not clear to all the general public. The town of Krško (coordinates: 45°57'58.9"N 15°29'05.1"E) lies in a half-opened basin surrounded by several chains of hills. The inner hills are low and not steep, the more distant outer hills are high and steep.

4 Conclusions

The paper presents our new sunflower diagram which is suitable both for basic meteorological parameters, as well as the analysis of daily concentrations of air pollutants. The motivation of our work was to find a solution for a graphical presentation and analysis of the diurnal characteristics of large amounts of data for atmospheric variables. Most of the atmospheric variables have a distinctive diurnal cycle, but its statistical presentation was a challenge. The sunflower diagram is our novel solution for this problem. The cases presenting different possible usages with an emphasis on air pollution data are presented. This varies from the air-pollution flowers for each month of the year or for each day of the week to show its applicability for short or very-long time-series analysis. The air-pollution flower diagram shows it at its best when comparing patterns for different variables or patterns for the same variable but for a different location. Such a clean and clear graphical presentation of the diurnal-pattern's characteristics is the main novelty of the presented diagram.

Acknowledgement

The authors acknowledge that the projects (ID L1-4154, ID L2-5475, ID L2-6762, ID L7-8268 and L2-8174) were financially supported by the Slovenian Research Agency. We are grateful to the Slovenian Environment Agency for the measurement data from the towns of Zagorje and Celje.

References

Božnar, M. Z., Grašič, B. and Mlakar, P. (2014) 'The problem of limit values exceedances detection in complex terrain using measurement and models', *16th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, 8-11 September 2014, Varna, Bulgaria.* BATCHVAROVA, Ekaterina (ur.), KIROVA, Hristina (ur.), HRISTOVA, Elena (ur.). HARMO 16: proceedings. [Sofia]: National Institute of Meteorology and Hydrology, Bulgarian Academy of Sciences, 5 pages

Božnar, M. Z., Grašič, B., Mlakar, P., Soares, J. R., de Oliveira, A. P. and Costa, T. S. (2015) 'Radial frequency diagram (sunflower) for the analysis of diurnal cycle parameters: solar energy application', *Applied energy*, vol. 154, 592-602, doi: 10.1016/j.apenergy.2015.05.055.

Cleveland, W. S., and McGill, R. (1984) 'Graphical perception: Theory, experimentation, and application to the development of graphical methods.', *Journal of the American statistical association* 79.387: 531-554.

Garland, J. A., and Derwent, R. G. (1979) 'Destruction at the ground and the diurnal cycle of concentration of ozone and other gases.', *Quarterly Journal of the Royal Meteorological Society* 105.443: 169-183.

Mayer, H. (1999) 'Air pollution in cities.', *Atmospheric environment* 33.24 (1999): 4029-4037.

Mlakar, P., Božnar, M. Z., Grašič, B., and Tinarelli, G. (2012) 'Zasavje canyon regional online air pollution modelling system in highly complex terrain - description and validation', *International Journal of Environment and Pollution*, Volume 50, Numbers 1-4, December 2012, pp. 22-30

Mlakar, P., Grašič, B. and Božnar, M. Z. (2014) 'Validation of small scale prognostic air pollution modeling system in highly complex terrain.' *Air Pollution Modeling and its Application XXII*. Springer Netherlands, 2014. 581-585.

Sandradewi, J., A. S. H., Prévôt, E., Weingartner, R., Schmidhauser, M.,Gysel, and Baltensperger, U. (2008). 'A study of wood burning and traffic aerosols in an Alpine valley using a multi-wavelength Aethalometer'. *Atmospheric Environment*, 42(1), 101-112.

Thunis, P., Georgieva, E., Pederzoli, A. and Pernigotti, D. (2011) 'The DELTA tool and Benchmarking Report template. Concepts and User's Guide', Join Research Centre, Ispra, 31.

Zelenko, B. and Lisac, I. (1994) 'On the Statistical Wind Rose Analysis'. *Theor. Appl. Climatol.* Vol. 48, Issue 4 : pp. 209-214, DOI: 10.1007/BF00867051.

Figure 1. The left picture presents the sunflower of the measured half-hour data for January 2015, for the global solar radiation at the Pustice station in the continental part of Slovenia. The right picture presents the same analysis for July 2015.

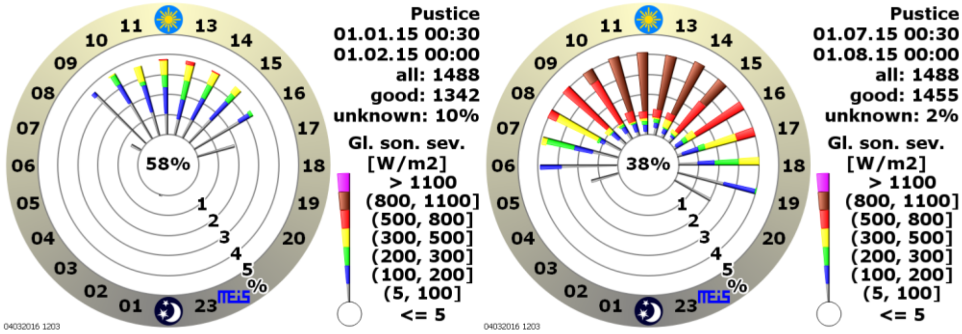

Figure 2. Air-pollution flowers for PM_{10} measured values (dust) for Zagorje in central Slovenia. 30-minute measured data for a period of one year (2013). Air-pollution flowers are made for each individual month.

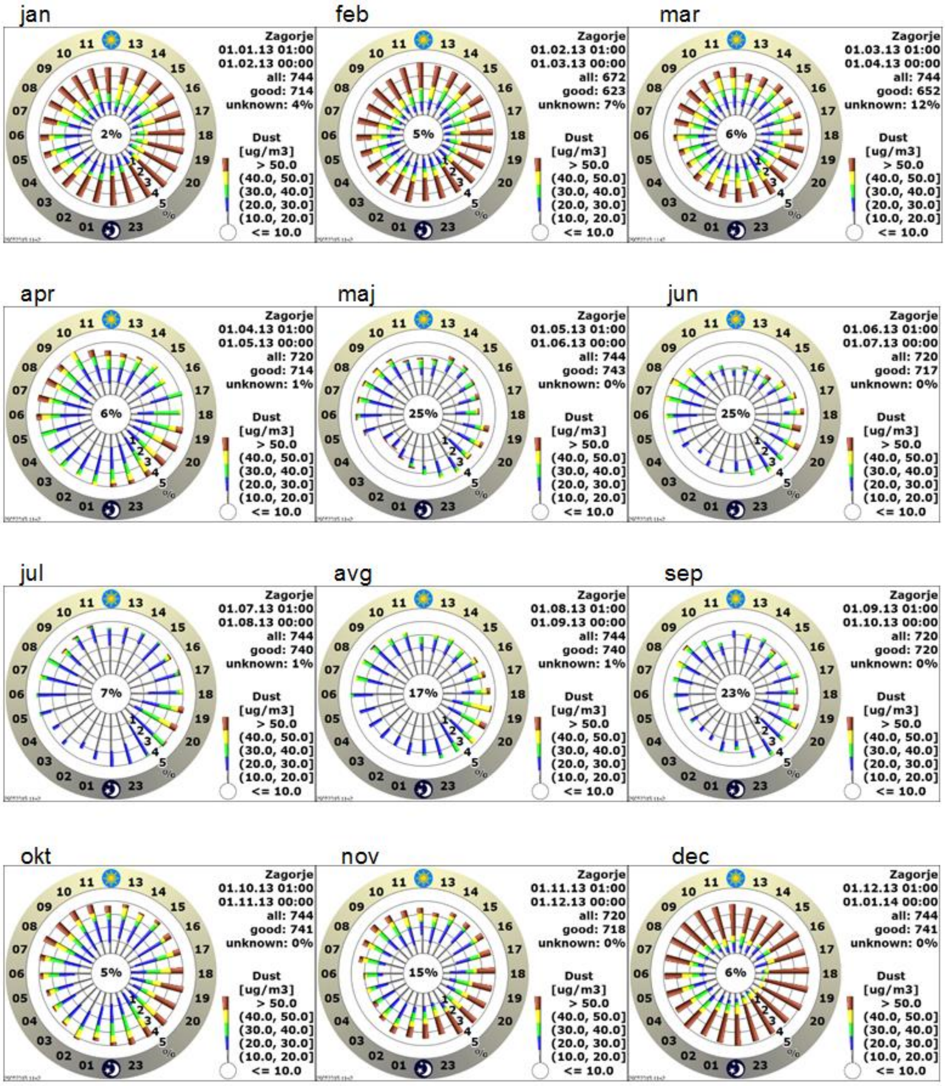

Figure 3. Air-pollution flowers based on the measurements for 8 individual consecutive days for PM_{10} in the town of Celje (22–29 November 2015.

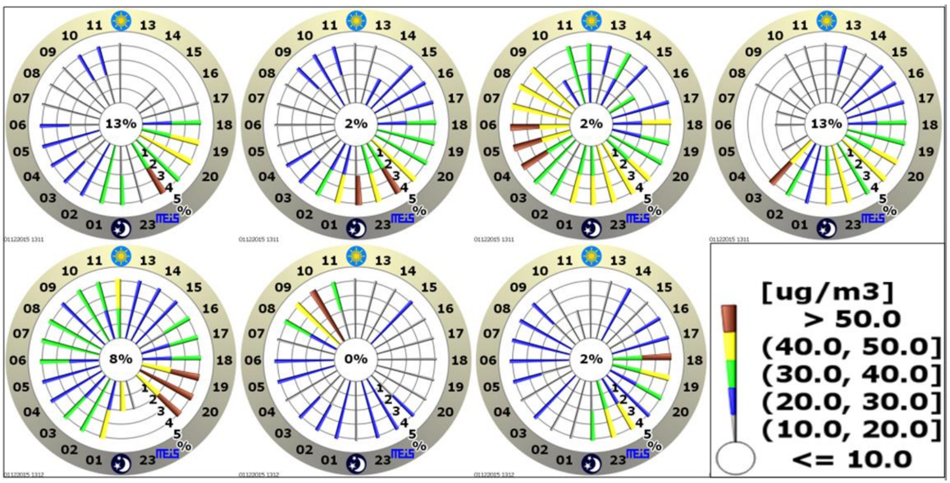

Figure 4. Air-pollution flowers for ozone and NO_2 for two selected months, in the town of Celje; the station is in the town centre.

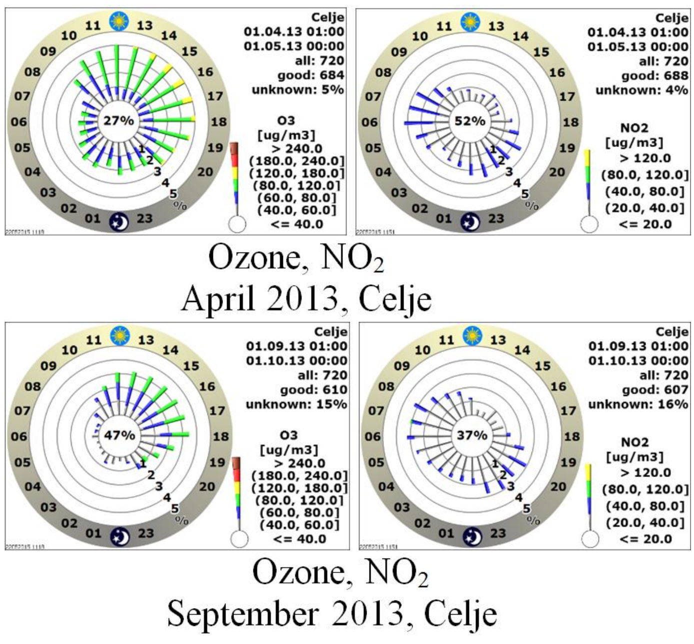

Figure 5. An example of a weather-flower diagram; a presentation example of a selected location at Krško for eight days (previous day in a grey colour, current day, the following 6 days)