20th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
14-18 June 2020, Tartu, Estonia

LINKS BETWEEN THE CONCENTRATIONS OF GASEOUS POLLUTANTS MEASURED IN DIFFERENT REGIONS OF ESTONIA

Aare Luts[1], Marko Kaasik[1], Urmas Hõrrak[1], Marek Maasikmets[2] and Heikki Junninen[1]

[1] Department of Environmental Science, Tartu University, Estonia
[2]Estonian Environmental Research Centre, Estonia

Abstract: The factors that determine the concentrations of air pollutants (NO_2, SO_2, ozone), measured in 10 monitoring stations (4 rural background, 3 urban and 3 industrial) in Estonia, are studied, applying the factor analysis. The multiple regression analysis of ozone concentration in one rural background station was performed, based on the ozone concentrations in other stations and NO_2 concentration in the same station. It was found that the ozone concentration is rather well predictable (determination coefficient R=0.79), using only the concentrations from another rural station that is about 200 km away. Adding all the other stations in regression, the ozone concentrations were predicted somewhat better (R=0.89). Surprisingly, the NO_2 concentration in the same station did not appear a good predictor for ozone (R=0.27). The factor analysis reveals differences between rural and urban stations, also maritime-continental difference for NO_2 and ozone and specific industrial impact in case of SO_2.

Key words: *nitrogen dioxide, sulphur dioxide, ozone, factor analysis, multiple regression***.**

INTRODUCTION

The composition and concentrations of trace gases in the ambient air depend on many factors, e.g. on natural and anthropogenic (pollutant) emissions and transport, air chemistry, planetary boundary layer mixing state *etc*. In this study we are interested to find out the links between the patterns of the gas composition measured in various parts of Estonia.

We have used the measured two-year concentrations of hourly averaged values of O_3, NO_2 and SO_2 from the national air quality monitoring stations: background stations Lahemaa, Vilsandi and Saarejärve, urban stations Tallinn-Liivalaia, Tallinn-Õismäe, Tartu, industrial stations Kohtla-Järve, Kunda and Narva (national monitoring system, operated by Estonian Environmental Research Centre), and also the measured concentrations from the Tahkuse rural air monitoring station operated by the University of Tartu. We performed factor analysis (Varimax rotated) to search for the factors that determine the gas concentrations. Our special interest was to examine the regressions to model and predict the ozone concentration at Tahkuse, using the measured data from the same and from the other monitoring stations.

THE LINKS BETWEEN CONCENTRATIONS IN STATIONS

The NO_2 links

The links determined by the factor analysis are depicted in the Figure 1 and in the Table 1. NO_2 concentrations are essentially determined by two factors (cumulatively about 60% of the all variance in the concentration values, Figure 1), adding three more factors makes it possible to determine about 82% of the all variation. All the factor scores are calculated by realstatistics packet (2021). The concentration at Tahkuse behaves similarly to the one at Saarejärve and in certain extent to the ones at Tartu and Lahemaa. The Vilsandi station is quite unique. Kohtla-Järve, Narva, Kunda and in some extent Tartu depict rather similar variations, and the same is valid for the Tallin urban stations Liivalaia and Õismäe.

The first factor (Table 1) is easy to interpret as the factor of urban transport pollution, which is most related to the street station Liivalaia in the center of Tallinn city and with urban background station Õismäe in Tallinn. Factor 2 distinguishes clearly the continental rural background stations Tahkuse, Lahemaa and Saarejärve from others. Factor 3 is almost solely determined by maritime station Vilsandi on a small remote island. Factor 4 represents the small industrial town Kunda (cement factory, marine port) and also the Lahemaa background station, located rather close (30 km) to Kunda. Factor 5 may be related to urban-industrial conditions with minor impact of vehicles.

Table 1. The scores of the most impotrant 5 factors that determine NO_2 concentration variations at specific stations

Station	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Tahkuse NO_2	-0,02	**-0,50**	-0,03	-0,24	0,06
Lahemaa NO_2	0,02	-0,32	0,09	**0,47**	**-0,42**
Vilsandi NO_2	0,03	0,19	**-1,04**	-0,02	0,01
Tartu NO_2	-0,03	-0,05	0,00	-0,29	**0,55**
Saarejärve NO_2	-0,08	**-0,57**	0,17	-0,16	0,02
Kohtla-Järve NO_2	-0,19	0,15	-0,01	0,34	0,33
Narva NO_2	-0,17	0,09	0,00	-0,04	**0,59**
Liivalaia NO_2	**0,76**	0,03	-0,02	-0,19	-0,22
Õismäe NO_2	**0,55**	0,06	-0,03	0,05	-0,16
Kunda NO_2	-0,11	0,19	-0,01	**0,75**	-0,10

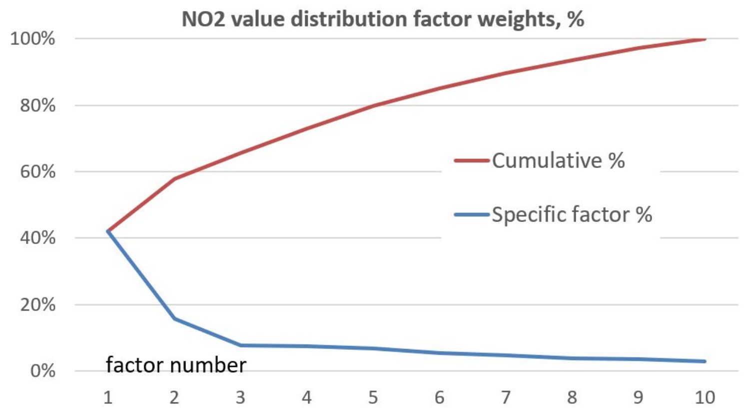

Figure 1. The relative importance of the factors that determine the observed NO_2 variations

The SO_2 links

The main results of factor analysis are presented in the Table 2 and in the Figure 2. SO_2 concentrations are almost evenly determined by many factors, the most relevant five factors determine about 67% of the all variation. To determine above 90% of all the variations, the 8 factors should be taken into account. Therefore, the SO_2 concentrations are rather location-specific. Mainly Kohtla-Järve and Narva can be grouped together according to the similar variations in the SO_2 concentrations and the same is valid for the Tallinn urban stations Liivalaia and Õismäe. Vilsandi and Tartu are similar according to the factor 5, but distinct according to the other factors.

Table 2. The scores of the most important 5 factors that determine SO_2 concentration variations at specific stations

Station	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Tahkuse SO_2	0,03	-0,04	**0,45**	-0,03	0,18
Lahemaa SO_2	**0,44**	-0,07	0,06	0,10	-0,12
Vilsandi SO_2	0,20	0,18	-0,14	-0,13	**0,69**
Tartu SO_2	-0,27	-0,16	0,02	-0,01	**0,61**
Saarejärve SO_2	-0,06	0,04	**0,80**	-0,09	-0,19
Kohtla-Järve SO_2	0,01	**-0,65**	-0,05	0,02	-0,10
Narva SO_2	0,04	**-0,59**	0,04	-0,04	0,06
Liivalaia SO_2	-0,06	0,03	-0,04	**0,53**	-0,05
Õismäe SO_2	-0,13	0,00	-0,07	**0,58**	-0,07
Kunda SO_2	**0,70**	-0,01	-0,06	-0,18	0,01

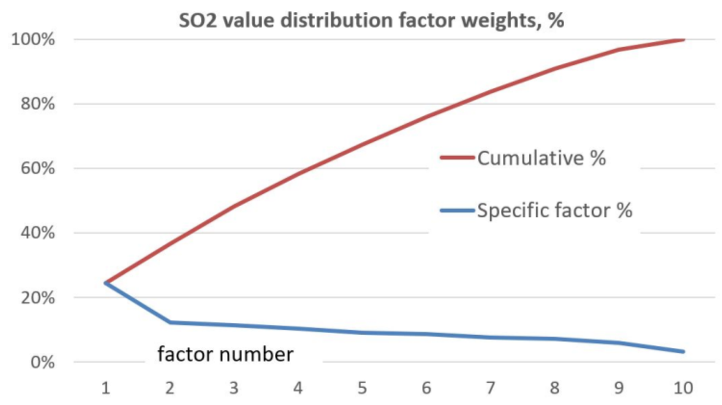

Figure 2. The relative importance of the factors that determine the observed SO_2 variations

The factors, that determine the SO_2 concentrations, could be as follows: 1) industrial pollution from Kunda cement factory, factor 1 is large at Lahemaa and Kunda; 2) factor 2 has large negative values at the stations located within the oil-shale based industrial area in North-Eastern Estonia; 3) the factor 3 could represent the residential heating because the scores of the factor 3 are large at Tahkuse and Saarejärve stations where the pollution from local heating can be significant, especially at Saarejärve; 4) the factor 4 could represent the urban pollution, because the sores are large at the stations within Tallinn city; 5) the factor 5 could be somewhat similar to the factor 3 (local pollution), because it is large at Tartu and Vilsandi whereas Tartu has many small houses with local heating systems. However, this factor is also different in some measure, because at Saarejärve the score of this factor is negative whereas at Vilsandi the factor 3 is negative. It is remarkable that first two factors are most probably driven by large industrial emissions of SO_2, respectively by Kunda cement factory and oil-shale industrial complex.

The ozone links

The links determined by the factor analysis are depicted in the Figure 3 and in the Table 3. Ozone concentrations are essentially determined by the first factor (about 67% of the all variance in the concentrations), adding four more factors makes it possible to determine above 90% of the all variation. According to the factor analysis the ozone concentration behaviour at the monitoring stations can be allocated to three closely bounded groups and also to certain groups where the links are weaker.

The concentration at Tahkuse behaves similarly to the one at Saarejärve. Also Kohtla-Järve and Narva depict rather similar variations, and the same is valid for the Tallinn urban stations Liivalaia and Õismäe. The weaker but still remarkable links are between the concentrations at the stations Vilsandi and Kohtla-Järve (factor 1). Vilsandi and Kohtla-Järve are not similar in the environment (remote marine environment vs industrial one), so the links between these locations have somewhat unexpected character.

Table 3. The scores of most impotrant 5 factors that determine ozone concentration variations at specific stations

Station	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Tahkuse ozone	**0,69**	0,02	-0,06	-0,14	**-0,51**
Vilsandi ozone	-0,34	-0,10	-0,27	**-1,20**	-0,13
Lahemaa ozone	0,14	-0,12	0,11	0,18	0,20
Tartu ozone	**0,45**	0,17	-0,12	-0,02	-0,18
Saarejärve ozone	**0,76**	-0,07	-0,19	0,21	-0,14
Kohtla-Järve ozone	-0,30	0,00	-0,16	0,07	**0,79**
Narva ozone	**-0,47**	0,06	-0,21	-0,04	**0,88**
Liivalaia ozone	-0,22	0,00	**0,89**	0,17	-0,26
Õismäe ozone	-0,13	-0,08	**0,64**	0,06	-0,14
Tahkuse NO2	0,07	**-0,98**	0,12	-0,14	0,01

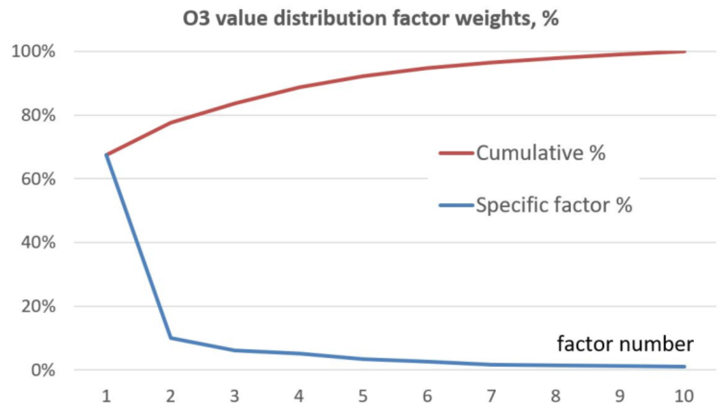

Figure 3. The relative importance of the factors that determine the observed ozone variantions

Ozone is generated by chemical reactions initiated by solar UV radiation, therefore the main factor could be rural background mixed by urban-industrial conditions. Tartu seems similar to rural stations according the factor, which may be due to absence of high traffic streets and a lot of residential heating nearby the station. It is known that ozone and NO_2 participate in the same chemical reactions, therefore according to the Table 3 the factor 2 could be the responsible for the concentration of NO_2 and/or somewhat the concentrations of the pollutant gases in general. The factor 3 could correspond to (overall) urban pollution, because its values are large at the Tallinn urban stations. The factor 4 is almost solely determined by Vilsandi, thus most probably represents maritime conditions. The factor 5 could correspond to some other (industrial) pollution, because the values of it are large at the industrial region stations Kohtla-Järve and Narva.

THE MODEL TO PREDICT THE OZONE CONCENTRATIONS AT TAHKUSE

We made attempts to calculate the Tahkuse ozone concentrations by multiple linear regression analysis using the data from other stations as independent variables. When to consider the simple regressions with only one argument, the highest coefficient of determination R corresponds to the regression where the ozone data from Saarejärve are used as independent variables (R=0.797). In the case of the multiple linear regressions where the data from Vilsandi, Tartu and Lahemaa were added, coefficient R values become equal to 0.854, 0.872 and 0.883, respectively. In the case of the multiple regression where the ozone data from the all stations and also the NO_2 data measured at Tahkuse station were used as independent variables, the determination coefficient R=0.889. The ozone prediction results for the case with four regression arguments (ozone concentrations from Saarejärve, Vilsandi, Tartu and Lahemaa) are depicted in Figures 4 and 5.

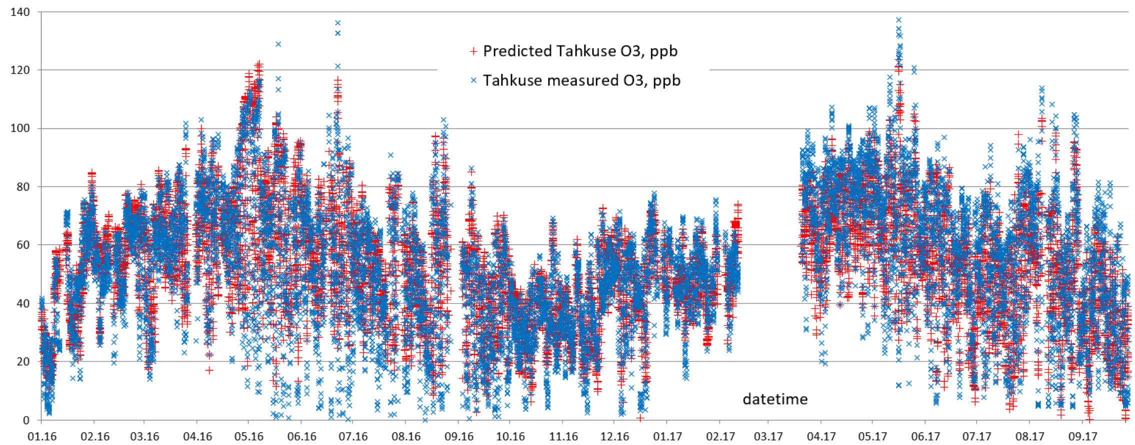

Figure 4. The observed and predicted ozone concentrations

Because ozone and NO_2 participate in the same chemical reactions, a remarkable regression between the quantities of these parameters could be expected, but actually in the case when only NO_2 concentrations were selected as regression independent variable, the determination coefficient was only 0.27 and also the combinations of NO_2 concentrations with the concentrations of ozone, measured at the other stations, only yield a negligible enhancement in the determination coefficient (enhancement in the determination coefficient value driven by NO_2 concentrations within the argument list is below 0.005), therefore we omitted NO_2 concentrations from the independent variables.

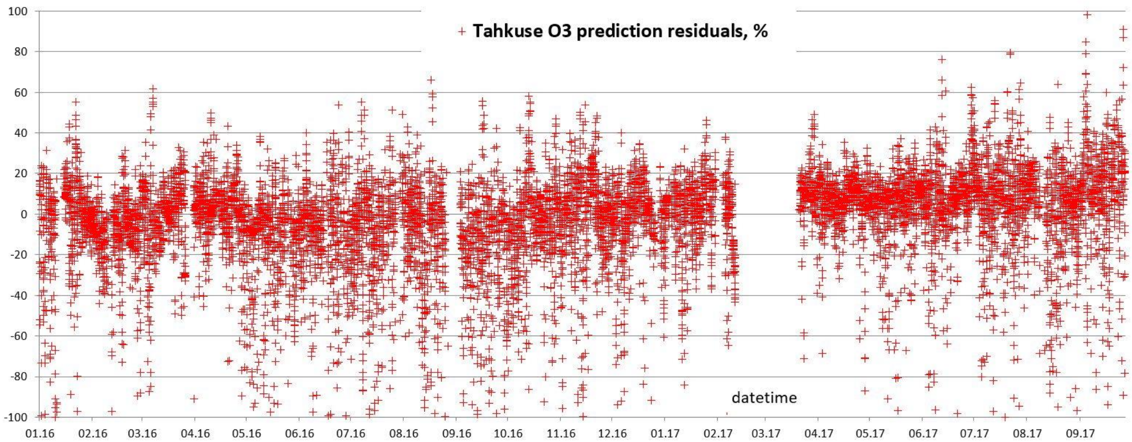

Figure 5. The residuals between observed and predicted ozone concentrations, in per cents

CONCLUSIONS

We studied the factors that determine the variations in the ambient trace gas O_3, NO_2 and SO_2 concentrations measured at several monitoring stations in Estonia and in particular examined the regressions to predict (model) the ozone concentrations for the Tahkuse station. The SO_2 concentration variantions are quite location-specific, the most important five factors only enable to determine about 67% from the overall variations. The ozone concentrations in general are determined mainly by the first factor that can be roughly rural background mixed with urban-industrial difference (determination coefficent R=0.67), but four more factors can also be distinguished. The ozone concentration at a specific location (Tahkuse) can be predicted using the multiple linear regression model where the measured ozone concentrations at Saarejärve, Vilsandi, Tartu and Lahemaa are taken as the independent variables. In this case the determination coefficient R=0.883 whereas the median of the residuals is equal to about 1.2% (the predicted concentrations are slightly overestimated). Nevertheless, at times the residuals can reach several tens of percents. Unexpectedly we detected that the NO_2 concentrations can not be used as an effective argument to predict the ozone concentration, the corresponding determination coefficient was only about 0.27.

Acnowledgements

This work was supported by the European Regional Development Fund (project MOBTT42) under the Mobilitas Pluss programme, Estonian Research Council (project PRG714), Estonian Environmental Observatory (KKOBS, project 2014-2020.4.01.20-0281).

REFERENCES

realstatistics, 2021: https://www.real-statistics.com/multivariate-statistics/factor-analysis/factor-scores/.