Statistical examination of the relationship between PM10 emissions and road traffic

Götz Wiegand, Volker Diegmann, Florian Pfäfflin
IVU Umwelt GmbH · Burgweg 10 · D-79350 Sexau · Germany
URL: www.IVU-Umwelt.de · eMail: GW@IVU-Umwelt.de

Presented at: Workshop "Particles in the size of 2.5 to 10 microns in urban areas"
4.-6.11.2002 FU Berlin

1 Introduction

In 1998 and 1999 the BUWAL[1] studied the relevance of road traffic for PM10 and PM2.5 concentrations. The investigation was principally based on chemical analyses of average daily values of dust samples, and receptor- and dispersion model analyses. In addition, time resolved measurements of PM10 concentrations were undertaken in Zurich over 2 periods. This measurement programme has not been optimised to statistical analysis. Nevertheless a statistical analysis is done with this dataset because it is one of the most comprehensive datasets available.

The objective of the study has been the application of statistical methods on the time resolved measurement data set from 15.1.1999 to 1.3.1999 in order to determine the extent to which the PM10 concentration is influenced by traffic loads and meteorological variables.

It would be advantageous if one could find out more about PM10 concentrations by statistical means, because the generation of data for analysis of the chemical composition of dust constituents is expensive.

2 Input data

2.1 The EMPA data set

Total and background-subtracted PM10 concentrations, obtained at Zuerich's Wiedikon measurement station, are used as dependent variables. The Wiedikon measurement site is located in the vicinity of a crossroad south-west of Schimmelstrasse, a through road with an average daily traffic load of approx. 30,000 vehicles.

The direction of the wind that transports the pollutants from Schimmelstrasse to the Wiedikon measurement site is unclear, as the flow field appears to be considerably complicated by the effect of the crossroad.

The PM10 concentration measured at a distance of about 600 m in a north-north-east direction is used as background measurement (see Fig. 1).

[1] Hüglin, C.: Anteil des Straßenverkehrs an den PM10- und PM2.5-Immissionen, Chemische Zusammensetzung des Feinstaubes und Quellenzuordnung mit einem Rezeptormodell, Report C4, *Bundesamt für Umwelt, Wald und Landschaft* (BUWAL) [the Federal Office of the Environment, Forests, and the Countryside], Berne 2000

The following independent variables are available:

- Wind-direction (°), wind-speed (m/s),
- Car load (number), truck load (number),
- Relative humidity (%), temperature (°C), air pressure (hPa), precipitation (mm),
- Global radiation (W/m²), radiation balance (W/m²) and
- Weekday (1 - 7).

Fig. 1 Map of Zurich with the "Wiedikon" road measurement site under investigation and the "Kasernenenhof" background measurement station approx. 600m in a north-north-west direction.

2.2 Vehicle speed

The velocity of vehicles is not measured but could be an important variable for PM10 concentrations near roads. Therefore a simple vehicle speed model is employed. It has a similar structure to that used in the Canyon Plume Box model[2] (CPB). CPB vehicle velocity is proportional to road capacity and the vehicle load at any particular time.

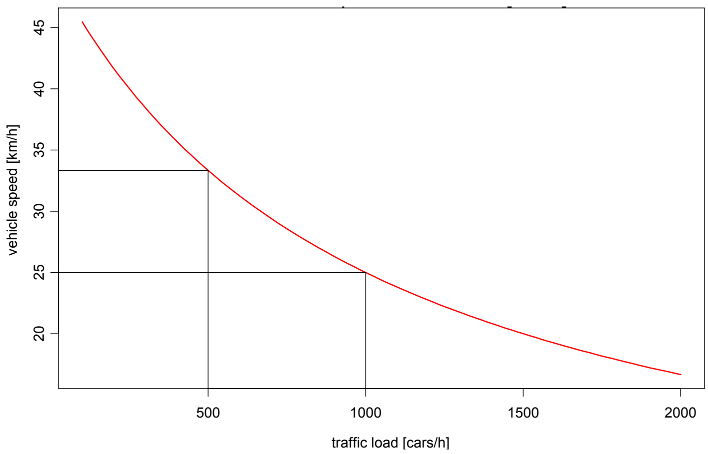

Fig. 2 Simple model of the relationship between vehicle velocity and hourly traffic load with a road capacity of 1000 vehicles/h and a undisturbed flow of the traffic of 50 km/h. The speed of the cars is reduced to one half when the traffic load equals the capacity of the street.

2 Yamartino, R. J., Wiegand, G.: Development and Evaluation of Simple Models for the Flow, Turbulence and Pollutant Concentration Fields within an Urban Street Canyon; Atmospheric Environment Vol.20, No.11, pp 2137-2156; 1986

2.3 Wind-directions

A plot of the wind-directions- and –speeds is shown in Fig. 3. Fig. 4 shows the map with the classification of the wind-directions and the heading of the street. One can see, that the direction, in which the background measurement station is situated (north-east), is underrepresented in the windrose of the dataset.

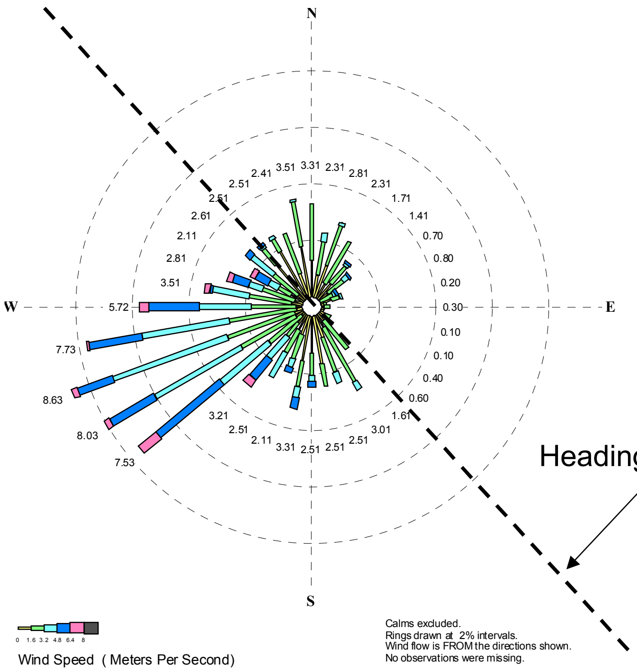

Fig. 3 Wind-direction and speed distribution for Wiedikon during the period from 15.1.1999 to 1.3.1999. See text for more details.

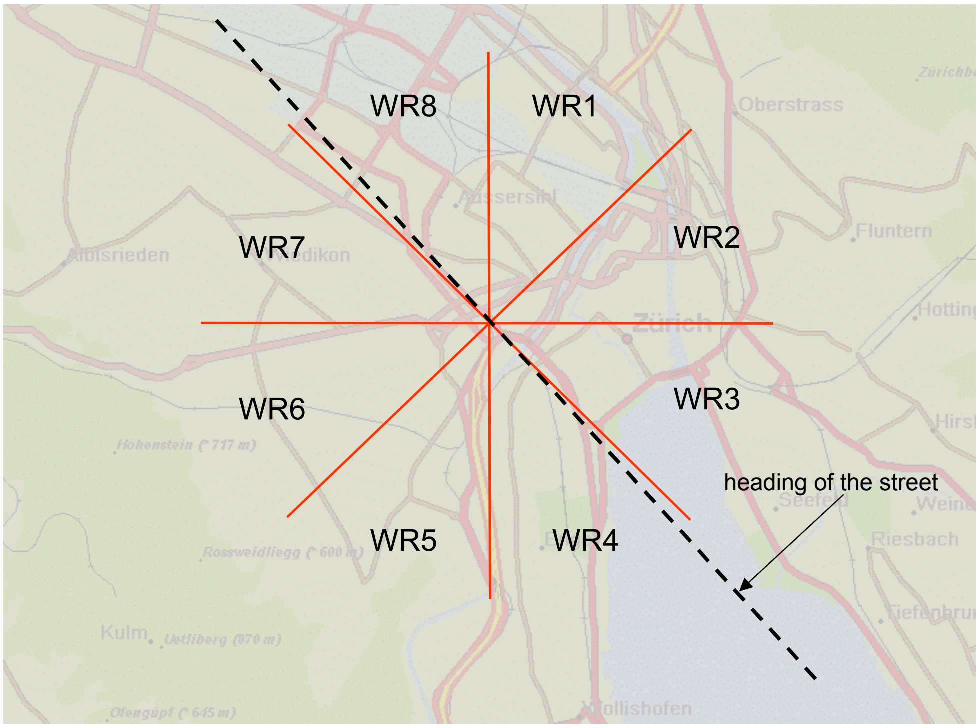

Fig. 4 The wind-direction is divided into 8 sections named WR1, …, WR8 and dummy variables are defined for each wind-section. Along the street are WR7/WR8 and WR3/WR4.

2.4 The data base scatter plot matrix

The scatter plot matrix of selected variables in Fig. 5 provides 2-dimensional sections through the data set. In the first column or first line one finds the PM10 concentration in relation to the meteorological or traffic variables. The matrix is symmetrical.

The extremely high PM10 concentrations are found about noon and with wind-directions from the north or south. Wind-speeds are low, air pressure is extremely high, car and truck loads are high and the modelled vehicle speed is low.

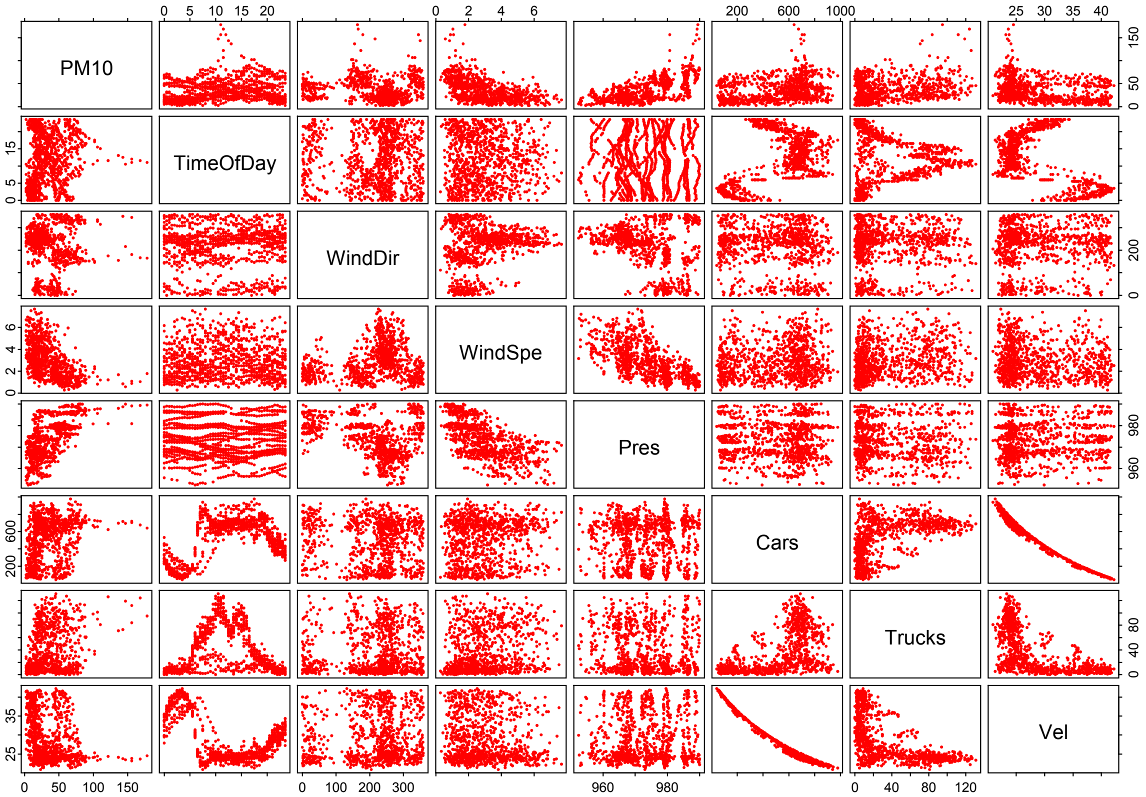

Fig. 5 The scatter plot matrix for selected variables: PM10 = PM10 concentration in µg/m³ at Wiedikon, TimeOfDay = time of day in hours, WindDir = wind-direction in °, WindSpe = wind-speed in m/s, Pres = air pressure in hPa, Cars = number of cars per half hour, Trucks = number of trucks per half hour, and Vel = vehicle velocity. See text for more details.

3 Statistical model

3.1 General remarks

It is difficult to separate the influence of traffic data from the effects of meteorological variables, as both groups of variables are subject to similar daily cycles. Statisticians call this property "multicollinearity". The high covariance of the independent variables leads to a "nearly" singular data-matrix, i.e. the determinant of the data-matrix is near zero. The inverse of the data-matrix appears in the denominator of the estimators of coefficients and their variances. That leads to problems with the interpretation of the coefficients, because the estimated variances of the estimated coefficients become inflated and the model becomes unstable.

Multicollinearity is a problem of the data and not a problem of the violation of regression-model-assumptions. Strongly correlated variables do not contribute unique information to the model. Through inspection of the variance-inflation-factors[3] one can identify the variables which caused the multicollinearity problem. Those variables were omitted from the regression model.

The regression models in this study are estimated with the S+ (Version 4.52) procedures "lm" and "step". All the variables mentioned are significantly different from zero, and the assumptions of linear regression models are fulfilled.

3.2 Total PM10 concentration

3.2.1 Meteorological variables

Variations in total PM10 concentrations in the data can be explained to 66% using meteorological data. The regression is carried out with standardised variables, allowing to conclude the importance of the variables directly from the coefficient values (so-called β-coefficients). The β-coefficients in the order of their importance are:

1: 16% atmospheric pressure
2: 11% wind-direction 4
3: 9% wind-direction 5
4: 9% wind-direction 8
5: 7% wind-direction 6

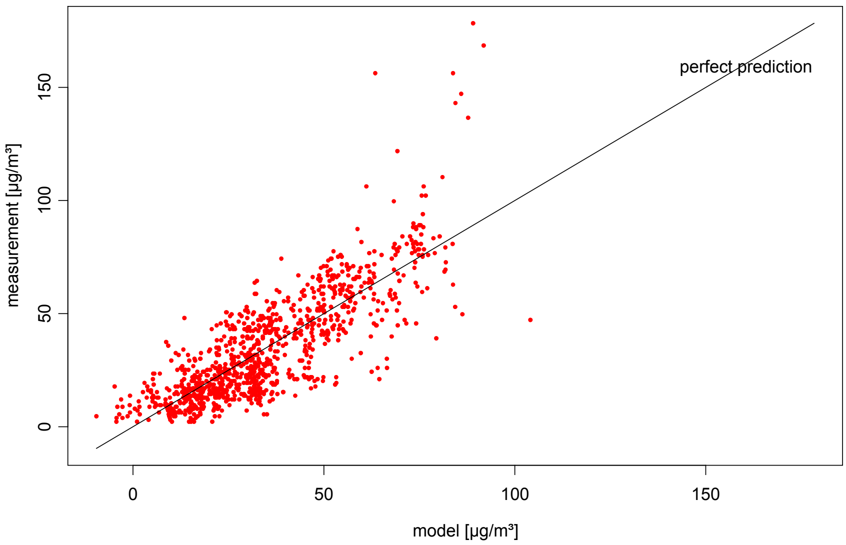

Fig. 6 Scatter plot of total PM10 concentrations against the statistical model with meteorological variables. The coefficient of determination value is 66%. See text for more details.

Air pressure has the greatest influence here – though this is not immediately explicable. It is not clear to the authors what direct influence air pressure could have on PM10 concentrations, maybe the influence is indirect, over an unobserved variable.

Although the extreme PM10 concentrations above 100 µg/m³ in Fig. 5 can be explained in tendency by the regression model, they attract attention as extreme values.

[3] see Wooldridge, J.M., Introductory Econometrics: A Modern Approach, South Western, 2000

3.2.2 Meteorological and traffic variables

The coefficient of determination rises to 71% when traffic load variables are added. As the explanatory variables are interdependent, the difference in the coefficient of determination (5% in this case) does not equal the weight of the traffic influence. The β-coefficients in the order of their importance are:

1: 15% atmospheric pressure
2: 10% wind-direction 4
3: 9% wind-direction 5
4: 8% wind-direction 6
5: 8% wind-direction 8
6: 6% wind-direction 7
7: 6% temperature
8: 5% cars
9: 5% trucks

The traffic variables "cars" and "trucks" only appear in 8^{th} and 9^{th} position respectively. They make an approximately equal contribution towards explaining the variations in the PM10 concentrations, together about 10%.

3.3 Background-subtracted PM10 concentrations

3.3.1 Meteorological variables

The time series of the total and background-subtracted PM10 concentrations is shown in Fig. 7. One can see that the background-subtracted time series is very much smoother than the time series of the total PM10 concentration. One can also recognise the rhythm of the traffic cycle far more clearly.

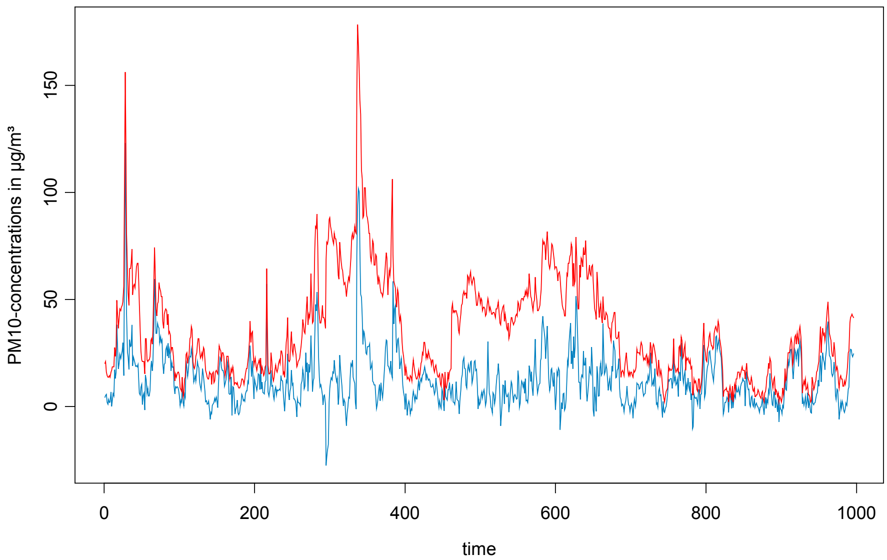

Fig. 7 Time series of total (red) and background-subtracted (blue) hourly PM10 concentrations. A considerable amount of the variation in the total PM10 series is filtered out when the background is subtracted, and the rhythm of the traffic cycle becomes clearly visible.

The variations in background-subtracted PM10 concentrations in the data can be explained to 29% with the meteorological data obtained at the same time. The β-coefficients in the order of their importance are:

1: 13% wind-direction 4
2: 11% wind-direction 5
3: 10% radiation balance
4: 9% wind-direction 6
5: 7% wind-direction 7

The air-pressure is not dominant any longer, that means that the air pressure seems to be a variable describing effects (direct or indirect) far away from the source. While variables which act at a short distance become more relevant with background subtracted PM10-concentrations.

3.3.2 Meteorological and traffic variables

When the traffic load variables are included in the regression, the coefficient of determination rises to 40%. As the explanatory variables are interdependent, the difference in the coefficient of determination (11% in this case) does not equal the strength of the traffic influence. The β-coefficients in the order of their importance are:

1: 12% wind-direction 4
2: 11% number of trucks
3: 10% wind-direction 5
4: 10% wind-direction 6
5: 9% number of cars

The number of trucks ranks at 2nd place just below wind-direction 4, while the number of cars is in 5th position. Together, the two variables contribute approximately 20% towards explaining the variations in the background-subtracted PM10 concentrations.

4 Conclusions

As is known from chemical analyses of dust constituents[4], the traffic component contributes about 50% to PM10 concentrations on a busy road in an urban setting.

This study, however, has shown that while traffic load variables do indeed have a significant statistical influence on PM10 concentrations, it is far less than the influence of meteorological variables. The reasons for this somewhat surprising result could be:

- The effect of traffic is measured using numbers for cars and trucks, and perhaps these numbers are not the best measure for traffic-based PM10 emissions.
- Only linear models were considered, although there were obviously non-linear relationships between the variables (see Fig. 5).
- The wind-directions from the road to the measurement site are significantly under-represented (27%) in the data set (see Fig. 3).
- The background measurement site is on the "wrong" side of the road investigated (see Fig. 1), in consideration of the wind-direction-distribution.
- The background measurement site is too far away to represent the real background for the "Wiedikon" measurement site, at least for low wind-speeds, where measurements at Wiedikon are no longer contemporaneously to measurements at Kasernenhof.

There are currently two measurement programmes in progress (see Table 1) that appear to be considerably better suited for statistical analysis than the "Wiedikon" measurement programme which was, as mentioned above, designed to answer a different question.

Table 1 Overview of variables measured in two measurement programmes in progress

	Up- and downwind measure-ments in main WD	Traffic variable	Meteorological variables	PM10	PM2.5	PM1	PM10	PM2.5	PM1
				average daily values			time-resolved		
UMEG Karlsruhe	✓	✓	✓	✓	✓	–	✓	–	–
EMPA Zuerich	✓	✓	✓	✓	–	✓	✓	–	✓

[4] See e.g. Hüglin, C., Anteil des Straßenverkehrs an den PM10- und PM2.5-Immissionen, Chemische Zusammensetzung des Feinstaubes und Quellenzuordnung mit einem Rezeptormodell, Bericht C4, Bundesamt für Umwelt, Wald und Landschaft (BUWAL), Berne 2000
The results of this work are in line with other literature.