Analysis and Prediction of Particulate Matter PM10 for the Winter Season in Graz[1]

Siegfried Hörmann, Brigitte Pfeiler, and Ernst Stadlober

Graz University of Technology, Austria

Abstract: We summarize the investigations and results of recent empirical analysis on particulate matter PM10 in Graz. The influence of meteorological as well as anthropogenic factors is presented and discussed. Moreover we introduce a prediction model using current information and meteorological forecasts to predict the average concentration of particulate matter PM10 for the next day. Finally, we report on experiences with a test run carried out from December 16, 2004 until April 15, 2005.

Zusammenfassung: Wir berichten über unsere umfangreichen Untersuchungen und empirischen Resultate im Zusammenhang mit dem Grazer Feinstaubproblem. Neben Analysen des Einflusses von meteorologischen Bedingungen und von Menschen verursachten Faktoren, wird ein Prognosemodell entwickelt, das den aktuellen Stand der PM10 Belastung mit meteorologischen Vorhersagen für den nächsten Tag verknüpft, um daraus den PM10 Mittelwert des Folgetages zu prognostizieren. Schließlich wird über Erfahrungen aus einer Testphase, die vom 16. Dezember 2004 bis 15. April 2005 abgelaufen ist, berichtet.

Keywords: Particulate Matter, Empirical Analysis, Prediction Model, Multiple Linear Regression.

1 Introduction

The stringent load of particulate matter in Graz has caused remarkable attention in recent time and is a very present and explosive topic in media and politics. Mainly in the winter season the threshold value of PM10 (Particulate Matter) is exceeded regularly. This has caused public authorities to assign the Institute of Statistics to analyze the influence of meteorology and traffic to PM10. Since 2003 a collection of data has been analyzed on behalf of the Umweltamt Graz. This is documented in Stadlober und Pfeiler, 2004a, 2004b. In spring 2004 we started with the development of a statistical model in order to predict the average concentration of PM10 of the next day by using the information available at 12am of the current day (see Hörmann und Stadlober, 2004). Provided that the first stage — a test run from December 16, 2004 until April 15, 2005 — leads to suitable results, the second stage will be implemented: the prediction model as monitoring and decision tool starting at October 1, 2005. Similar investigations are planned for the cities of Klagenfurt and Bozen which are our partners in the EU-project KAPA GS (Klagenfurts Anti PM10 Action Programme in cooperation with Graz and South-Tyrol: project duration

from July 1, 2004 to June 30, 2007). Its aim is to investigate practical steps for achieving a sustained success in the reduction of `PM10`.

The data under investigation are provided from the FA 17C of the Styrian government (Mag. Schopper). They contain air quality and meteorological data measured on 8 different gaging stations in and around Graz:

- Graz-Mitte, Graz-Don Bosco, Graz-Süd (*traffic area*);

- Graz-Nord, Graz-Ost (*residential area*);

- Graz-Platte (*green area*);

- Kalkleiten (*control station for measurements of air temperature*).

The exploratory data analysis has been carried out with the help of the statistical package SPSS 12.0.

2   Facts on Particulate Matter

Particulate Matter, (PM) is a mixture of midget dust particles. We distinguish between the measures `PM10`, `PM2.5` and `PM1`, where the number denotes the diameter of the particle. E.g., `PM10` are particles with a diameter less than $10 \mu m$. The concentration in the air is measured in units of $\mu g/m^3$, with a threshold value of $50 \mu g/m^3$ for the daily `PM10` average. According to *Immissionsschutzgesetz – Luft (IG-L)* (1997) this limit must not be exceeded on more than 30 days of the year (valid for the years 2005–2009); a reduction to 25 days in 2010 is envisaged. Additionally, the annual average of `PM10` should be below $40 \mu g/m^3$. E.g., in 2003 the station Graz-Mitte registered even 136 exceedances, and the annual average amounted to a rather high value of $50 \mu g/m^3$.

2.1   What causes Particulate Matter?

Besides the size of the particles we have to characterize their origin and the chemical composition. The particles may be directly exhausted by burning processes or arise from mechanical abrasion of tyres, brakes, tarmac etc. But there are also natural sources like pollen or crushing and grinding rocks and soil (primary particles). Contrarily, there are particles which arise from aerially pollutants (secondary particles). For a more detailed description we refer e.g. to the URL `http://www.airinfonow.org/html/ed_particulate.html`.

The coarse particles (`PM10`) are smoke, dirt and dust, the fine particles (`PM2.5`) are toxic organic compounds or heavy metals. Automobiles with diesel engines eject very fine carbon-particles (`PM1`) which pose a serious threat to health. Estimates about the proportion of the PM polluters differ widely. In a Swiss study of the Bundesamt für Umwelt, Wald und Landschaft (Kropf, 2001) 36% of `PM10` emission is assigned to traffic and nearly the same amount to industry and private households. However, such figures have to be treated critically since they give an overall average based on a long time period for a large region and do not differentiate between urban or rural environments.

2.2 Health Effects of Particulate Matter

The effects on the health caused by PM depend on the size of the particles. In general we can say the smaller the particles are, the more toxic they are. The reason is that particles with a diameter of $10 \mu m$ or more are filtered when inhaled through the nose. While bigger particles remain in the throat, the smaller ones (PM2.5 or PM1) get in the trachea and in the respiratory system. Besides, they are also more toxic since they often consist of heavy metals and cancer causing organic compounds. For interested readers we refer to the documentation of Swiss medical doctors (Umweltschutz, 2003).

3 The Situation in Graz

Unfavorable weather conditions in the area of Graz cause high concentrations of particulate matter: low wind velocities, rare days with precipitation (rain or snow), many days with so called temperature inversion. Our empirical analysis showed that these factors have the strongest meteorological impact on PM10, which shall be discussed in the subsequent sections.
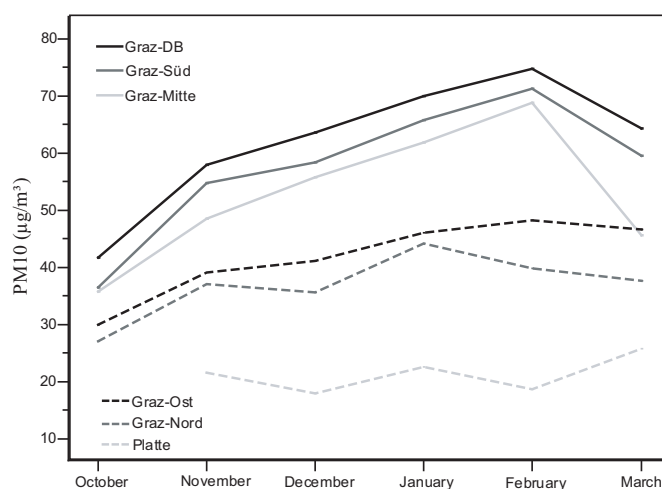

Figure 1: Monthly means of PM10 from October 2003 to March 2004 at six sites of Graz.

Figure 1 shows the monthly averages of PM10 from October 2003 to March 2004 of the 6 sites of Graz resulting in three different levels of concentration: high values at Graz-DB, Graz-Süd, Graz-Mitte (traffic area, industrial area, urban area), moderate values at Graz-Ost and Graz-Nord (residential areas) and low values at Graz-Platte (green area) where the value of October 2003 is missing.

3.1 Temperature Inversion

In case of temperature inversion the air temperature on the ground is lower than the air temperature 200-600m above the ground. In that case there is no mixture of the cold

air on the ground with the warmer air at higher altitude. To investigate the influence of temperature inversion we divide our data in two groups: If the average difference of the air temperature of Graz (350m sea level) and Kalkleiten (710m sea level) is negative within a day (0:00-24:00) this indicates *inversion* and if it is positive this indicates *no inversion*. Figure 2 exhibits the box plots of the daily averages `PM10` at the three sites Graz-Nord, Graz-Mitte and Graz-DB categorized by inversion. The `PM10` medians of the days without inversion are between 54% and 65% of the medians corresponding to inversion days.
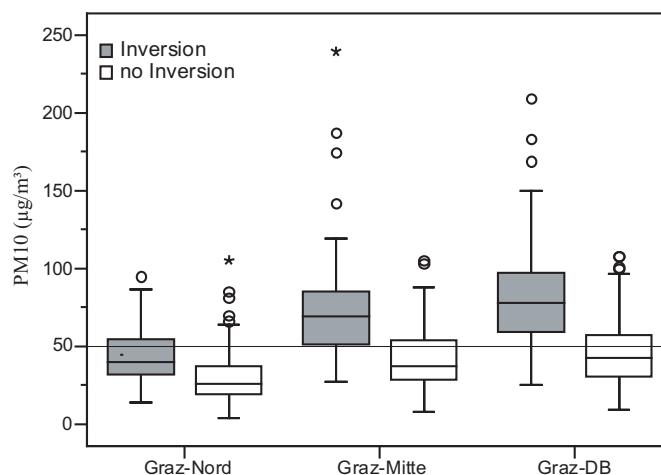

Figure 2: Box plot series of daily `PM10` concentrations for days with and without inversion at three sites of Graz.

The ZAMG Regionalstelle Stmk. (Podesser et al., 2004) investigated one typical inversion day (March 17, 2004, from 6am to 8pm) by measuring `PM10` concentrations and air temperatures at different altitudes. This was done by probes with captive balloons. Figure 3 displays the `PM10` concentration as function of time and altitude. It demonstrates that the dissolution of inversion at noon time corresponds to a lower level of `PM10` on the ground.

In order to analyze the daily load curves we examine the mean values of `PM10` on six consecutive 4 hour (4h) intervals. This reveals the fact that on average there is a decrease of `PM10` in the time period 12am to 4pm. This may be caused to a large extent by the meteorological fact of heavier inversions in the morning and in the evening than at noon time. Figure 4 shows this characteristic at the five sites of Graz on the ground (≈ 350m sea level), but not at Graz-Platte (661m sea level) which is most of the time not influenced by inversion because of its higher sea level.

3.2   Urban Traffic

The concentration of `PM10` is highly influenced by human factors. In urban areas it is mainly influenced by traffic, factory fumes and domestic fuel. Especially in Graz traffic seems to be the most important factor. In this region it may be responsible for 50% to 70% of the human caused `PM10` pollution. It is connected with carbon-particulate emissions of

PM10 concentration [μg/m³] on 17.03.2004
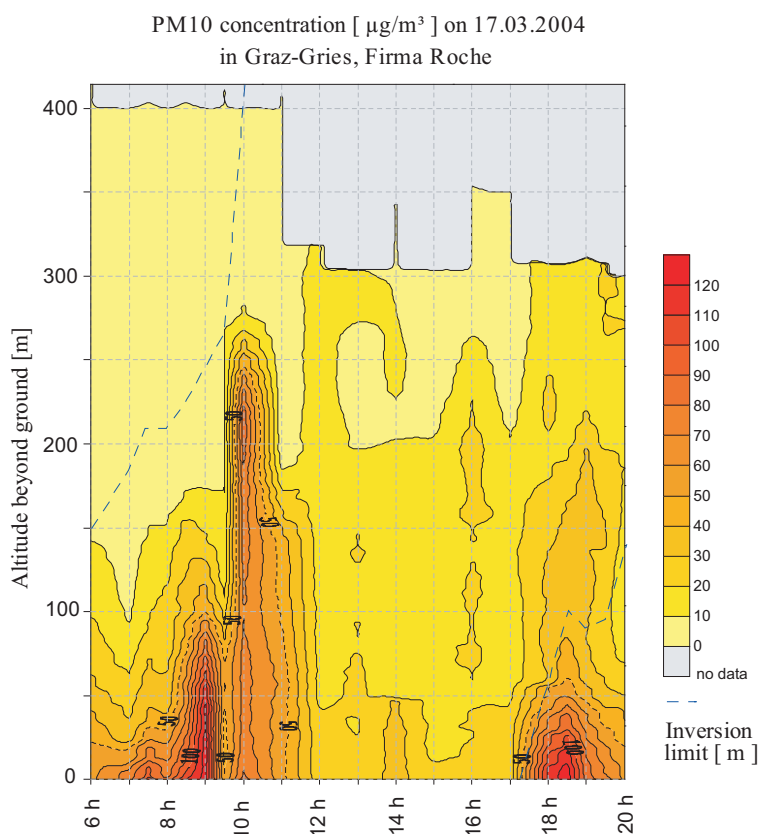in Graz-Gries, Firma Roche

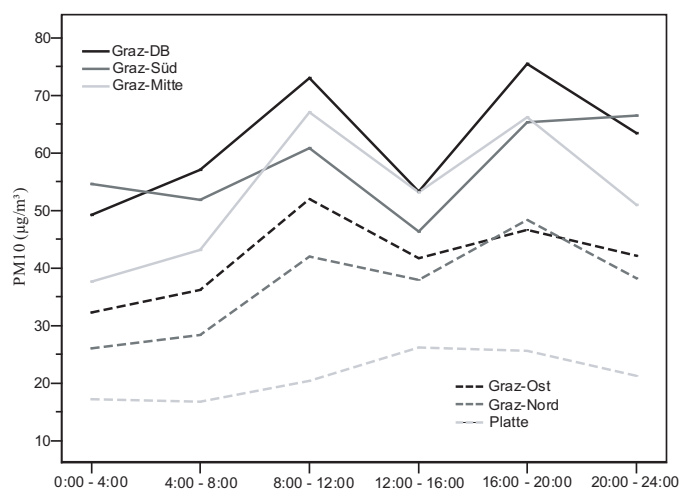Figure 3: PM10 as a function of time and altitude (Abb. 10 in Podesser et al., 2004).

Figure 4: 4h mean values of PM10 at six different sites of Graz.

diesel engines from passenger cars and trucks (very fine particles PM1), abrasion of tyres, brakes, coupler and tarmac, and the re-dispersion of particulate matter PM10. We studied the level of PM10 under different weather and traffic conditions to get some indications to what extent the pollution level of PM10 depends on traffic frequencies.
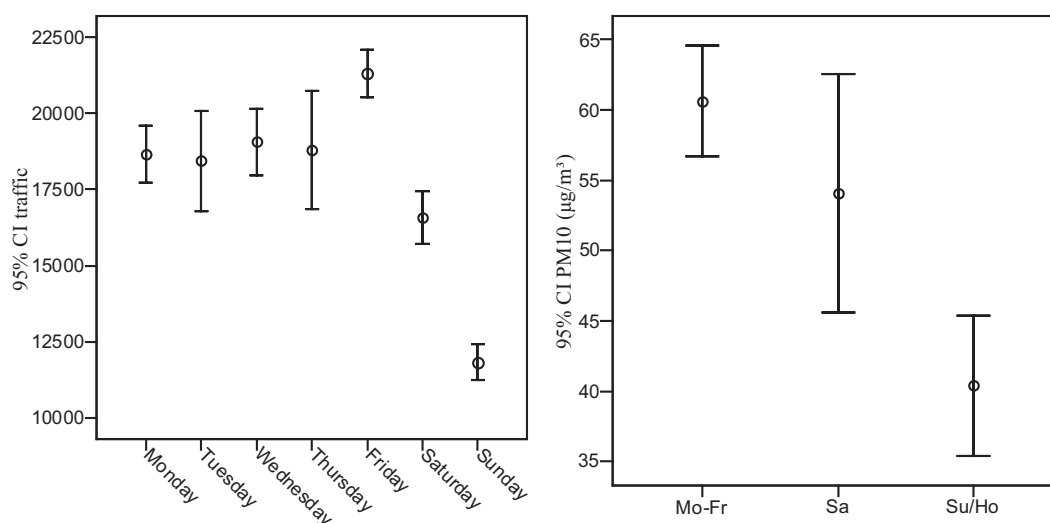
Figure 5: Error bars of traffic frequencies and of `PM10` at Graz-Mitte depending on the weekday.

We utilized measurements of traffic frequencies obtained by detectors (without differentiating the type of the vehicles) near the sites Graz-Nord, Graz-DB, Graz-Mitte and Graz-Ost (period October 2003 to March 2004). At a first glance we compared the frequencies on Fridays with the corresponding frequencies on the subsequent Sundays. On average the frequency on Sundays is less than 57% of the frequency on Fridays. Figure 5 shows the mean values of `PM10` and of the traffic frequencies for different weekdays at Graz-Mitte: a decrease of the traffic frequencies on weekends is associated with a decrease of `PM10`.

The box plot series in Figure 6 exhibits the typical characteristics of traffic frequencies on workdays and on Sundays/holydays. In Graz-Mitte there is much lower variability than at the traffic spot Graz-DB. The largest differences between the traffic frequencies of workdays and holidays are in the three 4h periods within 4am and 4pm. To deepen the exploratory analysis we split the traffic frequencies for each site separately at the quartiles and get four categories of frequencies: light, moderate, normal, heavy. Figure 7 displays box plot series of `PM10` for these four categories of traffic frequencies, grouped according inversion. Obviously, *inversion/heavy traffic* leads to very high concentrations (medians at 101 and $93\mu g/m^3$), *no inversion/light traffic* to the lowest ones (medians at 28 and $26\mu g/m^3$). At Graz-DB there is some saturation of `PM10` in the traffic category *heavy*. This may be caused by the fact that in case of high density of trucks fewer units are counted than at lower density levels of trucks, and it is well known that the impact of one truck to `PM10` corresponds to the impact of nearly 10 passenger cars.

3.3 Precipitation, Wind

Precipitation in form of rain or snow, and wind favor low concentrations of `PM10`. However, in Graz we have most of the time rather low wind velocities, and, e.g., in Winter
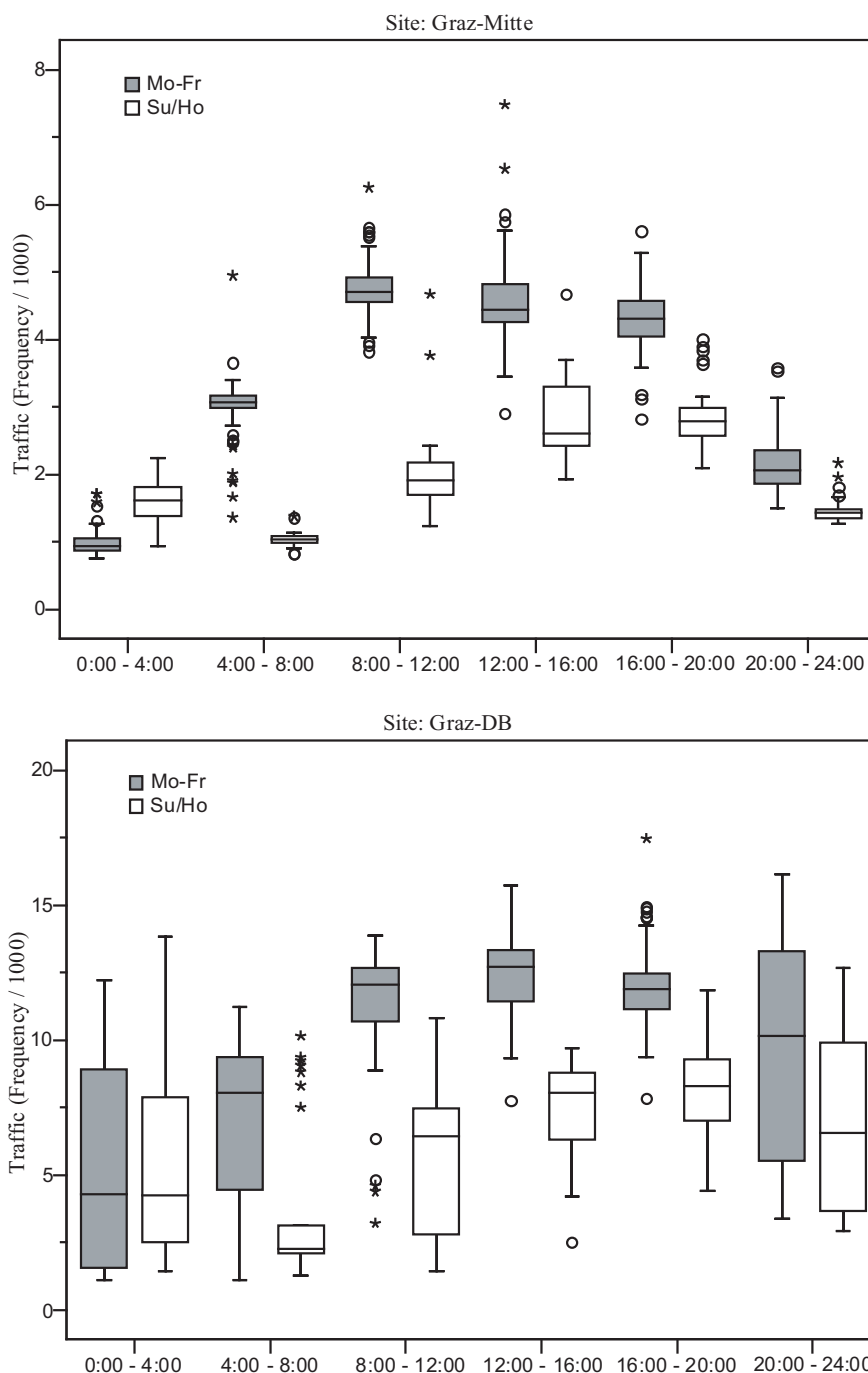
Figure 6: Box plot series from 4h-means of traffic frequencies on weekdays and on Sundays/holidays at Graz-Mitte and Graz-DB.

2003/04 only 14% of the 4h-periods registered precipitation. To get an impression we look at combinations of two wind categories (*easy/strong wind*: average wind speed below/above 0.65m/s) with two precipitation categories (*no precipitation/precipitation*) and exhibit corresponding box plots of the 4h-mean values of PM10. Figure 8 shows the cru-
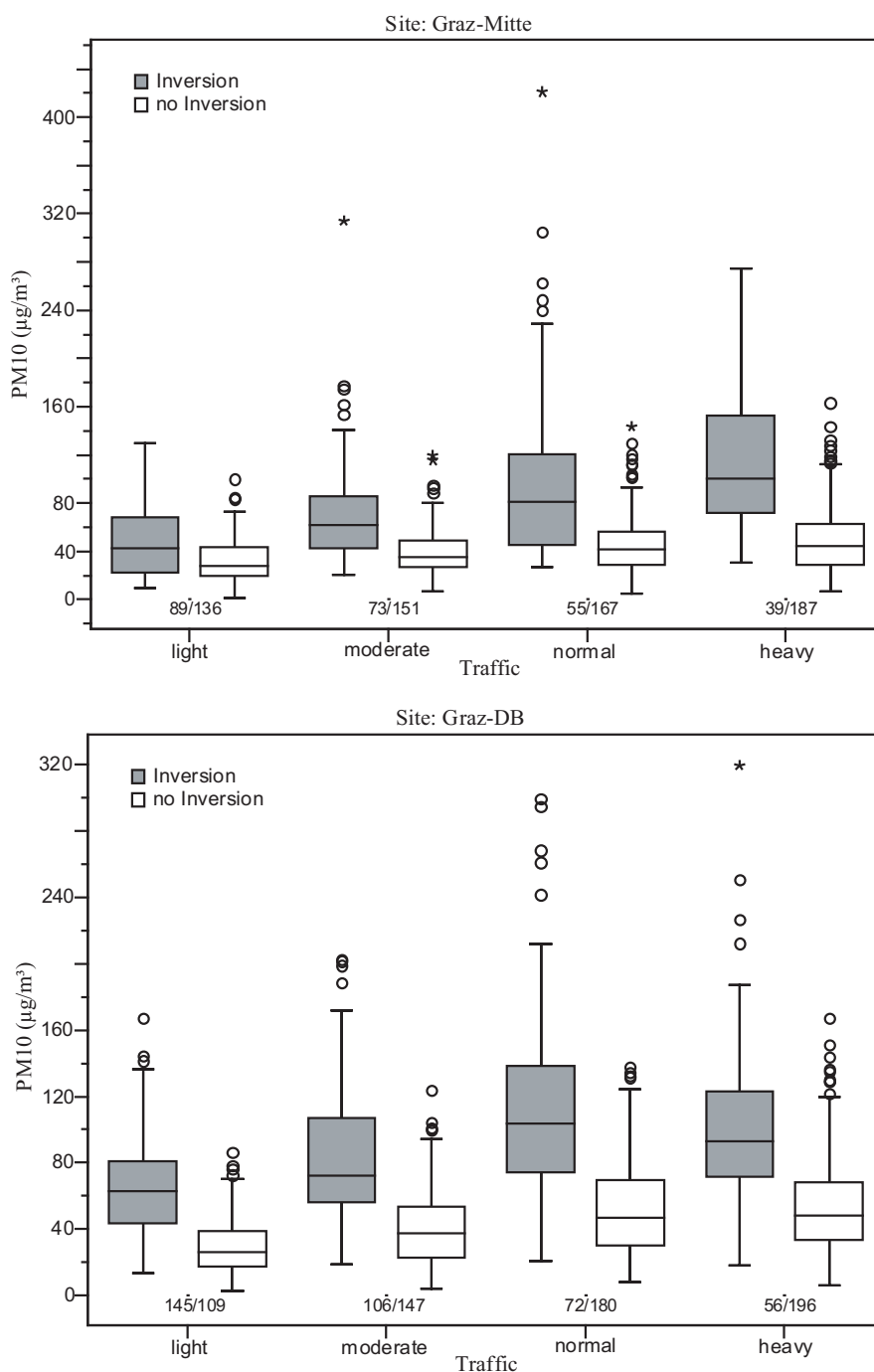
Figure 7: Box plot series of 4h-means of PM10 in four traffic categories, grouped according inversion, at Graz-Mitte and Graz-DB.

cial impact of both factors.

Finally, we consider the combined influences to the PM10 concentrations of (i) wind and traffic frequency, and (ii) precipitation and traffic frequency.

The combination *strong wind/heavy traffic* (median: $34\mu g/m^3$) leads to small loads
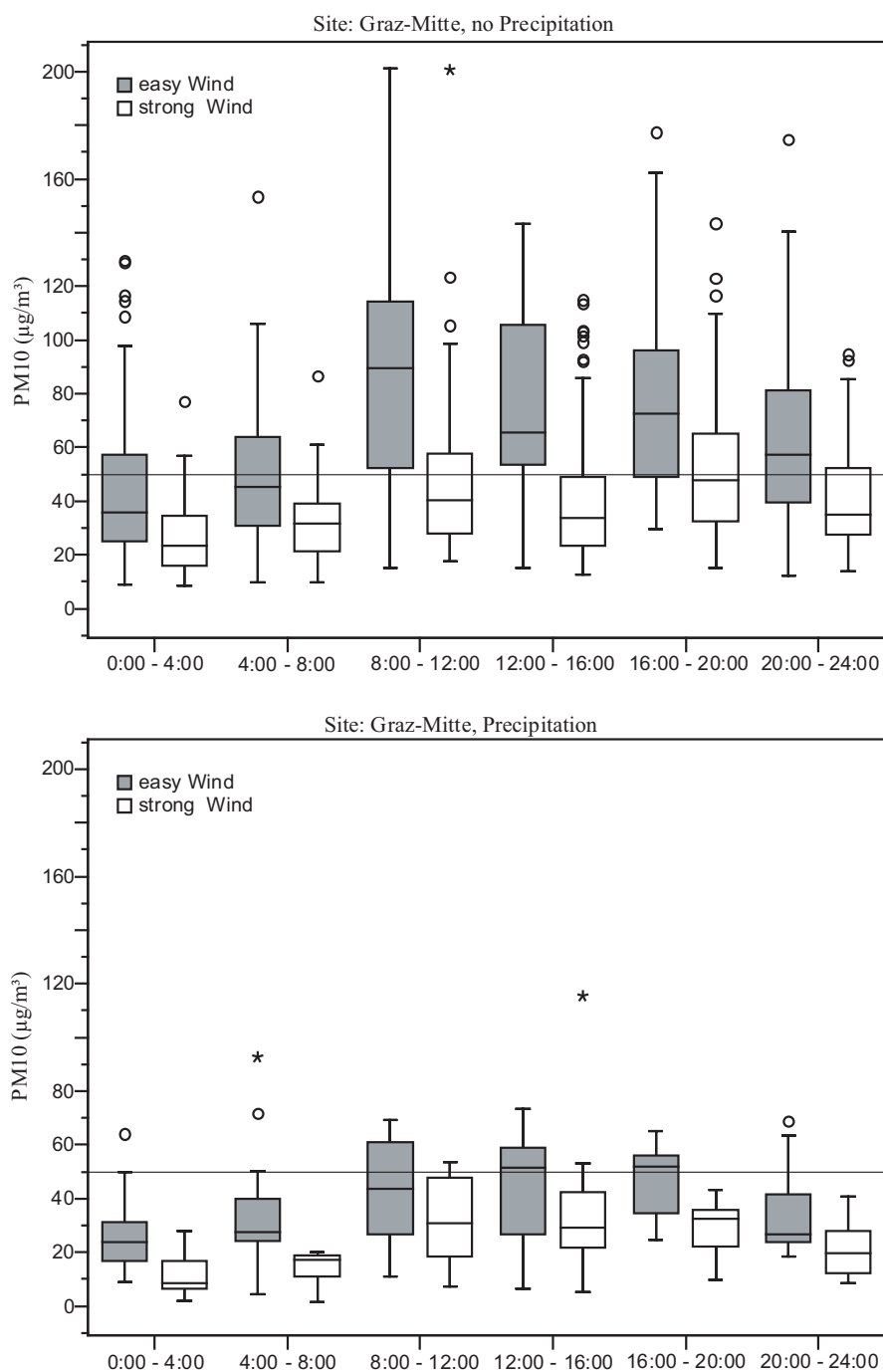
Figure 8: Box plot series of the 4h-means of PM10 in two wind categories, grouped according precipitation, at Graz-Mitte.

comparable with *easy wind/light traffic* (median: $38\mu g/m^3$); *precipitation/heavy traffic* (median: $30\mu g/m^3$) corresponds to *no precipitation/light traffic* (median: $34\mu g/m^3$). This may lead to the interpretation that the influences of wind and precipitation to PM10 are of similar magnitude (see also Table 2, t-values).
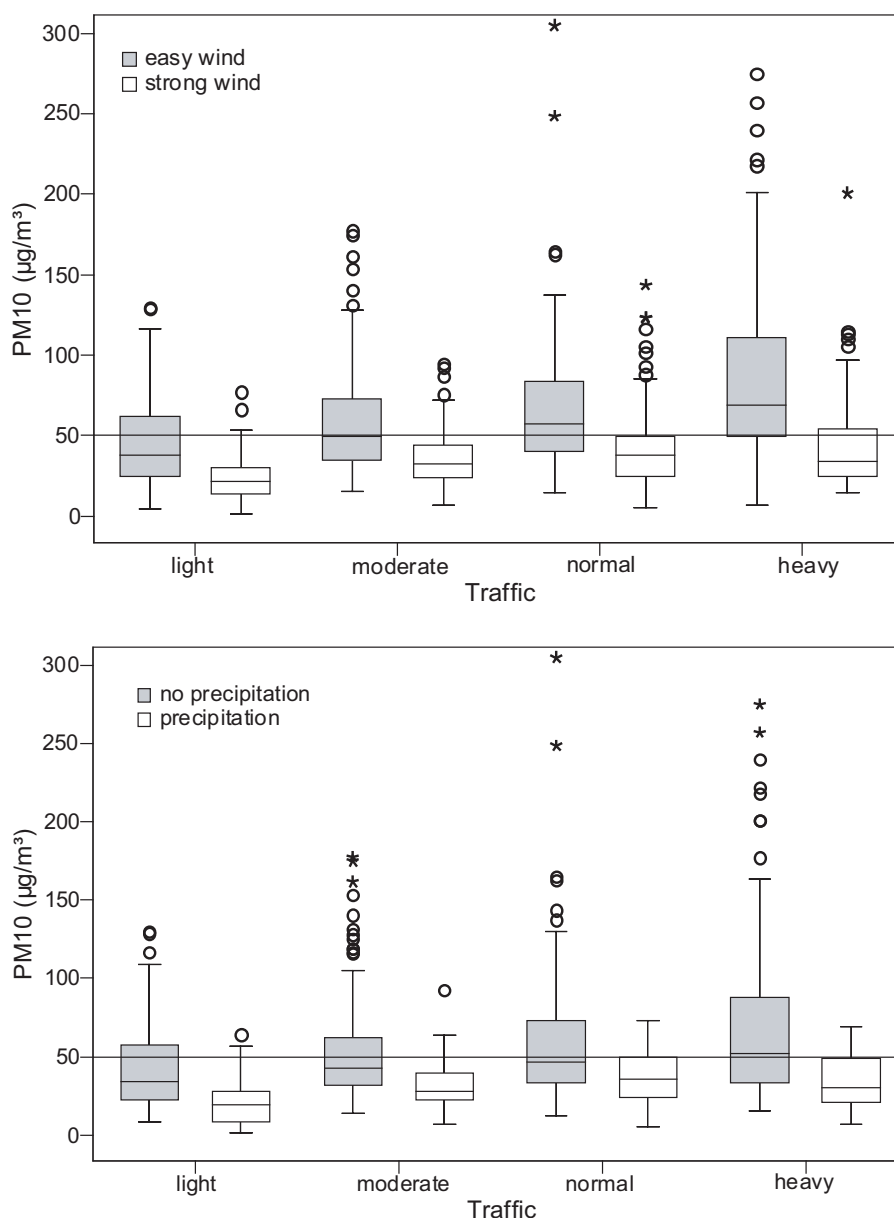
Figure 9: Box plot series of the 4h-means of PM10 in four traffic categories, grouped according wind (above) and precipitation (below), at Graz-Mitte.

4 Prediction Model for the Daily PM10 Concentration at Graz-Mitte

The aim is to predict the PM10 concentration for the following day $x + 1$, based on the information of the current day x (today's PM10 concentration and meteorological forecasts for day $x + 1$). The calculated prediction value should afterwards (*i*) be classified (e.g., low, medium or high concentration) or (*ii*) it should be the base for the calculation of a

probability for exceedance of a certain concentration level. The idea is to make a decision rule available which can be used to implement short-term actions, which may lead to a reduction of the PM10 value. For instance the rule could be expressed as follows:

If the level of $75\mu g/m^3$ is exceeded with a probability of more than 50%, the actions listed in . . . become effective.

4.1 Data Material

As shown in the empirical analysis of Section 3 the PM10 concentration differs from station to station according to the local exposure (traffic area, residential area, green area). Thus a prediction model can only be employed at a specific place. For us the most important gaging station is Graz-Mitte, because it is near the highly frequented pedestrian zones and it reflects to which level of PM10 concentration urban residents are exposed to. Thus, the data material of Graz-Mitte constitutes our basis for the calculation of the prediction model. For this we can refer to observations of the winter seasons 2002/03 (182 days) and 2003/04 (183 days). Measurements of the daily PM10 average are available within the following periods:

1. Winter season 2002/03:
   (178 observations)

   - 1.-13. October
     3 missing values
   - 17. October - 15. November
     1 missing value
   - 17. November -31. March

2. Winter season 2003/04:
   (171 observations)

   - 1. October - 10. March
     12 missing values
   - 23. March - 31. March

For the calculation of the model we need the measurements of all (meteorological) parameters, so the data basis reduced to 157 days in 2002/03 (86%) and to 169 days (92%) in 2003/04.

4.2 A Multiple Linear Regression Model

As was demonstrated in the empirical part of the paper the meteorological factors *inversion, wind* and *precipitation* have a great impact on PM10. While the effect on PM10 of inversion (measured by temperature difference Graz-Kalkleiten) and wind depends on the intensity, the influence of precipitation seems to be nearly independent from the magnitude. Hence our model will contain precipitation as $0/1$ variable only. Other available meteorological variables (e.g., humidity, high inversion, wind direction, etc.) were not included in the model. The reason is twofold. First of all our investigations showed that they do not improve the model essentially since their effects are already described quite well by the included parameters. By virtue of the advantage of a simple and manageable model we included those parameters which have the strongest impact. Second we are forced to include only variables which are realistic to be forecasted by meteorologists. Besides we shall emphasize once more that the purpose of our model is not to give pre-

cise point estimates. It should simply recognize high exceedances of the threshold value or give all-clear if no exceedance takes place.

Human caused effects arise from the different PM10 loads on workdays and holidays. The model will contain a categorical variable DAY with the three categories *workday, Saturday, Sunday/holiday* describing the anthropogenic factor. We should emphasize that because of the close relation of traffic and the variable DAY (see Table 1) no traffic forecasts are needed for our model.

Table 1: Comparison of traffic frequency in three categories and type of day.

	Traffic	low	medium	high	Total
	(cars/day)	(< 15.000)	($15.000 - 18.000$)	(> 18.000)	
Mo-Fr	Count	1	9	65	75
	% within Category	1.3%	12%	86.7%	
Sa	Count	3	13	5	21
	% within Category	14.3%	61.9%	23.8%	
Su/Ho	Count	25	1	0	26
	% within Category	96.2%	3.8%	0%	
Total	Count	29	23	70	122
	% within Category	23.8%	18.9%	57.4%	100%

Furthermore we observe high correlations between the PM10 loads on subsequent days. Thus a prediction model for the PM10 average on day $x + 1$ should contain the PM10 value of the day x. In praxis the predictions for day $x + 1$ should stand by at a certain time tt of the day x (12am say). It is thus necessary to consider the 24h-average PM10 from tt on day $x - 1$ to tt on day x.

Our prediction model is based on a multiple linear regression model for $\sqrt{\text{PM10}}$ and contains the parameters above which are summarized in Table 2 below. The square root transformation is necessary to assure a constant error variance (see Ssection 4.3).

Regression Model *Let p_1, p_2, and p_3 be the observed values of wind, inversion and precipitation (categorical value) on day $x + 1$. Let v_1 be the category of day $x + 1$ and v_2 be the 24h PM10 average from tt o'clock day $x - 1$ to tt o'clock day x. Then we assume that the following regression model for $\sqrt{\text{PM10}}$ holds*

$$\sqrt{\text{PM10}} = \sum_{k=1}^{3} \alpha_k p_k + \sum_{k=1}^{2} \beta_k v_k + \epsilon, \quad \text{where} \quad \epsilon \overset{iid}{\sim} N(0, \sigma^2). \tag{1}$$

The idea of our prediction model is quite simple. In order to predict the PM10 average of day $x + 1$ we use instead of the observed values of p_1, p_2, and p_3 their meteorological forecasts. The values of v_1 and v_2 are already available on day x.

Prediction Model *Let $\hat{p}_1, \hat{p}_2$, and $\hat{p}_3$ be the forecasts of wind, inversion and precipitation (categorical value) for day $x + 1$. Let v_1 be the category of day $x + 1$ and v_2 be the 24h PM10 average from tt o'clock day $x - 1$ to tt o'clock day x. Then the prediction of PM10 for day $x + 1$ is given by*

$$\hat{\text{PM10}} = \left(\sum_{k=1}^{3} a_k \hat{p}_k + \sum_{k=1}^{2} b_k v_k \right)^2, \tag{2}$$

where the constants a_k ($k = 1, 2, 3$) and b_k ($k = 1, 2$) are the estimates of α_k and β_k in regression model (1).

We should note that our predictions are slightly biased but this is negligible for our purpose.

Table 2: Parameters in the regression model.

var.	(LABEL)	description	unit	remarks
$\hat{p}_1$	(WIND)	average wind velocity	m/s	prediction for day $x + 1$
$\hat{p}_2$	(TDGK)	temperature difference between Graz and Kalkleiten	C°	prediction for day $x + 1$
v_1	(DAY)	Mo-Fr, Saturday, Sun/Holiday	$1 = $ Mo-Fr $2 = $ Sa $3 = $ Su/Ho	categorical value on day $x + 1$
$\hat{p}_3$	(PREC01)	precipitation	$0 = $ no pc $1 = $ pc	prediction for day $x + 1$
v_2	(MEANtt)	PM10 average of the last 24 hours, calculated at tt o'clock	$\mu g/m^3$	measured value on day x

4.3   Parameter Estimation and Statistical Properties

Our aim is to estimate the parameters of the regression model (1) and to verify the underlying normality assumptions. For the estimation of the parameters we use the statistic package R. As data basis we used the observations of the two winter seasons 2002/03 and 2003/04.

```
             Estimate  Std. Error  t value   Pr(>|t|)
 (Intercept)   9.0478      0.2697    33.55     0.0000
 WIND         -1.6447      0.2434    -6.76     0.0000
 TDGK         -0.3211      0.0330    -9.73     0.0000
 DAY          -0.8112      0.0766   -10.59     0.0000
 PREC01       -0.8323      0.1335    -6.23     0.0000
 MEAN12        0.0186      0.0022     8.49     0.0000

Adjusted R-squared: 0.6695;   Residual standard error: 1.078
```

The regression model seems to fit quite well. All variables have a significant impact on the response $\sqrt{\text{PM10}}$ and the model explains 67% of the variance. In order to check the assumptions of the regression model we have to analyze the residuals. Referring to Figure 10 we see that there is no indication of a violation of our assumptions.

4.4   Validation of the model

By applying the prediction model the main focus is not to get a nearly precise PM10 value, but rather a value whose deviations from the observed values are within a narrow range. Needless to say that there is a natural error which cannot be described by the model. To get
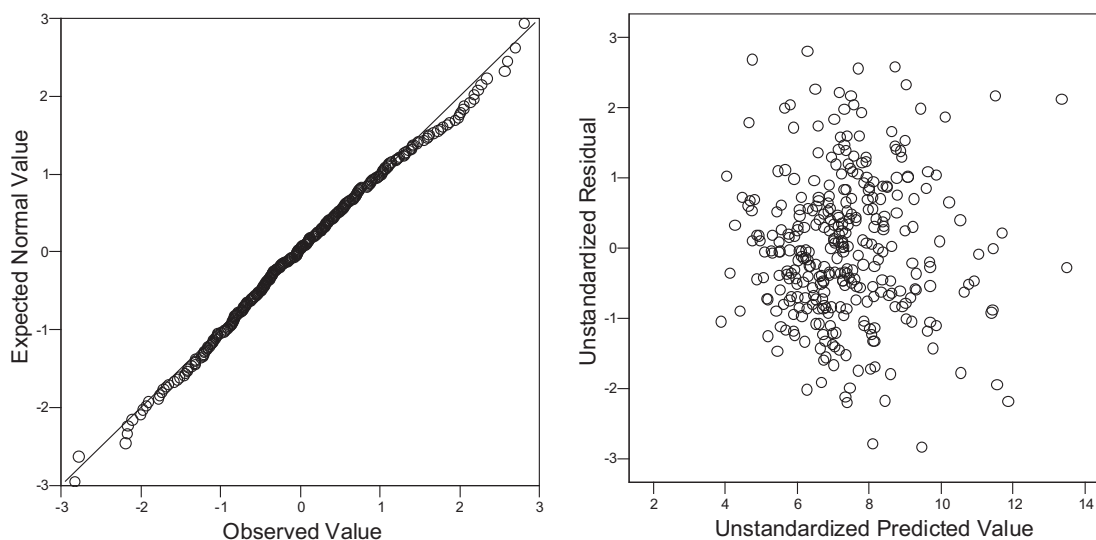
Figure 10: Q-Q-Plot of residuals vs. normal quantiles and scatter-plot of predicted values vs. residuals.

an impression for the order of deviation of the prediction from the observation we list the percentiles of the differences in Table 3. In about 50% of the cases the prediction differs from the observation less than $10\mu g/m^3$. In 80% the difference is less than $20\mu g/m^3$.

Table 3: Percentiles of observation–predictions; both seasons.

			Percentile			
5	10	25	50	75	90	95
-23.9	-16.7	-9.8	-0.8	9.4	23.9	32.1

An important feature additional to the fitting properties is the stability of the model against seasonal trends. Can we expect that our predictions have a similar quality when applied to data from a future winter season? In reality we have the problem that we work with meteorological forecasts, which cause an additional error. But at the first step let us assume that the forecasts are precise. What can we expect about the mean or variability of the errors? In order to answer this question we calculate two prediction models by means of the same method as above. The first model is based on the data of 2002/03 (prediction model 1) and the second one on the data of 2003/04 (prediction model 2). Then we compare the models in the following way. We 'predict' the PM10 averages of the winter season 2003/04 by prediction models 1 and 2 and analyze their differences (Figure 11). Apart from a slight shift (predictions of model 1 are about $5\mu g/m^3$ to high), the predictions vary not that much (the interquartile range of the differences is $5.9\mu g/m^3$). Such a shift may arise from some differences in the basic conditions for the two seasons: nearby road works, more abrasive material due to unfavorable meteorological conditions, etc.
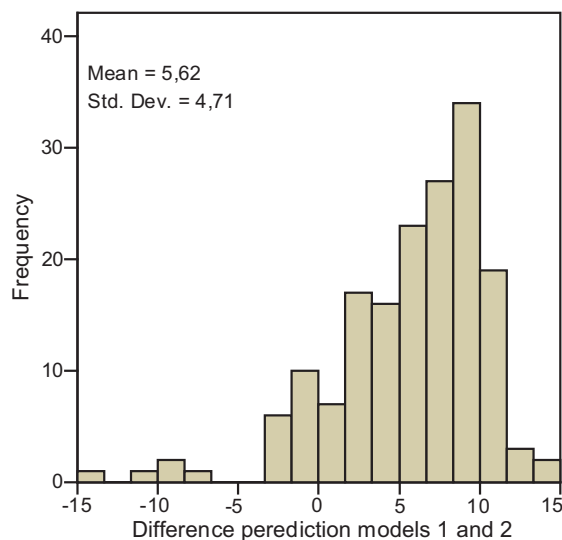
Figure 11: Prediction model 1–prediction model 2 for the winter season 2003/04.

Comparing the observations of the winter season 2003/04 with the predictions of prediction model 1 (based on data from 2002/03) we get also quite reasonable results. In about 80% the error of the prediction is less than $20\mu g/m^3$ (see Table 4).

Table 4: Percentiles of observation–predictions via prediction model 1; season 2003/04.

			Percentile			
5	10	25	50	75	90	95
-27.0	-22.7	-15.6	-5.8	3.3	17.6	26.2

4.5 Decision Rules

The prediction model aims at least at two applications. The first task is its use as monitoring and information tool. Common people should be able to look at an online prediction of the `PM10` concentration of the subsequent day. Such an information tool is already available in the internet (see `http://www.feinstaubfrei.at/htm/ampel_f.htm`) where a simple traffic light system may be observed. If the prediction is less than $50\mu g/m^3$ a green light is displayed. Values from 50-100$\mu g/m^3$ are indicated with orange and predictions of more than $100\mu g/m^3$ are signified with red light. (Since green light may be misconceived as an incitation for taking the car, perhaps in future in this case no light will be displayed).

The second application aims at a simple decision rule for short term actions. It is planned to force people to avoid unnecessary driving or even to initiate sporadic traffic stops. In this application the prediction becomes a very sensitive topic. In the sequel we discuss a possible decision rule based on our model and show its consequences.

Decision Rule:

• *If the prediction is less than $75\mu g/m^3$ no actions are implemented.*

• *If the prediction is higher than $75\mu g/m^3$ actions have to be taken.*

As can be seen in Table 5 the model predicted 14 times a value higher than $75\mu g/m^3$, but the observed value was below this limit (error rate of 5.7%). In three cases the model predicted less than $75\mu g/m^3$ although the observation was even higher than $100\mu g/m^3$ (error rate of 10.0%). As might be expected the difficulties become visible when the PM10 value is between $75\mu g/m^3$ and $100\mu g/m^3$. For instance a prediction of $70\mu g/m^3$ might be good if the observation is $80\mu g/m^3$. But in this case our rule would lead to a false classification.

Table 5: Comparison of PM10 in three categories and the prediction in two categories.

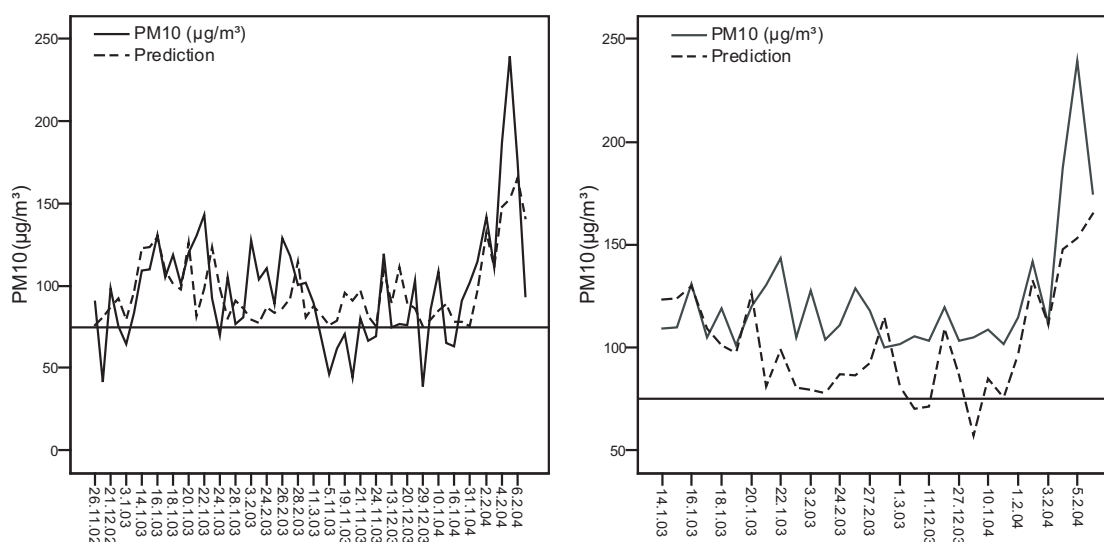|	Prediction	$<75\mu g/m^3$	$\geq 75\mu g/m^3$	Total
PM10 $\leq 75\mu g/m^3$	Count	233	14	247
	% within Category	94.3%	5.7%	
$75\mu g/m^3 <$ PM10 $\leq 100\mu g/m^3$	Count	33	15	48
	% within Category	68.8%	31.2%	
PM10 $\geq 100\mu g/m^3$	Count	3	27	30
	% within Category	10.0%	90.0%	
Total	Count	269	56	325
	% within Category	82.8%	17.2%	100%

Figure 12: In the left graph we selected days where the prediction was more than $75\mu g/m^3$. The right graph shows PM10 and the corresponding prediction on days where PM10 exceeds $100\mu g/m^3$.

However, Figure 12 shows that nearly on every day where the observed `PM10` value exceeded $100\mu g/m^3$ (actions are strongly indicated) the prediction was at least higher than $75\mu g/m^3$, hence the right decision was proposed. On the other hand, for prediction values higher than $75\mu g/m^3$ the observed `PM10` values were at least close to $50\mu g/m^3$ or beyond the threshold value. Again, in this case the decision rule is also reliable.

4.6 Test Run

From December 16, 2004 to April 15, 2005 the prediction model has been tested in a real life environment. In the course of the morning the ZAMG Steiermark (contact person Dr. Sudy) provided forecasts for *inversion* (temperature difference Graz-Kalkleiten), *wind* and *precipitation*. The 24h-mean `PM10` at 12am was picked from the online server of the Styrian government. Finally all these basic data were recorded in an excel sheet from which the corresponding `PM10` predictions were calculated.

During the test run we did some 'fine tuning' and made modifications if necessary. Within our test period we had the Christmas holidays with their atypical volume of traffic. This might be one possible explanation for the larger prediction errors of the daily `PM10` averages during this time. We mention also the extremely high `PM10` load on the New Years day. Obviously, this stems from the fireworks in the night where the half-hour average increased from $75\mu g/m^3$ at 11:30pm, 31.12.2004, to nearly $350\mu g/m^3$ during the hour after midnight. This is an example how sensitive `PM10` values can react in case of special events.

In addition, the wind forecasts posed some problems. Since the gaging station is in a build-up area, on many days the forecasts were considerably higher than the average wind velocities measured. Hence, at the beginning of February we changed to categorical predictions of the average wind velocity. The meteorologist on duty had to give us an estimation on a scale from 1 to 4 (1 = no wind, 4 = strong wind). We replaced his estimation by the mean value of the wind velocities in the corresponding quartile range of the empirical wind distribution. With this the prediction of wind velocity was improved, but some potential for further improvements still remains. On the other hand, the forecasts of inversion and precipitation were suitable for our purposes (see Table 6 and Figure 13).

Table 6: Comparison of forecasted and measured precipitation.

		(Measured Value)	no prec.	prec.	Total
Forecast	no prec.	Count	70	10	80
	prec.	Count	13	18	31
Total			83	28	111

According to our decision rule we made only seven errors with undesirable effects (Table 7). In one out of 121 days we predicted less than $75\,\mu g/m^3$ though the measured value was higher than $100\mu g/m^3$. Six times we predicted more than $75\mu g/m^3$, although the observed value was below this limit. However, Figure 14 shows, that in these cases the corresponding `PM10` values were still beyond the threshold value of $50\mu g/m^3$.
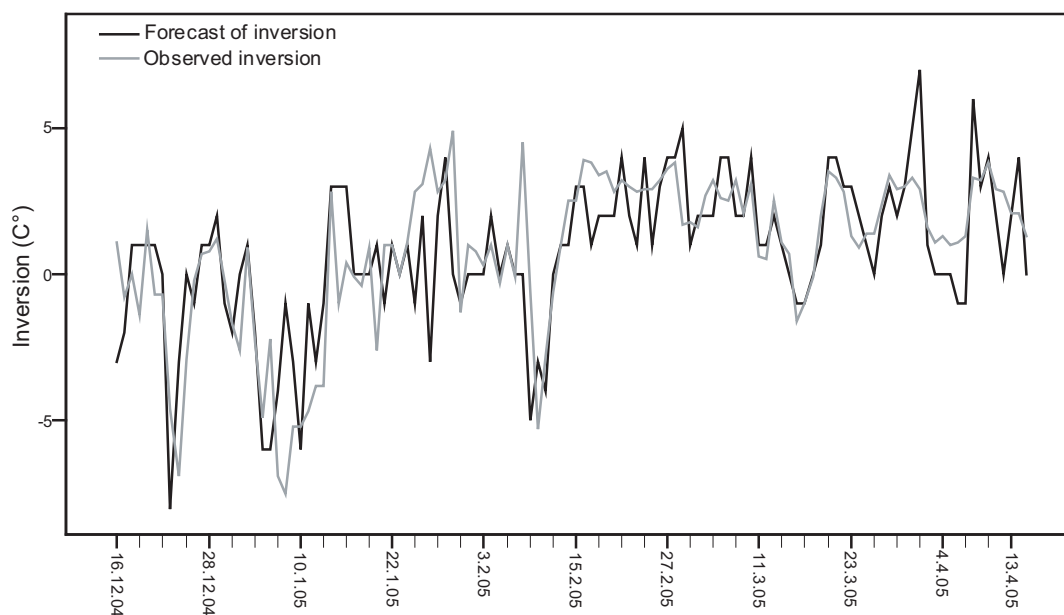
Figure 13: Time series representation of forecasted and observed inversion.

Table 7: Comparison of `PM10` in three categories and the forecasts in two categories.

	Prediction	$<75\mu g/m^3$	$\geq 75\mu g/m^3$	Total
PM10 $\leq 75\mu g/m^3$	Count	87	6	93
	% within Category	93.5%	6.5%	
$75\mu g/m^3 <$ PM10 $\leq 100\mu g/m^3$	Count	13	9	22
	% within Category	59.1%	40.9%	
PM10 $\geq 100\mu g/m^3$	Count	1	5	6
	% within Category	16.7%	83.3%	
Total	Count	101	20	121
	% within Category	83.5%	16.5%	100%

5 Conclusions

Our empirical analysis shows that the `PM10` concentration in Graz is highly influenced by three meteorological factors *inversion, precipitation, wind* and by human impacts like *traffic, industry, households*. The human influence is described by a categorical variable defining the type of day (working day, Saturday, Sun/holyday). First of all our comparison of working days with Sun/holydays demonstrates that the decrease of `PM10` is strongly associated with the decrease of the traffic frequencies. But what about the other human factors, factory fumes and domestic fuel? It can be assumed that domestic fuel remains nearly constant over the weekend and it is well known that in Graz factories contribute less than 5% to the `PM10` emissions. This leads us to the conclusion that the variable *type of day* to a great extent corresponds to the factor *traffic*.
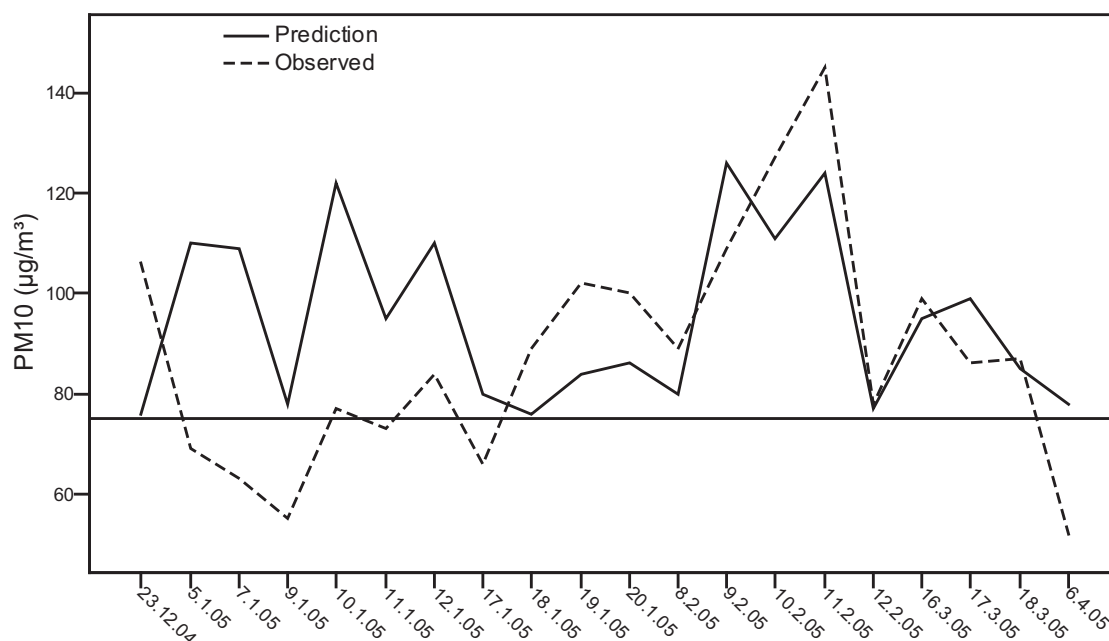
Figure 14: Days when the prediction was higher than $75\mu g/m^3$.

Our prediction model for the average `PM10` concentration of the subsequent day includes three meteorological forecasts (inversion as temperature difference, wind velocity, precipitation as $0/1$ variable), type of day and the 24h mean of `PM10` until 12am of the current day. This model has been tested for 121 days (December 16, 2004 to April 15, 2005) and we obtained suitable predictions except in cases of (i) inaccurate forecasts of the wind velocity, (ii) atypical periods (Christmas holidays) and (iii) special events (Silvester day).

References

Hörmann, S., und Stadlober, E. (2004, December). *Modellierung und Prognose der Feinstaubkonzentration in Graz-Mitte für den Zeitraum Herbst/Winter* (Tech. Rep.). Institut für Statistik, TU-Graz. `http://www.stat.tugraz.at/archiv/feinstaub_files/publications.htm`. (Endbericht im Auftrag der Stadt Graz)

Immissionsschutzgesetz – Luft (IG-L). (1997). Bundesgesetzblatt I, Nr. 115/1997, Novelle 2001: Grenzwerte für PM10.

Kropf, R. (2001). *Massnahmen zur Reduktion der PM10-Emissionen* (Tech. Rep.). Bundesamt für Umwelt, Wald und Landschaft (BUWAL). (`http://www.umwelt-schweiz.ch/imperia/md/content/luft/fachgebiet/d/grundlagen/um_136_d.pdf`)

Podesser, A., Wölfelmaier, F., und Rieder, H. (2004, April). *Kombinierte Feinstaub-Meteorologiesondierung am 17.03.2004 in Graz* (Tech. Rep.). ZAMG

Regionalstelle Steiermark. (`http://www.innovation.steiermark.at/cms/dokumente/10023151/6fe2dbba/Feinstaub_Ballonsondierungen.pdf`)

Stadlober, E., und Pfeiler, B. (2004a, April). *Explorative Analyse der Feinstaub-Konzentrationen von Mai 2003 bis Dezember 2003* (Tech. Rep.). Institut für Statistik, TU-Graz. `http://www.stat.tugraz.at/archiv/feinstaub_files/publications.htm`. (Endbericht im Auftrag der Stadt Graz)

Stadlober, E., und Pfeiler, B. (2004b, November). *Explorative Analyse der Feinstaub-Konzentrationen von Oktober 2003 bis März 2004* (Tech. Rep.). Institut für Statistik, TU-Graz. `http://www.stat.tugraz.at/archiv/feinstaub_files/publications.htm`. (Endbericht im Auftrag der Stadt Graz)

Umweltschutz Ärztinnen und Ärzte für. (2003). *Feinpartikel (particulate matter pm).* Mediendokumentation, 2003. (`http://www.umwelt-schweiz.ch/imperia/md/content/luft/fachgebiet/d/grundlagen/pm10_aefu_2003.pdf`)

Authors' address:

Siegfried Hörmann
E-mail: `shoermann@tugraz.at`

Brigitte Pfeiler
E-mail: `bpfeiler@gmx.at`

Ernst Stadlober
E-mail: `e.stadlober@tugraz.at`

Institute of Statistics
Graz University of Technology
Steyrergasse 17/IV
8010 Graz