

Correcting air pollution concentrations for meteorological conditions; an extended regression-tree approach

Authors: H. Noordijk, H. Visser

Abstract

Regression-tree analysis is a rather unknown method in environmental sciences. The method is examined for its use in time series analysis of air pollution data. The regression-tree approach is used to define meteorological classes which describe the influences of meteorology on air pollutant concentrations. This description is used to remove the influence of meteorological fluctuations from the concentrations. Some pitfalls of the method and their solutions are discussed. The technique is illustrated on time series of SO₂, NO_x and PM₁₀ over the Netherlands. In these cases, original and corrected trends are nearly identical, except that year-to-year fluctuations are removed.

Key words index

Multiple-regression analysis, trend analysis, SO₂, NO_x and PM₁₀.

1 Introduction.

When analysing long-term trends in air pollution components, one has to take into account three major causes of changes. In general, the main reason to analyse these trends is to detect changes in human pressure on the atmospheric environment. Concentrations may then be linked to emissions, either entirely anthropogenic or from both human and natural sources. Besides changes in emissions, changing weather conditions and differences in measurement techniques may be reflected by the detected trends. A reliable method which removes influences of changing weather conditions will clarify the relations between concentrations and emissions. Also, it helps to identify artefacts due to changes in measurement conditions, as it removes weather-driven variations which may hide those artefacts in the data.

Over the last decades, a variety of procedures have been used to remove weather induced noise from air-pollution data. Traditionally, multiple-regression models dominate this field. This domination can be understood, as this technique is widely known and easy to use. However, multiple-regression techniques have some notorious disadvantages. The investigator has to make *a priori* assumptions on the shape of the relations between pollutant concentrations and weather parameters, as well as assumptions on interdependencies between parameters. The technique does not allow circular relations as wind direction. Also, it does not allow discrete changes such as limit values above which effects start to arise. Therefore, the results obtained with these techniques also depend on the knowledge and choices of the investigator and on the degree up to which reality fits the basic rules underlying multiple-regression descriptions. Extreme values are often a problem in multiple-regression, as regression coefficients are modeled to limit the overall variability in the data.

Recently, neural networks have been applied frequently to correct trends for influences of meteorology. With this method, the dependence on *a priori* knowledge and choices of the investigator is minimal. Neural networks are not hindered by the previously mentioned disadvantages of multiple-regression. A serious disadvantage of neural networks, however, is that it works as a black box. One cannot verify the plausibility of the underlying relations, as they are entirely hidden.

In this study, a different method is used: regression-tree analysis (Breiman *et al.*, 1984). This method combines the transparency of multiple-regression models with the flexibility in relation modelling of neural networks. The handling of extreme values is superior to that of multiple-regression. Regression-tree analysis is in the medical field widely known as a useful statistical modelling tool. Its use for environmental trend analysis has been very limited. For air-pollution trends, the specific power and pitfalls of the methodology have not been well documented. In the vision of the authors, trend analysis by regression trees has preference over multiple-regression techniques or neural networks. Several pitfalls of this method and options for improvement will be discussed. Some results of the method on PM₁₀, SO₂ and NO_x will illustrate the power of the method.

2 Methodological requirements.

In order to remove weather induced noise from the data, we think that the correction method should meet the following criteria:

- The method must correct only influences of meteorology and missing data. The trends induced by changing emissions or by changes in measuring methods should not undergo artificial changes by the correction method.
- The method must be objective, which means that the researcher has little choices left which may influence the correction procedure and the resulting trends.
- The method must be robust. Its results has to be independent of complications in the data, such as missing concentrations, missing meteorological data or extreme values.
- The method must be relatively simply applicable for a wide range of situations and substances.
- The method must be transparent. One has to be able to prove the reliability and scientific plausibility of the results in different stages of the method.
- Finally, the derived models should be parsimonious, explaining the air pollution data with a small number of meteo-parameters. These models should be consistent with the data and with prior knowledge.

In the following, a regression-tree approach will be described which meets these requirements. First, a general overview will be presented, followed by extensions to the standard methodology.

3 Meteo-correction through Regression Tree analysis: available methods.

The method can be subdivided in two steps. The first step is to classify the meteorological influences on the data by regression-tree analysis. This classification, specific for each location and compound, will be used to define a "standard meteorological year". This standard year is based on the expected normal frequency of the meteorological conditions which cause variations in air pollutants. The second step is to calculate weight factors per year and per class by use of the derived meteorological classes and the defined standard meteorological year. With these weight factors, statistical indicators such as yearly averages, percentiles or exceedances above a threshold can be corrected for meteorology.

3.1 Regression-tree analysis.

Tree-based modelling is a statistical exploratory technique for uncovering structures in the data. Outside the environmental field, it is increasingly used for devising prediction rules, screening variables and assessing the adequacy of linear models. In our application, the technique explains concentrations of an air pollutant by the use of a number of meteorological variables. The technique results in a number of "leaves", in our case meteorological classes, for which the variance in the concentrations due to meteorology is minimized. When modelling the day-to-day variations, these concentrations can be daily averages, daily maxima, etc. Meteorological variables can be the daily average, the minimum or maximum of the temperature, wind speed and wind direction, relative humidity, etc.

In the tree-modelling procedure, the concentration dataset is first divided into two subsets; two meteorological classes. The technique splits the data according to a specific value of one meteorological variable. This is done in such a way that the variation in the concentration data in the two subsets is minimised. Next, the procedure splits one of the two subsets, again in order to minimise variation in the three resulting subsets. This splitting technique is repeated again and again, until the size of the growing regression tree (Figure 1) meets limiting values. The first separation of the data in Figure 1 is at a daily mean temperature of 2.62 °C. The ellipses are subsets which are further split up, and rectangles denote the final meteorological classes, the 'leaves' of the tree.

As with multiple-regression, the tree model explains the data better as more parameters are included. To meet the requirement that the model should be parsimonious, a cost-complexity measure is added (Breiman *et al.*, 1984). The more complex the regression tree, the higher the value of the cost function. The model is then optimised, by minimisation of the variation, including the cost-complexity measure. In this way, a number of less important subsets are excluded from the model; the regression tree is "pruned".

Regression tree analysis may be hindered by the presence of trends in the data. Furthermore, the sampling time may influence its outcomes. These aspects have led to a few extensions to the method, which will be discussed.

3.2 Correction for meteorological conditions

After defining the meteorological classes, the measured air pollutants can be corrected for influences of meteorology. Stoeckenius (1991) derived meteo-corrected trends by calculating averages or exceedance probabilities per meteo-class, multiplying them by the frequency of occurrence of these meteo-classes. Thus, meteo-corrected annual averages or annual percentiles can be calculated.

The method followed by Stoeckenius has a few drawbacks. First, it is limited to the correction of annual averages and percentiles. Other interesting values, as the number of days in excess of a standard, cannot be corrected in his approach. Further, meteo-classes are sometimes “empty”. An empty class is present when for the meteorological conditions of that class no measurements are available within a specific year. In the case of correcting annual averages, these empty classes are not really of concern, as they are usually classes with rather few measurements under specific conditions. However, when correcting more extreme values, for example the excess number of days or hours above a standard, such “empty classes” may have a considerable influence, especially when the class defines those conditions which cause exceedances of the standard.

Here we will present 2 new approaches to derive meteo-corrected trends after regression tree analysis.

3.3 Validation.

The regression-tree method, as implemented in our software, is tested on several artificial concentration series. An example of an artificially derived series is the situation with a constant concentration of 50 µg/m³ for a daily averaged wind direction between 5 and 55 degrees, and a constant background concentration of 20 µg/m³ for all other directions. In all tests, the method produced the expected results.

A cross-validation of the derived regression trees is executed before pruning them. An important aspect here is the stability of an estimated tree. This stability can be low when the tree contains too many explanatory variables or if two or more explanatory variables are highly correlated (multicollinearity). The best way of validation and pruning includes the use of an independent set of data for validation. The software splits the data in ten parts. We use nine parts to grow the tree and the tenth part to validate it (Breiman *et al.*, 1984). This is repeated in ten different ways, so that every subset is used for cross-validation. The final tree is the average of the 10 cross-validated trees.

The next step in the validation procedure is the comparison with other statistical models. As competing models, the overall mean predictor μ , the simple persistence model (the prediction for tomorrow is the measurement of today) and multiple-regression are used.

Breiman *et al.* (1984) defines the following *mean squared error* $R_{RT}^*(d)$ for a Regression Tree (RT) with prediction rule d :

$$R_{RT}^*(d) = \frac{1}{N} \sum_{i=1}^N [y_i - d(x_i)]^2 \quad (8)$$

The asterisk is used to denote the fact that R_{RT} is *estimated* on the data. We now define the performance of our regression-tree model, relative to that of the rival model μ , as:

$$RE_{\mu}^*(d) = \frac{R_{RT}^*(d)}{R_{\mu}^*} \quad (9)$$

The same formula is used for the other rival models. The performance can also be presented as a percentage, following:

$$P_{\mu}^*(d) = [1 - RE_{\mu}^*(d)] * 100 \quad (\%) \quad (10)$$

4 Extensions to standard regression-tree analysis

4.1 Trend removal.

The regression-tree method, as introduced by Stoeckenius (1991), has one severe pitfall. If there is a real trend in the data, either downward or upward, it will influence the regression tree in such a way that the meteo-corrected trend will be artificially leveled off. The reason for this is, that the method itself does not distinguish between "older" and "newer" measurements in the series. The method needs time-independent data with a constant variation over time.

Figure 2 shows a bias in Regression-Tree predictions, when compared to measurements with a strong, downward trend. These measurements are monthly averaged concentrations of SO₂ from the station Posterholt of the Dutch National Air Quality Monitoring Network (Van Elzakker 2001). If there is a strong downward trend, the method tends also to put the earlier high measurements in separate classes as that will reduce the variability in the total data set. Those measurements are concentrated in separate meteorological classes which define a false standard meteorological year, biased to the real meteorology causing variations in concentrations. The meteo-correction procedure, now based on shifted classes, tends to spread out the influence of early high measurements over all the years. In the same way, the entire concentration series will be influenced by low measurements at the end of the period. In other words, the method will tend to diminish an existing trend.

A solution for this problem is to remove the trend in the concentrations *before* analysing the data with the regression-tree method. In principle, one could subtract or divide the measurements by the estimated trend by:

$$y'_t = y_t - \text{trend}_t \quad (1)$$

or

$$y'_t = y_t / \text{trend}_t \quad (2)$$

To divide by the trend has an extra advantage. In the period with higher concentrations, the variability in the data is also higher. Differences in data variability over the entire period are normalised when one divides by the trend. For this reason, we have applied formula (2). The trend is assessed by an n-degree polynomial

$$y_t = a_0 + a_1t + \dots + a_nt^n \quad (3)$$

with a maximal value of 3 for n.

As the trend is removed from the data, the average class concentrations in the subsets are dimensionless data, varying around 1.0 (Figure 3). This tree gives a true image of the meteorological influences on the concentrations. As expected, there are clear differences between this tree and the tree based on concentrations without trend removal (Figure 1). Both trees start with a first division by a temperature of nearly 3⁰C. Thereafter, the trend-corrected tree of Figure 3 makes again subdivisions according to temperature. Regression-tree analysis on the original concentrations, however, divides the data in the next steps by precipitation and Pasquill class. Further subdivisions on lower levels are entirely different between both trees. Figure 4 presents the Regression-Tree predictions after trend-removal. Now the predictions fit the data very well.

When the meteorological classes are properly defined by the use of trend-removed concentrations, we return to the original concentration data to correct them for meteorology. In this way, trend-correction is only used to derive a realistic division in meteo-classes. As the correction itself is performed on the original concentration data, the trend removal cannot lead to artificially induced disturbances of the meteo-corrected trends.

Figure 5 shows the effect of trend-removal on the meteo-corrected trends of SO₂ for Posterholt. As expected, the meteo-corrected trend based on the simple regression tree without trend removal is smoothed down when compared to the original measurements. The fluctuations caused by meteorology are removed. The meteo-corrected trend, based on trend-removed concentrations, shows exactly the same downward line as the original data, which is in line with

emission data. Also for the period 1980-1992 there is good agreement between emission reductions and concentrations, corrected for meteorology (Van Jaarsveld 1995).

4.2 Influence of sampling time.

A serious point of concern is the sampling time according to which air pollution measurements and meteorological data are used. The causes of variation in the data depend on the order of the time-step. One can distinguish between variation caused by the rhythm of day and night, variation by the accidental changes in weather patterns and variation by seasonal effects. For meteorology-correcting air pollution trends, variation due to diurnal and seasonal patterns is not the subject of this study as these patterns are supposed to be more or less constant over the years. If one chooses a time-step of one hour, the total variation in the data is caused by all the factors, whereas a time-step of one day excludes the variation caused by diurnal patterns. The averaging over days will result in a better trend-correction for meteorology, as the regression-tree method will be more focused on the causes of day-to-day variation instead of dissipating its power over less relevant diurnal factors.

A time-step of one day is in general better than smaller time-steps, as diurnal patterns introduce a large amount of variability which can be avoided by a proper choice of one daily indicator. However, there is no distinct preference for time-steps larger than one day. In our study, we focus on the difference between days and months for temporal resolution. In both cases, the variation will be caused by accidental and seasonal weather patterns. In general, the focus on larger time-scales will also be a focus on larger spatial scales as weather variations on small time-scales tend to be limited to smaller spatial scales. However, the accidental weather patterns contained in monthly data also tend to approximate seasonal patterns.

The choice for larger time-scales will be illustrated by the meteorology-correction of PM_{10} -concentrations. Measurement and modelling of fine dust particles is complicated, as these particles originate from a variety of extremely different sources and substances. The concentration of PM_{10} on a certain location and time in the Netherlands is the result of mainly North West European emissions of fine dust particles, SO_2 , NO_x , ammonia and the natural contribution of sea salt and crustal material, as well as their transport and the individual deposition and chemical conversion characteristics of each of these compounds and their subsequent oxidation products. The transport, deposition and conversion of these compounds and their decomposition products are all influenced by the meteorology over the entire North West European region. How well can a method, based on local Dutch meteorology, describe the meteorological influences on such a complicated substance?

The answer is: not very well. The first exercises on rural PM_{10} -concentrations could explain only 35% of the variation in the data. For this poor explanation, a large number of meteorological variables and classes are necessary, sometimes more than 20 classes. Generally, variables selected are wind speed, wind direction, the amount and duration of rainfall, temperature, air pressure, relative humidity and the amount of radiation. Already the first variable by which a tree is split up, differs widely between different PM_{10} -monitoring stations. As expected, local meteorology can only poorly explain the variability in PM_{10} concentrations. Local concentrations are the consequences of a scatter of causes on a much wider spatial scale. A model which can explain only one-third of the variation, is of little value for removing meteorological variations from the trend.

On the basis of monthly values, however, 70% of the variation is explained by the regression-tree method. The tree is limited to only four or five meteorological classes, the small number of 12 concentration values per year does not allow any further details. The main meteorological factors are rainfall, temperature and wind direction. These factors are interdependent, as rainfall in the Netherlands comes with westerly winds and temperatures tend to be more extreme with continental (easterly) winds. The better performance of the method may be caused by a closer relation between monthly values for local meteorology and meteorological behaviour on larger spatial scales which drives the formation, transport and removal of PM_{10} in the atmosphere. One should, however, bear in mind that the variation in daily data of PM_{10} is much higher than the variation in monthly data, so that the reduction in variation between the two series is not directly comparable. Also, by use of temperature the method tends to model the seasonal behaviour, which is not the focus of this study.

5 Extensions to the meteo-correction method.

We derived 2 meteo-correction procedures, each different from the method of Stoeckenius (1991).

Method 1

The first method avoids the artifacts and limitations which have been mentioned for Stoeckenius' method. When the tree with meteorological classes is available, weight factors per year and per meteo-class W_{kj} are derived according to:

$$W_{kj} = \frac{N_k}{n_{jk}} \quad (4)$$

with

N_k = long-term annual mean number of days with meteorology in class k

n_{jk} = actual number of days with meteorology in class k for year j

This definition has the advantage that, instead of only calculating a meteo-corrected annual mean concentration, a variety of meteorological corrected statistics can now be derived. Here we present three different quantities.

The meteo-corrected annual mean concentration C_{av} :

$$C_{av} = \frac{\sum (C_i \times W_{kj})}{\sum W_{kj}} \quad (5)$$

with C_i the average concentration at time i.

The meteo-corrected percentile values (here P_{98}):

$$P_{98} = C_i \text{ at } \sum W_{kj} = 98\% \quad (6)$$

The meteo-corrected number n of days above a standard:

$$n = \sum W_{kj} \text{ with concentration above the standard}$$

In fact, the only difference with normal concentration statistics is that normal statistics are calculated with $W_{kj} = 1$ for each day.

A pitfall in the general procedure is the occurrence of years with "empty" meteorological classes. In these cases, a specific meteorological class did not occur during the year, or the concentration measurements during that meteorology are absent. A correct procedure needs a complete concentration distribution per year, including all the meteorological classes. Empty classes occur more often when they contain less data over the years, which may be the case for important but very specific meteorological conditions. An example of this is the occurrence of extreme low temperatures with unusual high SO_2 concentrations. In many winters, however, these extreme low temperatures do not occur in the Netherlands. Another cause of empty meteorological classes is a large time-scale for averaging. For example, if the modelling is based on monthly averages, only twelve measurements per year are available. Already with a small number of meteorological classes there will be years with empty classes.

Our approach, when empty classes can not be avoided, is to fill them with artificially derived data. Measurements in the same meteorological class in the years before and after an empty class are averaged. These class averages are then linearly interpolated. This interpolated class average is then multiplied by correction factors depending on the average number of measurements and their standard deviation in his class over all the years. The result is a set of artificial data for the empty class, with a mean value derived from interpolation and with a standard deviation equivalent to the value over other years for that class.

This correction may be important for statistical endpoints which focus on larger time-steps or extreme values, such as 98 percentiles or the number of exceedances above a threshold value. However, in the case of annual averages, derived from data on a daily scale, this correction for

empty classes hardly affects the results.

Method 2

Here we will present only trends based on annual averaged concentrations. For this simple statistical endpoint, an easier method is developed which does not take effects by empty classes into account. The method can, however, also be applied for other values, such as percentiles or excess values.

The approach estimates daily or monthly meteo-corrected concentrations. The prediction for a day t by the regression tree model is given by \hat{y}_t , which is the mean concentration within the meteorological class of that day. The effect of meteorology for that day is then quantified by the difference between \hat{y}_t and μ_y , where μ_y is the mean concentration over all measurements. Now, the meteo-corrected concentration $y_{\text{corr},t}$ on a day is given by the difference between the measured concentration y_t and the effect of meteorology:

$$Y_{\text{corr},t} = Y_t - (\hat{Y}_t - \mu_y) \quad (7)$$

For those time series which need trend-removal prior to regression tree analysis, the trend-removed concentrations are corrected with trend-removed effects of meteorology, based on the difference of trend-removed values of \hat{y}_t with μ_y . Finally, the meteo-corrected concentrations are derived by applying the converted trend-removal procedure on the trend-removed meteo-corrected data.

This method is validated for a number of individual measurement locations by comparison with results from the reference method 1. In our examples, only very limited differences between the outcomes of both methods are observed. Even in the case of PM_{10} , with only 12 monthly values per year, corrections of both methods differ no more than a few percent.

6 Applications.

6.1 Strong trend: SO_2 (Figure 6)

Prior to the analysis of 22 rural SO_2 -sites, the strong downward trend was removed from the data, covering the years 1989-2001. As a first step, two rural stations regression trees are grown both on daily and monthly data. For both time-steps, temperature appeared to be the main explanatory variable. As in many European regions, Dutch SO_2 concentrations peak at unusual low temperatures. The regression-tree approach was able to explain 46 and 39% of the variation, based on daily data. For monthly data, however, 80 and 65% of the variance was explained. For the meteo-corrected trend-removed annual averages, 40% of the variance was removed by analysis on a daily basis, and more than 60% was removed on a monthly basis.

As the performance for monthly data was clearly better, the remaining 19 Dutch rural stations were only analysed on a monthly basis. The explained variance ranged from 59 to 80%, on average it was 68%. Next to temperature, wind direction and the amount of rain are often recurring explanatory variables. As expected, the relatively high peak in 1996 was caused by meteorology. This winter was unusually cold. The measured downward trend over the entire period was not influenced by meteorology, as the meteo-corrected trend falls in the same way. The trend in the concentrations matches well with the trend in northwest European emissions. The presented emissions are weighted for their contribution to Dutch concentrations before adding them up.

6.2 Complex substance: PM_{10} (Figure 7)

As discussed before under "influences of sampling time", rural concentrations of PM_{10} depend on many factors over a much wider area than the Netherlands. Maybe due to that reason, analysis based on daily values failed, as for SO_2 .

On the basis of monthly averages over the period 1992 to 2001, approximately 70% of the variation was explained as these data tend to reflect processes on larger spatial scales. As expected, the relatively high values in the years 1996 and 1997 are due to the cold winters of these years. Rainfall, temperature and wind direction mainly influence PM_{10} behaviour. The meteo-corrected trend shows a smoother decline than the original concentrations, except for the first years. The overall reduction in measured PM_{10} -concentrations over these years is after

meteo-correction not significantly different. The reduction in the concentrations is rather limited when compared to the reduction in northwest European emissions, weighted for their contribution to Dutch concentrations and corrected for natural emissions (sea salt and soil dust).

6.3 Local effects: NO_x (Figure 8)

As an example of correction for local conditions, the trend in NO_x -concentrations at 11 traffic sites of the Dutch monitoring network will be discussed. In order to focus on local emissions and weather conditions, daily averaged NO_x -concentrations measured at rural monitoring sites nearby are subtracted from the street measurements.

On average, 54% of the variation is explained by the regression trees, ranging from 37% up to 75%. For all sites, wind direction is an important variable. As expected, the selected wind directions bear clear relations to the direction of the street and the influence of local buildings. In central parts of the Netherlands, the wind velocity is the second important meteorological parameter. In the southern sites, temperature is the second parameter. Meteo-corrected concentrations and measurements, averaged over all the traffic sites, differ only slightly from each other as the measured trend already shows very little ups and downs due to meteorology (Figure 8). The emissions of NO_x by city traffic show a less smooth trend. In the last period, reported emissions are fairly constant whereas the concentrations remain declining. This obvious discrepancy may be caused by the methodology underlying emission data and the fact that concentrations are measured in larger cities, whereas the emissions refer to all Dutch cities and villages.

7 Conclusions.

The meteo-correction procedure proposed here, corrects for both meteorological variations and missing data. The method is able to derive concentration trends which could be measured if standard meteorological conditions would occur with 100 percent monitoring data availability. Such corrected trends aid the interpretation of monitoring data and the comparison with emissions.

For SO_2 , the steep downward trend in the data is retained in the corrected data, where the effects of exceptional meteorological conditions are removed. This trend is in agreement with model calculations and emissions. The large variability in PM_{10} -data, due to meteorology, is reduced. The downward trend for PM_{10} over the years is, however, not influenced by meteorology. In the case of NO_x , the method confirms the continuing decline in measured concentrations, due to local traffic, despite the stagnation in reported emission data.

References

- Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J., 1984. Classification and regression trees. Wadsworth, Belmont, California.
- Elzakker, B.G. van, 2001. Monitoring activities in the Dutch National Air Quality Monitoring Network in 2000 and 2001. RIVM report 723101055. Bilthoven, the Netherlands.
- Jaarsveld, J.A. van, 1995. Modelling the long-term atmospheric behaviour of pollutants on various spatial scales. Thesis, University of Utrecht, the Netherlands.
- Stoeckenius, T., 1991. Adjustment for ozone trends for meteorological variations. In Tropospheric Ozone and the Environment 3. Papers from an international speciality conference, Berglund, R.L., Lawson, D.R. and Mckee, D.J. (ed.), Air & Waste Management Association, Pittsburgh PA.
- Visser, H. and Noordijk, H., 2002. Correcting air pollution time series for meteorological variability. With an application to regional PM_{10} -concentrations. RIVM report 722601007/2002. Bilthoven, the Netherlands.

Acknowledgements

The authors want to thank Laurens Brandes (RIVM) for providing the emission data, presented in Figures 6 to 8

Figures

Figure 1. Regression tree of SO₂ for station Posterholt, based on trend-removed monthly-averaged concentrations. Temperature is expressed in °C, precipitation in mm, radiation in W/m², Pasquillbyday is Pasquill stability by day, winddirSector6 is the fraction of time with wind directions in sector 6 (300 - 360 (N) degrees).

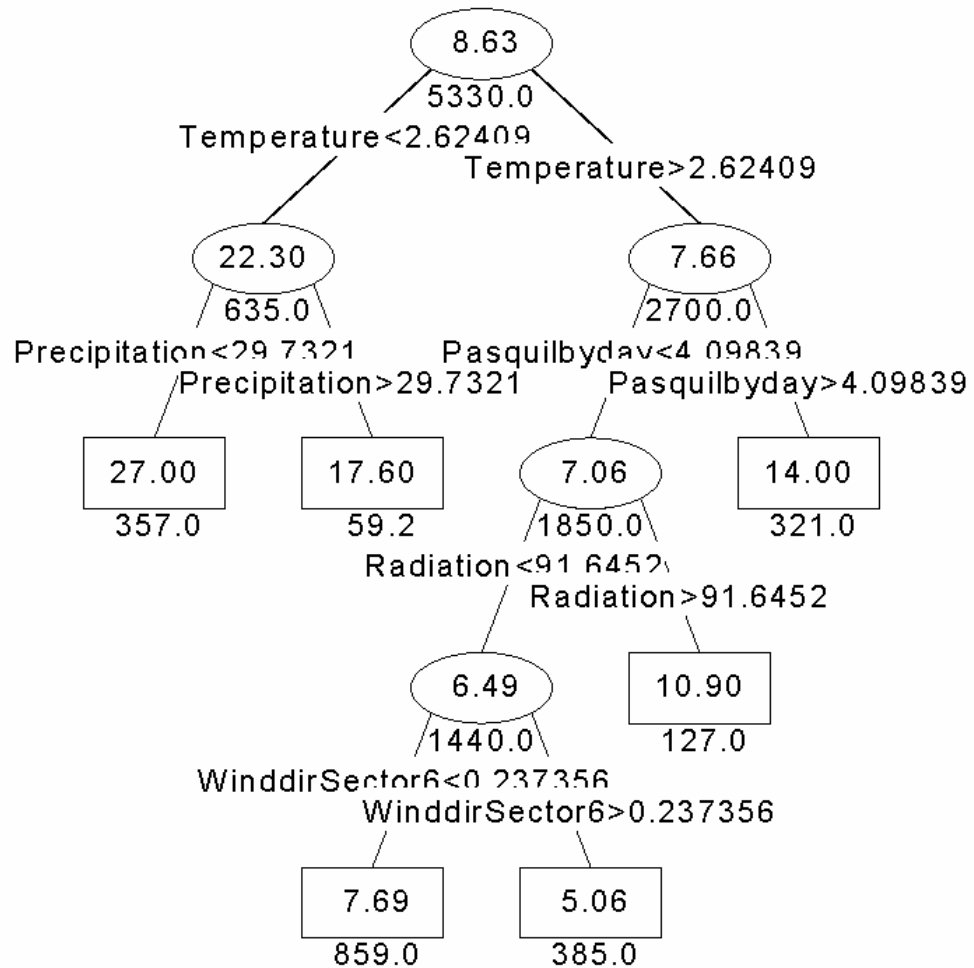


Figure 2. Monthly concentrations ($\mu\text{g}/\text{m}^3$) of SO_2 for station Posterholt, no trend-removal applied. Measurements versus Regression-Tree predictions.

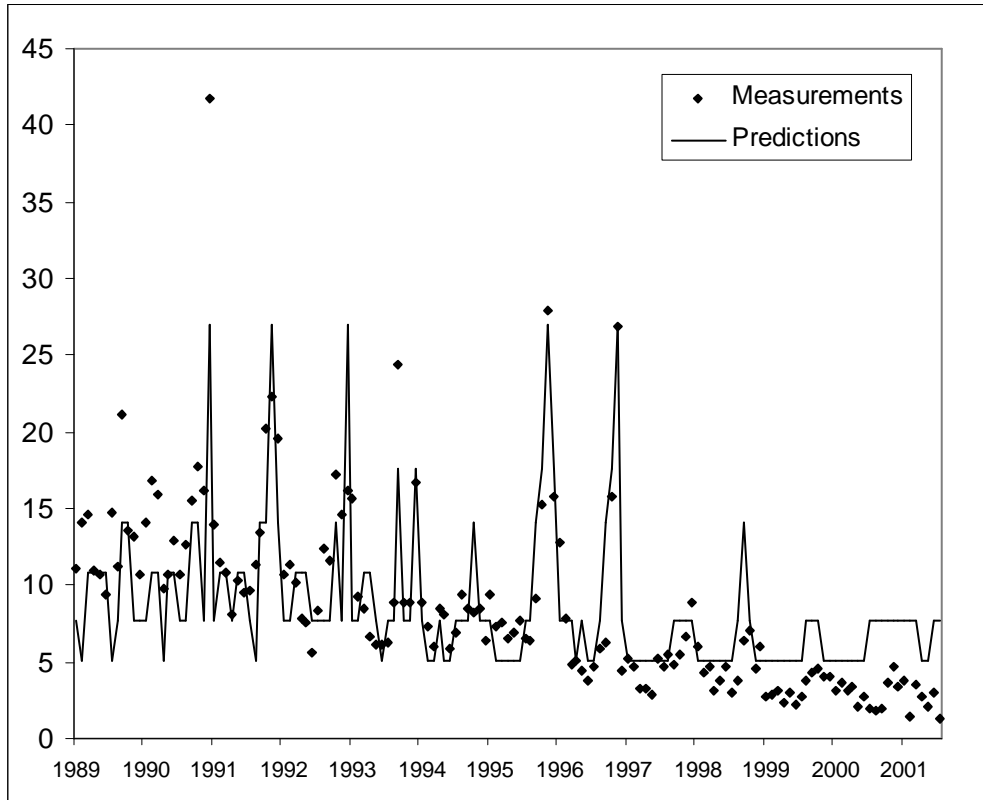


Figure 3. Regression tree of SO₂ for station Posterholt, based on trend-removed monthly-averaged concentrations. Temperature is expressed in °C, precipitation in mm, wind speed in m/sec.

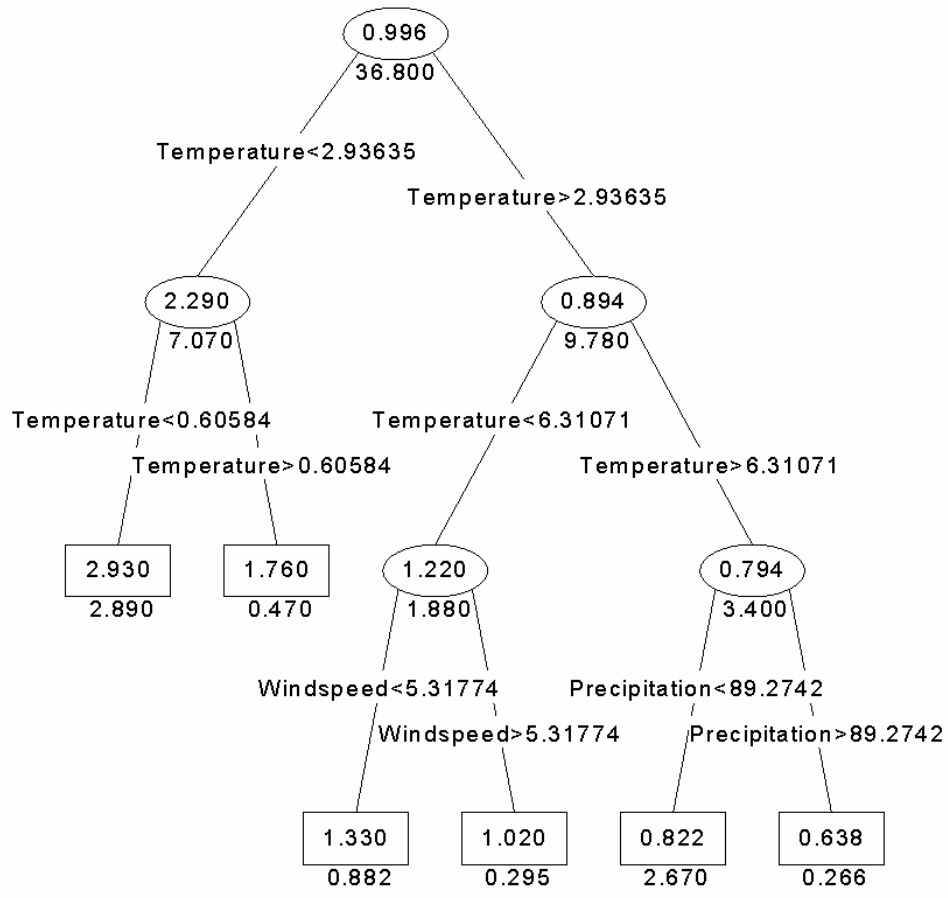


Figure 4. Monthly concentrations ($\mu\text{g}/\text{m}^3$) of SO_2 for station Posterholt, trend-removal applied. Measurements versus Regression-Tree predictions.

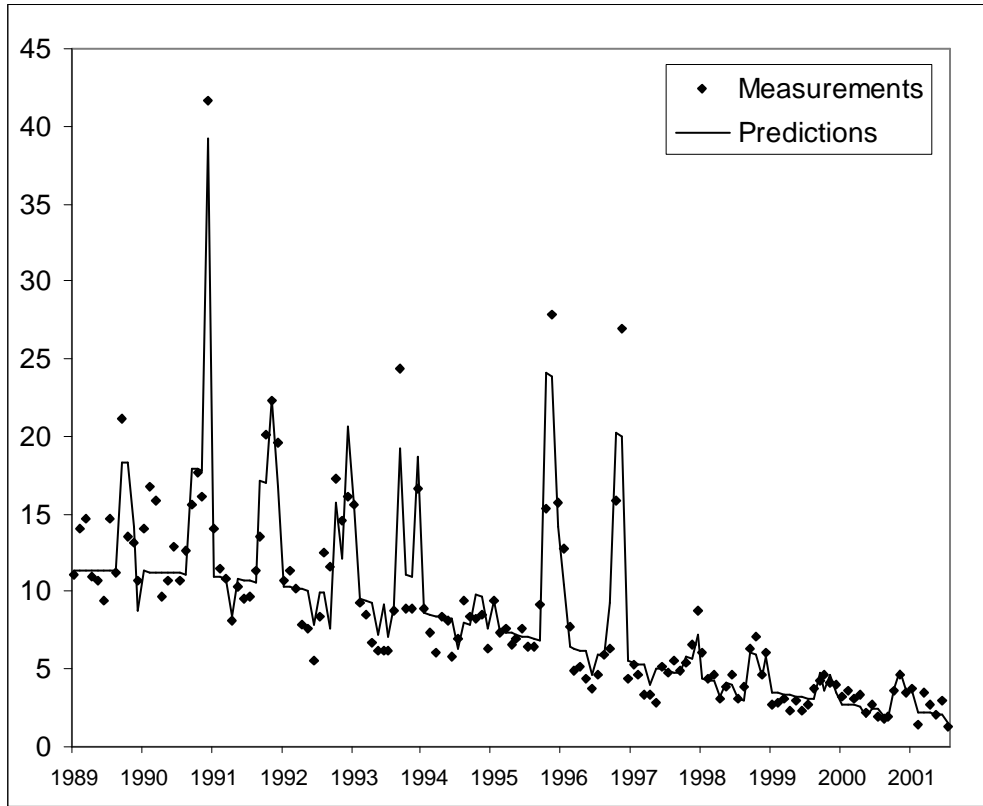


Figure 5. Annual averaged concentrations ($\mu\text{g}/\text{m}^3$) of SO_2 for station Posterholt. \diamond Measured concentrations, \circ meteo-correction based on original concentrations, and \blacksquare meteo-correction based on trend-removed concentrations.

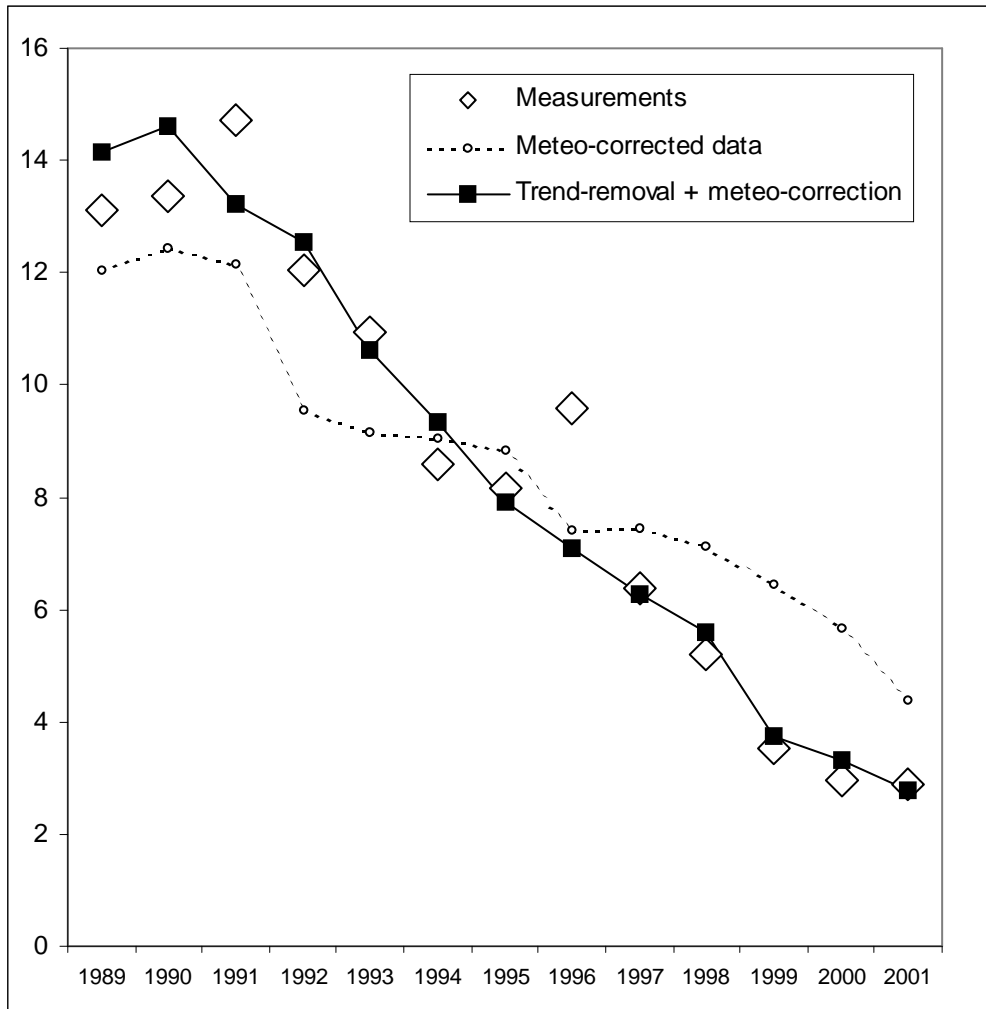


Figure 6. Annual rural concentrations of SO₂, averaged over 22 stations. ◇ Measured concentrations (µg/m³), ■ meteo-corrected data (µg/m³), and Δ weighted and indexed emissions of northwest Europe.

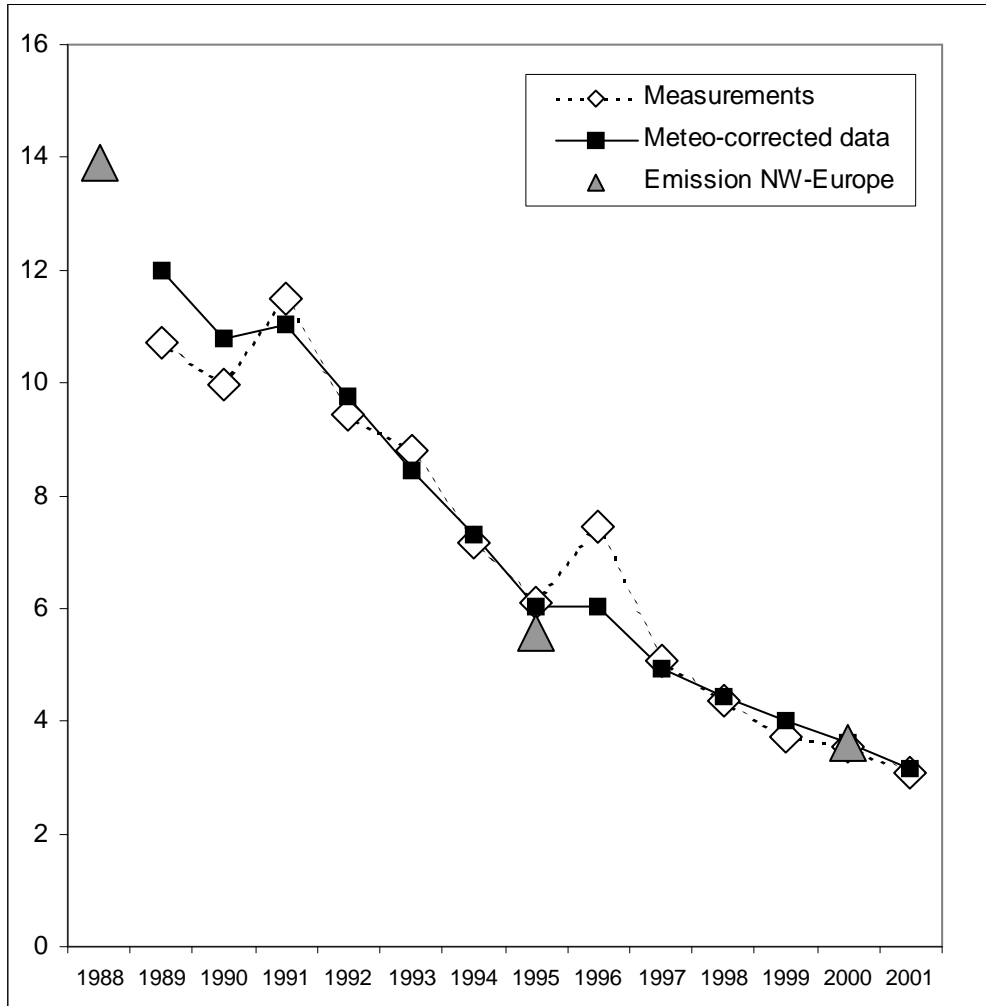


Figure 7. Annual rural concentrations of PM₁₀, averaged over 9 stations. \diamond Measured concentrations ($\mu\text{g}/\text{m}^3$), \blacksquare meteo-corrected data ($\mu\text{g}/\text{m}^3$), Δ weighted and indexed emissions of northwest Europe.

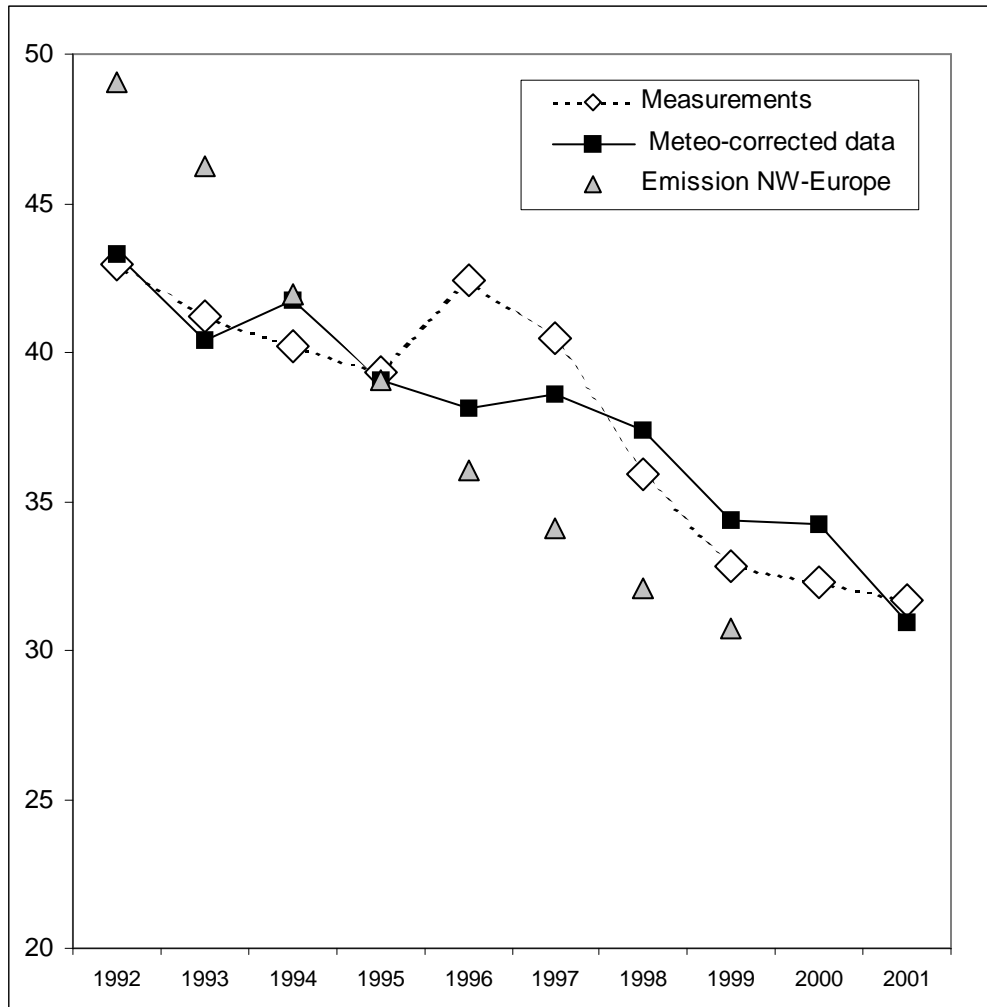


Figure 8. Annual traffic-related concentrations of NO_x , averaged over 11 stations. \diamond Measured concentrations ($\mu\text{g}/\text{m}^3$), \blacksquare meteo-corrected data ($\mu\text{g}/\text{m}^3$), \triangle indexed urban traffic emissions.

