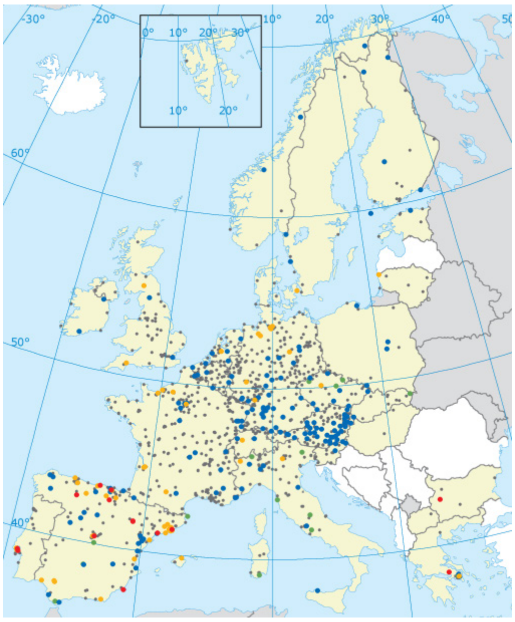
Discounting the impact of meteorology to the ozone concentration trends



ETC/ACM Technical Paper 2015/9
March 2015

Sverre Solberg, Augustin Colette, Cristina Guerreiro

European Topic Centre
on Air Pollution and
Climate Change Mitigation

The European Topic Centre on Air Pollution and Climate Change Mitigation (ETC/ACM)
is a consortium of European institutes under contract of the European Environment Agency
RIVM Aether CHMI CSIC EMISIA INERIS NILU ÖKO-Institut ÖKO-Recherche PBL UAB UBA-V VITO 4Sfera

Front page picture:
Annual changes in annual mean of the maximum daily 8-hour O_3 concentrations in the period 2001–2010 (Source: EEA report 4/2012)

Author affiliation:
Sverre Solberg, Cristina Guerreiro, Norwegian Institute for Air Research (NILU), Norway
Augustin Colette, Institut National de l'Environnement Industriel et des Risques (INERIS), France

DISCLAIMER

This ETC/ACM Technical Paper has not been subjected to European Environment Agency (EEA) member country review. It does not represent the formal views of the EEA.

© ETC/ACM, 2015.
ETC/ACM Technical Paper 2015/09
European Topic Centre on Air Pollution and Climate Change Mitigation
PO Box 1
3720 BA Bilthoven
The Netherlands
Phone +31 30 2748562
Fax +31 30 2744433
Email etcacm@rivm.nl
Website http://acm.eionet.europa.eu/

Contents

1 Introduction

Several questions have been raised regarding ozone trends and their underlying reasons; Is there a discrepancy between the observed long-term development in European ozone levels and the development in ozone ambient air concentrations expected from the regulations of precursors emissions? Alternatively, is the effect of the emission abatement policies simply masked by the large year-to-year variability in ozone? Or, is the hemispheric baseline level of ozone rising and compensating the reduction due to European emission reductions?

At present, these questions seem somewhat unsettled and the topic has raised controversy in the scientific community and among policy makers. The 2013 air quality report from EEA states that "*…there is a discrepancy between the past reductions in emissions of O_3 precursor gases in Europe and the change in observed average O_3 concentrations in Europe"* (EEA, 2013).

Measurements indicate that in certain regions of Europe the decline in ozone is less than expected based on the reported emission reductions (Colette et al., 2011; Solberg et al., 2009; Jonson et al., 2006). One possible reason for this discrepancy is that the changes in ozone caused by the drop in European precursor emissions are masked by the influence from inter-annual variations in meteorology. The level of ozone is the result of a complex interaction of photochemical and meteorological processes which are interconnected, and various studies have tried to separate and quantify the influence of meteorological variation on the trends in ozone. This "detrending" of the meteorological impact is the topic of the present report. A number of other possible reasons for such a discrepancy between observed ozone and precursors trends exist: 1) changes in the hemispheric baseline ozone levels transported into Europe; 2) changes in the ozone chemistry due to changes in the ratio of NOx: VOC emissions; or 3) additional precursor emissions not taken into account in the standard assessment of ozone such as forest fires event; etc.

A wide range of European ozone trend studies has been published the last years and the next chapters give an overview and categorization of these studies. The meteorological impact on the ozone trends is either not considered in these papers, or it is handled through various types of screening of data or by comparison with chemical transport models (CTMs).

Due to the complex interrelations between ozone, meteorology, emissions and chemistry there is no simple way to apply a meteorological detrending of ozone. Strictly speaking, it is not really possible to separate the effects of meteorology from all other processes, although certain methods have shown to give valuable results. A comprehensive review of statistical methods for so-called "meteorological adjustments of tropospheric ozone" has been provided by Thompson et al. (2001).

The review by Thompson et al. (2001) is extensive and thorough, albeit limited to North American studies and based on links between ozone and local (*in-situ*) meteorological parameters. A major weakness with such methods is the inherent assumptions of *in-situ* relationships between meteorological data (e.g. temperature) and ozone. Since photochemical formation of ozone is occurring over the timescale of hours to days, the ozone measured at a monitoring station is the net result of temperature (and all other parameters) experienced by the travelling air mass the last one – five days. Thus, the statistical links between temperature and ozone at a single point in time and space is scientifically not very robust, although found to be statistically significant in many studies.

As noted by Parrish et al. (2012), a major problem in ozone trend analyses is deriving the long-term change in the presence of short-term (inter-annual and shorter) variability that is much larger than long-term changes. Further, the studies presented in the scientific literature differ in the time period covered, and in the degree to which varying regional influences may obscure baseline trends.

The basic question from the ozone abatement policy side is: "what kind of effects on ozone has the NOx and VOC emission regulations lead to?"

This raises a number of other questions, such as

1. *Do we see any long-term changes in ozone?*
2. *Do these changes agree with what we expect?*
3. *Can we separate the effects of reduced precursor emissions from other effects?*

Each of these questions in turn raises new issues, like:

1. *Do we see any long-term changes in ozone?*
    - *In which ozone parameter (mean/max/other)?*
    - *At which stations/countries/regions?*
2. *Do the results agree with what we expect?*
    - *What do we expect?*
        - *What do CTMs and surrogate models used to inform air quality regulation predict?*
        - *Is it possible to identify empirical relationships?*
    - *If they don't agree, why?*
        - *Systematic errors in models/understanding?*
        - *Measurement data not representative or of insufficient quality?*
        - *Incorrect emission data?*
        - *Systematic differences in agreement between regions?*
        - *Masked by the year-to-year variation in weather?*
3. *Can we separate the effect of reduced precursor emissions from other effects?*
    - *How?*
    - *Do the various methods agree?*

Thus, the simple question above raises issues on all aspects of atmospheric ozone, from instrumentation, QA/QC of measurement data, siting conditions, monitoring history, data screening and selection of indicators to the application of more or less advanced statistical tools and chemical transport models.

A conceptual view (a "flow chart") of an ozone trend assessment is given in (Figure 1). On the left side is the CTM and measurement data, respectively. Often in the literature the modelled and measured trends are compared directly without any screening or filtering of the data and the estimated trends typically don't agree very well. As a refinement to this simple approach, one can apply various kinds of screening of the data, like QA/QC-inspection of the monitoring data, screening by region or meteorological regime, etc, furthermore the data could be normalised in some way and the seasonal amplitude could be removed etc. In addition, the selection of ozone index (hourly values, monthly means, 8-h maximum, etc) could be essential for the result, and finally all kinds of statistical tools for trend calculations could be applied, ranging from the most basic linear approach to advanced methods like generalized additive models etc.

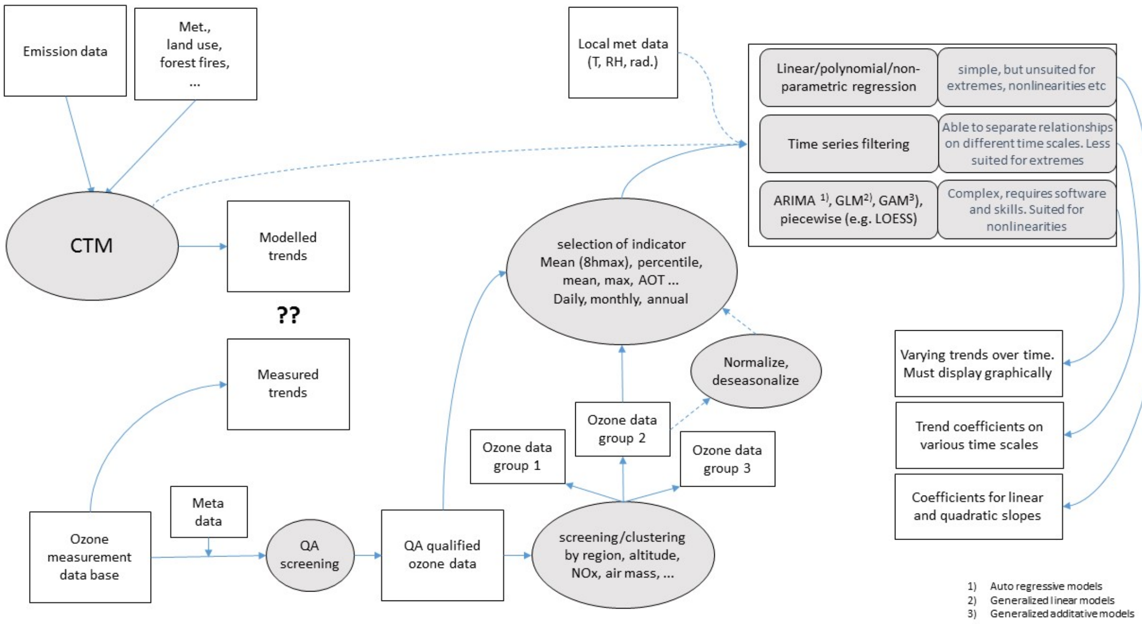


Figure 1: Conceptual view of an ozone trend assessment.

This scoping paper addresses the first main output expected from Task 1.1.2.6 - Air pollution policy effectiveness analysis, namely: “Explore the possibilities and outline approaches for de-trending the impact of meteorology to the ozone and particulate matter pollution trends“. Chapter 2 presents some basic principles of the most commonly used statistical methods for trend analysis of air pollution. Chapter 3 gives an overview over the different types of ozone trend studies available in the scientific literature and Chapter 4 summarises the finding from the review of these studies. The conclusion from the review of ozone trend studies is that the best method for discounting the impact of meteorology on ambient air concentration trends is to use chemical Transport models (CTMs). Finally some discussion and recommendations are given in Chapter 5.

2 Statistical methods

Various statistical/mathematical methods have been applied for ozone trends studies in the scientific literature, and the review paper by Thompson et al. (2001) gives a detailed overview of the topic. Below we just mention a few basic principles.

Among the most popular methods applied in the literature is linear (1) and quadratic (2) regression using a least squares approach. The advantage of the quadratic vs linear regression is the possibility of identifying trends that increase or decrease with time. Several studies (e.g. Parrish et al., 2012) have found a steady increase in ozone over Europe during the 1990s followed by a levelling off and even reductions in the 2000s, and thus a quadratic regression have proven more apt than a linear approach. In principle, higher order polynomial regression could be used, but is rarely seen in the literature. Alternatively, one could apply linear methods for sub-periods of the data. A requirement for using these regression methods is that the residuals are unbiased and fairly normally distributed, and this prerequisite is sometimes overlooked in the reported trend studies.

Linear regression: $Y = Y_0 + \alpha x + \varepsilon$ (1)

Quadratic regression: $Y = Y_0 + \alpha x + \beta x^2 + \varepsilon$ (2)

Where:
Y is the seasonal mean,
x is the time,
ε is residual noise, and
α and β are the linear and quadratic trend terms.

Equally popular is the Mann-Kendall non-parametric test (Annex 1) which often is combined with a Sen's slope estimation (Gilbert, 1987). The main advantage of the Mann-Kendall method is that it relies on the ranks of the data and not on the data values themselves. The method is therefore not sensitive to outliers and the data don't need to be normally distributed. The basic principle is to inspect all pairs of data in a time series consisting of the ranks of the data (e.g. monthly mean values). The Mann-Kendall method then tests the hypothesis (H_0) that there is no trend in the temporal development of these ranks. If H_0 is rejected, the Sen's slope method could be used to estimate the linear trend with confidence intervals.

Some studies adapt the measurement data to a sine curve (3) simulating the seasonal cycle or a sine/trend curve simulating a seasonal cycle with a superimposed linear trend. It should be said, though, that a sine fit sets a very strict and most often not very appropriate seasonal variation on the data. As seen e.g. in Colette et al. (2011) the seasonal cycles in ozone are often far from a sine function, and furthermore, the timing of the amplitude (the seasonal maximum) could be changing, as discussed by Parrish et al. (2013), and this is not captured in a simple sine fit function.

Sine fit to seasonal cycle: $Y = Y_0 + A \sin(x-\varphi)$ (3)

Where:
Y_0 is the annual average,

A the amplitude (seasonal cycle),
x the month, and
φ the phase shift

As an alternative to a sine-fit assumption, several studies instead apply a deseasonalization to the data prior to any further trend assessment. This could be done in various ways, e.g. simply by subtracting the average monthly mean from the data values.

Numerous other, more advanced, mathematical methods are used to estimate ozone trends in more rigorous manner, but these are not outlined in more detail here. Methods like generalized linear models (GLM) and generalized additive models (GAM) are available through the statistical R package and have been applied in some ozone trend studies. Piecewise methods (like LOESS or LOWESS) have also been used. The piecewise methods basically applies a local regression to a sub-set, i.e. a moving-window with a given length, of the full time series in which the data points are given weights that are inversely proportional to the distance (in time) from the local regression point. One drawback with these approaches is that the calculated trends are not easily communicated in terms of single numbers like the linear or polynomial methods.

3 Different types of ozone trend studies

A large variety of European based ozone trend studies have been published in the scientific literature. They can be grouped into certain categories:

- *Trend analyses based on measured ozone time series from one or more sites only*
- *Trend analyses based on measured ozone time series "filtered" by additional parameters like air mass origin (based on trajectories), altitude, NAO-index, local meteorological data etc*
- *Trend analyses based on measured and modelled ozone data in combination*

In addition to this classification based on methodology, the published trend studies could be divided into two categories: baseline/background and European. The former refers to air masses at the inflow boundaries of the European boundary layer, unperturbed by fresh European emissions, typically observations from the western coast or the Atlantic or from high elevation mountain sites, whereas the latter refers to observations from inside the European continent. Furthermore, several studies (e.g. Parrish et al., 2012) make a distinction between "background", referring to the unperturbed pre-historical conditions and "baseline", referring to the present northern hemispherical conditions, far away from major emission sources. In the following we also adopt this terminology when discussing the ozone baseline.

3.1 Trend studies based on measured ozone time series alone

These trends studies include various kinds of statistical/mathematical analyses of the ozone observational data from one or more monitoring platforms. The platforms could be regular surface monitoring stations, vertical soundings or observations from airplanes. The statistical/mathematical analyses range from the most basic approach using a simple linear least squares fit to more advanced statistical methods using non-linear regression models, LOESS regression or time series tools (e.g. ARIMA[1]) sometimes combined with a grouping/categorization of the monitoring sites by statistical clustering, PCA[2] or EOF[3] analyses.

The only explanatory variable in these studies is time itself, like hour, month, season and year or a combination of these. Furthermore, many kinds of ozone metrics (the dependent variable) are used, like monthly means, seasonal means, daytime max, the number of exceedance days above certain thresholds etc.

Additionally, some studies apply a deseasonalization, log-transformation or other kinds of conversion of the ozone data to remove the seasonal cycle and prepare a data set that is closer to a normal distribution and thereby more robust to analyse statistically.

---

[1] Autoregressive integrated moving average
[2] Principal component analyses
[3] Empirical orthogonal functions

3.2 Trend studies based on measured ozone time series "filtered" by additional data

The basic of these methods is a screening of the ozone data by certain additional parameters, like local meteorology, air mass origin, local pollution etc. A typical example is ozone data from a coastal station like Mace Head screened for Atlantic, inflow air masses by use of a backward trajectory model, or a mountain site screened for free tropospheric air masses by picking measurements at certain times of the day only. The screening could also be based on local meteorology, like temperature, wind and humidity measured at the station or on the large-scale circulation type as classified by NAO[4] index, GWL[5] or similar indicators. The screening implies that a pre-selection of ozone data is applied prior to the subsequent statistical trend calculations.

Alternatively, the additional data are included directly in the trend calculations: the measured temperature, wind speed and humidity or NAO-index are included as explanatory variables in a multiple regression procedure where the aim is to isolate the influence of each of these variables as well as any long-term trend on ozone (the dependent variable). Methods with different levels of complexity ranging from plain linear regression to more sophisticated non-linear methods (Wood, 2006), including generalized linear models (GLM), generalized additive models (GAM) and also Bayesian approaches have been applied (Zheng et al., 2007).

As in part 1.1, these methods could be applied to hourly, daily, monthly or seasonal ozone data, or to derived parameters (number of exceedances), and the ozone data could be subject to log-transformation and other adjustments.

3.3 Trend studies based on measured and modelled ozone data in combination

In these type of studies, the ozone observational data are compared to and analysed together with results from multi-year chemical transport model (CTMs) runs. A basic method is to apply a linear regression to the measured and modelled data, separately, and compare the estimated trends. Another standard procedure is to do two sets of CTM runs, one with the assumed real year-to-year changes in emissions and another with fixed emissions. The level of agreement between the trends estimated from the observations and these two CTM scenarios could be used to evaluate the validity of the emission data and other model assumptions.

[4] North Atlantic Oscillation
[5] Grosswetterlage (German phrase for weather type)

4 Trend studies in the literature

A vast number of papers and individual research projects have been aimed at tropospheric ozone trends during the years, and it's beyond the scope of this report to provide a complete review of all these. The following gives an overview of the most relevant studies with particular emphasis on publications from the last years. The various studies are subjectively allocated into the three categories defined above and the main aim is to present the statistical methods and main findings.

4.1 Unfiltered observational based studies

4.1.1 Baseline studies

A comprehensive European ozone baseline study by Logan et al. (2012) investigated different monitoring types (surface sites, sondes, aircrafts etc) with focus on observations above 2 km of altitude. They did not apply any filtering to the data. Ozone trends were calculated by linear regression (Eq. 1) based on the monthly and seasonal (3-months) means separately, for every 1 km altitude. In addition, for the seasonal means, a quadratic polynomial (Eq. 2) was applied for each season separately to look for cases with a change in the slope over time.

The study showed that at least some of the ozone trends reported in the literature could be ascribed to problems with instrumentation, or were inconsistent in some way with other data. They recommended that ozone assessments should be based on alpine and MOZAIC time series and the Payerne and Hohenpeissenberg sondes data from 1998 onwards. They estimated mean trends at around 3 km altitude of 6.5-10 ppb for 1978—1989, 2.4-4.5 ppb in the 1990s and a decrease of 4 ppb during the summer months in the 2000s, but with no significant trends in other seasons that decade. They found no indications that the trends could be explained by changes in stratospheric ozone.

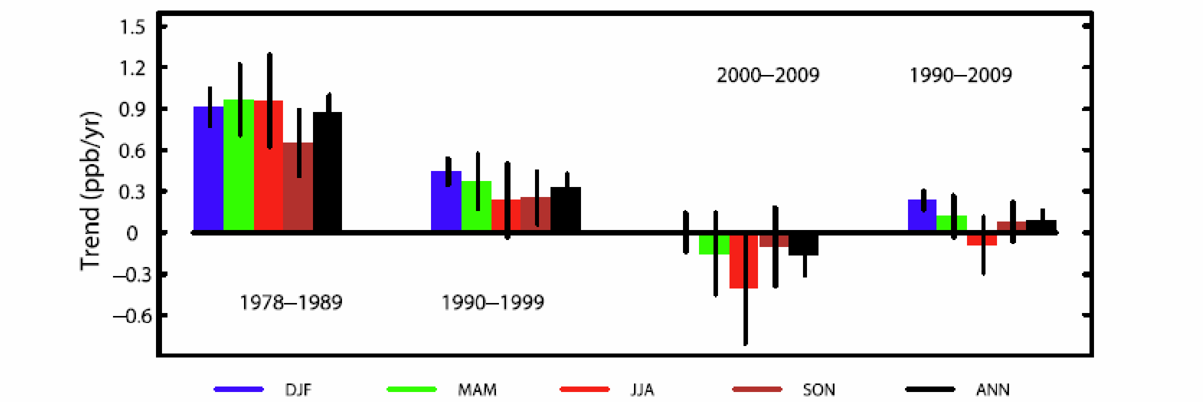


Figure 2 **(adopted from Logan et al., 2012) Seasonal and annual trends in ozone (ppb/year) for Zugspitze (1978–1989) and for the mean alpine time series for the more recent decades and for 1990–2009 as indicated.**

In their Mace Head study, referred in more detail in 4.2, Derwent et al. (2013) applied a non-linear regression to a polynomial (Eq. 2) based on annual mean concentrations using all data from 1987-2012. They found a linear coefficient of 0.65 ± 0.14 ppb/year and a quadratic coefficient of -0.0187 ± 0.006 ppb/year^2, indicating a steep rise the first part of the period, followed by a gradual down-slowing eventually levelling off, in good agreement with similar studies (Logan et al., 2012; Parrish et al., 2012), even though the annual mean concentration may not be a well-suited parameter for trend assessment due to the significant seasonal differences (Parrish et al., 2013).

A review of published global tropospheric ozone trends, including baseline studies and results from a few European sites were recently given by Cooper et al. (2014). Table 1 lists some relevant baseline studies of ozone trends.

Table 1. Baseline studies

Ref.	Ozone data type	Time period	Method	Findings
Logan, 2012	2 km-tropopause MOZAIC HPB, PAY sondes	1980s-2000s	Linear regr. of monthly and seasonal means + quadratic poly.	Increase in 1990s Decrease in 2000s in summer
Parrish et al 2009	Surface MH	1950-present	Linear regr. On seasonal means Pollution filter Air mass origin	Increase in MBL inflow to NA
Parrish et al 2013	Surface MH	1950s-2000s	Sine funct to monthly data Air mass filtering	Shift in seasonal cycle "models have difficulties reprod. trends"
Parrish et al 2012	Surf + free trop MH, HPB, mountain sites	1980s-2000s	Seasonal means quadr. Fit and lin regr.	Sign. Negative accel. -> rate of increase has slowed, some places decrease
Jenkin 2008	Surface Strat Vaich + UK sites	1990-2006	5-95 percentiles	Increase in perc, but decrease in N180. Narrowing od dist.
Pausata et al 2012	17 EMEP sites	1990-2005	Monthly anomalies	Increase in 1990s linked to NAO
Cooper et al 2012	Surface sites USA	1990-2010	Daytime data percentiles	Decrease, summer O3 east NA No trend or increase in W NA
Oltmans 2013	Surface + mtn. MH, ZUG, IZA	1970-2010	Aut-reg models, cubic poly. Fit on monthly values	Shift in max from May-> April NAO shift at IZA
Hess & Zbinden 2013	Whole trop. NA and Europe	1990s-2000s	Normalized, smoothed monthly data	Strat decisive for surface sites, incl MH

4.1.2 European studies

Wilson et al. (2012) analysed ozone data from 158 European rural background monitoring stations during 1996-2005. This was based on the so-called harmonized data set prepared as part of the EU FP6 project GEOMON in which a selection of stations was made based on site locations and sufficient data coverage during this decade. This selection lead to a data set biased to central and northern Europe.

They applied several statistical trend methods on the 5- and 95 percentiles and monthly means (based on hourly ozone values) using the R software with the Openair package (Carslaw and Ropkins, 2012; Carslaw, 2014). LOESS trends were applied to identify certain types of trend behaviour whereas the non-parametric Mann-Kendall method with a Sen-Theil slope estimate was used to quantify linear slopes in the data. All test were applied to the sites individually. Prior to the Mann-Kendall calculations the data were deseasonalized by built-in methods in the Openair package. Linear trends were also estimated by log-transformed data.

They found a significant ($p<0.1$) positive annual trend in the ozone mean, 5- and 95-percentile at around 50 % of the sites for the 1996-2005 period. However, the estimated slopes turned out to be very sensitive to individual years like the 2003 “heat-wave year” and they conclude that alternative statistical approaches capable of removing the meteorological influence should be used. A comparison with CTM calculations was also done and this is discussed in the chapters below

.

In the framework of the CityZen FP7 project, Colette et al. (2011) investigated ozone urban, suburban, and rural observed trends over 705 airbase stations passing quality screening over the 1998-2007 decade. They apply a similar deseasonalised Mann-Kendall and Sen-Theil slope method with the more stringent criteria for detecting significant trends of a p-value smaller than 0.05. The estimated average daily mean ozone trend at UB, SB and RB sites was 0.37, 0.27 and 0.05 µg m^{-3} yr^{-1}, respectively. The proportion of sites where the ozone trend was found to be significantly positive was 30.8% when considering daily means but this number dropped to 18.5% when considering O_3 daily peaks, reflecting the larger efficiency of mitigation strategies for ozone peaks than for baseline ozone.

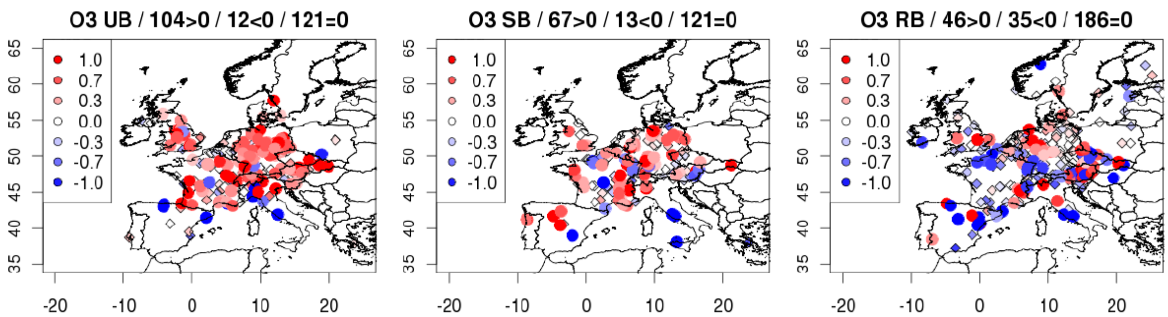


Figure 3: Ozone trend in µg/m³/year at Airbase stations over the 1998-2007 period. The heading indicate the number of stations where the trend is statistically increasing, decreasing, or un-significant using a p-value of 0.05.

Sicard et al. (2013) investigated Airbase ozone data from background stations (rural, suburban and urban) over the period 2000-2010 from the Mediterranean region. They applied a Mann-Kendall regression with Sen slope estimation based on various annual statistics: 24-h mean concentration, median, 98th percentile, average daily maximum and hourly peak maximum. No filtering of data was applied. They found a decrease (-0.43 %/year) in the annual mean concentrations at rural sites and an increase at suburban and urban sites. A decrease in the annual maxima and 98th percentiles was estimated for around 75 % of the sites, including all three categories with downward trends of the order of 0.77 %/year to 1.14 %/year for the P98 and annual max, respectively.

Jenkin (2008) applied linear regression to annual ozone percentiles (ranging from the 5^{th} to the 95^{th}) from UK sites for the period 1990-2006 and found downward trends in the annual maximum and highest percentiles and increasing trends in the lowest percentiles, in qualitative agreement with reduced NOx emissions. For the most remote sites in the north, however, a significant upward trend was estimated for all percentiles, which was linked to a gradual rise in the hemispheric baseline level.

Derwent and Hjellbrekke (2013) studied the trend in annual maximum 8-h average ozone mean concentration at EMEP sites by use of Mann-Kendall regression. They found that for the sites with the longest monitoring history, i.e. running through 1980-2009, highly significant downward trends were found at two British sites whereas downward, non-significant trends were found at three German sites. For the period 1990-2009, 40 sites were analysed and 16 of these sites showed a statistically significant downward trend. They furthermore found that the magnitude of the reductions were correlated with the peak values at the start of the period, i.e. that the trends were most negative for the sites with the highest 8-h maximum in 1990.

Tørseth et al. (2012) investigated differences at EMEP stations in the O_3 99^{th} percentile, AOT40 and the frequency distribution of hourly values between two 10-yr periods, 1990-1999 and 2000-2009. They found indications of reduced 99^{th} percentiles in some regions (England, Benelux, Germany, Czech republic) and no trend in Switzerland and Austria. The differences in the mean frequency distributions showed a mixed pattern with the expected narrowing of the distribution at some sites (UK and the Netherlands) and no apparent change at other sites (mainly Germany, Switzerland and Austria).

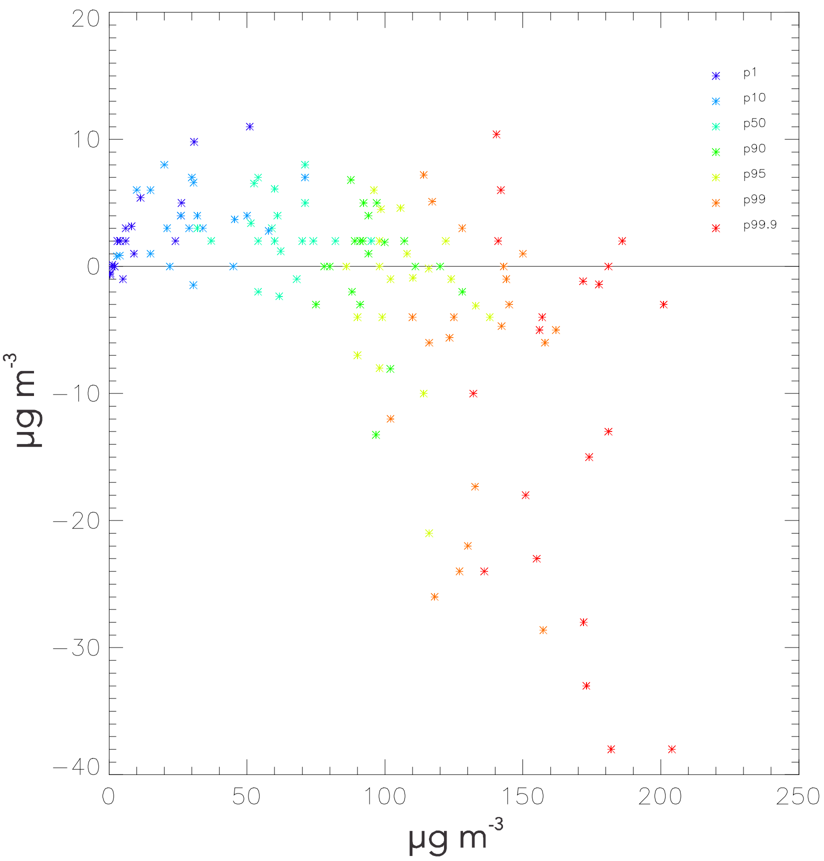


Figure 4 (adopted from Tørseth et al.) The change in ozone percentiles from the 10-year period 1990–1999 to 2000–2009 as a function of the percentile values in 2000–2009 for 20 EMEP stations. The colour codes indicate the type of percentiles.

In an ozone review paper Simpson et al. (2014) applied an analysis similar to that in Tørseth et al. (2012) for the change in ozone percentiles. They found similar differences between regions in Europe as reported by Tørseth et al. (2012). Reductions in the highest percentiles (P^{99} and $P^{99.9}$ and to a less extent P^{95} based on hourly data) were however seen at almost all the sites inspected.

Table 2 lists some relevant European studies of ozone trends.

Table 2. European studies

Ref.	Ozone data type	Time period	Method	Findings
Wilson et al 2012	Surface GEOMON sites	1996-2005	Obs. and mod. Linear trend, perc., annual + seasonal Deseasonalise, loess reg. mod, linear Mann-Kendall	Sign pos trends at a number of sites, all perc., particularly P5 in winter, spring, summer
Sicard et al 2013	Surface Airbase sites, Mediterr.	2000-2010	Daily data, perc. Mann-Kendall	Neg. trends in P98 and hmax at many sites
Colette et al. 2011	Surface CityZen Airbase sites	1998-2007	Deseasonalize Mann-Kendall Monthly data	Red. of NO2 Increase in dmaxO3 UK, Benelux, DE
Vautard et al 2006	Surface 37 EMEP sites	1990-2002	perc	ObsP90 > ModP90 in UK ObsP90 < ModP90 in DE, centr Eur
Jonson et al., 2006	EMEP sites	1990-2004	Linear trend in mean daily max summer ozone	Discrepancy between model and obs. Could be due to background.
Derwent et al 2013	Surface MH	1987-2012	Quadratic. Poly. Pollution sorting Air mass filtering	Peak in 2007 in unfiltered data Rise in Eur. Winter, decline in summer. For baseline: 17evelling off decrease in seas. Means.
Ordonez et al 2005	Surface, low-elevation 12 sites in the Alps in CH	1992-2002	ANCOVA vs dmaxO3	Downw. Trends in the summer P90 of dOx or O3max at 6 ind.sites, not in any rural. Increase in winter.
Akritidis et al. 2014	74 rural EMEP sites	1996-2006	Mann-Kendall, deseasonalized data	Pos trends. Model underestimates trends
Solberg et al. 2009	Background rural Airbase sites	1995-2005	Linear trend, annual ozone metrics	Neg. trends in the NW, else mixed. Mismatch between models and obs.
Fernandez-Fernandez et al., 2011	EMEP Iberian sites	2001-2007	Kendall rank correlation O38hmax vs T, Rad.	Significant pos. correlation, particularly in the NE
Tørseth et al., 2012	EMEP sites	1990-2009	Change in P99, freq. distrib etc between two decades (1990s-2000s)	Reductions in P99 and narrowing of freq dist in the NW. No change in AT, CH countries
Cristofanelli et al., 2015	Mt Cimone	1991-2011	Linear based on monthly anomalies for different times of day and separate seasons.	Negative acceleration in summer, no change in other seasons. Possible links with the occurrence of heatwaves.

4.2 Filtered observational based studies

Numerous studies have looked at trends in data screened by air mass characteristics, typically by use of back-trajectories. The monitoring station at Mace Head on the west coast of Ireland plays a special role in this respect since it is often regarded as representative of the inflow boundary conditions for Europe and is in fact used so directly by the EMEP unified model (Simpson et al., 2012). Thus, various studies have presented ozone trend analyses based on Mace Head data screened for certain types of atmospheric transport conditions.

Recently, Derwent et al. (2013) analysed the complete 25 years' time series of ozone (1987-2012) by use of a combined pollution filtering (co-incident measurements of anthropogenic tracers) and dispersion model filtering of the ozone data and calculated trends for baseline, European and southerly air masses, respectively. They applied the linear non-parametric Mann-Kendall method and the polynomial regression mentioned above (Eq 2.) to monthly mean concentrations. For the baseline, they found positive linear terms and negative acceleration terms in most months, pointing to constantly decreasing growth rates leading to decreasing levels the last few years, as for the unfiltered data mentioned in 4.1.1. For the period overall, they found stronger baseline increases in winter and spring than in summer.

For the European excess (the European subtracted the baseline) they found significant upward trends in December and significant downward trends in May, June and July in nice agreement with what to expect based on the European emission reductions this period.

Parrish et al. (2013) studied the trend in the seasonal cycle of baseline ozone based on filtered Mace Head data and unfiltered data from Jungfraujoch, Zugspitze and Hohenpeissenberg and used two methods. In the first approach they applied a sine fit (Eq. 3) to monthly average concentrations for running 5-years' periods and looked for significant changes in the sine fit coefficients during the whole time period. In the other approach they looked for significant changes in the date of the year when the cumulative ozone concentration reached 50% of the year's total.

Both methods indicated a shift of the ozone maximum concentration to earlier in the year. At the sites with the longest records, Zugspitze and Hohenpeissenberg, a significant trend of 5-7 days per decade was found, whereas for the baseline data from Mace Head a trend of 3 days/decade was estimated although not statistically significant.

Parrish et al. (2012) investigated long-term ozone data from 11 locations around the globe including baseline data from Mace Head filtered by dispersion modelling and unfiltered data from several other European sites. They applied linear and quadratic fits to 3-months seasonally averaged data and argued for this averaging period compared to e.g. monthly averages which may include too much variability and compared to a deseasonalization procedure with its inherent problems as also discussed by Parrish et al. (2009). They found marked differences between the filtered and unfiltered data from Mace Head (1989-2010) with larger magnitudes in the linear and quadratic coefficients in the baseline data compared to the unfiltered data, illustrating the importance of sorting baseline from non-baseline data. The unsorted data showed no significant slowing in the ozone growth rate.

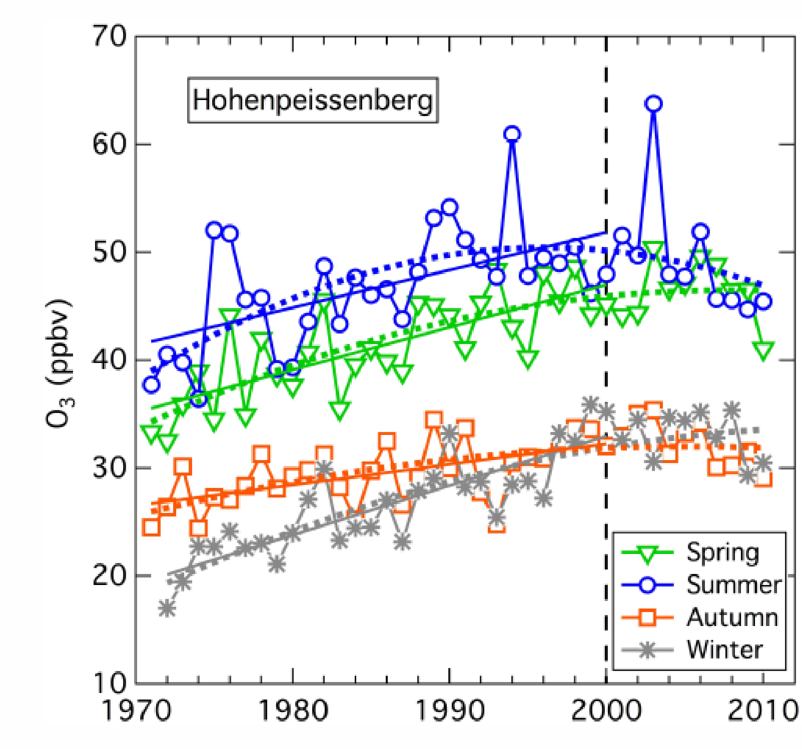


Figure 5 **(adopted from Parrish et al., 2012) Seasonal O_3 averages measured at Hohenpeissenberg Germany.The solid lines indicate linear regressions for the data up to and including year 2000. The dotted lines give the quadratic regressions for the entire data sets.**

They concluded that at least before 2000 the average increase of O_3 at each site was approximately the same at around 1%/year relative to 2000 in each season. They also showed evidence that the rate of growth had slowed, particularly over western and central Europe to the point that a recent decrease was observed in some seasons, particularly in summer. Furthermore, whereas they found relatively rapid increases in free tropospheric ozone over North America, particularly in air masses originating from Asia, no similar behaviour was seen downwind of North America.

A number of ozone trend studies has applied other kinds of data filtering or categorization than back-trajectories. Pausata et al. (2012) analysed the anomalies in monthly mean surface ozone concentrations at 17 European (EMEP) monitoring sites with respect to NAO[6] index in the period 1990-2005. They found a clear positive correlation between the plain NAO index (the NAOI defined as the pressure anomaly difference between two fixed locations, normally at Iceland and Portugal/the Azores) and measured O_3 anomalies in winter. In other seasons a similar relationship was not significant. They showed, however, that for spring and summer the first principal component (PC1) of the EOFs turned out to be a much more appropriate parameter than the NAOI due to displacements of the large-scale synoptic systems. Based on their results, they argued that the NAO behaviour and the associated

[6] North Atlantic Oscillation

surface O_3 anomalies could explain the variability reported over western and northern Europe during the 1990s and 2000s.

Another study looking for long-term trends and statistical relationships between large-scale dynamics and tropospheric ozone measurements was carried out by Hess and Zbinden (2013). They investigated the links between stratospheric ozone using the CAM-chem model and ozone monitoring data during 1990-2009. Although they focused on free tropospheric ozone data (sondes and air craft data), they also used a few surface monitoring sites, including Mace Head and Jungfraujoch. They based their trend analyses on "regional records of ozone variability", defined as regional (spatial) averages of the normalized 12-months running means of the deviations from the monthly means. Thus, the analyses were based on a highly smoothed (deseasonalized and normalized) data set. They found that the ozone trends at the surface sites had many of the same characteristics as for 500 hPa and even 150 hPa with a significant positive trend during 1990-2000 and no trend during 2000-2009. They concluded that the analyses strongly suggests that the stratosphere has been a strong source of variability at Mace Head.

The findings of Hess and Zbinden is in line with a similar study by Ordonez et al. (2007) looking at the linear correlation between the normalized and deseasonalized ozone data from Jungfraujoch and Zugspitze vs the lower stratospheric ozone.

Oltmans et al. (2013) looked at a number of surface and ozonesonde stations around the globe, including Mace Head, Hohenpeissenberg, Zugspitze and Izana in Europe. Their analyses were based on monthly ozone data constructed from either daily 24-h data or 8-h daytime or night-time data depending on the degree and type of diurnal ozone cycle at the site. They used a trend methodology developed by Harris et al. (2001) for total ozone consisting of an auto-regressive model with a cubic polynomial fit and the 500 hPa and 100 hPa temperatures as two of the explanatory meteorological variables. Ozone growth rates were estimated through a bootstrap technique as explained in Harris et al. (2001). Oltmans et al. (2013) also studied trends in the frequency distribution of ozone through two metrics (W_Low and W126) expressing the low and high part of the frequency distribution of hourly ozone concentrations.

For the European sites, they found strong signs of increasing levels during the 1990s followed by a decade with flattening or even reductions in the mean O_3 levels, in agreement with the other European baseline studies (refs). Differences were found in the long-term ozone development at Zugspitze (2980 m asl) and similar altitudes above Hohenpeissenberg (from sondes) without any obvious explanation apart from the general differences between ozone sondes and continuous monitors. They also reported an apparent increase in the W_Low and decrease in the W126 the last period, indicating that the entire frequency distribution is shifting downwards.

Cristofanelli et al. (2015) analysed trends in hourly ozone monitoring data from the WMO/GAW station at Mt. Cimone from 1991 to 2011. When avoiding the first years of data (when another site/instrumentation was used), a significant negative acceleration (using a quadratic polynomial) was found in summer, indicating a reduced or even negative growth rate in ozone. In other seasons no statistical trends were seen. As reported by e.g. Parrish et al. (2013) they found indications that the annual peak in ozone had shifted from summer to spring. They also looked for relationships to the occurrence of heat waves, episodes of stratospheric intrusions and NAO index and concluded that these processes could have partly explained a large positive ozone anomaly during 2005-2008.

The comprehensive review of methods for “meteorological adjustments” of surface ozone measurements by Thompson et al. (2001) mentioned earlier, concerns statistical methods of varying complexity for relating *in-situ* meteorological measurements to ozone monitoring data. The number of such studies for Europe is considerably less than what has been presented for North America. Some of these European studies is mentioned in the following.

Fernandez-Fernandez et al. (2011) investigated surface ozone trends at EMEP sites on the Iberian Peninsula and applied a Kendall rank correlation analysis for the daily max 8-h mean ozone concentration vs observed daily mean temperature and daily mean solar radiation. They found significant positive correlations for all sites and strongest signals in the northeast. This analysis was limited to the rather short period of 2001-2007.

In the paper by Ordonez et al. (2005) an analysis of covariance (ANCOVA) was used to derive the influence of the meteorological variability on the daily maximum ozone concentrations at 12 Swizz sites during 1992–2002. They applied the ANCOVA to each site and each season (spring = MAM, etc), separately. A backward elimination procedure by the R software was used to identify the most significant meteorological parameters.

The square of the afternoon temperature, the morning global radiation and the number of days after a frontal passage turned out to be the three most important explanatory variables in spring and summer. In winter parameters linked to mixing (wind speed and radiation) was the most important ones. They found no significant downward trends in the meteorologically adjusted summer seasonal medians or the 90th percentiles of daily O_3/Ox[7] maxima at six rural sites as opposed to what to expect from the general emission reduction. For stations in the industrialized region around Zurich, significant downward trends were found, however. They argued that an increasing trend in the ozone background levels could be compensating the effect of the reduced European precursor emissions.

4.3 Studies of measurements and models combined

Numerous studies have looked at European ozone trends by use of observational and modelling data combined the last years. In its most simple way, linear trends in certain ozone metrics (daily max, monthly means, percentiles etc) are computed from the CTM results and the monitoring data, respectively and then compared.

Akritidis et al. (2014) applied the RegCM3/CAMx model for the period 1996-2006 with two scenarios, using constant and yearly changing anthropogenic emissions, respectively. The model runs were compared to surface ozone monitoring data from 74 rural sites (below 1500 m asl) in the EMEP network (Tørseth et al., 2012). Annual trend calculations were based on Mann-Kendall regression of deseasonalized monthly mean concentrations and seasonal trend calculations on plain seasonal means. They found positive annual trends at most sites both in the observations and in the model run with changing emissions which they explained by the

[7] $Ox = O_3 + NO_2$

reduced NOx-titration of O_3 following reduced NOx emissions. The results indicated that the sign of the modelled trends agreed with the observed trends while the linear correlation was rather poor. Overall, they concluded that the model apparently underestimated the observed trends.

The multi-model study by Colette et al. (2011) analysed deseasonalized data from the 10-years period 1998-2007 and applied a Mann-Kendall method to estimate linear trends. They found a very good agreement between observed and modelled NOx levels in Europe, whereas they commented that "ozone trends turned out to be much more challenging to reproduce".

In a previous EEA study Solberg et al. (2009) analysed rural Airbase ozone data from 1990-2005 and EMEP model calculations for 1995-2005 with fixed and varying anthropogenic emissions, respectively. They applied linear regression to various annual ozone metrics like AOT40, SOMO35[8], N8h120[9] and MTDM[10]. They also used the model to quantify the impact of the meteorological inter-annual variability on ozone over Europe. They found the clearest observed downward trends in areas (UK, Benelux etc) where, according to the model, meteorological variability should be the largest and thus most likely to mask emission-induced trends. On the other hand, no observable downward trends in ozone were found in regions (Austria, Switzerland) where the model predicted significant reductions and meteorology should have the least influence. This apparent mismatch is similar to the findings reported by Colette et al. (2011). The short length of time (10 years) for these model-measurement comparisons could be a critical limitation, as discussed by Solberg et al. (2009). Additionally, the studies could be influenced by the change in the ozone baseline trend at the mid of the periods, going from an increase in the 1990s to a no-trend (or even decline) in the 2000s (Oltmans, Logan …).

Vautard et al. (2006) compared 10th and 90th percentiles based on EMEP ozone measurements with CHIMERE model calculations for the period 1990–2002. They found indications of reduced peak concentrations and an increase in the ozone baseline level. They argued that for central/northern Europe the emission reductions given in the official EMEP data could be too optimistic for this period.

Jonson et al. (2006) used the EMEP model and predicted reductions of some 5–10 ppb in mean daily maximum summer ozone concentrations (June to August) in large parts of Europe and up to 12 ppb in Germany for the period 1990–2004, clearly larger than seen in the measurement data. They argued that most of the monitoring sites were located in areas where NO_X emission reductions were expected to have less effect and in the areas most sensitive to changes in background concentrations (north-western Europe).

4.4 Other studies

A couple of studies from outside Europe is mentioned below as their methodological approach may be of relevance for future European based studies.

---

[8] Accumulated ozone concentrations in excess of 35 ppb.
[9] Number of days with a maximum 8-h running mean concentration exceeding 120 µg/m3.
[10] Mean of the 10 highest daily maximum concentrations (based on hourly data) during April–September.

Davis et al. (2011) applied various non-parametric regression methods including generalized linear models (GLM) and generalized additive models (GAM) using the R software to investigate the relationships between ozone and local meteorological data from both measurements and model calculations during 2002-2005. The study was based on observed ozone data from 74 cities in the eastern US and meteorological data from surface stations and vertical soundings as well as similar data from the CMAQ model. Their results showed that the daily maximum temperature (T_{max}) and the daily average relative humidity (RH_{avg}) were the two most important explanatory variables in the regressions, in agreement with similar purely observational based studies. More important, however, was the finding that the link between T_{max} and O_3, increasing from southwest to northeast, was stronger in the observational data than predicted by the model. Similarly, the link between RH_{avg} and O_3, increasing in the opposite direction (from northeast to southwest) was also stronger in the observational data then given by the model. Thus, a general underestimation of the met-O_3 sensitivities were identified in the CMAQ model which could be of concern e.g. for evaluations of future climate-change/air-quality relationships.

Chan (2009) presented a fairly comprehensive and extensive methodology for studying ozone trends based on data from 97 monitoring sites in North America for the period 1997-2006. They first applied a PCA or EOF methodology to group the sites into homogeneous regions in terms of ozone variability. Next, they classified the ozone data in these regions by use of air-mass trajectory clustering. Finally, the ozone trends were calculated for these regions using generalized linear models (GLM) with local meteorological data as explanatory variables and the trajectory clusters as a classification variable. They argued that this kind of multiple-site analyses should be more statistically robust than single-site studies.

In a somewhat similar study Zheng et al. (2007) investigated the performance of two statistical approaches, both using nonlinear regression methods for assessing ozone trends in the US from 1997 to 2004. As in Chan et al. (2009) they first applied a PCA to allocate the stations into homogeneous groups in terms of ozone. They then applied a GAM and DLM[11] for the regressions using daily 8-h maximum ozone concentrations and daily meteorological values. They found that the GAM and DLM methods produced very similar results. The GAM is recommended due to its higher simplicity and ease of use in various software packages, whereas the DLM is significantly more difficult to apply, as they are normally not included in standard statistical software.

[11] Dynamic Linear Model

5 Discussion and recommendations

Chapters 3 and 4 have given an overview over the vast number of studies related to ozone trend studies and meteorological influence in the scientific literature. Although many of the studies rely on similar methodologies, the differences in selection of sites, metrics, time periods, regions, etc give a rather diverse picture of the present knowledge status of European ozone trends. The use of data screening based on trajectories or on local meteorological data and the use of more sophisticated nonlinear regression methods like GLM and GAM add on to the variety.

For European baseline ozone, the most recent studies paint a fairly consistent picture of a rough doubling of O_3 from the 1950s in all seasons up to about the year 2000, followed by a decade with no growth or even reductions in O_3 at some sites in some seasons, particularly in summer (Simpson et al., 2014).

For the European surface monitoring sites, the picture is much more mixed than for the baseline studies, and the questions discussed in the Introduction remain open. A condensed summary of these studies is attempted in the following.

The signs of reduced peak ozone levels in Ireland, the UK and some sites in the Benelux region during the last 10-15 years are clear and strong and it is very likely that this is caused by reductions in European anthropogenic emissions (Derwent et al., 2013; Jenkin 2008; Derwent and Hjellbrekke, 2013). This is observed by all recent publications and manifests itself through a downward trend in summer ozone, particularly for the highest percentiles. An increase in winter-time ozone is also apparent, in agreement with reduced NOx-titration, thus leading to a general narrowing of the ozone frequency distribution over the year. These changes appear to be even larger than predicted by numerical models. The strongest impact is found in the areas closer to the European mainland, whereas in the outskirts the ozone levels appear to be most controlled by the baseline trends. Furthermore, the data indicate a shift in the time of the annual ozone maximum to earlier in the year (Parrish et al., 2013).

In other parts of Europe the ozone trends are more unsettled. Several of the studies that do account for the meteorological influence (either through CTMs or by statistical regression methods) are likely to be hampered by the short length of the ozone time series or by the actual period studied.

Many of the CTM based studies (Akritidis et al., 2014; Wilson et al., 2012; Colette et al. 2011, Solberg et al., 2009) investigated ozone time series for a length of only around 10 years which may be too short to be able to see significant changes (Simpson et al., 1997), given the inherent uncertainties related to inter-annual variations in emissions, inflow baseline conditions etc.

Furthermore, several of these and other modelling studies covered a time period from the start or middle of the 1990s to the start or middle of the 2000 which might be an additional reason for the problems detecting clear trends. As various studies have shown, the baseline data indicate increasing ozone trends during the 1990s followed by a period of no change or slight reductions. It also seems that the last few years of the 1990s peaked in baseline ozone possibly due to contribution from biomass burning events those years (Simmonds et al., 2005). Any downward trend in European photochemically formed ozone due to reduced emissions over this period may thus have been masked by the baseline ozone peaking in the middle of the period. The masking effect of the baseline was mentioned as a possible reason for the lack of downward ozone trends in Switzerland in the non-modelling study by Ordonez et al. (2005).

Future trend studies should benefit from the extended length of the ozone monitoring data, and they should investigate time series of at least 15 years of data preferably not centred at the baseline peak years 1998-99 unless the baseline conditions are handled in a robust manner in the analyses.

Moreover, it is crucial to take into account the inter-annual variations in meteorological influence on ozone when carrying out the trend analyses. Regression methods, linking certain ozone metrics to observed *in-situ* data of temperature, radiation etc have the advantage of being purely empirically based, avoiding all uncertainties necessarily embedded in CTMs. The empirical methods are on the other side seriously limited conceptually by the inherent assumptions of local relationships. Empirically based statistical methods could possibly be more suited at sites closer to the precursor emission sources in southern locations with a fast photochemistry and a short "ozone history", but this needs to be confirmed by CTM calculations. For rural sites in North Europe it is not likely that such methods would give particularly clear results. As noted by Jacob and Winner (2009) the correlation often seen between ozone and surface temperature is generally limited to polluted conditions (ozone > 60 ppb), whereas no correlation is found at lower ozone levels.

Very few studies have investigated the *validity* of empirically based regression methods. One exception is the mentioned paper by Davis et al (2011) who found systematic differences in the regression parameters between O_3 and local meteorological variables when based on a CTM vs when based on observational data for the eastern US.

An alternative to the locally based regression methods is to look at empirical relationships between observed ozone and certain indicators for the type of meteorological "regime", like the NAO index, EOFs, grosswetterlagen (weather pattern) or other parameters representing the large-scale atmospheric circulation type (Pausata et al., 2012; Hess and Zbinden, 2013). Intuitively, this may seem as a more appropriate method for European conditions. The limitation of this approach is that the classification of weather regimes could be too coarse to provide meaningful results when used in an ozone trend study. It should be said, however, that very few such studies have been presented in the scientific literature.

In the project COST733 (http://cost733.met.no/) a detailed procedure for atmospheric circulation type classification was formulated. The project also developed a comprehensive open-source software tool to calculate, visualize and evaluate these classification types (Philipp et al., 2014). In a study by Demuzere et al. (2011) various COST733 circulation classification methods were used to analyze ozone data from 130 rural Airbase sites from central Europe. Their conclusion was not too optimistic with respect to the capabilities of this method for ozone assessments: "… circulation types based on sea level pressure and (or) variables from the upper atmosphere have limited explanatory power with respect to peak surface ozone concentrations". The main problem as they saw it was that local-regional effects, like emission sources, topography etc was not captured by the circulation classes.

The limitations embedded in analyses of observed ozone data with respect to local meteorological variables or to parameters representing the large-scale circulation types suggest that use of CTMs is the tool with the largest capability to provide meaningful analyses of long-term ozone trends. Albeit their many uncertainties, the present CTMs represent the current state-of-the-art knowledge of atmospheric science built into a quantitative prognostic tool. Based on the published studies a number of remarks could be made with respect to the use of CTM for ozone trend studies:

- *A sufficiently long time period (~15 years or more) should be used.*
- *The change in baseline conditions needs to be simulated in a realistic way. Observational-based boundary conditions (e.g. using Mace Head observations) is one alternative, but there is a risk of double counting since such sites are not exclusively driven by imported ozone. The most appropriate strategy remains using large scale models, with the inherent uncertainties that they carry.*
- *A proper selection of ozone metrics is required. Typically, high percentiles (e.g. 95th, 99th) for the summer half year and low percentiles for the winter could be used. Monthly, season or annual means are less suited. A selection of time of day (or daily max) should be applied at sites where the diurnal cycle is locally determined.*
- *One should take into consideration that there may be a trend in the timing of the seasonal maximum, i.e. that the ozone level is now peaking earlier in the year than before.*
- *At least two sets of model runs should be carried out – with real, changing emissions and with fixed emissions.*
- *Various regression methods should be applied, at least with linear and quadratic fits, respectively, to account for changes in the trend with time.*
- *Besides local emissions and meteorological forcing, there are large uncertainties in the trend of the ozone (and precursors) inflow to Europe.*

None of the points above represent new and innovative ideas, but are part of the traditional approach. The limited number of years, selection of time period and type of ozone metric applied in various studies may have made it difficult to identify significant trends, though. As noted by Cooper et al. (2012), detecting robust ozone trends requires many years of data and their study focused on a 21-year period using only hourly average data reported between 11 and 16 local time.

One specific remark could be highlighted: This literature review indicates that rather few studies have explored *in-depth* the discrepancies between observed and modelled European ozone trends by use of the CTMs. Such studies might prove useful for providing a better understanding of strengths and weaknesses of present models and main processes controlling ozone trends.

A main advantage of the CTMs is their deterministic nature. The modelled ozone concentration at a point in time and space could be ascribed to the net effects of various processes, e.g., emissions, photochemical formation, dry deposition, although not in a linear way. Discrepancies between observed and modelled ozone, as discussed in this report, could thus in principle be linked to deficiencies in certain process formulations in the model, at least when studied over a certain length of time.

A joint model initiative known as Eurodelta-Trends looking at long-term trends in modelled and observed ozone and PM in Europe is ongoing and will be published in scientific papers and reports.

Errors in the model's biogenic emissions would give rise to one type of "signal", errors in dry deposition another, etc. This approach is basically what is called a sensitivity analyses. Our suggestion is to apply a sensitivity approach focussed on the mismatch between modelled and observed data and not on the model results directly. Furthermore, this kind of analysis should be applied to the data in a detailed, e.g. day-to-day or episode-wise manner. Ideally, ensemble model runs should be applied to provide a pdf (probability density function) of the modelled ozone trend. Although CPU-demanding, this type of analyses should be feasible to a certain extent. True Monte-Carlo runs with many variables are probably not possible to carry out for multiple years whereas sensitivity analyses based on Latin Hybercube Sampling (LHS) may be an alternative option.

Process based analyses of models/observations could also be combined with various clustering/filtering methods based on atmospheric circulation and/or trajectories like in Chan (2009) or based on certain conditions (biomass burning episodes, heat waves/drought etc).

Furthermore, one could also imagine using the CTMs to estimate the type of ozone metric most sensitive to anthropogenic NOx and VOC emissions (according to the model) and least sensitive to other conditions (like meteorology etc) and thus most suited in trend analyses. This metric could even be allowed to vary between the monitoring sites, with one type of metric in Scandinavia and another in southern Europe etc. To our knowledge, no studies have provided any quantitative evaluation of the various ozone metrics used in the literature with respect to robustness for long-term trend assessment.

Several studies have documented a change in the frequency distribution of ozone (e.g. Tørseth et al., 2012; Simpson et al., 2014), which is scientifically sound since reduced NOx emissions should lead to a decrease in both the lowest and highest peak ozone values. Superimposed on this is the change in the baseline ozone with indications of a recent decline in Europe. Thus, any model/observational trend study should preferably include consideration of the ozone pdfs and not only single ozone metrics.

Finally, although the topic of measurement data quality has not been raised in this report, it is a critical issue with respect to any robust trend assessment. Previous studies, like the EEA ozone trend study in 2009 (Solberg et al., 2009) revealed that a certain fraction of the observational data are likely to be hampered by errors in calibration, instrumentation, data format etc. A QA/QC screening of the observational data is thus a prerequisite for any long-term trend assessment. One could argue that the QA/QC criteria should be stricter for data included in long-term trend studies compared to other kinds of evaluations since small biases in time series could have a substantial impact of estimated slopes.

References

Akritidis, D., Zanis, P., Pytharoulis, I., Karacostas, Th. (2014) Near-surface ozone trends over Europe in RegCM3/CAMx simulations for the time period 1996-2006. Atmospheric Environment, 97, 6-18.

Carslaw, D.C. and K. Ropkins, (2012). openair — an R package for air quality data analysis. Environmental Modelling & Software. Volume 27-28, 52-61.

Carslaw, D.C. (2014) The openair manual — open-source tools for analyzing air pollution data. Manual for version 1.0, King's College London. http://www.openair-project.org.

Chan, E. (2009) Regional ground-level ozone trends in the context of meteorological influences across Canada and the eastern United States from 1997 to 2006, J. Geophys. Res., 114, D05301, doi:10.1029/2008JD010090.

Colette, A., Granier, C., Hodnebrog, Ø., Jakobs, H., Maurizi, A., Nyiri, A., Bessagnet, B., D'Angiola, A., D'Isidoro, M., Gauss, M., Meleux, F., Memmesheimer, M., Mieville, A., Rouil, L., Russo, F., Solberg, S., Stordal, F., Tampieri, F. (2011) Air quality trends in Europe over the past decade: a first multi-model assessment. Atmos. Chem. Phys. 11, 11657-11678. http://dx.doi.org/10.5194/acp-11-11657-2011.

Cooper, O. R., R.-S. Gao, D. Tarasick, T. Leblanc, and C. Sweeney (2012) Long-term ozone trends at rural ozone monitoring sites across the United States, 1990–2010, J. Geophys. Res., 117, D22307, doi:10.1029/2012JD018261.

Cooper, O. R., Parrish, D. D., Ziemke, J., Balashov, N. V., Cupeiro, M., Galbally, I. E., Gilge, S., Horowitz, L., Jensen, N. R., Lamarque, J.-F., Naik, V., Oltmans, S. J., Schwab, J., Shindell, D. T., Thompson, A. M., Thouret, V., Wang, Y., Zbinden, R. M. (2014) Global distribution and trends of tropospheric ozone: An observation-based review. Elementa: Science of the Anthropocene, 2: 000029,
doi:10.12952/journal.elementa.000029elementascience.org

Cristofanelli P., Scheel H.-E., Steinbacher M., Saliba M. Azzopardi F., Ellul R., Frohlich M., Tositti L., Brattich E., Maione M., Calzolari F., Duchi R., Landi T. C., Marinoni A., Bonasoni P. (2015) Long-term surface ozone variability at Mt. Cimone WMO/GAW global station (2165 m a.s.l., Italy). Atmospheric Environment 101, 23-33.

Davis, J., Cox, W., Reff, A., and Dolwick, P. (2011) A comparison of CMAQ-based and observation-based statistical models relating ozone to meteorological parameters, Atmos. Environ., 45, 3481–3487, 2011.

Demuzere, M., Kassomenos, P. and Philipp, A. (2011) The COST733 circulation type classification software: an example for surface ozone concentrations in Central Europe. Theor. Appl. Climatol., 105, 143–166.

Derwent, R. G., Hjellbrekke A-G. (2013) Air Pollution by Ozone Across Europe, in Viana M, ed., Urban Air Quality in Europe. Berlin Heidelberg: Springer. (The Handbook of Environmental Chemistry, vol. 26): pp. 55–74. doi:10.1007/698_2012_163

Derwent, R. G., Manning, A. J., Simmonds, P. G., Spain, T. G., O'Doherty, S. (2013) Analysis and interpretation of 25 years of ozone observations at the Mace Head Atmospheric Research Station on the Atlantic Ocean coast of Ireland from 1987 to 2012 Atmos. Environ., 80, pp. 361–368.

EEA (2013) Air quality in Europe - 2013 report, EEA Report No 9/2013, European Environment Agency (http://www.eea.europa.eu/publications/air-qualityin-europe-2013).

Fernández-Fernández, M. I., Gallego, M. C., García, J. A., & Acero, F. J. (2011) A study of surface ozone variability over the Iberian Peninsula during the last fifty years. Atmospheric Environment, 45, 1946–1959.

Hess, P.G., Zbinden, R. (2013) Stratospheric impact on tropospheric ozone variability and trends: 1990-2009. Atmos. Chem. Phys. 13, 649-674. http://dx.doi.org/10.5194/acp-13-649-2013.

Jacob, D. J., and Winner, D. A. (2009) Effect of climate change on air quality. Atmospheric Environment 43(1): 51-63.

Jenkin, M.E. (2008) Trends in ozone concentration distributions in the UK since 1990: local, regional and global influences. Atmos. Environ. 42, 5434-5445.

Jonson, J.E., Simpson, D., Fagerli, H., Solberg, S. (2006) Can we explain the trends in European ozone levels. Atmos. Chem. Phys. 6, 51-66.

Logan, J.A., Staehelin, J., Megretskaia, I.A., Cammas, J.-P., Thouret, V., Claude, H., De Backer, H., Steinbacher, M., Scheel, H.E., Stübi, R., Fr€ohlich, M., Derwent, R. (2012) Changes in ozone over Europe: analysis of ozone measurements from sondes, regular aircraft (MOZAIC) and alpine surface sites. J. Geophys. Res. 117, D09301. http://dx.doi.org/10.1029/2011JD016952.

Oltmans, S. J., et al. (2013) Recent tropospheric ozone changes – A pattern dominated by slow or no growth, Atmos. Environ., 67, 331–351.

Ordonez, C., Mathis, H., Furger, M., Henne, S., Hglin, C., Staehelin, J., Prevot, A.S.H. (2005) Changes of daily surface ozone maxima in Switzerland in all seasons from 1992 to 2002 and discussion of summer 2003. Atmos. Chem. Phys. 5, 1187e1203. http://dx.doi.org/10.5194/acp-5-1187-2005.

Parrish, D. D., et al. (2012) Long-term changes in lower tropospheric baseline ozone concentrations at northern mid-latitudes, Atmos. Chem. Phys., 12, 11485–11504.

Parrish DD, Law KS, Staehelin J, Derwent R, Cooper OR, Tanimoto H et al. (2013) Lower tropospheric ozone at northern midlatitudes: changing seasonal cycle. Geophys Res Lett 2013,40(8):1631-1636.

Pausata, F.S.R., Pozzoli, L., Vignati, E., Dentener, F.J. (2012). North Atlantic oscillation and tropospheric ozone variability in Europe: model analysis and measurements intercomparison. Atmos. Chem. Phys. 12, 6357-6376.

Philipp A., C. Beck, R. Huth and J. Jacobeit (2014) Development and comparison of circulation type classifications using the COST 733 dataset and software. International Journal of Climatology. DOI: 10.1002/joc.3920.

Sicard, P., De Marco, A., Troussier, F., Renou, C., Vas, N., Paoletti, E. (2013) Decrease in surface ozone concentrations at Mediterranean remote sites and increase in the cities. Atmos. Environ. 79, 705-715.

Simmonds, P.G., Manning, A.J., Derwent, R.G., Ciais, P., Ramonet, M., Kazan, V., Ryall, D. (2005) A burning question. Can recent growth rate anomalies in the greenhouse gases be attributed to large-sale biomass burning events? Atmospheric Environment 39, 2513-2517.

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, Á., and Wind, P. (2012): The EMEP MSC-W chemical transport model – technical description, Atmos. Chem. Phys., 12, 7825–7865, doi:10.5194/acp-12- 7825-2012.

Simpson, D., Arneth, A., Mills, G., Solberg, S., Uddling, J. (2014) Ozone — the persistent menace: interactions with the N cycle and climate change. Current Opinion in Environmental Sustainability 2014, 9–10:9–19.

Solberg, S., Jonson, J.E., Horalek, J., Larssen, S., de Leeuw, F. (2009) Assessment of Ground-level Ozone in EEA Member Countries, with a Focus on Long-term Trends (EEA Report No7/2009). European Environment Agency, Copenhagen.

Thompson, M. L., J. Reynolds, L. H. Cox, P. Guttorp, and P. D. Sampson (2001), A review of statistical methods for the meteorological adjustment of tropospheric ozone, Atmos. Environ., 35, 617 – 630, doi:10.1016/S1352-2310(00)00261-2.

Tørseth K, Aas W, Breivik K, Fjæraa AM, Fiebig M, Hjellbrekke AG et al. (2012) Introduction to the European monitoring and evaluation programme (EMEP) and observed atmospheric composition change during 1972–2009. Atmos Chem Phys 2012, 12(12):5447-5481.

Vautard, R., Szopa, S., Beekmann, M., Menut, L., Hauglustaine, D.A., Rouil, L., Roemer, M. (2006) Are decadal anthropogenic emission reductions in Europe consistent with surface ozone observations? Geophys. Res. Lett. 33, L13810. http://dx.doi.org/10.1029/2006GL026080.

Wilson, R.C., Fleming, Z.L., Monks, P.S., Clain, G., Henne, S., Konovalov, I.B., Szopa, S., Menut, L. (2012) Have primary emission reduction measures reduced ozone across Europe? An analysis of European rural background ozone trends 1996-2005. Atmos. Chem. Phys. 12, 437-454. http://dx.doi.org/10.5194/acp-12-437-2012.

Wood, S. N. (2006) Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC. ISBN 978-1-58488-474-3.

Zheng, J., J. L. Swall, W. M. Cox, and J. M. Davis (2007) Interannual variation in meteorologically adjusted ozone levels in the eastern United States: A comparison of two approaches, Atmos. Environ., 41(4), 705–716, doi:10.1016/j.atmosenv.2006.09.010.

Annex 1: The Mann-Kendall test

For analyzing a possible trend in observed time series the non-parametric Mann-Kendall test (Gilbert, 1987) has been used. This test is particularly useful since missing values are allowed and the data need not to conform to any particular distribution. Moreover, as only the relative magnitudes of the data rather than their actual measured values are used, this test is less sensitive towards incomplete data capture and/or special meteorological conditions leading to extreme values.

In the trend analyses a *consistent set* of stations is used. Requirements for a consistent set are:

- for each year within the time period a minimum data coverage of 75% is required;
- annual data is available for at least 75% of the years within the time period.

The Mann-Kendall statistic *S* is defined as:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \mathrm{sgn}\left(x_j - x_k\right) \qquad [1]$$

Where,

$$\begin{aligned} \mathrm{sgn}(x_j - x_k) &= 1 && \textit{if } (x_j - x_k) > 0 \\ &= 0 && \textit{if } (x_j - x_k) = 0, \\ &= -1 && \textit{if } (x_j - x_k) < 0 \end{aligned}$$

x_j is the observable (concentration, number of exceedance days, exposure) in year j; n is the available number of years with a valid measurement. In other words, *S* is the number of positive differences minus the number of negative differences. If *S* is a large positive number measurements taken later in time tend to be larger than those taken earlier in time. Similarly, if *S* is a large negative number, this indicates a downward trend. The Mann-Kendall statistic is only calculated for consistent sets of stations.

If a linear trend is assumed, the time series can be given as:

$$C_t = B + Q \cdot t + \varepsilon_t \qquad [2]$$

Where,
C_t is the concentration in year t,
B is the intercept,
Q is the slope, and
the residuals ε_t have a zero mean.

The slope is estimated by Sen's non-parametric procedure (Gilbert, 1987). For each time series with n valid measurements a set of slope estimates Q_{jk} is computed for each of the $n(n-1)/2$ data pairs:

$$Q_{jk} = \frac{x_j - x_k}{j - k} \qquad [3]$$

Sen's slope estimate equals the median of the $N=n(n-1)/2$ slope estimates.
To calculate the 95% confidence interval about the true slope, first

$$D_\alpha = Z_{1-\alpha/2}\sqrt{VAR(S)}$$

is calculated. *VAR(S)* is the variance of the Mann-Kendall statistics *S* and is, when no ties are present, calculated by:

$$VAR(S) = \frac{1}{18}\left[N(N-1)(2N+5)\right]$$

$Z_{1-\alpha/2}$ is obtained from the standard normal distribution ($Z_{0.975} = 1.96$).

Next $M_1 = (N-D_\alpha)/2$ and $M_2 = (N+D_\alpha)/2$ are calculated. The lower, Q_L, and upper limits, Q_U, of the confidence interval are the M_1-th largest and (M_2 + 1)-th largest of the *n* ordered slope estimates. If M_1 or M_2 is not an integer value, the lower or upper limit is interpolated. The standard uncertainty σ_i is given by:

$$\sigma_i = \frac{1}{4}\left(Q_U - Q_L\right) \qquad [4]$$

To obtain an estimate of *B* (the concentration at $t=0$), equation [2]) the *n* values of the differences $x_i - Q\,t_i$ are calculated. The median of these values gives an estimate of *B*.

The uncertainty of region-averaged trends are reported with 2σ errors calculated through error propagation from the uncertainties in trends at the individual stations (σ_i, equation [4]) by means of Eq [5] where N_R is the number of individual stations in region R:

$$\sigma_R = \frac{\sqrt{\sum \sigma_i^2}}{N_R} \qquad [5]$$

To test for homogeneity of trend direction at *M* multiple stations, compute the homogeneity chi-square statistic, χ^2_{hmg}, where

$$\chi^2_{hmg} = \chi^2_{total} - \chi^2_{trend} = \sum_{j=1}^{M} Z_j^2 - M\overline{Z}^2$$

$$Z_j = \frac{S_j}{\sqrt{VAR(S_j)}}$$

S_j is the Mann-Kendall trend statistic for the *j*-th station, and

$$\overline{Z} = \frac{1}{M}\sum_{j=1}^{M} Z_j$$

If the trend at each station is in the same direction, then χ^2_{hmg} has a chi-square distribution with M-1 degrees of freedom (df). To test for rend homogeneity between stations at the α significance level, the calculated value of χ^2_{hmg} is referred to the α critical value with M-1 degrees of freedom in a chi-square distribution table. If χ^2_{hmg} exceeds this critical value, the H_o hypothesis of homogeneous station trends is rejected. In this case no regional-wide statements should be made about trend direction. If χ^2_{hmg} does not exceed the α critical value in the chi-square distribution table, then the statistics $\chi^2_{trend} = M\overline{Z}^2$ is referred to the chi-square distribution with 1 df to test the null hypothesis that the (common) trend direction is significantly different from zero.