Trend Analysis of Ozone and Nitrogen Oxides in Sydney Using A Long Range Dependence Time Series Model

M. Azzi[a] and **H. Duc[b]**

[a] CSIRO, Division of Energy Technology PO Box 136, North Ryde, NSW 2113.

[b]Environment Protection Authority of New South Wales PO Box 29 Lidcombe, NSW 2141.

Abstract: This paper describes the use of a long-range dependence time series model to determine ozone and nitrogen oxides (NO_2 and NO_x) trends at a number of monitoring sites in Sydney. The technique removed seasonally, auto-regressive and moving average dependence from the time series. The trend was then modelled using the fractional long-term dependent component and can be determined or detected at a very small resolution of concentration level. Other recent techniques of finding long term trend by removing the effect of meteorology are also described. These techniques, such as the Rao-Zurbenco method, were shown to be equivalent to the long-range dependence method in term of the result of the trend analysis.

Keywords: *Long range dependence; ozone trend; Ozone precursor emission; KZ filter*

1. INTRODUCTION

Ozone is a photochemical pollutant generated in the troposphere in the presence of sunlight from the photochemical reactions between ozone precursors nitrogen oxides (NO_x) and volatile organic compounds (VOCs). Trend in ozone levels can be determined by performing various analyses on ozone data collected over many years at different monitoring sites. This can provide an insight as to whether the implemented control strategies have been effective in reducing ozone levels or if adjustments are required.

To detect changes in ozone precursor emissions, the meteorological effect should be removed from the ozone data. There are a number of methods, which were published recently, to find the trend of ozone by removing the meteorological effects. These methods (Rao and Zurbenco [1994]; Flaum , Rao and Zurbenco [1996]; Xu ,Yap and Taylor [1996]) mainly use different models containing various meteorological variables, which will be then removed to isolate the long-term trend component from the other components in the models of the time series.

Other recent advanced methods to filter out the different frequency scales to determine the long-term trend have been used on air quality, water and climate data with some success, such as the Kolmogorov-Zurbenco filtering approach (Rao, Zurbenco et al [1997]) and the wavelet transform (Whitcher [2000], Lau, Weng [1995]).

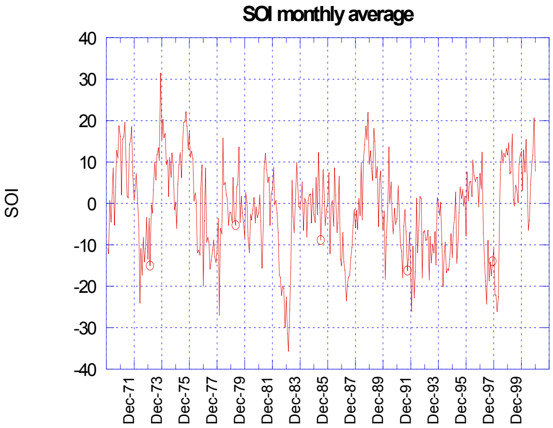
One particular useful method that can be applied to many different air pollutant time series is the Long Range Dependence (LRD) method. This method identifies the LRD component in the fractional ARMA model as the anthropogenic trend (Anh, et al. [1997]).

In this paper, the LRD model is applied to various air pollutant time series in the Sydney area to find the long-term anthropogenic trends. The emission from the motor vehicles in the Sydney basin plays an important role in interpreting the significance and implication of these trends.

2. EFFECTS OF METOROLOGY ON AIR QUALITY

As ozone level is highly dependent on temperature, it is important to look at the various mechanisms that can effect the regional temperature. Besides the diurnal, synoptic and seasonal time scales, there is another longer time scale: the global scale driven by the El-Nino Southern Oscillation (ENSO). This phenomenon is mostly strongest in the Southern part of the Pacific Ocean. Its effect is causing a spell of drought with high temperature in Eastern Australia and high rainfall, cooler climate in the West Coast of North America. As illustrated in Figure (1) the monthly average of the Southern Oscillation Index (SOI) from 1970 to 2000. High negative SOI values indicate the El-Nino effect while positive ones, the La-Nina phenomenon. The 4 recent El-Nino phenomena are the 1982-1983, 1987-1988, 1991-1994 and 1997-1998 periods with the 1982-1983 period as the strongest. For Sydney, there is a high degree of correlation between the number of days above goals with the SOI values in the 4 recent El-Nino periods (Figure 2).

Figure1-Monthly average of Southern Oscillation



Index (SOI) from 1970 to 2000

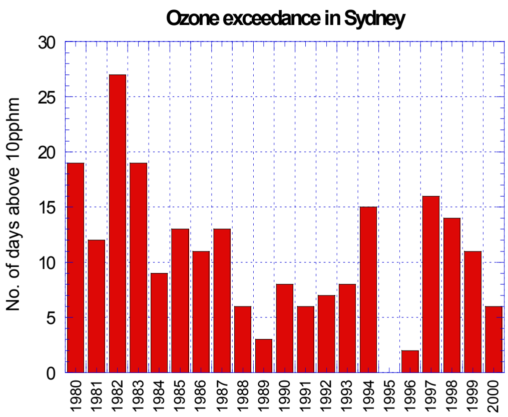


Figure 2–Number of days above ozone goal (10pphm) in Sydney

It is evident that removing the temperature effect in the ozone level will reduce significantly the meteorological influence due to natural processes. The remaining effect is then mainly due to anthropogenic sources. The removal of temperature effect was modelled and applied to ozone data by various authors using different techniques such as Rao, Zurbenco [1994], Anh, Duc, Azzi [1997], Milanchus, Rao, Zurbenco [1998], Flaum, Rao, Zurbenco [1996], Xu, Zap, Taylor [1996].

3. MODELLING ANTHROPOGENIC TRENDS USING RAO-ZURBENCO METHOD

Rao and Zurbenco [1994] used the ozone and temperature time series to find the meteorologically adjusted ozone trend by using filtering and regression techniques. A time series $X(t)$ is assumed to be represented as

$$X(t) = e(t) + S(t) + W(t) \quad (3.1)$$

Where $e(t)$ is the trend component, $S(t)$ the seasonal component and $W(t)$ white noise. The random variations $W(t)$ can be removed from the series by a simple iterative application of a moving average filter:

$$Y(i) = \frac{1}{m}\sum_{j=-k}^{k} X(i+j), \quad m = 2k+1, \quad (3.2)$$

where $Y(i)$, the output of the first iteration, then becomes the input for the next iteration of (3.2). The number of iterations (p) and the filter width value m are to be determined from the data to achieve noise-free series. This p-iterative application of a moving average filter of width m is called the Kolmogorov-Zurbenko filter, KZ(m,p).

The effect of meteorological variability on the air pollutant time series has to be removed prior to any trend analysis. In the Rao-Zurbenko method for daily ozone series, the meteorological effects are represented by the maximum daily temperature. Both the ozone and the temperature time series are first filtered to remove the noise using the Kolmogorov-Zurbenco filter, KZ(m,p). The meteorological effects can then be removed by using the regression technique.

To be specific, denote the filtered log of ozone concentrations by $O_{kz}(t)$ and the filtered temperature by $T_{kz}(t)$. Then the meteorological effects, represented by the seasonal component, are removed from filtered log of ozone by the linear regression:

$$O_{kz}(t) = a + b\,T_{kz}(t) + \varepsilon(t) \quad (3.3)$$

The noise term $\varepsilon(t)$ then represents changes in ozone attributable to changes in emissions.

The Rao-Zurbenco method can be applied to any pollutant, which exhibits temperature or seasonal dependence. However, another method called the Long Range Dependence (LRD) model as described in the next section can be applied to any long term time series.

4. MODELLING THE LRD COMPONENT

The recent literature on air pollution modelling has paid attention to the long-range dependence in air quality data. It has now been established that the LRD phenomenon is present in air quality, meteorological, hydrological and geophysical data (Beran [1992], Ooms and Franses [2001], Haslett and Raftery [1989]).

A stochastic process $X(t)$ is said to exhibit LRD if its spectral density has the form

$$f(\omega) = f_*(\omega)\ \omega^{-2\beta}, \quad \beta > 0,\ \omega \in \Re \qquad (4.1)$$

where $f_*(\omega)$ is slowly varying as $\omega \to 0$. The spectral density has an integrable singularity at the origin if $0 < \beta < ½$ with the characteristic effect that the autocovariance function of $X(t)$ decays to zero at a very slow rate so that the autocorrelation function is not absolutely summable.

The significant component at a very low frequency shows that the time series contains a slow varying trend, which is not easily detected and removed using standard time series analysis such as autoregressive and moving average (ARMA) or autoregressive integrated moving average (ARIMA). In fact, the presence of LRD invalidates many of the traditional methods of data description using autoregressive and moving average (*ARMA*) models (Beran [1992]).

A discrete stationary approximation of the LRD factor $\omega^{-2\beta}$ of (4.1) is

$$f(\omega) = \frac{\sigma^2}{2\pi}\frac{1}{\left|1-e^{i\omega}\right|^{2d}}, \quad \sigma^2 > 0,\ 0 < d < ½,\ \omega \in (-\pi,\pi]$$

(see Anh and Lunney [1995]).

Therefore the LRD and short-memory components of a discrete time series $X(t)$ can be modelled by a fractional *ARMA (p,d,q)*:

$$\begin{aligned}&(1-B)^d\left(1-\theta_1 B - \ldots - \theta_p B^p\right)X(t) = \\ &\left(1+\phi_1 B + \ldots + \phi_q B^q\right)\varepsilon(t)\end{aligned} \qquad (4.2),$$

Where B is the backshift operator $BX(t) = X(t-1)$, d is the LRD parameter, $\varepsilon(t)$ is white noise with variance σ^2.

The spectral density of the time series generated by model (4.2) is

$$f(\omega) = \frac{\sigma^2}{2\pi}\frac{1}{\left|1-e^{i\omega}\right|^{2d}}\frac{\left|1-\phi_1 e^{i\omega} - \ldots - \phi_q e^{iq\omega}\right|^2}{\left|1-\theta_1 e^{i\omega} - \ldots - \theta_p e^{ip\omega}\right|^2}, \omega \in (-\pi,\pi]$$

The Sydney pollutant series appear to have additive seasonally, suggesting a model of the form

$$X(t) = S(t) + R(t)$$

Where $S(t)$ is the seasonal component and $R(t)$ is the random component. Also due to large variations in the seasonal component, particular in the summer period, it is necessary to use a Box-Cox transform

$$Y(t) = \frac{X^\alpha(t)-1}{\alpha}, \ \alpha > 0 \quad (4.3)$$

to stabilise the variance. A special form of the Box-Cox transform is the logarithmic transform as $\alpha \to 0$.

The average of the Box-Cox transform of the daily maxima over all years for each day of the year is then regressed on a set of annual harmonics. Substraction of the estimated seasonal effect from the Box-Cox transform of the daily maxima then yields the seasonally adjusted series ready for trend analysis. Thus, the series is seasonally adjusted using the yearly profile of the transformed series. For ozone and nitrogen oxides data series in Sydney, the choice of $\alpha = 0.2$ based on (4.3) has been proved as appropriate (Anh, Duc, Azzi [1997]).

The Haslett-Raftery algorithm (Haslett and Raftery [1989]) can be invoked to estimate *d* and the *ARMA* coefficients of (4.2) simultaneously on the seasonally adjusted series. Removing the short-memory *ARMA* component from the estimated model (4.2) will then give the LRD component for trend analysis. The Haslett-Raftery algorithm and the associated computing program are readily available for use since its publication.

It has been proved that the trend as derived by using the Rao-Zurbenco method is the same as the LRD component of the series (Anh, Duc, Azzi [1997]) using the LRD method above for both the ozone and NO_2 series at a monitoring site in Sydney. The LRD component of the pollutant time series, corresponding to the low frequency component of the time series, gives the same result as the trend obtained by removing the effect of meteorology (with temperature as the dominating variable).

Another method of analysing the LRD time series is using the wavelet transform. Whitcher [2000] uses the Discrete Wavelet Packet Transform (DWPT) on the monthly CO_2 data series to estimate the parameters of a fractional seasonal long memory model called the seasonal persistent process (SPP) of the form

$$(1-2\phi B+B^2)^d X_t=\varepsilon_t \qquad (4.4)$$

where $\phi=\cos(2\pi f_G)$ and $\{\varepsilon_t\}$ is Gaussian white noise with variance σ_ε^2. The spectral density of time series X_t is given as

$$S(f)=\sigma_\varepsilon^2\left\{4(\cos(2\pi f)-\phi)^2\right\}^{-d}, \text{ for } -\frac{1}{2}<f<\frac{1}{2}$$

The spectral density has singularity at $|f_G|<1/2$. The process $\{X_t\}$ is stationary and invertible for $|\phi|=1$ and $-1/4<d<1/4$ or $|\phi|<1$ and $-1/2<d<1/2$

The SPP process includes the seasonal effect into the model rather than removing it before analysis. It is similar to the seasonally adjusted fractional ARIMA process described above. When ϕ=1, the SPP process is equivalent to the fractional AR process.

4.1. APPLICATION OF LRD MODELLING TO OZONE AND NO_x TIME SERIES

As discussed above, due to variability of both meteorological variables and pollutant data, the regression method is of limited use in finding the linear trend due to anthropogenic emission. Time series of monitoring data for Ozone and NO_x collected at a number of stations were used to study their trends using the LRD method. In 1998, the Sydney basin has 19 monitoring stations located throughout the region (Figure 3).

The daily maximum values for ozone and NO_x are used in the trend analysis. Missing data are either interpolated (less than 3 missing points) or replaced with average seasonally values in the series. As with ozone data, the NO_x and NO_2 time series were exhibiting high seasonally and temperature dependence. Therefore, these series data were transformed using Box-Cox transform to stabilise the variance before being analysed to find the trends.

The modelling of the LRD component on the ozone, NO_x and NO_2, after removing the seasonally variation, shows that each of the series can be represented by an autoregressive (AR) model of order 3, a moving average (MA) of order 1 and a long range dependence (as represented by a fractional coefficient) component.

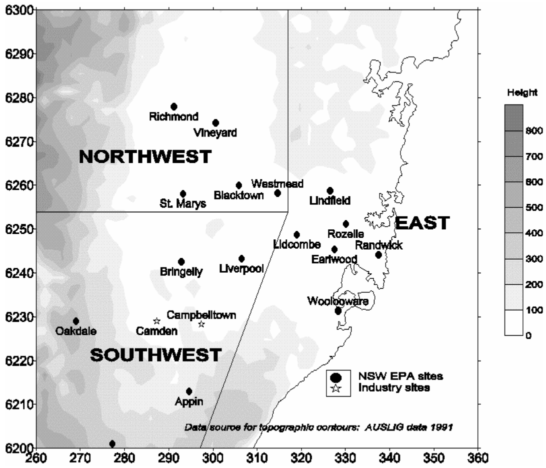


Figure 3. Sydney air quality monitoring network

The trend part of the time series is taken to be the LRD component. To see the best trend pattern, a smoothing process using Kolmogorov-Zurbenco (KZ) filter (Rao, Zurbenco [1994] and Rao, Zurbenco et al [1997]) with 450 data points, KZ(450,1), were applied to Ozone and NO_x trend components .

The smoothed trends of Ozone and NO_x for some sites are shown in Figure 4 and 5 below. In all the trend graphs, the values were obtained after an inverse transform of the transformed series. The trend values are relative to the long-term average value indicated by the unit value of 1.

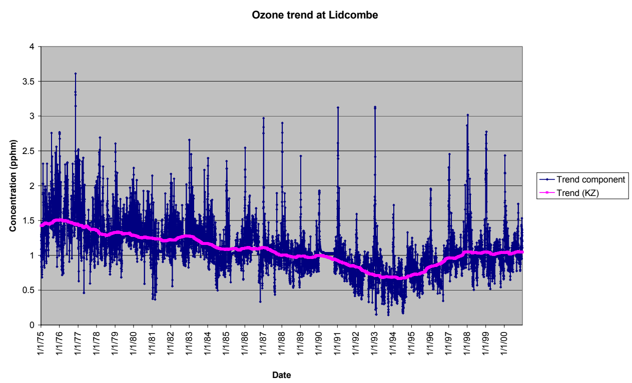


Figure 4 Ozone trend at Lidcombe (1975 - 2000)

The trend of ozone at a number of sites shows a consistent picture, an increase trend from about 1994 is observed at all sites (except Campbelltown, which has an increasing trend from 1996). The magnitude of the increase is largest at Lidcombe and smallest at Woolooware.

The overall trend from the ozone data representing all the sites in the Sydney region (1993-2000) is mostly stable from 1994 with larger increase from 1996 and a stabilising trend from 1998.

For NOx, the pattern is similar for sites in eastern Sydney (Lidcombe, Rozelle, Earlwood), except Woolooware with no trend. An increase level of nitrogen oxides from 1991 or 1992 to 1994 and then a decreasing trend from 1994 to 1998 is observed at these sites. From 1998 to 2000, the trend at these sites is stabilised. This decreasing trend from 1994 to 1998 at these sites could be due to the improvement in the emission of new vehicle fleet following from the introduction of catalytic converter from early 1989. But the rising number of motor vehicles could offset this gain since 1998 as shown by the trend after 1998. In the south west of Sydney, there is a significant increase (of about 1pphm) in the levels of NOx at Campbelltown (1991-1994) and Liverpool (1993-1995) compared to other sites before the trend is stabilised. A local nitrogen oxides source operating in the area could be the reason for this large increase.

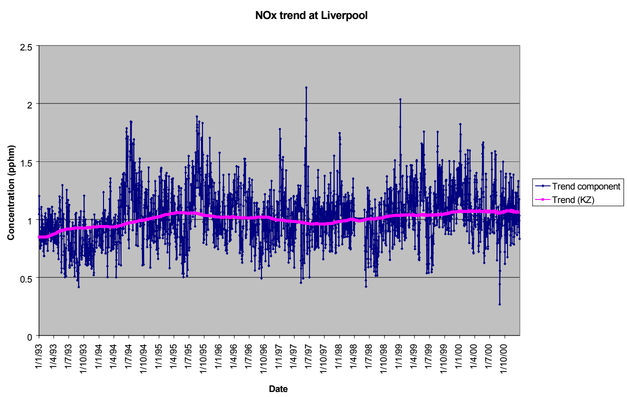


Figure 5 – NO_x trend at Liverpool (1993-2000)

4.2. DISCUSSION AND CONCLUSION

The trends of various pollutants at a number of sites in the Sydney area are presented. For ozone the trends are more consistent across all sites. It is probably due to the reason that this pollutant is more regional and well mixed with widespread, well-dispersed sources compared to others (Huang et al. [2000]). Trend of photochemical smog precursor, NO_x, can be corroborated with change in the emission inventory over the time period under consideration (Wolff et al. [2001]). However, the only known available emission inventory for the Sydney region is the 1991-emission data set. Such corroboration is therefore not possible.

Overall, it can be seen that the levels of nitrogen oxides are declining since 1994 and then stabilise toward the end of 1998. But for ozone, an increasing trend is observed since 1994 for all sites. From the photochemical point of view, this could be explained by an increasing emission of volatile organic compound (VOC) across the Sydney basin or a decreasing level of nitrogen oxides where the extent of reaction is less than an optimum value. Blanchard [2000], Blanchard and Stoeckenius [2001] have shown that following a NO_x control, increases in peak ozone concentration can happen in some areas where the extent is less than 0.6.

Since the introduction of unleaded petrol fuel, an increasing level of VOC is a strong possibility. Indeed, Bravo, Torres [2000] has shown that since the introduction of reformulated gasoline, the ozone level is worsened in Mexico City. To determine whether this is the case, the monitoring of VOC at a number of locations in the Sydney basin is necessary, as VOC data is not currently monitored continuously in Sydney.

There is limitation about the LRD method to find air quality trend due to anthropogenic sources free from the meteorological effects. In most situations, where the data period for analysis is usually about 10 (or > 10) years, it is effective in isolating and removing short-term climate variability on the seasonal and inter-annual scales. But for long-term climate changes (such as the global warming on inter-decade scale), it may not be possible to separate these sources of climatic variability as they are at about or below the lowest frequency range that can be resolved with the time window of the available data. In other words, they are at about the same lowest frequency that can be attributed to the anthropogenic sources.

These long-term climatic scales only affect certain air pollutants, such as ozone, which is strongly dependent on temperature for its production. In Sydney, analysis of the temperature daily maximum data from 1980 to 1998 at Mascot airport near the coast, using LRD model, shows that the series does not exhibit LRD ($d \approx 0$) and there is virtually no trend in temperature at this site. If this is also typical at other sites then the effect of the trend in long-term climatic change is too small or not detectable. The air quality trend of various pollutants, especially ozone and nitrogen oxides, described above are due entirely to anthropogenic sources.

The separation of the low frequency component, identified as the trend of air pollutant time series, from other meteorological oscillations in the frequency spectrum has been recently used extensively to find the trend of ambient air quality monitoring data (Anh, Duc, Azzi [1997], Milanchus, Rao, Zurbenco [1998], Rao et al. [1997], Kuebler, Bergh, Russell [2001], Porter et al [2001]). This frequency separation is necessary to detect the very small trend signal buried inside the much stronger natural forcing components in the monitored data (Porter et al. [2001]. Various methods have been developed and these include the Kolmogorov-Zurbenco (KZ) filtering, the LRD method and the wavelet method.

The Long Range Dependence model for air pollution time series has been shown to be useful in detecting the trend due to anthropogenic emission. It also has been shown that this method is equivalent to other methods for finding ozone trend but has the advantage that it can be applied to other pollutants as well. The method is applied to find the trend of Ozone and NO_x time series at a number of stations in the Sydney basin.

5. REFERENCES

Anh V., Kavalieris L., Long-range dependence in models for air quality, Statistics in Ecology and Environmental Monitoring, D.J. Fletcher and B.F.J. Manly (Eds.), University of Otago Press, Dunedin, pp. 199-209, 1994.

Anh V., Lunney K., Parameter estimation of random fields with long-range dependence, Mathematical & Computer Modelling 21: 67-78, 1995.

Anh V., Nguyen K.L., Duc H., Stochastic models for prediction of pollutant ground concentrations, Appita J., 48:33-36, 1996.

Anh V., Duc H., Azzi M., Modelling anthropogenic trends in air quality data, J. Air & Waste Management Association, 47:66-71, 1997.

Beran J., Statistical models for data with long-range dependence, Statistical Science 4:404-427, 1992.

Blanchard, C., Ozone process insights from field experiments – Part III: extent of reaction and ozone formation, Atmospheric Environment, 34 (2000), pp. 2035-2043, 2000.

Blanchard C., Stoeckenius T., Ozone response to precursor controls: comparison of data analysis methods with the predictions of photochemical air quality simulation models, Atmospheric Environment, 35 (2001), pp. 1203-1215, 2001.

Bravo, H.A, Torres, R.J, The usefulness of air quality monitoring and air quality impact studies before the introduction of reformulated gasolines in developing countries. Mexico City, a real case study, Atmospheric Environment, 34 (2000), pp. 499-506, 2000.

Flaum J., Rao S., Zurbenco I., Moderating the influence of meteorological conditions on ambient ozone concentrations, J. Air & Waste Management Association, 46:35-46, 1996.

Haslett J., Raftery A., Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource, Applied Statistics 38:1-50, 1989.

Hogrefe C., Rao S. T., Zurbenco I., Porter P. S., Interpreting the information in ozone observations and model predictions relevent to regulatory polices in the Eastern United States, Bulletin of American Meteo. Soc., Vol. 81, No. 9, pp. 2083-2106, Sept. 2000.

Huang Yu-li, Batterman S., Residence location as a measure of environmental exposure: a review of air pollution epidemiology studies, J. of Exposure Analysis and Environment Epidemiology, Vol. 10 (2000), pp. 66-85. 2000.

Rao S.T., Zurbenko I.G, Detecting and tracking changes in ozone air quality, J. Air & Waste Management Association, 44:1089-1092, 1994.

Milanchus M., Rao S., Zurbenco I., Evaluating the effectiveness of ozone management efforts in the presence of meteorological variability, J. Air & Waste Management Association, 48:201-215, 1998.

Lau K. M., Weng H., Climate signal detection using wavelet transform: how to make a time series sing, Bull. of American Meteorological Society, Vol. 76, No. 12, pp. 2391-2402, Dec. 1995.

Ooms M., Franses P., A seasonal periodic long memory model for monthly river flows, Environmental Modelling & Software, 16(2001), pp. 559-569, 2001.

Porter P., Rao S., Zurbenco I., Dunker A., Wolff G., "Ozone air quality over North America: Part II – An analysis of trend dectection and attribution techniques", J. Air & Waste Management Association, 51(2001), pp. 283-306, 2001.

Rao S. T., Zurbenco I. G., Neagu R., Porter P.S, Yu J.Y, Henry R.F., Space and time scales in ambient ozone data, Bull. of American Meteorological Society, Vol. 78, No. 10, pp.2153-2166, Oct. 1997.

Whitcher, B., Wavelet analysis of seasonal long memory. In V. Núñoz-Antón and E. Ferreira (Eds.) *Statistical Modelling*, Proceedings of the 15th International Workshop on Statistical Modelling, pp. 276-281, Bilbao, Spain: Servicio Editorial de la Universidad del País Vasco, 2000.

Wolff G., Dunker A., Rao S., Porter P., Zurbenco I., Ozone air quality over North America: Part I – A review of reported trends, J. Air & Waste Management Association, 51(2001), pp. 273-282, 2001.

Xu D., Yap D., Taylor P., Meteorologically adjusted ground level ozone trends in Ontario, Atmospheric Environment, 30:1117-1124, 1996.