The West Glamorgan asthma study: A statistical analysis.

Beer, Helen

Beer, Helen (1998) *The West Glamorgan asthma study: A statistical analysis..* thesis, Swansea University.
http://cronfa.swan.ac.uk/Record/cronfa42691

THE WEST GLAMORGAN ASTHMA STUDY
A STATISTICAL ANALYSIS

by

HELEN BEER

Thesis submitted in fulfilment of the requirements of the
University of Wales
for the degree of
Master of Philosophy

European Business Management School
University of Wales Swansea
Singleton Park, Swansea, SA2 8PP
United Kingdom

September 1998

This Work is Dedicated to Mum, Dad, Martin and Janet

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed..(candidate)
Date......September 1998...............

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by explicit references. A bibliography is appended.

Signed..(candidate)
Date......September 1998...............

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and inter-library loan, and for the title and summary to be made available to outside organisations.

Signed..(candidate)
Date......September 1998...............

ACKNOWLEDGEMENTS

I wish to express my warmest gratitude to my supervisor Dr. Alan M. Sykes for his guidance, encouragement and APL expertise.

I wish to thank Iechyd Morgannwg, Swansea City Council, West Glamorgan County Council and Neath Borough Council for providing the data for this research and the European Business Management School for providing the facilities.

Many thanks to Iechyd Morgannwg, Swansea City Council, West Glamorgan County Council, Neath Borough Council and the European Business Management School for providing the financial assistance during the course of this research.

Finally I would like to thank my family and friends, especially my mum, for their continuous support. Also a big thank you to Mary whose APL expertise was much appreciated.

ABSTRACT

This thesis is based largely on the analysis of a large data set collected by Iechyd Morgannwg, West Glamorgan County Council and Neath Borough Council, with the aim of discovering whether air borne pollution caused any significant health effects on the lung function of school children aged between 8 and 11 years.

Individual time series of daily Peak Expiratory Flow Rate measurements of school children at four different sites in West Glamorgan are related to changes in air borne pollutants such as Nitrogen Oxides, Sulphur Dioxide, Carbon Monoxide, Ozone and PM10 (particulate dust).

Little work has previously been conducted on the acute health effects of air pollution, particularly with respect to children. Recent studies have attempted to model such data, each offering various methodology. This research, however, looks at the longitudinal nature of the data and offers an insight into ways to model it. All modelling was done using the statistical programming language APL.

Unlike most other studies the data, upon which this research was based, was plentiful. Many pollutant levels were measured as well as the weather conditions. The children's peak expiratory flow rate readings were accompanied by a questionnaire about the child's previous health symptoms and their home environment. Daily symptom diaries for each day of the study were also kept. This allowed a full and comprehensive analysis to be undertaken, where all confounding factors could be considered.

Longitudinal Statistical methods, including the use of Restricted Maximum Likelihood (REML), are applied to investigate the effects of the different pollutant variables.

LIST OF FIGURES

Figure

LIST OF TABLES

Table

TABLE OF CONTENTS

1. Introduction

1.1 Asthma and Air Pollution.

Asthma is a life-threatening disease which affects 10% of the population. The cost to Britain every year due to asthma has been quoted to be [1];

- 4 million working days a year
- £400m in lost productivity
- £350m in prescription costs

Asthma kills about 2000 people a year, and in 1994 prescriptions accounted for 10% of all NHS spending on drugs (£381 million) [2].

Asthma symptoms are the immune system's response at attempts to eliminate foreign matter from the body. Cells in the airways leading to the lungs react with these irritants causing swelling and narrowing which results in wheezing and breathlessness.

It is important in order to control respiratory illnesses, especially asthma, to understand what triggers an asthma attack. Diet, allergies (e.g. pollen, pets), stress, excitement, viral infections (e.g. throat and chest), pollutants, Environmental Tobacco Smoke (ETS), exercise and cold weather are all believed to be possible triggers of asthma, (there are over 200) [1].

A report published in October 1995 by the Committee on the Medical Effects of Air Pollution (COMEAP), states that avoiding exposure to tobacco smoke, furry pets or house dust mite during infancy may reduce the chances of developing asthma [3]. The use of central heating and closed windows leads to the build up of air and hence allergens accumulate in the home [1].

There is a link between poverty and severe asthma. People living in poorer houses, in particular in inner city areas, could well experience damp, mould, inadequate heating or even cockroaches which all trigger asthma and other related symptoms. There is usually more traffic pollution in these areas and occupants tend to smoke more. These factors not only trigger asthma but reduce the effectiveness of asthma medication [4].

For asthmatics, avoiding the above mentioned triggers in later life could reduce symptoms. The National Asthma Campaign believes that air pollution may be the most important trigger of asthma attacks. The recent hot summers have meant a build up of air pollution which can irritate the airways of susceptible people [5].

Asthma is on the increase, particularly in children. There has been some debate as to whether this increase is genuine or whether it is because of increased willingness for doctors to diagnose patients as asthmatic. 10% of British primary schoolchildren are diagnosed as asthmatic. A further 5-10% have the recurrent respiratory symptoms which suggest asthma but remain undiagnosed. American research suggests that two thirds of children with asthma will continue having attacks as adults [6].

In recent years much publicity has been given to air pollution and possible health effects. Air pollution is thought to play a significant part in an increased incidence of reported asthma cases in recent years [1]. In particular, Dr Mark Temple (then a General Practitioner), noticed an increase in the number of cases of acute episodes of asthma in the Glyn Neath area consistent with an increase in open cast coal mining activities.

Many studies have attempted to identify links between PM10s, (very small dust particles emitted from diesel vehicles), and the incidence of respiratory diseases including asthma [7]. Disease and deaths from particulate pollution are on the increase [8]. There have also been calls for tougher limits on exhaust emissions, and for local governments to alleviate traffic congestion in urban areas [9], but governments have found themselves under pressure from the motor industry. The European Union agreed to postpone tough new limits on emissions from some of the worst polluting lorries, but are in the process of trying to tighten those limits instead. Rather than simply limiting the number of peak concentrations of PM10, it had been suggested that there should be progressive lowering of the average concentrations in our cities throughout the year. Motor vehicles account for most of the PM10 produced during the winter [10].

A study was set up in West Glamorgan involving various local council authorities, based upon the papers [11, 12, 13], to assess the impact of air pollution on the lung function of schoolchildren in and around the Swansea area. A report was prepared as an immediate result of this study, based upon the work contained in this thesis. The report was completed and handed to West Glamorgan Health Authority on 7th May 1997.

1.2 An Overview of the Study.

The Basic Questions Requiring Answers

1. Does daily variation in ambient air pollution levels induce daily variation in lung function or other health symptoms?

2. Which pollutants are implicated?

3. Does the relationship vary

- by season?
- by presence or absence of lung disease (e.g. asthma)?
- by other individual factors (e.g. sex, age, height, ethnic group, location)?

4. Do meteorological conditions exaggerate any pollution effects?

5. Does the relationship between air pollution and lung function involve a time lag?

6. Are these lag effects similar

- for different pollutants?
- for different locations?

7. Is the effect of a pollutant more pronounced in the presence of other health symptoms (e.g. Cough, Sneeze, Wheeze)?

8. Does exposure to Environmental Tobacco Smoke (ETS), gas fires, pets or other environmental factors exaggerate the effects of air pollution?

The Children

426 children living in three areas in West Glamorgan were involved in this study, ranging from 8 to 12 years of age. Every child who agreed to take part in the study was given an identity code to preserve anonymity in the study databases. Their school, sex, age, form, height and ethnic group were noted so that investigations of any pollutant effects could be contrasted for different groups of children.

The Geography

There were three areas of study as outlined below;

- Area 1 (Site 1) = Bishopston Primary School.
- Area 2 (Site 2) = Cwmnedd Primary School, Glyn Neath & (Site 3) = Ysgol Gynradd Gymraeg Cwm Nedd, Glyn Neath.
- Area 3 (Site 4) = Terrace Road School, Swansea City Centre.

The two schools in Glyn Neath were situated next door to each other and were used so that the total number of children in both of these schools together was of the same order as the totals from the other areas.

Appendix 2.1 presents appropriate maps showing the regions involved in the study.

The Questionnaire

A two-page questionnaire, (see Appendix 2.2), was sent out to each child's parents. It gathered information regarding past symptoms such as wheezing or breathing difficulties, especially those suffered in the past 12 months. More general information on their home environment e.g. the presence of gas cookers, the method of heating the home, Environmental Tobacco Smoke (ETS), and pets was also requested.

The Pollution Measurements

Daily pollution measurements were made in the school yards of Swansea, Bishopston and one of the Glyn Neath schools. Daily pollen counts were to have been recorded but are largely missing. The following pollutant levels were measured at 15 minute intervals.

- NOx, NO, NO2 - The Nitrogen Oxides
- SO2 - Sulphur Dioxide
- CO - Carbon Monoxide
- O3 - Ozone
- PM10s - Dust Particulates

The dates of collection of the pollution data were:-

The winter study: 1st January 1995 - 28th February 1995
The summer study: 1st June 1995 - 1st August 1995

All the data, with the exception of the pollen counts, were entered into tables using the Paradox database. Each individual observation was coded as to its quality (e.g. bad data, missing data, negative readings, incorrect due to equipment failure or power failure). All data of bad quality, for whatever reason, were not used. The daily pollution averages were calculated in Paradox using only the remaining good data. (See Chapter 4.2 for the data available.)

Bishopston is an area west of Swansea relatively pollution-free (except for high ozone levels during the summer months) and was selected as a control site. Glyn Neath is situated in a valley, is close to an open cast coal mining site and has a major trunk road (A465) passing through the town. Sulphur Dioxide levels are high in Glyn Neath from coal burning and traffic emissions. Swansea City Centre has pollution problems caused by the large amount of traffic movement throughout the city every day. Industrial sites such as the BP Oil Refinery at Llandarcy may potentially effect all sites depending on the wind direction.

The Weather Measurements

Weather data was gathered for Glyn Neath and Swansea. Given the close proximity of Bishopston and Swansea it was thought that the climatic differences between these two sites would be minimal, and hence the Swansea weather data would also apply to Bishopston. For both sites therefore the recorded variables were

- Relative humidity
- Temperature
- Wind Speed
- Wind Direction

Measurements were recorded every 15 minutes, for all days as specified above. Any suspicious readings were recorded and subsequently ignored in the analysis.
For the Swansea site, additional measurements were also recorded for:-

- Hours of sunlight
- Amount of rainfall
- Air Pressure

People Involved in the Study

Dr Ronan Lyons, Dr Hilary Fielder and Mr Martin Heaven
Department of Public Health Medicine, West Glamorgan Health Authority

Mr Huw Morgan and Mr Philip Govier
Environmental Health Department, Swansea City Council

Mr Adrian Erikson
Environment and Highways Department, West Glamorgan County Council

Mr Peter Davies and Mr Martin Hooper
Environmental Health Department, Neath Borough Council

Dr Alan Sykes and Ms Helen Beer
Statistics Section of the European Business Management School
University of Wales Swansea

Funding for the study

Funding for this study was received from:-

1. Department of Public Health Medicine, West Glamorgan Health Authority
2. Environmental Health Department, Swansea City Council
3. Environment and Highways Department, West Glamorgan County Council
4. Environmental Health Department, Neath Borough Council
5. British Coal Opencast
6. European Business Management School, University of Wales Swansea.

All contributors are gratefully acknowledged.

2. Literature Review

There are many theories about what actually causes asthma. Poverty [14, 15], gas stoves [16], maternal smokers [17, 18], house dust mite [19, 20], sulphur dioxide [15] and even the measles vaccination given in childhood [21] have all been implicated. A British survey carried out by the team from the primary health research centre in London, into child health (1994), has found a statistical link between the pertussis vaccine, for whooping cough, and childhood asthma. Results summarised in the Journal of the American Medical Association demonstrated that a child who received the pertussis vaccine was 5.43 times more likely to suffer from asthma [21]. The effects of coal-dust [22, 23, 24] and environmental tobacco smoke [25, 26, 27] have also been shown to cause significant reductions in lung function, as would perhaps be expected.

The concerns over air pollution causing health effects, and in particular respiratory disease, have risen over the past years. With the number of vehicles on the road increasing, greater public awareness has prompted many questions and demands regarding the safety of the traffic pollution. In turn this has prompted the need for more studies to be carried out, and various air pollutants (both from traffic and from industry), to be examined in greater detail.

Air is said to be polluted if it “contains substances which may have a harmful effect on the environment and on health or cause a nuisance” [28]. Air pollution can cause immediate discomfort by eye irritation or cough as well as more long term effects, such as bronchitis, or premature death.

Individuals have different responses to air pollution, young children are susceptible to toxins which are absorbed quickly into their bodies. Elderly people and those suffering from respiratory disease are more sensitive to pollutants and therefore are more at risk. Asthma attacks could be triggered by increased levels of pollutants. When people exercise they inhale more deeply and therefore allow the pollutants to penetrate deeper into the lungs, putting joggers and cyclists at risk from the air pollution.

How the pollutants damage our health. [28]

Asbestos is a naturally occurring mineral, which is released into the air when the fibres are damaged or disturbed. It is easily caught in the lungs causing scarring and thickening of the tissue. The effects are cumulative leading to chronic illness and lung cancer (fatal). Smokers are much more at risk.

Lead is a bluish-grey soft metal released from exhausts of vehicles using lead petrol and from some industries [29]. It is a cumulative poison which causes headaches, tiredness and loss of concentration. In large amounts it causes severe damage to the central nervous system (although this is now rare).

Carbon Monoxide (CO) is a colourless, odourless, tasteless gas which is slightly lighter than air [29]. It is produced by the incomplete combustion of fuels and tobacco, unflued gas heaters or cookers. CO stops our blood from picking up oxygen and carrying it around the body to our major organs, leading to oxygen starvation. This causes headaches, tiredness, vomiting and even death in large enough quantities.

Nitrogen Oxides (NOx) is the term collectively used to refer to nitric oxide (NO) and nitrogen dioxide (NO2) [29]. It is produced from the combustion of fuels (cars and industry), indoor sources are unflued gas heaters or cookers, paraffin stoves and cigarette smoke. NO2 causes throat and eye irritation and in large enough quantities is highly toxic. A big problem is photochemical pollution, for example the effect of sunlight on NO2 is the creation of ozone [30].

Volatile Organic Compounds (VOCs) are particles which contain carbon and can evaporate easily. They are produced in exhaust fumes, cigarette smoke and industry, they are involved in the formation of ground level ozone.

Ozone (O3) is a highly reactive secondary pollutant [29], and a constituent of smog. At ground level it reacts very easily with biological materials. It causes runny eyes, throat irritations and breathing difficulties at low levels and reduces the lungs' resistance to disease. High concentrations of ozone may cause reversible reduction in lung function. Asthmatics may be sensitive to ozone and increase their sensitivity to allergens [6]. Previous studies in the United States suggested that prolonged exposure to ozone may result in tolerance to the pollutant. However a study conducted on children in Mexico City disagrees and found evidence to the contrary [31].

Sulphur Dioxide (SO2) is a colourless, acidic gas which is soluble in water [29], thus making it a constituent of winter smog. American research shows that acidity in the air irritates asthmatics' lungs and causes chest tightness especially if exercising, hence asthmatics are more at risk [6]. In Los Angeles drastic measures are being introduced to control the emissions which cause photochemical smog (for example banning barbecue lighter fuel and reducing the numbers of cars on the roads).

Suspended Particulate Matter (TSP) is produced from combustion processes. Particulates are measured either by black smoke (**BS**) or by measuring the number of particles which have a diameter of less than 10 micrometers (**PM10s**). The body stops the larger dust particles from entering the body, by our nasal hairs for instance, and hence these cause no long term effects. There is particular concern about **PM10s** as these are small enough to be inhaled and these smaller particles can then penetrate deeply into the lungs [6].

Benzene is a fairly stable colourless clear liquid which evaporates easily [29]. It is present in petrol and in high concentrations may cause anaemia. It is carcinogenic and, in particular, has been linked to leukaemia.

Formaldehyde is a colourless clear gas present in things such as chipboard, foam, some fabrics, car exhaust fumes and cigarettes; it creates indoor air pollution. It may cause eye, skin and throat irritations, nausea and allergic reactions, in high doses it increases the chances of developing lung disease.

Polynuclear Aromatic Hydrocarbons (PAHs) are emitted from coal burning and car exhaust fumes, some PAHs are carcinogenic.

Sulphur dioxide, nitrogen dioxide and ozone interact and increase the effect of allergens causing hayfever in many people [6]. Air pollution can affect asthma in many ways:

i. Direct irritant effect on hypersensitive airways.
ii. Generation of non-specific and allergen-specific airways hyperreactivity.
iii. Direct effect of inflammation of airways either triggering existing asthma or causing asthma in susceptible individuals.
iv. Modification of the immune system's response causing it to become a trigger for asthma.
v. Increase in the susceptibility to develop respiratory infections.

Smog is made up primarily of ground-level ozone, NOx and VOCs (from cars) and is more acute on hot summer days. It's not just an urban problem, but a rural one as prevailing winds move the pollutants. Smog damages agricultural crops and vegetation. Rain may clear the smog, but the rain then becomes acidic and continues to cause damage [32]. Little is known about the effects of smog on health. It is thought that combinations of pollutants may cause more harm [33]. Smog irritates the nose and throat, causes coughing and painful breathing. Scientific studies indicate that after a few days of continuous exposure, even though the respiratory systems disappear, the lungs continue to be damaged. Asthmatics and elderly people are considered to be most at risk, as well as children who tend to spend most of their time outdoors being physically active.

The dramatic smog of December 1952, in London, caused 4000 additional deaths. High pollution concentrations may have reduced people's resistance to diseases such as influenza, even after air quality improved [28]. The London smog of December 1991, had the worst NO2 pollution recorded since 1971, it is estimated that around 400 died because of this episode [6].

In Canada in November 1990, a management plan was adopted, designed to reduce smog producing emissions. Control initiatives implemented included car inspection, recovery of gasoline vapours at service stations, new car emission standards and control of commercial emission sources which involve paint applications, printing and dry cleaning. A public awareness and education program is being implemented at the national level. It is promoting alternative ways of commuting to work by using public transport, walking and cycling as well as ride-sharing, driving within the speed limit, proper maintenance of cars and reducing the number of trips made. In the home, fans and energy-saving lightbulbs are encouraged. Road traffic is blamed for causing eye, throat and skin irritations [34] as well as respiratory disease and visual intrusion. In Britain (1996), various measures were considered in an attempt to reduce harmful vehicle emissions. New incentives were discussed such as cutting road tax by five hundred pounds for lorries which had particulate traps fitted to their exhausts and making ultra-low sulphur diesel cheaper [35]. Sulphur is believed to be most polluting [36].

Local Air Quality Management (LAQM) schemes for local authorities have been set up. Further options could include [37]:

- Traffic management; diverting traffic and closing certain roads at peak periods.
- Powers to detect and test polluting vehicles.
- Giving major employers and schools incentives to encourage the reduction of numbers of single occupancy car commuter trips.
- Positive promotion of mass transit systems, cycling and walking.
- Changing planning regulations to reduce parking provision or tax it heavily.
- Introducing alternative fuels and encourage their use.
- Taking powers to reduce ozone emissions from respraying paint shops for instance.

Until these measures are enforced, air pollution will continue to cause adverse health effects.

The carcinogenic and polluting properties of Benzene led to catalytic converters being introduced, but the trouble with these is that for the first five miles of any journey they don't actually work. Research is underway to try to improve them [6].

There are quite often health warnings given out in the news informing the public of high pollution levels and, in the summer particularly, of high pollen counts. This enables the public to keep out of the pollution and avoid any harmful health effects. People who are outside are more aware of the pollution and have the choice to exercise less.

Recent research shows a conflict in beliefs as to whether air pollution has any harmful effects on health or not. In Britain, studies have repeatedly failed to link conclusively asthma and respiratory symptoms to traffic pollution [38]. In fact some research [39] revealed highly significant spatial variations in symptoms in some cities. In a study involving asthmatics in Tower Hamlets, London, it found that asthmatics of all ages were no more likely to live near heavily trafficked roads than the controls.

In America however, the new Environmental Protection Agency believes that air pollution is the main environmental health threat to children. Their agenda highlights the link between air pollution and asthma. It identifies ozone, sulphur dioxide and nitrogen dioxide as the pollutants which need to be tackled [40]. There is evidence that particulate pollution in winter is responsible for asthma hospital admissions [38]. In the Netherlands, respiratory illness has also been linked to particulates (PM10s) and to iron from the steel works [41].

In Europe, the large APHEA project [42] looked at the short term health effects of air pollution in 15 European cities. Below is a summary of the pollutants which were found to be statistically linked to mortality and hospital admissions for respiratory disease and the countries in which the results occurred.

Mortality:

PM13 : France.
BS : Poland, Greece, Spain.
TSP : Italy.
SO2 : France, Poland, Greece, Germany, Spain, Italy.
CO : Greece.
NO2 : Spain (elderly and cardiovascular mortality).
O3 : Spain (elderly and cardiovascular mortality).

Greece had its strongest effects during the winter. In Poland two cities found that SO2 and black smoke (BS) were linked to mortality, in the other two cities however this was contradicted.

Hospital Admissions:

PM13 : France.
BS : France, Finland (<14 years).
TSP : Finland (<14 years), Italy.
SO2 : France, Finland (>15 years), Italy.
O3 : Netherlands (>65 years), Finland (<14 years), London (summer only).

In France PM13s were found to have a significant effect upon the number of admissions for respiratory disease, these effects were stronger on cold days. In Finland the higher temperatures led to more hospital admissions.

Numerous other studies have reported how the occurrence of various pollutants have been statistically linked to:

a) Decreased peak expiratory flow rates

- NOx, NO2 [14, 43, 44, 45, 46]
- SO2 [44, 47, 48, 49]
- Ozone [48, 50, 51]
- PM10 [24, 47, 48, 52, 53, 54, 72]
- Black smoke [47, 48, 54, 55]
- Smoking [25, 27, 51]
- Pollen [45]
- Temperature [47]

b) Asthma attacks

- NOx, NO2 [45, 56]
- SO2 [57]
- Ozone [58, 73]
- Cold weather [44, 45, 57, 58]
- Pollen [57, 58]
- Smoke [73]

c) Hospital Admissions

- NO2 [45, 50]
- SO2 [45, 63]
- PM10 [65]
- Black smoke, TSP [45, 63, 66]
- All pollutants [64]

d) Mortality

- SO2 [67, 68]
- CO [68]
- Black smoke, TSP [68, 69]

e) Asthmatic Symptoms

i. Cough

- NO2 [59]
- SO2 [59, 60]
- Ozone [60]
- PM10 [24, 53, 60]
- Gas stove [16]

ii. Prolonging of symptoms

- NO2 [62]
- TSP [62]
- Rain and cold weather [51]
- Smoking [15]
- Coal Dust [22, 23]

iii. Wheeze

- SO2 [15, 23, 48, 60]
- Ozone [51]
- PM10 [22, 48]
- Black smoke [48, 55]
- Gas stove [16]
- Industry [61]

iv. Use of medication

- SO2 [48]
- PM10 [48]
- Black smoke [48]

These effects are well documented to happen even when the pollutants are below their current national air quality guidelines, which are considered to be safe levels [50, 53, 55, 60, 70]. Apart from the pollutants themselves, other factors have been found to be statistically important. An individual who has had a recent asthma attack is more likely to have another one in the near future [58]. A high intake of asthma medicine combined with higher temperatures significantly reduces lung function [44]. The effects of temperature and relative humidity vary between people [57]. Symptomatic adults and children have the strongest associations with particulates [15, 51, 53, 71]. Allergens are more harmful when covered in pollutants [50], and the pollutants are most harmful when combined [33, 50, 56].

Ongoing research is attempting to clarify the effects of air pollution on our health, but all the evidence so far suggests the need to drastically reduce pollution. In March this year (1997), the government committed itself to achieving new air quality objectives throughout Britain by the year 2005 [74]. The European Commission has also called for "urgent and radical measures" to combat the causes of pollution and in particular those which cause acid rain and ground level ozone. It hopes to limit the amounts of sulphur dioxide, nitrogen oxides and volatile organic compounds which a country would be allowed to emit after 2010 [75]. These measures have been introduced with the aim of reducing health effects now and in the future, but it may be too late; the damage could already have been done.

3. Statistical Procedures

There are various methodological issues which must be remembered and considered when analysing data regarding the health effects of air pollution.

3.1 Gathering the pollution data

- It is important to position the data gathering equipment in a location which realistically records the levels of pollutants to which the study population are actually exposed. For example, there is little point siting equipment in the central reservation of a busy main road, where it is believed the highest levels of pollutants will be recorded, since most people tend to walk on the pavement at the side of the road and will experience different levels of pollutants from those recorded.
- Data on a range of pollutants should be collected as there could be an interaction between them, i.e. the effects of a pollutant could be different in reality to those experienced in a laboratory, as outside people are exposed to a mixture of pollutants.
- Considerable resources for monitoring air pollution are required due to the necessarily long duration of any study. Some investigations have suffered from a lack of such resources [6].
- Air pollution monitoring equipment often malfunctions creating missing values in the data.
- The pollution data which is needed may not be gathered. It is commonly believed that because PM10s are inhalable they should be measured. A Mexican study [77] found that PM2.5 is actually a much better predictor than PM10.

3.2 Gathering personal data

Gathering data about an individual's home life, and previous history of asthmatic symptoms is very important as a way of assessing their susceptibility to pollution. Exercise tolerance tests are often used to assess an individual's undiagnosed asthmatic tendency [12, 13]. However, it is argued that exercise testing, (enhanced by cold air), is not a good basis for the diagnosis of asthma and adds little information to a well designed respiratory questionnaire in community studies of asthma in childhood [78].

Daily diaries of respiratory symptoms are a powerful technique for detecting acute health effects of air pollution exposure, although they can be difficult to analyse [60]. More problems arise when parents forget or misclassify their children's symptoms [16] and children themselves will have different perceptions as to how they report their symptoms. For example one child may report every little detail and another may only report the severe symptoms. This all makes the analysis more difficult. Children generally report more symptoms than adults, and Mondays have the most symptoms recorded [52].

The arrival or departure of individuals during a study period can also affect the prevalence figures of a symptom. Certain variables such as smoking, diet, occupation, fitness, genetic background, home environment, medication use and illnesses all confound the symptom reporting [6]. Some of these could be identified through a detailed questionnaire, but this kind of detailed response is time-consuming to obtain and the risks of inaccurate responses increase as the questionnaire length increases.

All of the above data are discrete variables, some of which are binary. However it is also possible to measure continuous variables such as the daily Peak Expiratory Flow Rates (PEFR) which were gathered in this study.

3.3 Modelling problems

One of the biggest problems is that of avoidance behaviour where people may stay indoors, avoid heavy exercise or use air-conditioning when pollution (and pollen) levels are high [52]. This makes it statistically difficult to find any connection between high pollution levels and detrimental health effects. Pollution levels peak at certain times of the day, affecting daily average levels, but the study population may well not be outside experiencing these pollution peaks.

Different sources of pollution and different combinations of pollutants have various health effects [6], thus making it very difficult to pinpoint significant interactions. The weather is well documented as having an effect on pollution levels [6]. Wind speeds and direction, temperature inversions, relative humidity and seasonal variations all affect pollution levels as do geographical locations of a site (e.g. coastal or in a valley). Ozone levels generally peak at noon, and the presence of a land-sea breeze creates an accumulation of oxidants during the morning [79]. Ozone depends on temperature and sunshine. Interdependent relationships occur between nitrogen monoxide, nitrogen dioxide, ozone and carbon monoxide [80].

There may be a time delay [45] between exposure to pollution and the onset of symptoms, which itself may be affected by the weather or by confounding issues such as whether the person smokes, or is a passive smoker.

More asthma attacks are reported to happen on Saturdays than on Sundays. The largest significant predictor variable was the presence or absence of an attack on the preceding day [58].

3.4 Basic Terminology Introduced

Let y_{it} = Observation at time *t* on subject *i*

x_{jt} = Observation of the *j*th pollutant variable at time t

The response variable y_{it} could be either discrete or continuous. The discrete case occurs when y_{it} is the binary answer to a field on the questionnaire or in the daily symptom diaries such as the presence (or absence) of a cough. Other examples of a discrete response variable are the Poisson counts of the number of asthma attacks in a given amount of time or daily mortality rates. However, as is the case in this study, the response variable y_{it} is continuous when looking at how peak expiratory flow rates change over time for each individual.

Pollutant effects may not be identified if the wrong model has been used. The lack of studies conducted in this area, means that there is little information to draw on when analysing such data. However, there are some useful ideas to be gained from the published accounts of previous studies. The statistical techniques vary enormously depending on whether y_{it} is discrete or continuous. Even though the focus of this research is on continuous data, for completeness the methodology for the discrete case is also considered.

3.5 Basic Models for the discrete case

3.5.2 Linear Regression $y_{it} = \Sigma \beta_j x_{jt}$

This is conceived to be the most straight forward approach for modelling discrete data. Separate regressions are usually needed for each six month warm and cold season. Stepwise and ordinary regressions risk letting the noise choose the significant pollutant. It is recommended to fit each pollutant separately, if more than one pollutant seems to be associated with the outcome, then pollutant interactions are investigated. The APHEA project used linear regression with log transformed dependent variables and parsimony is recommended [43].

3.5.1 Poisson Regression $y_{it} \sim P(e^{\sum \beta_j x_{jt}})$

Poisson regression is recommended when looking at counts of daily mortality rates [54] and when modelling incidence rates of symptoms. There are problems with this model:

i) Daily symptom rates are highly correlated and it has been suggested that incidence rates, (new occurrence of a symptom which was not there on the previous day), should be used rather than prevalence rates, (symptom occurs, whether it was there previously or not) [60].
ii) Multivariate analyses conducted on asthma mortality rates in Philadelphia [81] found that death rates were significantly higher among ethnic minority groups. This

suggests that ethnicity should be considered when looking at heterogeneous subgroups of the study population [54].

iii)Autoregressive models are recommended to account for the autocorrelation in the data [60]. Schwarz et al. [82] recommend using lagged prevalence rates to remove this autocorrelation, i.e. using the subject's symptom status on the previous day as the covariate.

The Poisson regression has a built in heteroscedasticity in the variance. Over a short time interval a linear trend seems adequate, but as the time period extends, there is a greater need for a non-linear model. To model for the heterogeneity, Schwarz et al. [82] assume that each individual is identical and combine the incidence rates from all subjects present on each occasion. This reduces the amount of computing work which needs to be done, since it now only depends on the number of response occasions (days) rather than on the total number of responses. A count is obtained for the number of incidences during the whole period for the *i*th subject, and they proceed by using Poisson regression techniques.

The APHEA project [43] found the need to control for seasonality and weather (possibly non-linear). This study looked at daily counts of mortality and hospital admissions. The authors assumed a homogeneous risk for the population with various risk factors, such as age and cigarette smoking, assumed not to vary from day to day, and hence not influencing the number of deaths. A canonical Poisson regression was used, which is a relative risk model, for example if the population size doubled, then the number of deaths and hospital admissions was also assumed to double.

3.5.3 Ordinary Logistic Regression $y_{it} = \frac{e^{\sum \beta_j x_{jt}}}{1 + e^{\sum \beta_j x_{jt}}}$

Diary studies are good at assessing the impact of short-term changes in the environment on human health. An individual's health status is assessed by the presence or absence of several symptoms each recorded as a binary outcome. There are dependencies among responses on consecutive days (autocorrelation) and among an individual's responses on any two different days (heterogeneity). The use of incidence rates [82], (new occurrence of symptom), can reduce autocorrelation in the data, for instance the autocorrelation caused by many individuals having a symptom on one day and then others developing the same symptom the next day. However, using incidence rates reduces the amount of information available. Variability among individuals can be caused by passive smoking, gas stoves and the presence of allergic conditions. Individuals have different thresholds, susceptibilities and reporting behaviour and their autocorrelation and heterogeneity must be considered when modelling the data.

Since the pollutants which may cause the symptom are not necessarily the same as those which increase its duration, the incidence rates on different days are independent. These could be analysed by using ordinary logistic regression. Health variables can be biased by people spending more time outdoors because of the warm weather, where symptoms may be caused by pollen say.

Liang and Zeger [83, 84] describe methods for fitting logistic models to the symptom rates, whilst accounting for the correlation, which yield robust variance estimates, even for a misspecified covariate.

In the European PEACE study [55], logistic regression models with binomial (for incidence) or normal (for prevalence) residuals, corrected for one day autoregression, were fitted when using daily incidence or daily prevalence rates, as the dependent variable. No relationships were found between the symptoms and the air pollution. It is suspected that this lack of association could be caused by:

i. Confounding infectious diseases.
ii. Incorrect trend fitting specification, (day-of-the-week effects were fitted).
iii. Threshold existence.
iv. Composition of particles.
v. Different subjects needed (PEACE study only used symptomatic children).

Korn and Whittemore [85] suggest that individuals may vary in their sensitivity to pollutants. Usual linear regression requires observations to be statistically independent with constant variability about the mean which in practice doesn't often happen. For example asthmatics are more likely to have an attack on days following an attack. Another problem is that of missing data, which can cause spurious associations. So they propose a two-stage analysis. For each individual, a separate ordinary logistic regression is calculated to obtain coefficients β_i (which we assume are multivariate normal) which vary for each of the i subjects about a population mean. Then for a selected homogeneous group, for each variable, a weighted average coefficient can be calculated, weighted according to the amount of data that the ith subject contributes. Subgroups can then be easily looked at and compared. The problems with this method:

i. The assumption of normality only holds for a sufficiently large number of observations.
ii. This method involves a large amount of computation.
iii. The model is not often appropriate for low response rates (as is often the case, especially when working with incidence rates).

Alternatively we assume regression coefficients β_i are constant across subjects (fixed effects). The common β_i assumes that successive responses are uncorrelated (possibly not true with real data). Each person's periods of sickness are compared only to their own periods of good health. However the intercept term is random, which allows for individual heterogeneity. This can cause computational difficulties, so we assume that the intercepts follow a distribution (normal say). This method still involves masses of computations. Susceptible subgroups can be easily identified by high regression coefficients on age or sex, and then looked at, as can lag effects, day-of-the-week effects and interactions of the covariates. Models can easily be tested for adequacy using plots of standardised residuals over time or response probability or levels of a particular covariate.

3.5.4 Autoregressive and Moving Average Models

Serial correlation is where two observations closer together in time are more alike than two which are randomly chosen. This serial correlation could be due to:

i. the fact that an individual's repeated measurements are usually correlated and not independent.
ii. omitted covariates such as epidemics.
iii. imperfectly controlled covariates such as weather.

Serial correlation does not bias the regression coefficients but affects the estimated standard errors. Autoregressive and moving average models are efficient for dealing with this type of trend in the data.

Day of the week and holiday effects also bias time series regressions. The APHEA project [43], suggests using dummy variables for season and trend. In Gaussian data, weighted moving averages (kernel smoothing) are recommended to smooth season and trend, where the most recent events have the most weight. This filtered data is also Gaussian. But when there are fewer observations and a Poisson distribution is assumed, it should be noted that the filtered Poisson data is not Poisson. They suggest putting the moving average filter into the regression model itself or using a generalised additive model, which can be generalised to multivariate smoothing. Too much filtering can result in the loss of cumulative effects. A semiparametric approach is to use regression spline functions, where the variable is split up into intervals, (this requires consideration and care), and fitting a cubic polynomial to each interval which all join smoothly at the interval boundaries. A parametric approach is to use sinusoidal terms to fit the long wave-length pattern, (for example the annual pattern), in the data. The problem is that this method assumes that the seasonal peak is the same height and occurs at the same time every year. For longer time series, different sinusoidal terms will have to be fitted for different periods.

To check the validity of creating each model, correcting for the confounding variables, such as seasonality and day-of-the-week effects use;

- Diagnostic plots
- Plots of the predicted outcome over time
- Statistical tests for goodness of fit
- Cross correlations

The Durbin-Watson statistic determines whether seasonality has been controlled for. The APHEA project [43] removed any remaining autocorrelation by using autoregressive error models. To investigate the lag structure, multiple lags could be modelled simultaneously. But since the variables are serially correlated these give unstable estimates, hence a moving average constraint is often used. Exploring too many lag structures risks identifying non-causal relationships which may have occurred by chance.

3.6 Basic Models for the continuous case

3.6.1 Linear Regression Model With a One Day Autoregressive Error Structure

In the European PEACE study [55], two panels of susceptible children selectively chosen by a questionnaire, in urban and non-urban locations, were followed and air pollution concentrations measured. Children chosen were those who had suffered from wheeze without colds, asthma attacks, dry cough at night during the past year or had ever been diagnosed as asthmatic by their doctor. Morning and evening peak flow rates were measured and the presence or absence of any respiratory symptoms recorded along with medication use.

A child's daily deviation, $\mathrm{DEV}_{ai} = x_{ai} - \bar{x}_a$, for child *a* and day *i*. In the analysis, the dependent variable was the population deviation dev_i, for a particular day, calculated by averaging all the individual deviations for all the children who were present that day. The pollutants' (PM10, black smoke, sulphur dioxide and nitrogen dioxide) 24 hour averages were used as the independent variables. Their statistical model was a linear model with normal residuals which had a one day autoregressive correction.

Various lags of up to two days were fitted. Significant negative associations were found in the control panel between lung function and PM10 and also black smoke.

3.6.2 Multilevel Mixed Linear Model Analysis

Goldstein [86] investigated multilevel mixed linear model analysis. When repeated measurements are made on an individual, we are presented with the problem of how to deal with this longitudinal data which requires specialised techniques to be employed in order to correctly interpret the parameters obtained. Multilevels occur for example when there are; Schools, classrooms within schools and children within classrooms. Goldstein assumes that simple random sampling occurs at each level and a linear model is set up at each level, capable of incorporating any interactions of explanatory variables. He suggests the use of iterative maximum likelihood to calculate parameters. This type of model allows for random coefficients at any level between groups but not within groups. For simplicity a constant variance is assumed. Variance heterogeneity could be introduced into a model by allowing a variance to be a function of time or age, without incorporating the same function into the fixed part of the model. This type of modelling procedure can be extended to:

i. Complex sample surveys involving clustering.
ii. Longitudinal data.
iii. Multivariate multilevel mixed effect models (including the usual multiple regression model with missing data).
iv. Maximum likelihood estimates which give efficient estimates for normal and other distributions.

3.6.3 Longitudinal Models

The methodology for analysing longitudinal data, (repeated measurements over time and over different individuals), was presented by Diggle, Liang and Zeger [87]. They showed the need to take individuals into account. There are variables which may affect the overall levels of peak flow rates, (site, age, sex and height), and variables which may affect peak flow rates over time (pollution). For longitudinal data the repeated measurements on an individual are usually autocorrelated. Diggle et al. recommend the use of the variogram to investigate the nature of this autocorrelation. A variogram is a function, which is directly related to the autocorrelation function, that describes this association among repeated measurements and is easily estimated with irregular observation times. Various aspects of an individual's behaviour may show stochastic variation, and if this is ignored, then the coefficients β_i in the model are incorrectly interpreted. The 'residuals' described in Diggle et al. are obtained by subtracting from the measurement the ordinary least squares estimate of the corresponding mean response: $r_{ij} = y_{ij} - \hat{\mu}(t_{ij})$. Time plots, scatterplots and empirical variogram plots of these residuals can help identify the underlying structure. In particular, the variogram can identify the underlying covariance structure and be used to formulate an appropriate model.

When analysing longitudinal data, missing values are common, but it is useful to fill in these gaps. Methods such as linear interpolation, the Kalman filter, fixed interval smoothing, kernel smoothing and spline estimation can be used. The details of Diggle, Liang and Zeger's methodology are shown more extensively in chapters 6 and 7 of this thesis.

4. The Data Gathered

4.1 The Personal data gathered

The 426 children contributing to the West Glamorgan study were chosen from four schools on the three sites of Bishopston, Glyn Neath and Swansea. The children living in Bishopston are thought to have a better quality of life due to, among other things, the lower levels of pollution. This in itself makes the house prices higher and hence the population are generally of a better socio-economic background than those living in Glyn Neath or Swansea, where poverty and poorer housing conditions are more common.

Table 4.1 presents basic site-sex frequencies for the 426 children in this study. Most of the children were of white ethnic origin. In the case of Swansea, there were substantially more female than male children (approximate ratio 3 to 2). Two neighbouring schools at Glyn Neath were used to make numbers comparable.

Table 4. 1 Basic distribution of children

Site	**MALE**	**FEMALE**	**Total**
1 = Bishopston	72	64	136
2 = Glyn Neath	53	52	105
3 = Glyn Neath	25	31	56
4 = Swansea	54	75	129
Total	204	222	426

AGE		**ETHNIC GROUP**	
8	74	Bangladeshi = 1	18
9	145	White = 2	406
10	149	Other = 3	1
11	57	Missing = 4	1
12	1		
Total	426		426

4.1.1 The Child.

The original questionnaire which was distributed to parents is presented in Appendix 2.2. Questions 1-9 sought information on the child's past symptoms.

Table 4.2 presents basic response frequencies for these nine questions, categorised by site and sex. Some important features emerge from these figures:-

- Children at site 2, Glyn Neath have a particularly high frequency of ever having had wheezing or whistling in the chest in the last 12 months [Q2] (24% of females at site 2, Glyn Neath, and 21% of males)
- Glyn Neath, site 2 similarly had the largest frequency of wheezing severe enough to limit the child's speech [Q5].
- A smaller proportion of children (24%) at site 4 (Swansea) reported attacks of wheezing compared to the other sites, (35.3% at Bishopston and 31% at Glyn Neath).
- For sites 1 to 3 the percentage of children who had ever been diagnosed as asthmatic [Q6] varied between 21 and 24% whilst the equivalent figure for Swansea, 13%, is significantly less (p-value $<10^{-6}$).
- Site 2 (Glyn Neath) reported the highest percentage of wheezing during or after exercise [Q8] - 19%.
- Approximately 15-25% of children reported a dry cough at night [Q9], though for females at site 2, the figure was closer to 40%.

These suggest a possible lack of awareness of children's conditions in Swansea, and even under reporting of symptoms.

Table 4. 2 Responses to Questions 1-9 of the Questionnaire

Site	Sex	Q1: Ever wheezed		Q2: Wheezed in last year		Q3: How many attacks of wheezing in last year			
		NO	YES	NO	YES	None	1-3	4-12	>12
1	M	47	25	59	13	59	7	4	2
	F	41	23	56	8	56	4	2	2
2	M	38	15	42	11	42	6	4	1
	F	34	17	39	12	40	6	5	1
3	M	15	10	20	5	20	3	0	2
	F	24	7	27	4	27	1	1	1
4	M	39	15	46	8	46	6	2	0
	F	59	16	63	12	63	6	1	5

Site	Sex	Q4: How often has wheezing disturbed sleep			Q5: Has wheezing stopped speech		Q6: Ever had asthma	
		Never	<1	1+	NO	YES	NO	YES
1	M	63	8	1	70	2	53	19
	F	58	4	2	64	0	54	10
2	M	45	6	2	48	5	42	11
	F	43	3	6	50	2	38	14
3	M	22	1	2	24	1	18	7
	F	28	1	1	29	1	25	6
4	M	49	4	1	52	1	47	7
	F	66	6	3	73	2	65	10

Site	Sex	Q7: Have they received treatment		Q8: Wheezy chest after exercise		Q9: Dry cough at night	
		NO	YES	NO	YES	NO	YES
1	M	67	5	61	10	57	15
	F	62	2	57	6	50	13
2	M	47	6	40	10	42	11
	F	49	2	43	9	32	20
3	M	23	2	22	3	17	8
	F	28	3	28	3	23	8
4	M	53	1	49	5	43	11
	F	70	5	63	12	56	19

4.1.2 The Child's Environment.

Questions 10 to 14 of the questionnaire extracted information on the child's home environment, with responses as follows:-

Table 4.3 Responses on the home environment

Site	Sex	Q10: Use gas for cooking		Q11: Calor bottle gas heaters	
		NO	YES	NO	YES
1	M	39	33	48	24
	F	30	34	48	16
2	M	35	18	46	6
	F	41	11	36	16
3	M	18	7	19	6
	F	16	15	23	8
4	M	14	40	20	33
	F	20	55	38	37

The most important points to note are:-

- The percentage of families using gas for cooking varied considerably between sites - 49% for Bishopston, 27% and 39% for the two Glyn Neath sites, and 73% for Swansea.
- In Swansea 55% of children lived in homes which used calor type bottle gas heaters [Q11].
- Only five homes in total used paraffin heaters [Q12], three of which were in Glyn Neath.
- In Glyn Neath and Swansea, over half the homes contained a smoker [Q13], whereas in Bishopston, this figure was less than a quarter.
- Finally more than 50% of the children had a furry pet in the home [Q14]. For Bishopston, the figure was 60%.

Table 4.3 Continued.

Site	Sex	Q12: Paraffin heaters		Q13: Smoker at home		Q14: Furry Pet	
		NO	YES	NO	YES	NO	YES
1	M	71	1	49	23	29	43
	F	64	0	50	14	25	39
2	M	51	2	26	27	22	31
	F	52	0	22	30	28	24
3	M	25	0	10	15	10	15
	F	30	1	15	16	14	17
4	M	53	0	26	28	26	28
	F	73	1	37	38	35	39

4.1.3 Daily Maximum Peak Expiratory Flow Rates.

The daily Peak Expiratory Flow rates (PEF) were collected throughout the periods:-

The winter study: **9th January 1995** to **17th February 1995**
The summer study: **5th June 1995** to **14th July 1995**

Every weekday morning for six consecutive weeks (in each study period) each child participating in the study was supervised where possible blowing into a peak flow meter three times. The maximum value, the child's Peak Expiratory Flow Rate (PEF), for that day was recorded by the child, correct to the nearest 10 units. Hence for each study period there are a maximum of thirty readings for each individual.

4.1.4 Attendance and quality of data.

Naturally not all recordings were available for each child, due to absences which were recorded in the school attendance variable:-

School Attendance Absent due to illness = 0
Present = 1
Absent on holiday = 3
Absent with a reason = 4
Absent no reason = 8
Data missing = 9

Table 4.4 records absence percentages across site and season.

Table 4.4 Percentage of Absences

	Bishopston		Swansea	
Absence	**Winter**	**Summer**	**Winter**	**Summer**
Illness (0)	6.3%	4.4%	10.7%	0.2%
Present (1)	93.2%	85.7%	81.8%	69.9%
Holiday (3)	0.5%	8.3%	0.5%	0.6%
With a reason (4)	0.0%	0.0%	0.0%	0.0%
Absent (8)	0.0%	0.1%	0.0%	15.3%
Data missing (9)	0.0%	1.6%	7.0%	14.0%

	Glyn Neath (2)		Glyn Neath (3)	
Absence	**Winter**	**Summer**	**Winter**	**Summer**
Illness (0)	9.6%	0.5%	5.4%	5.5%
Present (1)	89.1%	82.9%	94.2%	87.9%
Holiday (3)	0.4%	0.7%	0.4%	6.6%
With a reason (4)	0.0%	0.0%	0.0%	0.0%
Absent (8)	0.0%	11.1%	0.0%	0.0%
Data missing (9)	0.9%	4.8%	0.0%	0.0%

It should be noted that Swansea and Glyn Neath (2) have the highest percentages of absence due to illness during the winter months and that the percentage substantially drops during the summer months. This is perhaps to be expected, given the socio-economic status of each of the areas.

Not all PEF readings were judged to be reliable and hence data quality was recorded:-

Data quality Good = 1
Suspect because of faulty technique or known poor reader = 2
Suspect because of unusual result = 3
Suspect because of digit preference = 4
Data missing = 9

Only peak flows considered and coded as "Good" were used in the final analysis. Table 4.5 presents the overall percentage of usable data for each site.

Table 4.5 Percentage of Usable data for each site

	SEASON	
SITE	**Winter**	**Summer**
Bishopston (136 children)	87.9%	75.5%
Glyn Neath (105 children)	88.7%	80.8%
Glyn Neath (56 children)	96.6%	84.0%
Swansea (129 children)	79.2%	66.6%

4.1.5 Daily Symptom Diaries.

For every day of the study, each child would fill in their daily symptom diary. This registered whether, on that day, the child had suffered from any of the following symptoms; cough, wheeze, sneeze, sore throat or runny nose. However, no information was gathered as to whether asthmatic children had used their inhalers (or other medication).

Table 4.6 below shows the percentage of time in each study period that the symptom was present for, (e.g. the sum over all children of the number of days that a cough was present, divided by the total number of child days and expressed as a percentage). The numbers of children at each site are given in brackets. The main points to note are:-

- More symptoms were present during the winter study period as expected.
- The numbers of symptoms at Glyn Neath during the winter study did not decrease in the summer by as much as the numbers at the other sites.
- In particular Glyn Neath (3) shows only a marginal decrease in the summer months, the number of reported wheezes is actually greater in the summer.
- During the summer study, both Glyn Neath sites reported the most symptoms.
- Runny noses were the most frequently reported symptom, with coughs following next. Wheeze was the least reported symptom.

Table 4.6 Presence of symptoms

	Bishopston [136 children]		**Swansea [129 children]**	
Symptom	**Winter**	**Summer**	**Winter**	**Summer**
Cough	36%	21%	37%	17%
Wheeze	7%	3%	10%	7%
Sneeze	14%	9%	8%	6%
Sore throat	26%	14%	24%	12%
Runny nose	54%	31%	45%	29%

	Glyn Neath 2 [105 children]		**Glyn Neath 3 [56 children]**	
Symptom	**Winter**	**Summer**	**Winter**	**Summer**
Cough	37%	26%	29%	28%
Wheeze	7%	6%	12%	13%
Sneeze	10%	8%	17%	15%
Sore throat	21%	17%	20%	21%
Runny nose	44%	34%	55%	46%

4.1.6 Exercise Tolerance Test.

An Exercise Tolerance Test (ETT) was arranged to identify those children who suffered from an exercise-induced peak flow drop. This is a standard test which has been used as a means of indicating individuals with asthmatic tendencies. An individual blows into a peak flow meter and their initial peak flow is recorded. The child then takes part in a hard physical activity (e.g. cycling) for up to 6 minutes.

Immediately following this, another peak flow is taken and recorded. Current medical practice classifies those children who suffer a 10 to 15% drop from the first to the second peak flow rate as having 'exercise-induced asthma'. In our study only 23 out of the 426 children had this significant drop in peak flow rates. Given the small number of such children it is not possible to attach much significance to these results. Note however that overall almost twice as many females as males had exercise-induced asthma.

Table 4.7 Children with Exercise-Induced Asthma

	10-15 % exercise induced asthma		
SITE	**Male**	**Female**	**Total**
Bishopston (1)	4	5	9
Glyn Neath (2)	2	4	6
Glyn Neath (3)	1	1	2
Swansea (4)	1	5	6
Total	8	15	23

4.1.7 Definitions of Asthmatic and Wheezy children.

For the purpose of this study, an *asthmatic* is defined to be any child who has ever had asthma [Q6 on the questionnaire]. In our study of 426 children, 84 were classed as asthmatic. This is almost a fifth of the children.

A child is said to be a *Wheezer* if they had reported any of the following asthmatic tendencies:

- GP diagnosis of asthma
- Ever wheezed
- Wheezing at night
- Wheezing after exercise
- Exercise induced peak flow drop
- Dry cough at night

Table 4.8 presents the frequencies reported for each symptom. Consequently, 186 children (43.3%) in our study group were classed as wheezers. Restricting the above categories to symptoms suffered in the last 12 months, reduces the figure to 130 children (30.2%).

Table 4.8 Number of children reporting each symptom

SYMPTOM	**Number of children reporting symptom**	
Ever suffered from asthma	84	19.5%
Wheezing at any time in past	128	29.8%
Wheezing in last year	73	17.0%
Wheezing after exercise in last year	58	13.5%
Wheezing at night in last year	19	4.4%
Dry cough at night in last year	105	24.2%

Table 4.9 presents frequencies and percentages of the number of non-asthmatics and asthmatics across site and sex.

Table 4.9 Distribution of Asthmatics and Non-asthmatics

	MALE		**FEMALE**	
Site	**Asthmatic**	**Non-asthmatic**	**Asthmatic**	**Non-asthmatic**
Bishopston	19	53	10	54
Glyn Neath 2	11	42	14	38
Glyn Neath 3	7	18	6	25
Swansea	7	47	10	65
Total	44	160	40	182

- In our study there are more females (222) than males (204), but there are more male asthmatics than female asthmatics. However the proportion of male asthmatics is not significantly different from the proportion of female asthmatics.
- The percentages of asthmatics were Bishopston (21%), Glyn Neath sites (24%) and Swansea (13%). As reported earlier, the proportion of reported asthmatics in Swansea is statistically significantly less than those of the other areas, whereas the difference in proportions between Bishopston and the Glyn Neath sites is not significant (p-value >0.1)
- Only 11 out of 84 (13%) of asthmatics proved positive in the Exercise Tolerance Test.

- There were five eight-year-olds who had an exercise induced peak flow drop but they were classed as non-asthmatic. This could be due to the fact that they are so young, they have not had a GP diagnose Asthma yet.
- Swansea has the least number of children affected by exercise.

4.2 Weather and Pollution Data

4.2.1 Introduction.

In order to relate peak-flow measurements to variation in atmospheric pollution, monitoring equipment was sited in the three school yards, namely Bishopston Primary School, Cwm Nedd School in Glyn Neath (covering for both schools in Glyn Neath) and Terrace Road School in Swansea. The pollution and weather data was deliberately collected for a substantially longer period than the intended period of monitoring peak expiratory flow rates, so as to facilitate statistical models involving lagged effects.

The following **pollution measurements** were recorded at 15 minute intervals throughout the periods given in Table 4.10.

- NOx, NO and NO2 - The nitrogen oxides
- SO2 - Sulphur Dioxide
- CO - Carbon Monoxide
- O3 - Ozone
- PM10s - Dust Particulates

Weather data was also gathered but only for the two geographical sites Glyn Neath and Swansea, as it was felt that the Swansea meteorological data was also appropriate for Bishopston. For both the sites,

- Relative humidity
- Temperature
- Wind Speed and Direction were measured every 15 minutes.

Some daily pollen counts were noted during the summer period, but unfortunately this data is very sparse and so pollen count data was not used. The periods of collection for both the weather and the pollution data were larger than the periods used for monitoring peak flows to allow statistical modelling using lagged explanatory variables.

Table 4.10 Dates of data collection

SEASON	Peak flow data collected in the periods	Weather and pollution data collected in the periods
Winter	9th January 1995 - 17th Feb. 1995	1st January 1995 - 28th Feb. 1995
Summer	5th June 1995 - 14th July 1995	1st June 1995 - 1st August 1995

All the data, with the exception of the pollen counts, were entered into tables in the Paradox database. For each variable, a quality variable was also included which indicated data points which were missing, bad, incorrect due to equipment failure or power failure, these were ignored. Some readings were negative and these too were not used when daily averages were calculated in Paradox.

In Swansea, additional measurements for the number of sunlight hours, amount of rainfall and air pressure were taken. Table 4.11 presents the available explanatory data. The small number of isolated cases of missing pollution data were filled in by linear interpolation.

Table 4.11 Atmospheric Data Available

Weather data - Both study periods:

SITE	Relative Humidity	Temperature	Wind Speed	Wind Direction	Sun	Rain	Air Pressure
Bishopston	✓	✓	✓	✓	✓	✓	✓
Glyn Neath	✓	✓	✓	✓			
Swansea	✓	✓	✓	✓	✓	✓	✓

Pollution data - Winter study:

SITE	NOx	NO	NO2	SO2	CO	O3	DUST
Bishopston	✓	✓	✓	D		✓	✓
Glyn Neath	✓	✓	✓	D		✓	✓
Swansea	✓	✓	✓	✓	✓	✓	✓

Pollution data - Summer study:

SITE	NOx	NO	NO2	SO2	CO	O3	DUST
Bishopston	✓	✓	✓	✓	✓	✓	✓
Glyn Neath	✓	✓	✓	D		✓	✓
Swansea	✓	✓	✓	✓	✓	✓	✓

Key: ✓ = Data collected
D = Daily averages only available

Note that there were no CO levels recorded at Glyn Neath at all.

4.2.2 Summary Statistics and Graphs of Pollution Data.

Bishopston, Site 1, has relatively little pollution with only high ozone levels during the summer months and was specifically included in the study as a control site. Glyn Neath is situated in a valley, close to an open cast coal mining site and a major trunk road (A465). Sulphur dioxide levels (SO2) are high in Glyn Neath from coal-burning and traffic emissions. Swansea town centre has high pollution levels caused by the large amount of traffic passing through on a daily basis.

Table 4.12 presents average daily pollution levels and their standard deviations for both of the study periods.

NOx, NO and NO2 have lower averages in the summer than in winter with the reduction being of the order of 50%. In contrast SO2, CO and Dust have higher averages in the summer (especially Dust at Bishopston). Ozone averages are slightly lower in the summer, except at Bishopston where they rise. Standard deviations are lower in the summer (i.e. less variation in pollutant levels) for NOx, NO, NO2, CO and O3 (except at Bishopston), and are higher for SO2. Variation in Dust levels are higher in the summer except at Glyn Neath. Note that the high standard deviation at Glyn Neath during the winter, is due to one particular day (7th February 1995) where the Dust levels soared. This is a genuine reading caused by a malfunction of an oilfired boiler at nearby swimming baths. With this day removed the winter average remains high at 14.3198 whilst the standard deviation drops to 4.2066, lower than the other sites.

Table 4.12 Daily pollution averages and standard deviations

			WINTER		SUMMER	
Area	**Variable**	**Units**	**Average**	**Std Dev**	**Average**	**Std Dev**
Bishopston	NOx	PPB	6.24	4.35	2.56	2.98
	NO	PPB	0.92	1.43	0.46	0.75
	NO2	PPB	4.14	2.82	2.17	2.40
	SO2	PPB	0.72	0.64	2.28	3.43
	CO	PPM	....	....	0.18	0.13
	O3	PPB	17.56	3.69	31.59	10.74
	Dust	UG/M3	6.30	6.18	17.76	10.42
Glyn Neath	NOx	PPB	14.61	10.13	7.21	3.31
	NO	PPB	5.01	6.41	1.42	1.36
	NO2	PPB	9.77	4.51	6.10	2.55
	SO2	PPB	15.95	8.91	14.70	5.93
	CO	PPM	....	....	....	....
	O3	PPB	24.42	7.23	26.41	12.20
	Dust	UG/M3	15.68	10.99	16.74	6.65
Swansea	NOx	PPB	49.26	32.52	25.04	13.32
	NO	PPB	30.98	25.92	13.49	11.10
	NO2	PPB	16.71	7.78	10.31	5.07
	SO2	PPB	3.65	3.91	5.65	4.26
	CO	PPM	0.52	0.34	0.69	0.11
	O3	PPB	21.96	7.54	24.71	10.58
	Dust	UG/M3	21.10	5.68	25.35	13.28

Key: PPB = Parts Per Billion, PPM = Parts Per Million,
μg/M3 = Micrograms per cubic metre

Maximum Recommended Limits for Pollutants.

Various organisations perceive different levels above which a pollutant is thought to be detrimental to human health. The following table shows some maximum recommended limits for a 24-hour running average of each pollutant.

Table 4.13 Maximum Recommended limits for pollutants

Pollutant	**Maximum allowed level**	**Organisation**
Nitrogen Oxides (NOx)	78 ppb	WHO
Sulphur Dioxide (SO2)	80 ppb - 100 ppb	WHO, DoE
Carbon Monoxide (CO)	10 ppm (8 hr average)	WHO
Ozone (O3)	90 ppb	DoE
Dust Particulates	50 ug/m3	EPAQS

WHO : World Health Organisation

DoE : Department of the Environment

EPAQS: Expert Panel on Air Quality Standards

During the winter study; in Swansea NOx exceeded the safety limit on 10 occasions. In Glyn Neath dust levels became too high once. For the summer study, dust levels in Swansea exceeded the limit 4 times.

Correlations between the daily pollution averages at all sites.

Table 4.14 presents Pearson Correlation coefficients for each pollutant variable for each pair of sites together with associated significance level and the number of observations. The correlations for Ozone, O3 are large and positive for all pairs of sites. Correlations between Swansea and Bishopston tend to be small (ozone excepted) and non-significant for the winter period, whereas in summer, coefficients for NO2, SO2, O3 and Dust are positive and significant.

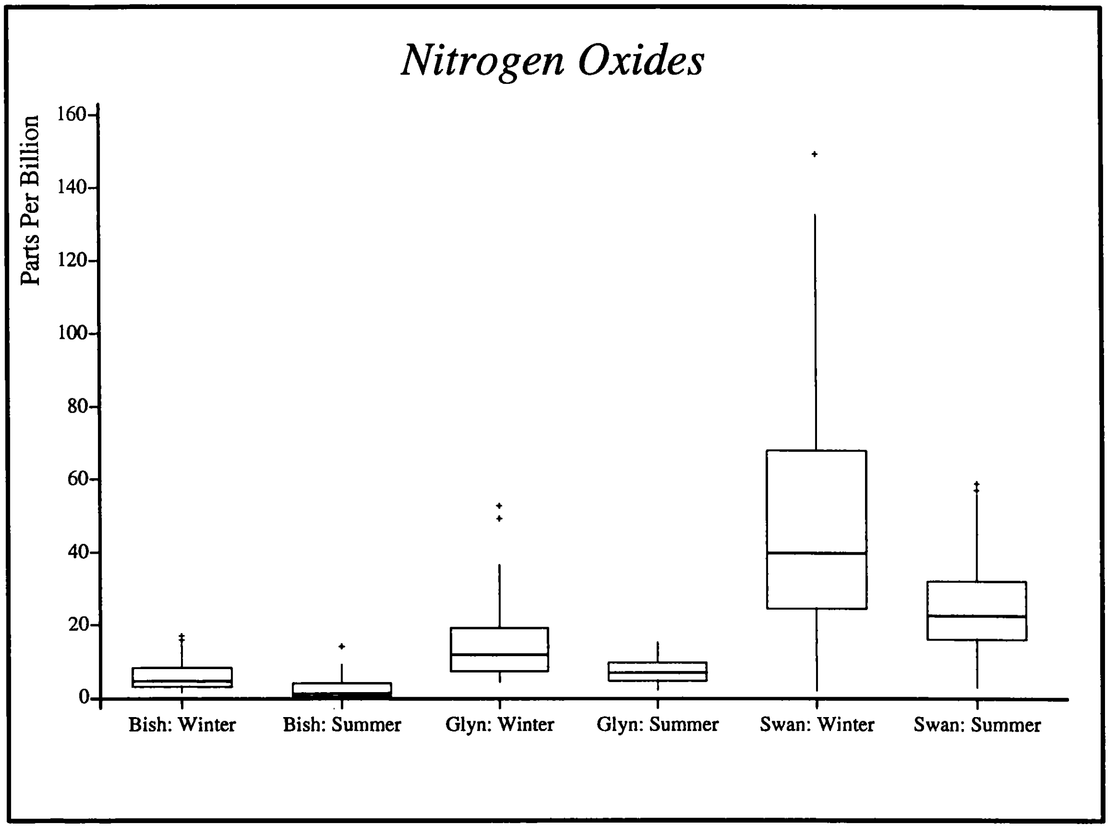
For Dust measurements, all pairs of sites correlate highly in the summer period. Note also that the Box and Whisker plot for dust levels (these follow) showed that all sites in the summer period had similar dust levels which may imply a common source of the dust. For pollutants NO2, SO2, the three sites show substantial correlations in the winter period, but not for the summer period.

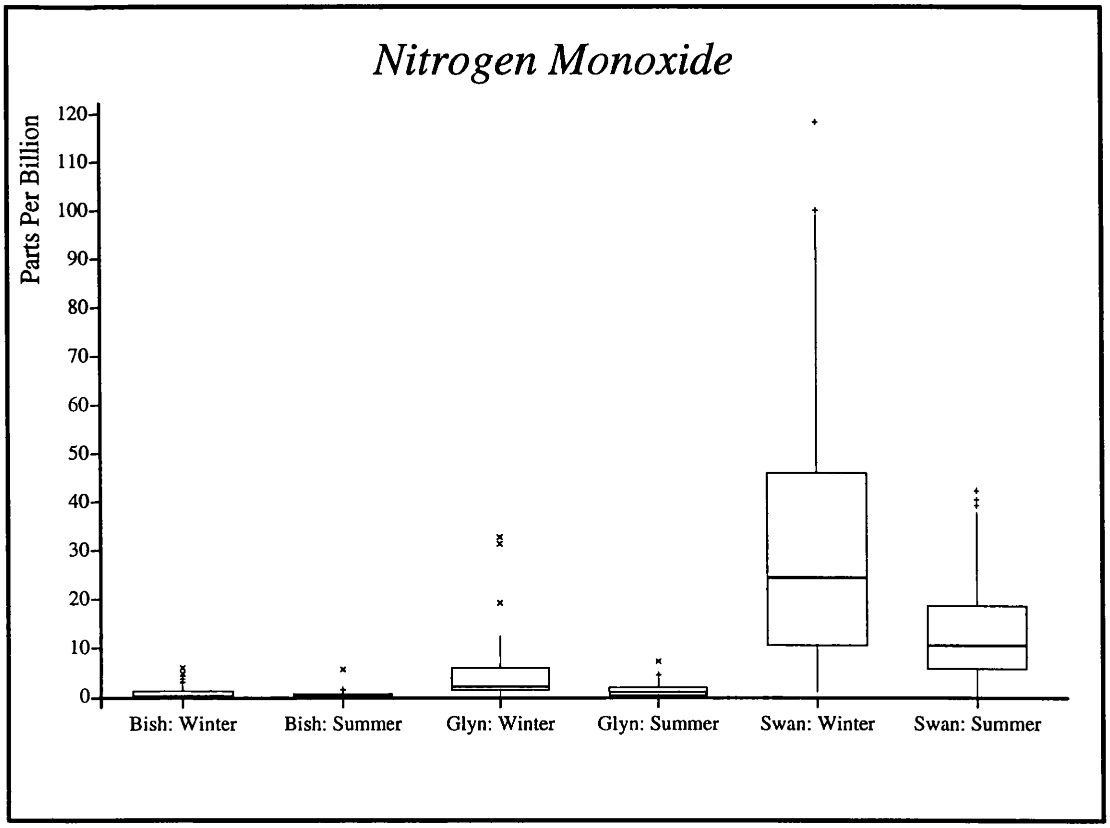
Table 4.14 Correlations between the daily pollution averages at all sites

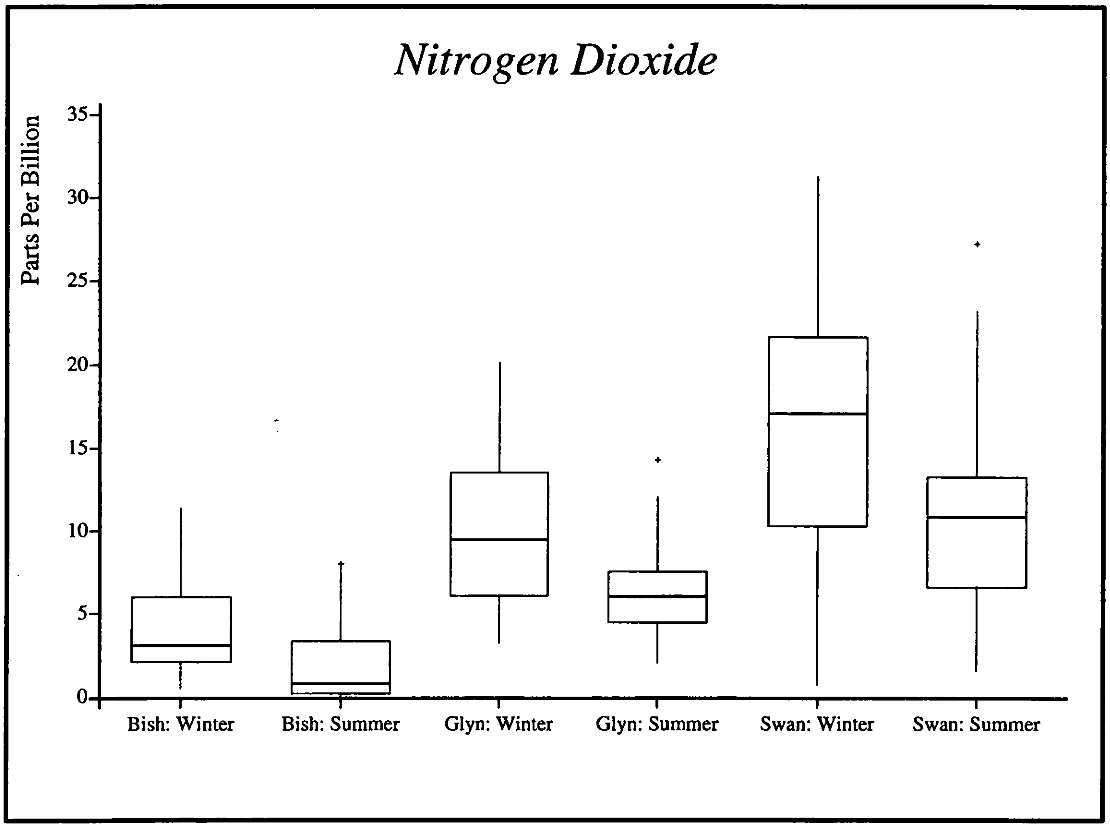
		WINTER			SUMMER		
Areas	**Pollutant**	**Correlation**	**Sig**	**N**	**Correlation**	**Sig**	**N**
Swansea	NOx	-0.11	0.50	42	-0.11	0.43	60
and	NO	-0.14	0.37	42	-0.078	0.55	61
Bishopston	NO2	0.12	0.44	42	0.44	0.0009	55
	SO2	-0.01	0.92	53	0.89	0.0001	58
	CO	....	....	....	0.08	0.63	43
	O3	0.79	0.0001	43	0.93	0.0001	61
	Dust	0.14	0.43	35	0.84	0.0001	61
Swansea	NOx	0.49	0.0001	58	0.12	0.40	49
and	NO	0.51	0.0001	58	0.19	0.20	48
Glyn	NO2	0.52	0.0001	58	0.14	0.34	49
Neath	SO2	0.70	0.0001	59	0.29	0.03	57
	CO	....	....	....	....	....	....
	O3	0.81	0.0001	58	0.59	0.0001	51
	Dust	0.39	0.0028	56	0.82	0.0001	51
Bishopston	NOx	0.51	0.0005	42	0.26	0.07	48
and	NO	0.43	0.0049	42	0.12	0.44	47
Glyn	NO2	0.49	0.0011	42	0.35	0.02	43
Neath	SO2	0.04	0.78	53	0.21	0.14	53
	CO	....	....	....	....	....	....
	O3	0.75	0.0001	43	0.59	0.0001	50
	Dust	0.01	0.97	35	0.88	0.0001	50

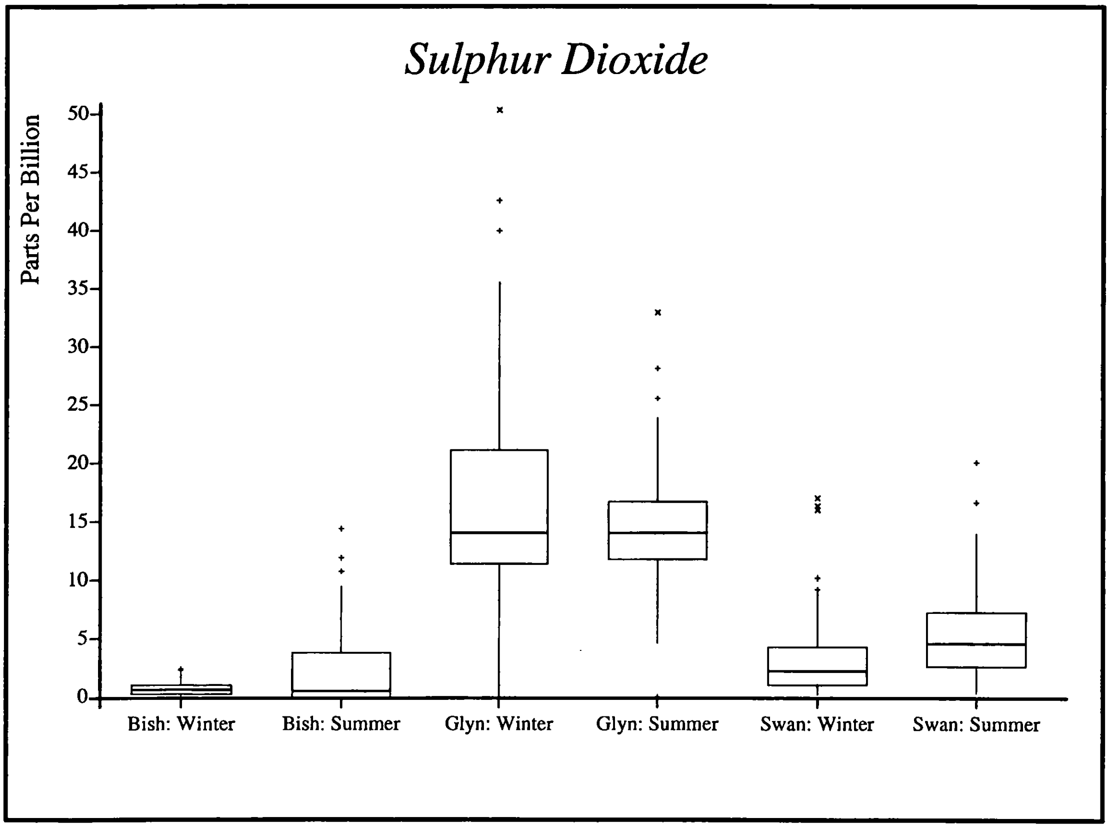
Box-and-Whisker Plots of Daily average pollution levels

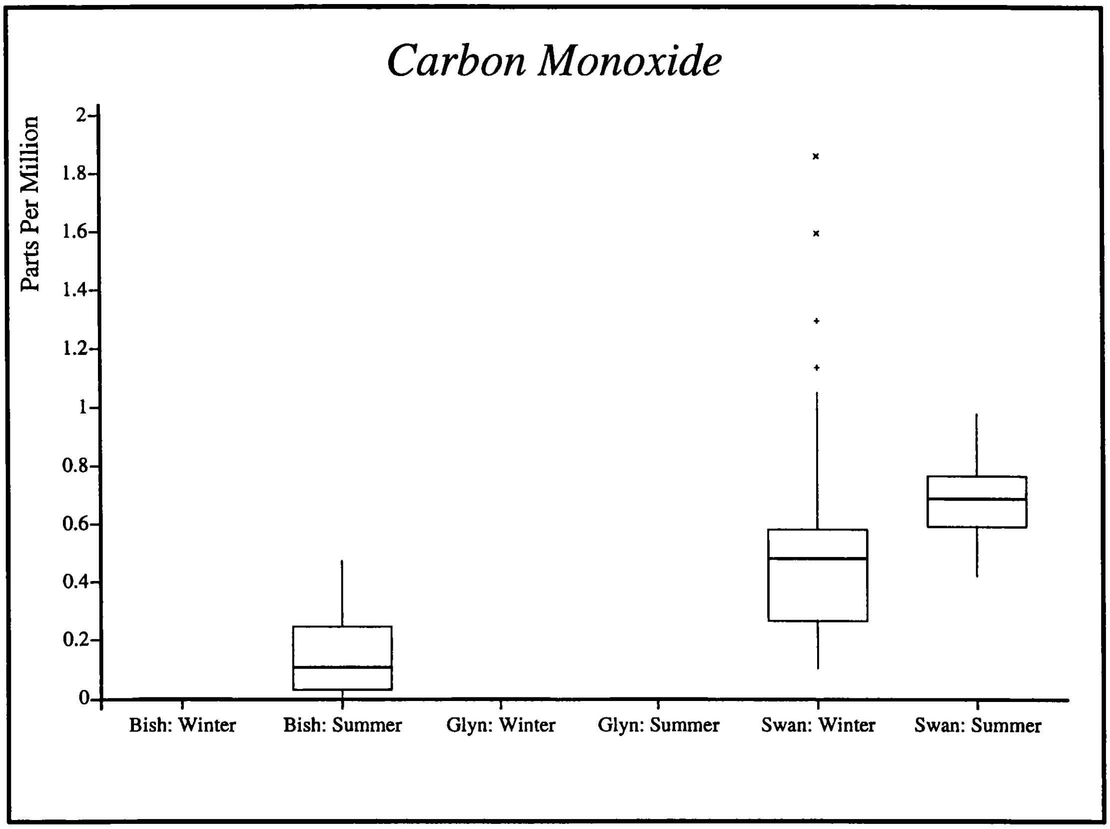
A graphical summary of the distributions of the daily pollution averages is provided in the following Box and Whisker plots for each site and season that measurements were taken. The horizontal lines in each box indicate the lower quartile, median and upper quartile of the distribution and the whiskers extend to the extremes of the distribution excluding outliers which are indicated individually. Dust in the summer in Bishopston and Swansea appear to be very similar.

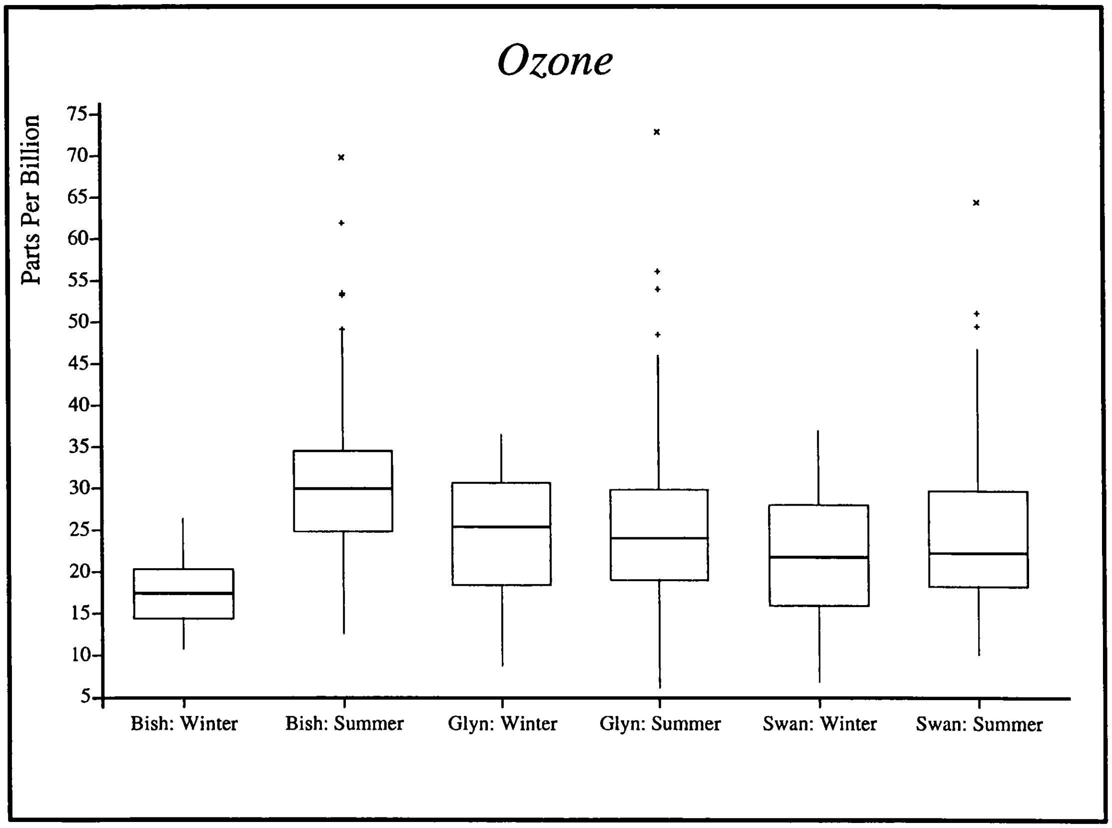

Nitrogen Oxides
Parts Per Billion
0
20
40
60
80
100
120
140
160
Bish: Winter
Bish: Summer
Glyn: Winter
Glyn: Summer
Swan: Winter
Swan: Summer

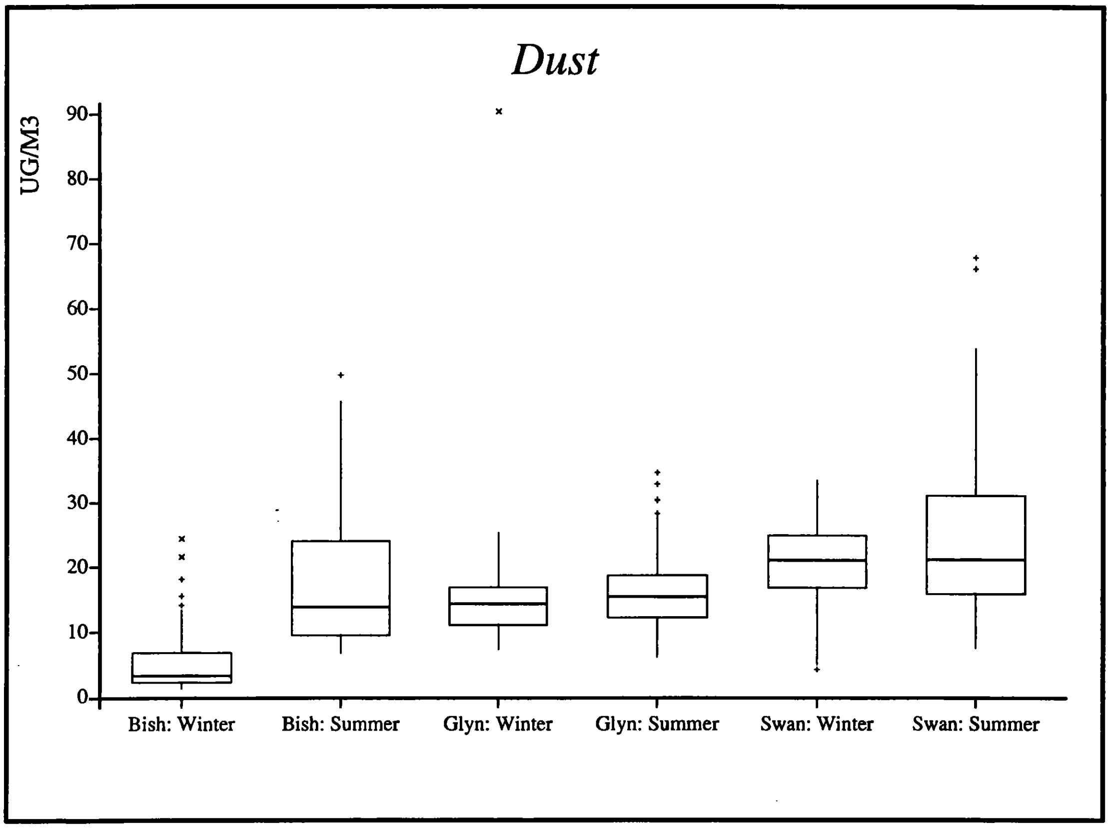

Nitrogen Monoxide
Parts Per Billion
0
10
20
30
40
50
60
70
80
90
100
110
120
Bish: Winter
Bish: Summer
Glyn: Winter
Glyn: Summer
Swan: Winter
Swan: Summer


Nitrogen Dioxide
Parts Per Billion
0
5
10
15
20
25
30
35
Bish: Winter
Bish: Summer
Glyn: Winter
Glyn: Summer
Swan: Winter
Swan: Summer


Sulphur Dioxide
Parts Per Billion
0
5
10
15
20
25
30
35
40
45
50
Bish: Winter
Bish: Summer
Glyn: Winter
Glyn: Summer
Swan: Winter
Swan: Summer


Carbon Monoxide
Parts Per Million
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Bish: Winter
Bish: Summer
Glyn: Winter
Glyn: Summer
Swan: Winter
Swan: Summer


Ozone
Parts Per Billion
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
Bish: Winter
Bish: Summer
Glyn: Winter
Glyn: Summer
Swan: Winter
Swan: Summer



Note that nitrogen oxide levels are much higher in Swansea than at the other sites. Winter levels are generally higher and more variable but Sulphur dioxide - SO2 levels in Bishopston and Swansea are higher during the summer with greater range. No CO data was recorded in Glyn Neath at all and none for Bishopston in the winter. The Ozone - O3 levels and ranges are similar for both seasons and all sites. Higher ozone levels are experienced in Bishopston during the summer months. Dust - PM10 levels in Bishopston are higher in the summer months whilst those in Swansea have more variability in the summer.

Time series plots of the pollution variables can be seen in Appendix 2.3.

4.2.3 Summary Statistics and Graphs of Weather Data.

Table 4.15 below is a summary of the daily average weather levels and their associated standard deviations, for both Swansea and Glyn Neath for both study periods.

Table 4.15 Daily weather averages and standard deviations

AREA	Variable	Units	WINTER Average	WINTER Std Dev	SUMMER Average	SUMMER Std Dev
Swansea	Rel. Humidity	%	88.65	6.85	78.51	11.02
	Temperature	Degrees C	7.57	1.85	16.47	3.15
	Wind Speed	M/ second	3.93	1.79	2.57	0.91
Glyn Neath	Rel. Humidity	%	88.55	5.42	79.92	10.64
	Temperature	Degrees C	7.03	1.88	15.82	3.49
	Wind Speed	M/ second	3.15	1.33	2.36	0.69

Graphs of the weather variables over time are displayed in Appendix 2.3.

Correlations between the daily weather averages at the Swansea and Glyn Neath sites.

Table 4.16 gives correlations between the daily averages for the weather data between the two sites. As expected all variables are highly positively correlated.

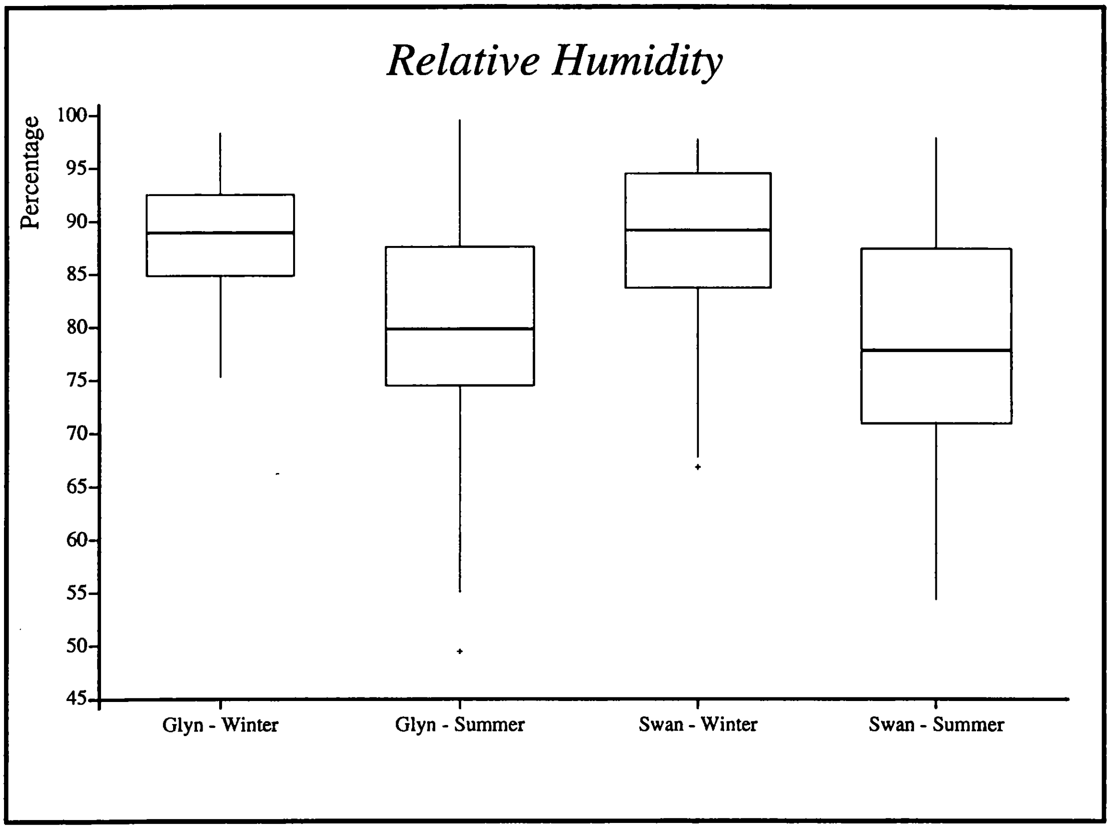
Table 4.16 Correlations between weather averages

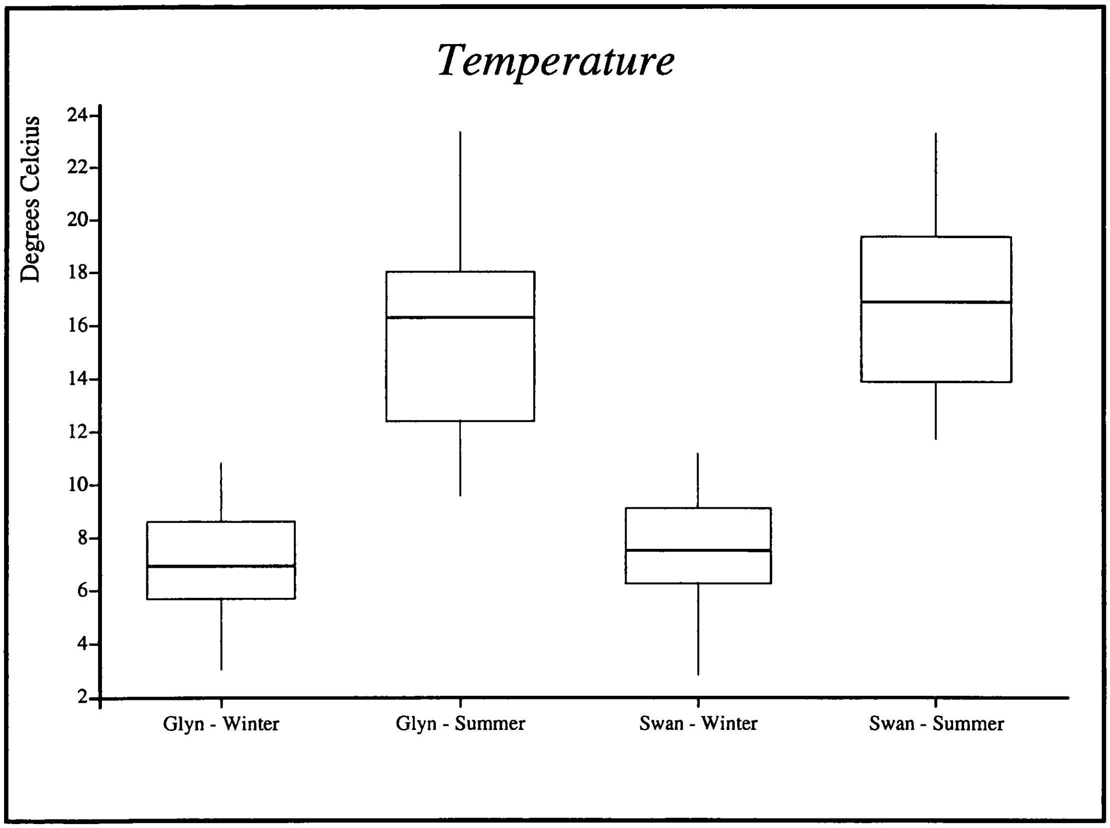
WEATHER VARIABLE	Correlation between Swansea and Glyn Neath		
WINTER:	**Correlation**	**Significance**	**Number**
Relative Humidity	0.78	0.0001	56
Temperature	0.91	0.0001	55
Wind Speed	0.87	0.0001	56
SUMMER:	**Correlation**	**Significance**	**Number**
Relative Humidity	0.87	0.0001	47
Temperature	0.98	0.0001	49
Wind Speed	0.45	0.0011	49

The correlations are significantly different from zero, and generally these correlations increase during the summer months.

Box -and Whisker plots of daily average weather levels follow. Relative humidity and Temperature were measured in Glyn Neath and Swansea only. Relative humidity levels are higher in the winter but less variable in the summer. Temperatures are higher (but more variable) in the summer as expected.

Box-and-Whisker Plots of Daily average weather levels





4.2.4 Correlations between variables within each site

Principal Component Analysis of the Nitrogen Oxides

With a view to further modelling in which correlations between predictor variables might cause problems, the relationship between the Nitrogen Oxide variables were investigated using principal component analysis. The results are presented in Table 4.17. Note that this principal component analysis was carried out on the raw data not standardised data. In the winter study the first principal component accounted for 98.5% of the total variation in the three variables (and 93.4% in the summer study). The first eigen vector, in both seasons, shows that NOx explains the most variation and the first eigen vector shows that NOx = NO + NO2, i.e. that NOx can adequately represent the other nitrogen oxide variables.

Table 4.17 Principal Component Analysis on the nitrogen oxide variables.

WINTER STUDY:		
Eigen values		
1261.6903	18.6208	0.5283
Eigen vectors		
0.7862	-0.2401	-0.5694
0.5882	0.5732	0.5704
0.1895	-0.7834	0.5919
SUMMER STUDY:		
Eigen values		
258.0187	17.8438	0.3987
Eigen vectors		
0.8084	-0.2003	-0.5535
0.5365	0.6377	0.5527
0.2423	-0.7438	0.6230

Correlations between the predictor variables.

The correlations between the predictor variables within each site were investigated. Tables 4.18, 4.19, 4.20 give correlations between potential explanatory variables for the three different sites.

Table 4.18 Swansea - Winter study

	Temp	Rhum	Rain	Air	NOx	SO2	CO	O3	Dust
Temp	1	0.68	0.05	-0.08	-0.44	-0.46	-0.48	0.56	0.28
Rhum	0.68	1	0.15	-0.11	-0.28	-0.11	-0.3	0.24	0.18
Rain	0.05	0.15	1	-0.45	-0.12	-0.15	-0.2	0.16	0.01
Air	-0.08	-0.11	-0.45	1	0.1	0.18	0.35	-0.27	0.02
NOx	-0.44	-0.28	-0.12	0.1	1	0.29	0.8	-0.71	0.09
SO2	-0.46	-0.11	-0.15	0.18	0.29	1	0.55	-0.7	0.22
CO	-0.48	-0.3	-0.2	0.35	0.8	0.55	1	-0.8	0.25
O3	0.56	0.24	0.16	-0.27	-0.71	-0.7	-0.8	1	0.02
Dust	0.28	0.18	0.01	0.02	0.09	0.22	0.25	0.02	1

Table 4.19 Bishopston - Winter study

	Temp	Rhum	Rain	Air	NOx	SO2	O3	Dust
Temp	1	0.52	-0.09	0.17	0.27	0.09	0.33	0.04
Rhum	0.52	1	0.01	0.29	0.51	0.25	-0.22	0.34
Rain	-0.09	0.01	1	-0.38	-0.2	-0.05	0.06	-0.21
Air	0.17	0.29	-0.38	1	0.33	0.52	-0.31	0.19
NOx	0.27	0.51	-0.2	0.33	1	0.4	-0.52	0.12
SO2	0.09	0.25	-0.05	0.52	0.4	1	-0.52	0.05
O3	0.33	-0.22	0.06	-0.31	-0.52	-0.52	1	-0.21
Dust	0.04	0.34	-0.21	0.19	0.12	0.05	-0.21	1

Table 4.20 Glyn Neath - Winter study

	Temp	Rhum	NOx	SO2	Ozone	Dust
Temp	1	0.29	-0.39	-0.5	0.41	0.21
Rhum	0.29	1	0.22	-0.02	-0.33	0.09
NOx	-0.39	0.22	1	0.27	-0.83	0.26
SO2	-0.5	-0.02	0.27	1	-0.44	-0.14
Ozone	0.41	-0.33	-0.83	-0.44	1	-0.09
Dust	0.21	0.09	0.26	-0.14	-0.09	1

During the winter study, few high correlations occur. Temperature and Relative Humidity are highly **positively** correlated at Swansea. Nitrogen Oxides (NOx) and Ozone (O3) are highly **negatively** correlated at all sites. This is possibly due to the complex chemical reactions taking place in the atmosphere, the formation of Nitrogen Oxides NOx for instance requires many Oxygen atoms. The correlation between SO2 and O3 is negative during the winter months and positive during the summer study.

Note that the negative correlation between SO2 and O3 is not as large as at the other sites - possibly because Glyn Neath has much higher SO2 values than the other sites.

Table 4.21 Swansea - Summer study

	Temp	Rhum	Rain	Air	NOx	SO2	CO	O3	Dust
Temp	1	-0.47	-0.07	-0.03	-0.34	0.7	0.05	0.57	0.74
Rhum	-0.47	1	0.4	-0.35	0.01	-0.61	-0.32	-0.44	-0.56
Rain	-0.07	0.4	1	-0.5	-0.11	-0.22	-0.24	0.05	-0.11
Air	-0.03	-0.35	-0.5	1	0.12	0.25	0.06	-0.29	0.04
NOx	-0.34	0.01	-0.11	0.12	1	-0.16	0.62	-0.45	-0.2
SO2	0.7	-0.61	-0.22	0.25	-0.16	1	0.21	0.5	0.87
CO	0.05	-0.32	-0.24	0.06	0.62	0.21	1	0.05	0.21
O3	0.57	-0.44	0.05	-0.29	-0.45	0.5	0.05	1	0.71
Dust	0.74	-0.56	-0.11	0.04	-0.2	0.87	0.21	0.71	1

Table 4.22 Bishopston - Summer study

	Temp	Rhum	Rain	Air	NOx	SO2	CO	O3	Dust
Temp	1	-0.46	-0.08	-0.03	0.54	0.57	-0.59	0.49	0.65
Rhum	-0.46	1	0.41	-0.35	-0.59	-0.56	-0.05	-0.51	-0.53
Rain	-0.08	0.41	1	-0.5	-0.21	-0.17	-0.09	0	-0.11
Air	-0.03	-0.35	-0.5	1	0.32	0.33	0.14	-0.25	0.07
NOx	0.54	-0.59	-0.21	0.32	1	0.88	-0.23	0.43	0.8
SO2	0.57	-0.56	-0.17	0.33	0.88	1	-0.28	0.37	0.73
CO	-0.59	-0.05	-0.09	0.14	-0.23	-0.28	1	-0.04	-0.36
O3	0.49	-0.51	0	-0.25	0.43	0.37	-0.04	1	0.7
Dust	0.65	-0.53	-0.11	0.07	0.8	0.73	-0.36	0.7	1

Table 4.23 Glyn Neath - Summer study

	Temp	Rhum	NOx	SO2	Ozone	Dust
Temp	1	-0.39	0.09	0.29	0.52	0.63
Rhum	-0.39	1	0.12	-0.2	-0.63	-0.38
NOx	0.09	0.12	1	0.16	-0.21	0.32
SO2	0.29	-0.2	0.16	1	0.11	0.23
Ozone	0.52	-0.63	-0.21	0.11	1	0.59
Dust	0.63	-0.38	0.32	0.23	0.59	1

Sulphur Dioxide, Ozone and Dust are highly positively correlated with temperature at Bishopston and Swansea. At Glyn Neath, Dust and Ozone are highly positively correlated with Temperature. The correlation between SO2 and NOx is inconsistent across the sites, being high and positive for Bishopston but not for the other two sites, suggesting perhaps that at Bishopston the low levels of NOx and SO2 have a common source, presumably exhaust emissions. The SO2 in Glyn Neath could come from either the heavy traffic using the A465 dual carriageway or it could come from local industry.

5. A Cross Sectional Analysis of Peak Flow measurements

5.1 Aims of the study

The main aims of this study were to assess whether the daily variation in air pollution levels induce daily variation in the lung function of the schoolchildren or other health symptoms (either creating them or prolonging them). If effects were found investigations attempted to discover which of the pollutants were implied. In this chapter we investigate the informative content of each child's average peak flow measurements.

A definite seasonal trend was found and so the winter and summer study periods were examined separately. Various subgroups were looked at to see if the effects of the air pollutants varied between site, sex, age, height, ethnic group or children with asthmatic tendencies. Since many previous studies, (see chapter 3), suggest the importance of the meteorological conditions on the effects of the air pollution, temperature and relative humidity were included in the regression models. The lag effects of air pollution on lung function were investigated and compared for the different locations and subgroups. These subgroups were:

- the different sites
- males and females
- wheezers and non-wheezers
- asthmatics and non-asthmatics
- passive smokers and non-passive smokers

5.2 Peak Expiratory Flow Averages

In this chapter we look at the variation in the average peak flow of each child and relate it to appropriate explanatory variables (e.g. height). Although this variation is not the main purpose of the study, the wealth of information present should not be ignored. Further there is the potential to investigate, given the large numbers in the samples, the significance or otherwise of potential home-environmental effects (e.g. passive smoking, cooking by gas, presence of pets) on the overall levels of peak flows.

5.2.1 Personal Peak Expiratory Flow Averages.

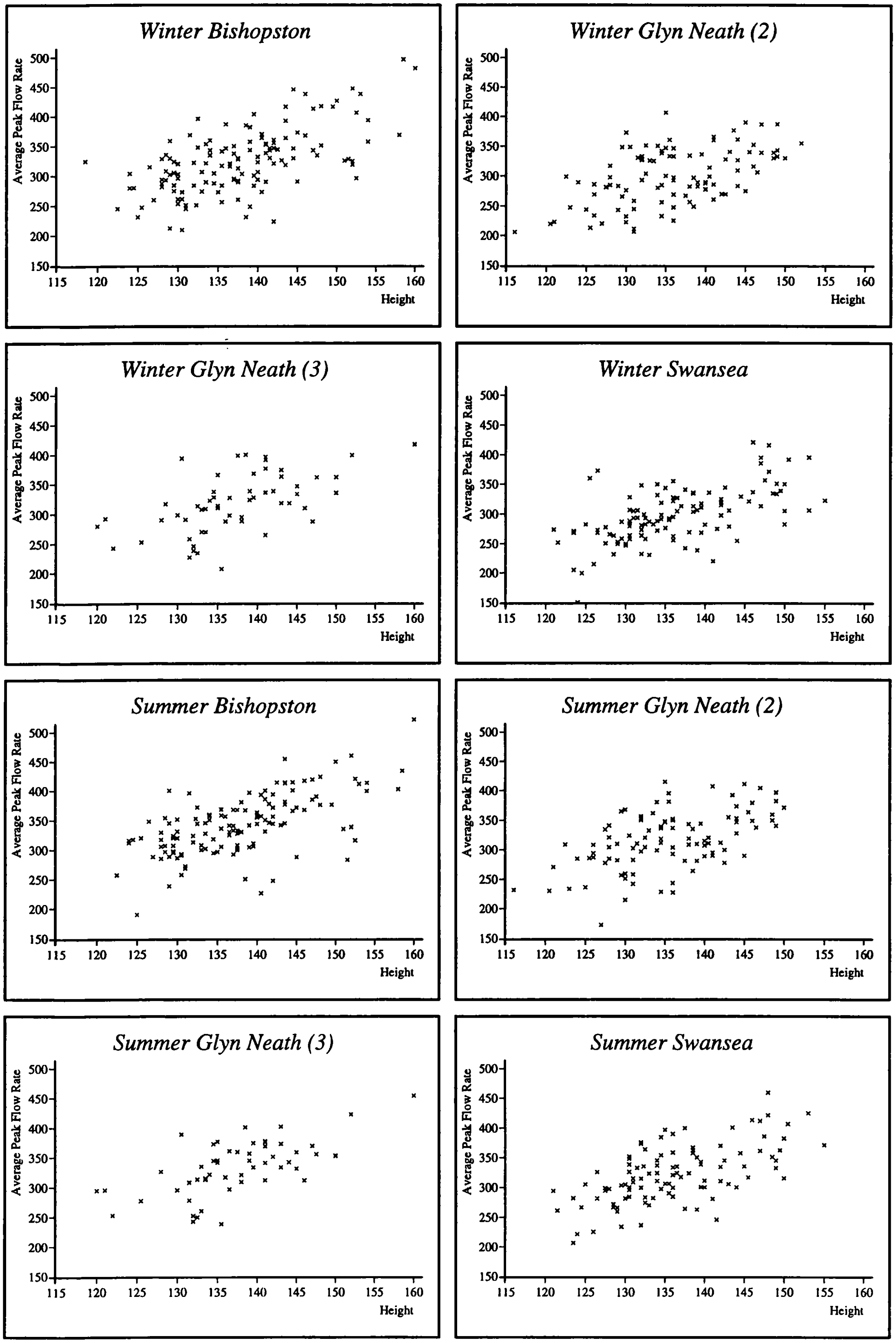
For both study periods, a maximum of 30 peak flows were recorded for each individual involved in the study, corresponding to Monday - Friday of six consecutive school weeks. The peak flow measurements are in litres per minute. Another file of data gathered recorded the quality of the peak flow rate measurement. If the peak flows were thought to have been made up by the children, then they were coded as suspicious.

For the purpose of this analysis, any suspicious readings were declared to be missing values. For each individual an overall average peak flow rate of admissible readings was calculated for each study period separately.

Some children left after the winter period, and others arrived before the summer period. Out of a total number of 426 children; for the winter there were 414 means (97.2%), and for the summer there were 396 means (93%).

Fig 5.1 shows the dependence of peak expiratory flow rate on height. For each site and season the peak flow rates increase linearly with height with a slope of 3.25 litres per minute per centimetre of height (see page 58). This linear trend appears to be consistent over all sites and for both seasons, (p=0.4518). Note however that no separate height measurement was made of children for the summer study period.

Fig 5.1 Dependence of peak flow rates on height



Winter Bishopston

Winter Glyn Neath (2)

Winter Glyn Neath (3)

Winter Swansea

Summer Bishopston

Summer Glyn Neath (2)

Summer Glyn Neath (3)

Summer Swansea

5.2.2 Analysis of Peak Expiratory Flow Rates with respect to the categorical data, allowing for height.

The most important variables collected were:

- Site
- Sex
- Age
- Ethnic Group
- Height
- Form in school
- Whether they were asthmatic
- Whether they were a wheezer
- Whether they were found to suffer from exercise-induced asthma

Other categorical / qualitative variables are discussed later in this section.

5.2.3 Regression analysis using the covariate data.

To make sensible inferences about average peak flow levels and their differences across sites, sexes and season, the first regression analysis model used involved factors: season, sex, site, ethnic group, whether the child was asthmatic, or had exercise-induced asthma, together with covariates: age and height. These predictor variables, together, accounted for 46% of the variation in average peak flows - see Table 5.1 for the results. In this table the intercept term (for zero age and height) assumes that the season is winter, the site is 1=Bishopston, sex is "male", the ethnic group is 1="Bangladeshi" for non-asthmatic children. As a consequence the parameters listed, e.g. site 2, indicate the adjustment to be made to the intercept term to obtain the correct average for site 2=Glyn Neath.

Table 5. 1 First Regression Model

Source	SS	DF	MS
Due to model	2.338E7	11	2.126E6
Residual	2.733E7	757	3.610E4
Total Corrected	5.071E7	768	6.603E4

Parameter	Estimate	Standard Error	t-statistic	p-value	
Intercept	-157.300	26.030	-6.0410	2.397E-9	**
season 2	20.600	2.720	7.5740	1.060E-13	**
asthmatic	-5.014	3.450	-1.4530	1.465E-1	
site 2	-16.000	3.620	-4.4210	1.128E-5	**
site 3	-3.184	4.255	-0.7485	4.544E-1	
site 4	-17.880	3.673	-4.8670	1.377E-6	**
sex 2	-15.560	2.739	-5.6800	1.920E-8	**
ethnic 2	-27.040	8.731	-3.0970	2.025E-3	**
ethnic 3	18.310	26.120	0.7011	4.835E-1	
pfdrop	9.327	5.875	1.5880	1.128E-1	
age	6.960	1.891	3.6810	2.2492E-4	**
height	3.250	0.224	14.5000	2.220E-16	**

** = Highly significant at the 1% level.

Consequently one can see that the adjustments for site (allowing for age, height etc.) are all downwards, significantly so for Site 2=Glyn Neath, and Site 4=Swansea where the adjustments are 16 and 17.88 litres per minute respectively, figures which are not only statistically very significant, but physiologically significant.

The ethnic group results were interesting too; Bangladeshi children had rather better peak flow rates than Ethnic Group 2 (White) whose peak flow on average was 27 litres per minute lower.

The effect of summer was significant (season 2), with peak-flows being 20.6 litres per minute higher on average in the summer.

Being asthmatic led to a lower, (but not significantly lower), peak flow than non-asthmatic children. Finally, it is interesting to note that the age variable was statistically significant even in the presence of the height variable though the effect (6 litres per minute per extra year) is not as high as the other effects we have discussed.

The following table, Table 5.2, shows the predicted estimates of the mean peak flow rate for a white child (ethnic group 2 who were the majority), aged 10 years and of average height 137 cm together with associated 95% confidence intervals

Table 5. 2 Predicted estimates for a 10 year old, white child, of height 137cm

WINTER STUDY	Male			Female		
	Avge	95% CI		Avge	95% CI	
Bishopston	330.6	324.1	337.0	315.0	308.6	321.5
Glyn Neath (2)	314.6	307.7	321.4	299.0	292.2	305.9
Glyn Neath (3)	327.4	318.9	335.9	311.8	303.7	320.0
Swansea	312.7	305.6	319.8	297.1	290.3	304.0

SUMMER STUDY	Male			Female		
	Avge	95% CI		Avge	95% CI	
Bishopston	351.2	344.6	357.8	335.6	329.0	342.2
Glyn Neath (2)	335.2	328.2	342.1	319.6	312.7	326.6
Glyn Neath (3)	348.0	339.4	356.6	332.4	324.1	340.7
Swansea	333.3	326.0	340.6	317.7	310.8	324.7

Note how, in terms of overall average peak flow rate, Glyn Neath 2 and Swansea appear to be on a par with each other, as do Bishopston and Glyn Neath 3. The reasons for this are not clear, but it may well have something to do with the quality of life provided for the child in his or her own home environment.

Including the questionnaire responses in the model:

Various responses to the questionnaire were included, one at a time, and tested for significance along with various interactions of categorical data variables. A site-sex interaction term was also included. This was necessary because the adjustment for being female, depended on which site the child was at, with the Glyn Neath sites requiring a reduction of 34 and 28 litres per minute respectively, whereas for Swansea girls, the reduction required was only 12.3 litres per minute.

Adding Q10, (*Do you use gas for cooking?*), was significant at the 5% level leading to an estimated drop of 5.8 litres per minute for cooking using gas. The second regression model (see Table 5.3) accounted for nearly 49% of the total variation in average peak flows. No other variables appeared to affect the overall averages significantly (not even the variable *pfdrop*).

Table 5. 3 Regression analysis including some interactions

Source	SS	DF	MS
Due to model	2.443E7	15	1.629E6
Residual	2.628E7	753	3.490E4
Total Corrected	5.071E7	768	6.603E4

Parameter	Estimate	Standard Error	t-statistic	p-value	
Intercept	-163.100	26.160	-6.2350	7.538E-10	**
season 2	20.490	2.674	7.6610	5.684E-14	**
asthmatic	-4.827	3.402	-1.4190	1.563E-1	
site 2	-0.804	4.917	-0.1635	8.701E-1	
site 3	10.600	6.167	1.7190	8.597E-2	-
site 4	-10.970	5.259	-2.0870	3.723E-2	*
sex 2	0.914	4.659	0.1962	8.445E-1	
ethnic 2	-27.780	8.589	-3.2350	1.271E-3	**
ethnic 3	17.690	25.790	0.6862	4.928E-1	
pfdrop	8.949	5.785	1.5470	1.223E-1	
age	6.594	1.866	3.5340	4.343E-4	**
height	3.287	0.226	14.5600	0.0000E0	**
q10	-5.823	2.922	-1.9930	4.664E-2	*
site 2 sex 2	-34.220	7.092	-4.8250	1.696E-6	**
site 3 sex 2	-28.240	8.409	-3.3590	8.229E-4	**
site 4 sex 2	-12.330	7.031	-1.7540	7.992E-2	*

Key: - = Significant at the 10% level.
* = Significant at the 5% level.
** = Highly significant at the 1% level.

Table 5. 4 New Predicted estimates for a 10 year old, white child, of height 137cm

WINTER	**Male**			**Female**		
	Avge	95% CI		Avge	95% CI	
Bishopston	325.4	317.7	333.1	326.3	318.4	334.3
Glyn Neath (2)	324.6	316.4	332.8	291.3	283.0	299.6
Glyn Neath (3)	336.0	324.9	347.1	308.7	298.5	318.8
Swansea	314.4	305.0	323.9	303.0	294.2	311.8

SUMMER	**Male**			**Female**		
	Avge	95% CI		Avge	95% CI	
Bishopston	345.9	338.0	353.7	346.8	338.7	354.9
Glyn Neath (2)	345.1	336.8	353.4	311.8	303.4	320.2
Glyn Neath (3)	356.5	345.3	367.6	329.2	318.9	339.4
Swansea	334.9	325.4	344.5	323.5	314.5	332.5

Site interacting with sex was also highly significant, especially with females at Glyn Neath having the lower averages.

5.2.4 Differences between the subgroups of children

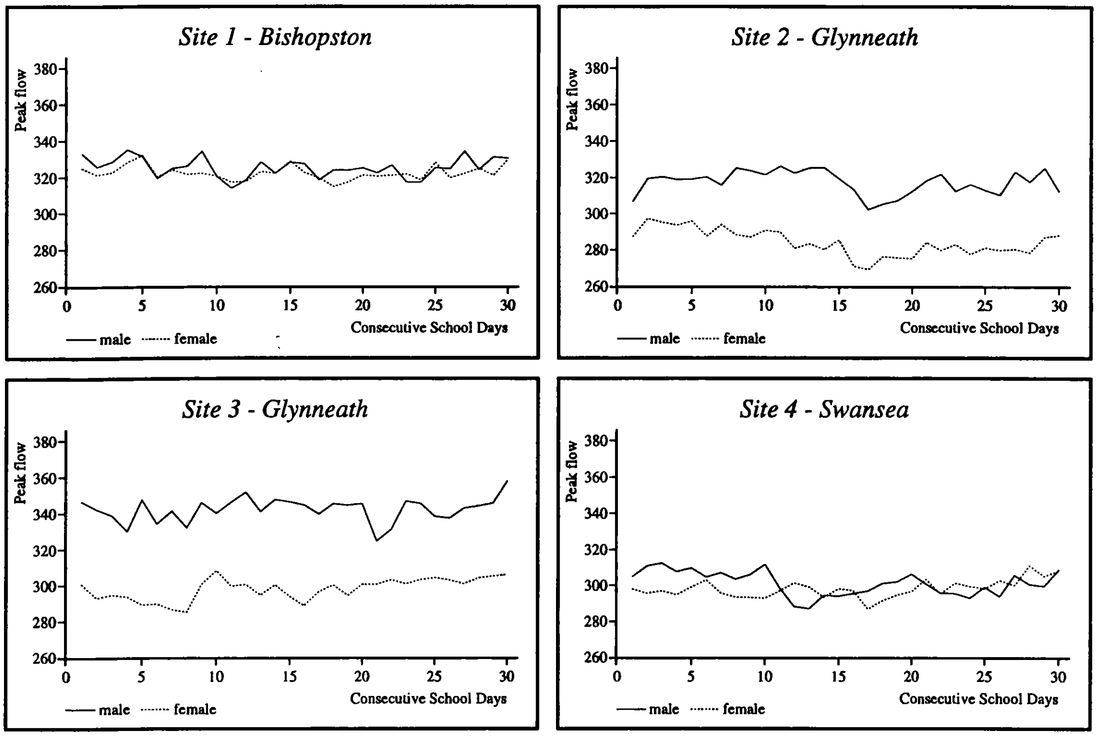
This effect, of site interacting with sex, anticipates differences that can be clearly seen when the daily average peak flow time series are plotted - see Fig 5.2 where similar patterns for both seasons emerge for each subgroup of children. The males and females at Bishopston and Swansea have similar peak flows, with males' peak flows being slightly higher. However at the Glyn Neath sites the females have much lower peak flow rates on average than the males. This sex effect is quite remarkable and since it is consistent across season, there is no need for a site-sex-season interaction.

Fig 5.3 shows the day-to-day trend for asthmatics and non-asthmatics separately at each site, the average peak flow rates for each subgroup being plotted for particular days. Again a noticeable difference in peak flow rates can be seen in the graphs between the two groups of children.

- The daily average peak flow rates for the asthmatics are much more variable for all sites.
- The levels at Bishopston were similar for asthmatics and non-asthmatics.
- At Glyn Neath (2) and Swansea, the asthmatics have a lower daily average peak flow rates, but at Bishopston and Glyn Neath (3) they have the higher averages.
- The asthmatics at Bishopston and Glyn Neath (3) show similar levels of peak flow rates and have higher average levels than Glyn Neath (2) and Swansea, who were also similar to each other.
- This difference between sites is more noticeable during the winter period.
- Non-asthmatics' graphs are similar across sites, with peak flow rates being slightly higher during the summer period.

Fig 5.2 Graphs showing the daily variation in average peak flow rates between sex, site and season.

Winter



Summer

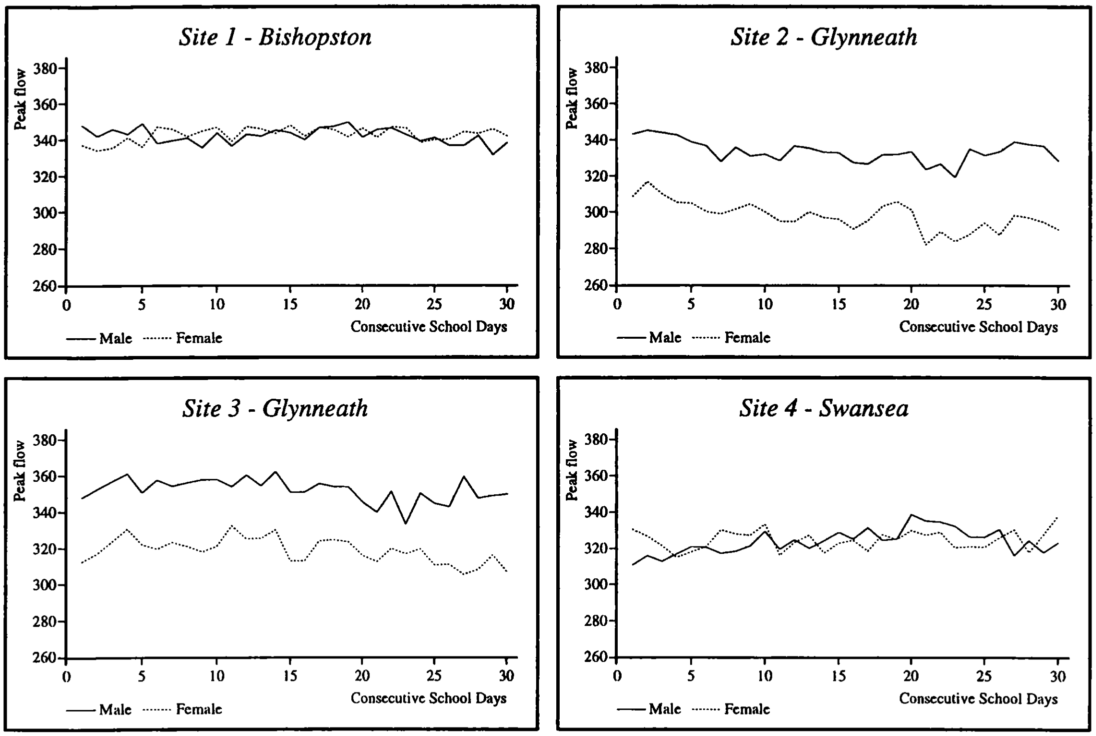
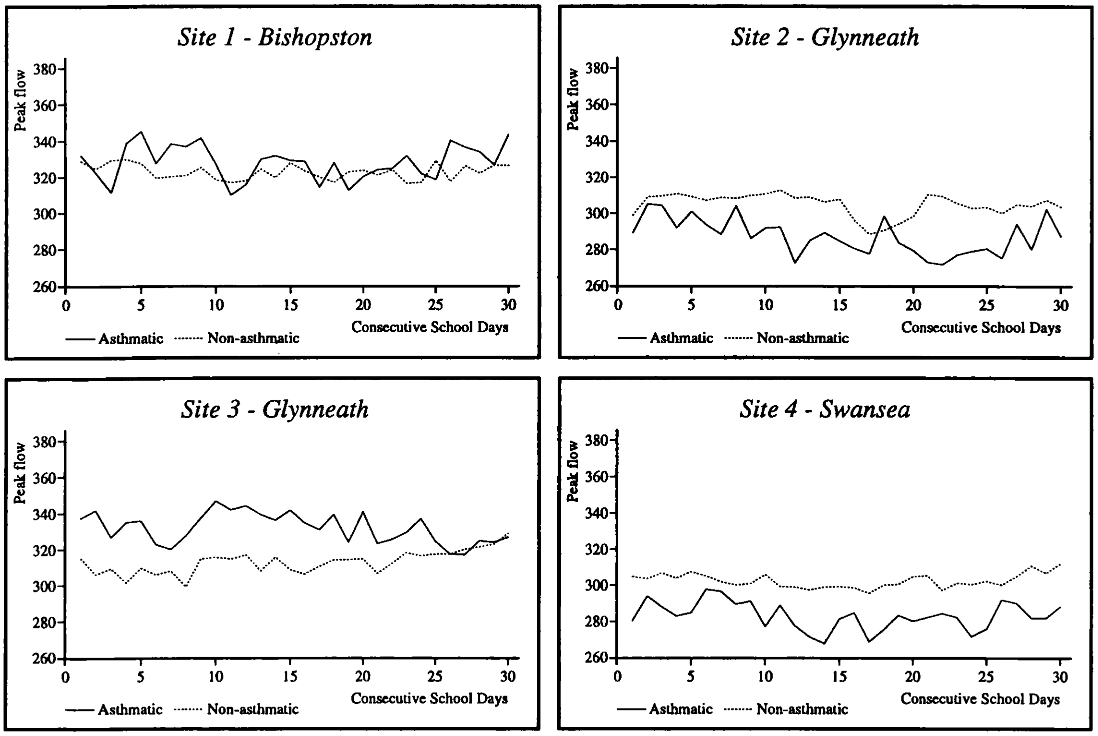


Fig 5.3 Graphs of daily variation in average peak flow rates for asthmatics and non-asthmatics.

Winter



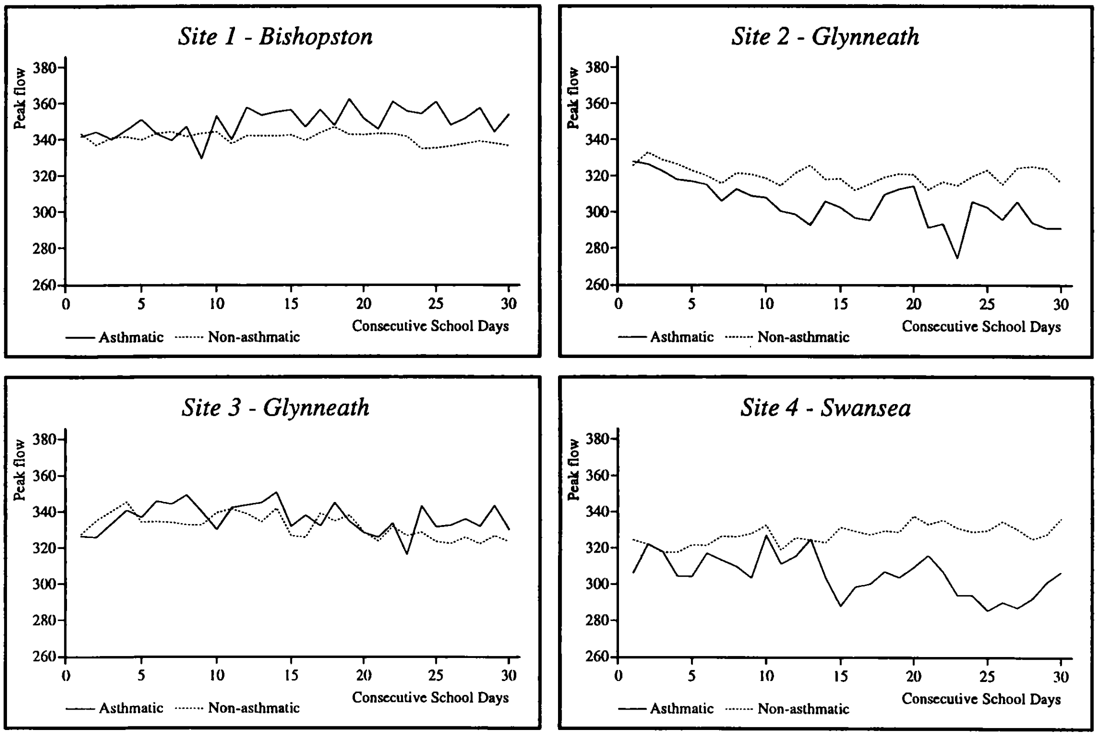
Site 1 - Bishopston

Site 2 - Glynneath

Site 3 - Glynneath

Site 4 - Swansea

Summer



Site 1 - Bishopston

Site 2 - Glynneath

Site 3 - Glynneath

Site 4 - Swansea

The daily variations in peak flow rates for the wheezers and the non-wheezers can be seen in Appendix 2.4. The graphs show similar trends to those for asthmatics and non-asthmatics as described above, except that the graphs do not fluctuate quite as much.

5.2.5 Seasonal Change in Peak Flow Rate.

301 children's overall average peak flow rate increased from the winter study to the summer study, and 72 decreased. Note that only 388 children had **both** a winter and a summer overall average, but only 373 of those answered **all** the questions on their questionnaire, these are the children that are considered below.

Generally, average peak flow rates increased by 20 units in the summer months. Note however that some of this increase could be attributed to growth during the six months, and since height was not measured again, this contribution cannot be allowed for. This increase could also be due to the warmer weather and the decrease in colds or viral infections. At Glyn Neath (2) and Swansea, more females than males, (but not significantly more), had a decreased peak flow rate in the summer. Note also the two Swansea males who had a decreased peak flow rate in the summer.

Table 5. 5 Children who increased their peak flow rate in the summer.

SITE	Male	Female	Total
1	52	49	101
2	40	31	71
3	14	25	39
4	42	48	90
Total	148	153	301

AGE	Male	Female	Total
8	24	26	50
9	60	49	109
10	46	59	105
11	17	19	36
12	1	0	1
Total	148	153	301

Table 5. 6 Children who decreased their peak flow rate in the summer.

SITE	Male	Female	Total
1	16	11	27
2	7	14	21
3	9	4	13
4	2	9	11
Total	34	38	72

AGE	Male	Female	Total
8	3	6	9
9	14	12	26
10	10	13	23
11	7	7	14
Total	34	38	72

The basic frequency tables above compare the numbers of children who increased or decreased their overall average peak flow rate during the year across site, age and sex.

To appraise the statistical significance or otherwise of these frequencies Binary Logistic Regression techniques were used. The dependent variable was equal to 1 if that child's average summer peak flow was less than his/her average winter peak flow and 0 otherwise.

This analysis found that the only variables which significantly affected the change in the seasonal peak flow averages were site, ethnic group and Q5 on the questionnaire. (Recall that Q5 is *In the last 12 months, has wheezing been severe enough to limit your child's speech to only one or two words at a time between breaths*?)

Table 5. 7 Binary regression

Variable	Estimate	Standard Dev	t-statistic	p-value	
Intercept	-0.2236	0.9106	-0.2455	0.8062	
Glyn Neath (2)	0.0372	0.3342	0.1114	0.9113	
Glyn Neath (3)	0.1927	0.3888	0.4958	0.6204	
Swansea	-1.1630	0.4518	-2.5730	0.0105	**
White children	-1.1170	0.8849	-1.2620	0.2078	
Other	9.6460	37.7300	0.2557	0.7984	
Q5	1.0500	0.6065	1.7320	0.0841	*

* = Significant at the 10% level

** = Significant at the 5% level

The negative sign for the Site=Swansea parameter and its statistical significance indicates that significantly fewer Swansea children reduced their average peak flows in the transition from winter to summer, when compared to Bishopston. Possibly the constant pollution present has caused the children to have consistently low peak flow rates all year around, whereas at the other sites they may have cleaner days in which their peak flow rates recover and fluctuate seasonally. The Glyn Neath sites showed that more children had lower averages in the summer, but this was not significantly more. The parameter estimate for Q5 (which would be significant at the 5% level for a one-sided test) indicates that the group of 13 children for which Q5=1 (see Table 5.7) had significantly more children with reduced average peak flows in the summer than one would have expected otherwise.

A wheezer was defined to be anyone who had answered **YES** to **any** of the following questions; Q1, Q4, Q6, Q8, Q9, or had an exercise-induced peak flow drop.

Overall children increased their peak flow by 20 units in the summer period, (see Table 5.3), some of which could be attributable to a physical increase in size.

There were significant decreases in peak flow rates, from winter to summer time, than expected amongst children who had:

(a) Experienced wheezing in the past (Q1).
(b) Had their sleep disturbed due to wheezing in the past 12 months (Q4).
(c) In the past 12 months, had wheezing which was severe enough to limit their speech to one or two words at a time between breaths (Q5).
(d) Asthmatics who had received treatment for asthma in the last 4 weeks (Q7).
(e) In the last 12 months, had a dry cough at night, apart from a cough associated with a cold or chest infection (Q9).
(f) Had exercise-induced asthma.
(g) Were classed as a wheezer.

- More wheezers decreased their peak flow rates (37 children) than expected (30 children).
- 7 children with exercise-induced asthma had decreased peak flow rates in the summer. The expected number was 4 children.
- 27 children who had experienced wheezing in the past (Q1) had a decreased average in the summer, only 22 were expected to decrease.
- 22 children who had experienced a dry cough at night (Q9) had a decreased average, only 18 were expected to decrease.

- More children who used gas for cooking in the home (Q10) and had calor bottle gas heating in the home (Q11) had increased their peak flow rate average (155 and 108 children) in the summer than expected (148 and 100 children respectively).

5.2.6 Change in seasonal average peak flow rate.

Out of the children who had both a winter and a summer overall average peak flow rate (388 children), the difference between the two was calculated (diffmean = summer mean - winter mean). Thus a value of diffmean > 0, indicates that the child's peak flow average increased in the summer study. Table 5.8 shows that some quite large differences could occur, particularly for some children at the Swansea site.

Table 5. 8 Number of children experiencing a large seasonal change in peak flow rate

SITE	Number of children	Number who increased or decreased by more than 50 units.	
Bishopston	128	10	(8%)
Glyn Neath (2)	92	9	(10%)
Glyn Neath (3)	52	3	(6%)
Swansea	99	18	(18%)

- 12% of asthmatics had a change of more than 50 units in their average peak flow rates.
- A third of those with exercise-induced asthma also had a change of more than 50 units in their seasonal averages.

Regression analysis was carried out on site, sex, ethnic group, age, height, exercise-induced asthmatics and the response to question 5 (*In the last 12 months, has wheezing been severe enough to limit speech to one or two words between breaths?*).This regression model accounted for only 7.88% of the variation of the change in seasonal average peak flow rate - see Table 5.9 for the results. Thus the predictor variables used have little predictive power.

In view of the poor level of explanatory power of this model it is not appropriate to attach much importance to the results. However note that:-

- the coefficient for Swansea, equal to 11.56 with standard error 3.402, confirms once again that the children in Swansea tended to significantly increase their average peak flow in Summer over the equivalent in Winter whilst children at the other sites did not.

Table 5. 9 Regression analysis for seasonal change in peak flow rate

Source	SS	DF	MS
Due to model	1.826E4	10	1826
Residual	2.134E5	360	592.7
Total Corrected	2.316E5	370	626

Parameter	Estimate	Standard Error	t-statistic	p-value	
Intercept	16.250	24.160	0.6729	0.5014	
site 2	0.878	3.385	0.2593	0.7955	
site 3	-3.022	4.024	-0.7511	0.4531	
site 4	11.560	3.402	3.3990	0.0008	**
sex 2	3.310	2.562	1.2920	0.1972	
ethnic 2	6.341	8.162	0.7769	0.4377	
ethnic 3	-34.600	25.570	-1.3530	0.1769	
age	-4.046	1.771	-2.2840	0.0229	*
pfdrop	2.434	5.433	0.4481	0.6544	
height	0.234	0.209	1.1170	0.2648	
q5	-15.140	7.010	-2.1600	0.0314	*

* = Significant at the 5% level.

** = Highly significant at the 1% level.

- the coefficient of height was not significantly different from zero as expected, but the coefficient of Age was negative and significant at the 5% level (p=0.02), a result which is difficult to interpret.
- the coefficient for Q5 was also negative and significant at the 5% level, (p=0.03), thus those children who had suffered sever wheezing in the last 12 months tended to decrease their average peak flows as they moved into the summer. This may be because they are allergic to summer allergens such as pollens.

5.2.7 Chi-Squared Tests.

For each day of the studies, the number of children at each site whose peak flow rate dropped from the previous day was counted. A simple Binomial test of the hypothesis p = probability of decrease = 0.5 versus p > 0.5 for that day was computed. Tables 5.10 and 5.11 show the days in each study where significantly more than 50% of the children had a drop in their peak flow rate from the previous day.

Table 5. 10 Winter Study

SITE	Day of study with respect to pollution data 1st January 1995 (Day 1) - 28th February 1995 (Day 59)
Bishopston	10, 16, 25, 27, 33, 44, 45
Glyn Neath (2)	10, 30, 37, 41, 44, 45, 47
Glyn Neath (3)	19, 32, 41
Swansea	48

Table 5. 11 Summer study

SITE	Day of study with respect to pollution data 1st June 1995 (Day 1) - 1st August 1995 (Day 62)
Bishopston	20, 23, 26, 34, 40, 43
Glyn Neath (2)	8, 12, 22, 34, 36, 41
Glyn Neath (3)	9, 19, 23, 28, 29, 33, 34, 36, 37, 41, 42
Swansea	16, 19, 28

Notice how Swansea has very few days when a significant number of pupils have a reduced peak flow from the day before. Glyn Neath (3), is particularly affected during the summer study. Site 1 - Bishopston and Glyn Neath (2) are affected similarly during both studies.

It seems surprising that Swansea, the site with overall highest pollution has fewest days where over half the children experience a drop in peak flow rate. This suggests the possibility that in areas with persistently high pollution levels, we may expect less evidence for day to day effect in lung function simply because children's peak flow never has chance to increase. This can be clearly seen in Figs 5.2, 5.3 and Appendix 2.4.

5.3 The Pollution Variables.

As discussed elsewhere in the report, measurements of NOx, NO, NO2, SO2, O3 and Dust (PM10) were available at all sites, CO at Swansea and during the summer at Bishopston. In addition temperature and relative humidity were also recorded. For our purposes, pollutant variables were averaged over the 24 hour period (midnight to midnight) to give a reading x_t of a predictor variable for peak flow reading y_t.

A relatively small number of missing values in the pollution data were imputed by linear interpolation. A number of gaps in pollutant measurements at the beginning and end of the study still remained, in particular, at the start of the winter period at Bishopston. This meant that for some analyses involving lagged pollutant variables, some early peak flow readings had to remain unused.

Elsewhere in the report, we have discussed the correlations between the pollutant variables. In particular we have noted the (expected) high correlation between NOx, NO and NO2. It seems sensible in view of these correlations to include only one variable from this group and NOx was chosen as being most appropriate in representing the initial harmful products from vehicles - see section 4.2.4.

5.4 Peak Flow Deviations from Each Child's Average.

Each individual child's peak flow average y_{it} (t=1,2,...,30) varies as we have noted and discussed in some detail in an earlier chapter. As an initial exploratory exercise we look at the daily deviations d_t of peak flow measurements from the child's average, $d_t = (y_{it} - \overline{y_i})$, and average these deviations to produce a measurement for a particular day, $\frac{1}{n}\left\{\sum_{i=1}^{n}\left(y_{it} - \overline{y_i}\right)\right\}$ to the extent that children were collectively below their mean or above it.

The daily average deviations were plotted as time-series together with pollutant time-series - See Appendix 2.5 for these graphs. Whilst it is difficult to relate the average deviation time-series with the pollutant time-series, it is clear that the average peak flow deviations in some cases, exhibit considerable trend:-

- Daily peak flow deviations for Swansea in the winter decrease to a low at day 17 (Tuesday of the fourth week) and then increase again.
- For Glyn Neath (2), in the winter period, a sudden dramatic drop occurred at the same time point (30th January 1995 to 3rd February 1995). The daily symptom diaries were examined for this period, revealing that over half of the children in the school reported a cough and a runny nose due, it appears, to a flu virus circulating at that time.

5.5 Fitting Crude Regression Models to the Average Peak Flow Deviations.

Exploratory regression models were used to extract information concerning the variation over time of the average peak flow deviations. Using the average peak flow deviations as the y-variable, first of all individual pollutants were used as predictor variables with varying lags. Table 5.12 shows the lags that gave the best predictions for all sites.

Table 5. 12 Best Lag times for pollutants.

Variable	**Winter**	**Summer**
Relative Humidity	3	5
Temperature	2	5
Rain Fall	2	M
Air Pressure	4	5
Nitrogen Oxides	5	5
Sulphur Dioxide	6	5
Carbon Monoxide *	4	4
Ozone	5	5
Dust Particulates	4	5

* = Carbon Monoxide was only collected at the Swansea site, and during the summer study at Bishopston.
M = Missing

It should be noted that a lack of rainfall data during the summer months, meant that it could not be included in the summer models.

With these lags, more detailed regression models were constructed involving the above variables, in some cases time variables, and for site 2 in winter, a dummy variable for the "flu epidemic" period.

With six or seven predictor variables, depending upon the site and season - see table 4.11, it is not surprising that the modelling was able to explain, in some cases, quite high percentages of the variation in the peak flow data - See table 5.13.

Table 5. 13 Results from model fitting.

WINTER SITE	Percentage of Variation accounted for	Residual Sum of Squares	Degrees of freedom
Bishopston	65.87%	35030	10
Glyn Neath (2)	89.82%	18510	19
Glyn Neath (3)	80.27%	13390	21
Swansea	77.66%	28840	17

SUMMER SITE	Percentage of Variation accounted for	Residual Sum of Squares	Degrees of freedom
Bishopston	88.76%	6704	15
Glyn Neath (2)	81.38%	36600	16
Glyn Neath (3)	59.98%	48370	17
Swansea	54.33%	28990	20

The following features emerged from the results.

Time Trends

Variables were included to measure any trend over time not captured by the pollution variables. Variables were also included to measure any day-of-the-week effect.

For the winter period, all sites other than Bishopston showed a significant drop in average deviations as winter progressed with a gradual increase after the first week of February. No such trend terms were significant for the summer study period.

Variation between days of the week was not apparent in the winter study. For the summer study however, there was such an effect, with Monday's peak flow deviations being lower on average than the subsequent days of the week. This effect was particularly pronounced for Bishopston and Glyn Neath (2) and less so for Glyn Neath (3). For Swansea there was no day-of-the-week effect whatsoever.

Atmospheric Variables.

The analysis did not lead to any helpful information regarding the effects of pollutant and atmospheric measurements. There was no consistency across sites and of the few coefficients of pollution variables that appeared to be statistically significantly different from zero and some had the wrong sign.

6. Longitudinal Models and REML

6.1 The Modelling of the Individual Peak Flow Time Series.

The previous analysis ignores the wealth of information in each child's individual peak flow record over time and a full analysis must ultimately recognise the longitudinal nature of the study.

The methodology presented by Korn and Whittemore [84] focused attention on a child's individual time series of peak flow measurements $y_{i1}, y_{i2}, ..., y_{in_i}$ which are the n_i peak flows recorded for child i at times $t_{i1}, t_{i2}, ..., t_{in_i}$. Denote a set of explanatory variables for child i by $x_{1t_{ij}}, x_{2t_{ij}}, \cdots, x_{pt_{ij}}$ where typically these might correspond to confounding variables, (e.g. temperature and relative humidity), and to potentially explanatory variables, e.g. NOx, SO2, CO, O3 and PM10 (Dust). We predict y_{ij} using the statistical linear model:

$$y_{ij} = \tilde{\beta}_0 + \beta_1 x_{1t_{ij}} + ... + \beta_p x_{pt_{ij}} + \varepsilon_{it_{ij}} \qquad i = 1,2,...,m;\ j = 1,2,...,n_i$$

This can be re-written in matrix form as $\underline{y}_i = X_i \underline{\beta} + \underline{\varepsilon}_i$. The intercept term is assumed constant for all individuals i at this stage, then β_i can be estimated using ordinary least squares techniques. However, an objection to this model is that the intercept term $\tilde{\beta}_0$ is certainly not common to all children, not even for a homogeneous group of children because, for one reason, the average peak flow depends on height which varies from child to child.

Korn and Whittemore [84] suggest the alternative two-stage random effects model:

$$\underline{y}_i = X_i \underline{\tilde{\beta}}_i + \underline{\varepsilon}_i \quad \text{where} \quad \underline{\tilde{\beta}}_i \sim MVN(\underline{\beta}, V)$$

and suggest an analysis which proceeds in two stages, by deriving $\hat{\underline{\beta}}_i$ and using these as surrogates for $\tilde{\underline{\beta}}_i$ to derive $\hat{\underline{\beta}}$ and $\hat{V}$. The $\underline{\varepsilon}_i$ process may not be "white noise", i.e. a sequence of random variables which are mutually independent and identically distributed with zero mean and variance σ^2. Because we are dealing with observations through time it is natural to assume, at the very least, that the error terms $\varepsilon_{it_{ij}}$ are auto-correlated, and we shall assume this auto-correlation is of lag at most 1, at least initially - we denote the auto-correlation parameter by ρ. An alternative is to use the ARIMA models of Box and Jenkins, the simplest of which is that the $\underline{\varepsilon}_i$ form an AR(1) process in which

$$\varepsilon_{it} = \rho\varepsilon_{i(t-1)} + a_{it}$$

where the a_{it} are now independently identically distributed $N(0,\sigma^2)$ variables. If one assumes that ρ also depends on *i*, then such a modified model can be analysed by performing separate "Cochran-Orcutt" regression models [86] on each peak flow process considered.

Once each child has a set of individual beta coefficients $\hat{\underline{\beta}}_i$, a subgroup of children (e.g. asthmatics) can be selected. Then, for that chosen subgroup, an overall coefficient can be calculated for each variable following Korn and Whittemore [84], by calculating an average coefficient which is weighted by the reciprocal of the individual's variance. Thus a set of parameters $\tilde{\underline{\beta}}_i$ can be found for each subgroup.

The methodology was programmed using the statistical package APL, details of this programming can be found in Appendix 1.1.

The different sites were explored, using one day lags for the pollutant variables. In the results which follow in table 6.1, bold figures indicate estimates which are significant at the 5% level.

Table 6. 1 Weighted average coefficients for each site using a one day lag

Bishopston	Mean (winter)	SDev (winter)	Mean (summer)	SDev (summer)
Rel. hum.	**0.1230**	0.0697	0.0390	0.0450
Temp.	**-0.6870**	0.2490	-0.1760	0.2150
NOx	0.0659	0.1415	**-0.0401**	0.2103
SO2	-0.4313	0.6828	**0.5238**	0.1927
CO			**-9.8452**	4.3004
Ozone	0.0608	0.0434	**0.1618**	0.0710
Dust	**-0.4067**	0.1866	**-0.1926**	0.0753

Glyn (2)	Mean (winter)	SDev (winter)	Mean (summer)	SDev (summer)
Rel. hum.	-0.0430	0.0800	0.0570	0.0430
Temp.	-0.3930	0.3290	-0.0451	0.1620
NOx	-0.0908	0.0720	**1.1610**	0.1590
SO2	-0.0380	0.0556	-0.0840	0.0569
Ozone	-0.0426	0.1174	**0.3693**	0.0772
Dust	0.0106	0.0286	**-0.5690**	0.0850

Glyn (3)	Mean (winter)	SDev (winter)	Mean (summer)	SDev (summer)
Rel. hum.	**0.2670**	0.0932	**-0.2320**	0.0480
Temp.	**-0.8540**	0.3796	-0.2460	0.1800
NOx	**0.1979**	0.0846	-0.0786	0.2010
SO2	0.0892	0.0652	0.0796	0.0652
Ozone	**0.4466**	0.1350	**-0.2145**	0.0872
Dust	0.0250	0.0331	0.0033	0.0974

Swansea	Mean (winter)	SDev (winter)	Mean (summer)	SDev (summer)
Rel. hum.	-0.0655	0.0710	-0.0060	0.0490
Temp.	-0.0560	0.3740	**0.9818**	0.2340
NOx	-0.0306	0.0262	-0.0279	0.0480
SO2	-0.2559	0.1622	-0.1820	0.2396
CO	-0.5376	3.0874	2.9917	4.4630
Ozone	**-0.2563**	0.1367	-0.0550	0.0760
Dust	**0.1575**	0.0899	0.0230	0.0810

The parameter estimates varied considerably for each individual, some were highly positive and others highly negative. This could explain why the weighted averages, often formed positive coefficients, some of these being significant at the 5% level. Interpretation was thus difficult. Various lag times were fitted, but the problem remained of interpreting significant positive parameter estimates.

6.2 Modelling consecutive differences in the time series

Consider two time series with the following structure

$$y_t - y_{t-1} = \mu_y + \varepsilon_t$$
$$x_t - x_{t-1} = \mu_x + \delta_t$$

If $\mu_y > 0$ and $\mu_x > 0$ then both time series $\{y_t\}$ $\{x_t\}$ will generally increase over time, and hence a regression of y_t on x_t will yield a significant positive correlation even when the error processes $\{\varepsilon_t\}$ and $\{\delta_t\}$ are statistically independent. For this reason, the modelling of relationships between time series usually begin by taking appropriate differences to render both $\{y_t\}$ and $\{x_t\}$ stationary. In our model, consecutive differences over time are considered, i.e. looking at $y_{it_{ij}} - y_{it_{i(j-1)}}$ associated with the time difference $t_{ij} - t_{i(j-1)}$ which may not equal 1 because of breaks in the time series at weekends (and occasionally at other times).

In an attempt to visualise the behaviour of these differences, the dependent variable (peak flow) differences were plotted against each independent variable (pollutant) differences in turn. Various lags times were plotted from one to six days and the graphs compared both visually and by the correlation coefficient calculated between the two plotted variables. Tables 6.2 and 6.3 summarise, for both seasons respectively, the lag which yielded the highest correlation between the two differenced time series. Also given in the table is whether there was a positive correlation (+) or a negative correlation (-) between the two differenced time series.

Table 6. 2 Winter Study Lags

	Bishopston		**Swansea**	
	Lag	Correlation	Lag	Correlation
Rel. humidity	1	–	4	–
Temperature	4	+	2	–
NOx	4	–	4	+
SO2	2	–	1	–
CO			4	+
Ozone	2	+	4	–
Dust	4	–	4	–

	Glyn Neath (2)		**Glyn Neath (3)**		**Glyn Neath Combined**	
	Lag	Correlation	Lag	Correlation	Lag	Correlation
Rel. humidity	2	–	4	–	4	–
Temperature	4	+	2	–	2	–
NOx	4	–	2	+	2	+
SO2	2	+	2	+	2	+
Ozone	4	+	2	–	2	–
Dust	4	+	4	+	4	+

Table 6. 3 Summer Study Lags

	Bishopston		**Swansea**	
	Lag	Correlation	Lag	Correlation
Rel. humidity	3	+	2	–
Temperature	3	–	4	–
NOx	4	–	3	–
SO2	4	–	4	–
CO	1	+	4	+
Ozone	3	–	2	–
Dust	1	–	4	–

	Glyn Neath (2)		**Glyn Neath (3)**		**Glyn Neath Combined**	
	Lag	Correlation	Lag	Correlation	Lag	Correlation
Rel. humidity	2	–	5	+	5	+
Temperature	5	–	3	–	5	–
NOx	5	–	2	+	2	+
SO2	1	+	1	+	1	+
Ozone	5	–	2	–	5	–
Dust	5	–	5	–	5	–

This was a purely exploratory exercise to try and locate strong trends in the data. For most of the pollutants at the various lags, the correlations fluctuated between positive and negative and so these figures should be treated and interpreted with great caution (especially so since, when these lags were fitted using the final modelling procedure, they were certainly not the best models).

Alternatively, consider y_t to be our peak flow process which we believe to be "driven" by a pollution variable x_t.

Box and Jenkins propose an impulse-response model of the form

$$y_t = v_0 x_t + v_1 x_{t-1} + \cdots + \varepsilon_t$$

with impulse-response coefficients $v_0, v_1, v_2, \cdots$ (possibly infinitely many of them). The problem of estimating a possibly infinite number of parameters of the impulse-response function is solved by the combined use of auto-regressive and moving average models. For example, if $v_i = \beta_1 \beta_0^i$, $i = 0,1,2,\ldots$ then $y_t = \beta_1 x_t + \beta_0 y_{t-1} + \varepsilon_t$. The general transfer model is of the form

$$y_t - \delta_1 y_{t-1} - \cdots - \delta_p y_{t-p} = \omega_0 x_t - \omega_1 x_{t-1} - \cdots - \omega_q x_{t-q} + \varepsilon_t$$

where the noise term ε_t itself can have an ARIMA structure.

Unfortunately, important though such models are conceptually, they indicate the plethora of ways in which one time-series may be "driven" by another time-series. The computations involved in applying transfer function modelling in our situation are not possible because of the absence of peak flow data at weekends. Imputation of missing values could be considered, but not without having some at least approximate model to use to generate the imputations.

6.3 The Proposed Methodology

We assume initially that an individual's peak flow process is related to a set of p explanatory variables and we ignore, for the time being, the problem of determining what these variables are. We know that any intercept term in a statistical model will be special to the child, but we assume that for an appropriately homogeneous group, any coefficients of the explanatory variables are common to all children in that group. Thus we assume

$$y_{ij} = \tilde{\beta}_0 + \sum_{k=1}^{p} \beta_k x_{kt_{ij}} + \varepsilon_{it_{ij}} \tag{1}$$

where the error process $\varepsilon_{it_{ij}}$ is assumed to follow an AR(1) process given by

$$\varepsilon_{it_{ij}} = \rho \varepsilon_{it_{i(j-1)}} + a_{t_{ij}} \tag{2}$$

in which $a_{t_{ij}} \sim N(0, \sigma^2)$ independently for all i,j.

Figure 6.1 presents a number of peak flow processes. Note that they appear to be non-stationary. However taking first differences seems to be sufficient to achieve stationarity. Hence we propose to consider the differences $z_{ij} = y_{ij} - y_{i(j-1)}$ which therefore satisfy the equation:

$$z_{ij} = y_{ij} - y_{i(j-1)} = \sum_{k=1}^{p} \beta_k \left(x_{kt_{ij}} - x_{kt_{i(j-1)}} \right) + \varepsilon_{it_{ij}} - \varepsilon_{it_{i(j-1)}} \tag{3}$$

Note that the random effects intercept term then disappears. Re-writing equation (3) in matrix terms gives

$$\underline{z}_i = \left(\underline{X}_{ij} - \underline{X}_{i(j-1)} \right) \underline{\beta} + \underline{\delta}_i$$
$$\underline{z}_i = D_i \underline{\beta} + \underline{\delta}_i$$

where $\underline{\beta}^T$ is now $\left(\beta_1, \beta_2, \cdots, \beta_p \right)^T$, and

$$D_i = \begin{pmatrix} x_{1t_2} - x_{1t_1} & \cdots & x_{pt_2} - x_{pt_1} \\ \vdots & \ddots & \vdots \\ x_{1t_{m_1}} - x_{1t_{m_1-1}} & \cdots & x_{pt_{m_1}} - x_{pt_{m_1-1}} \end{pmatrix}$$

$$\text{and } \underline{\delta}_i = \begin{pmatrix} -1 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & -1 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{it_1} \\ \varepsilon_{it_2} \\ \vdots \\ \varepsilon_{it_{n_i}} \end{pmatrix} = A_i \underline{\varepsilon}_i$$

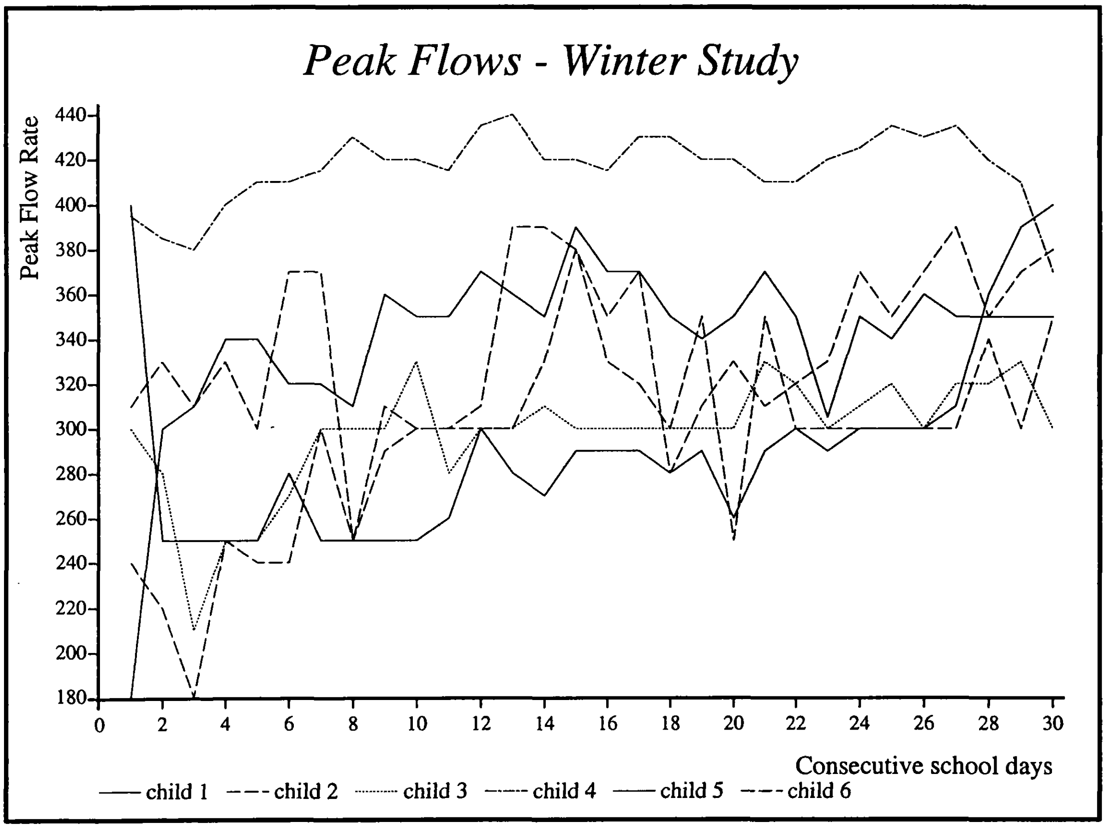
The assumption that the $\varepsilon_{it_{ij}}$ form an AR(1) process implies that if R_i is the correlation matrix of $\underline{\varepsilon}_i$, then

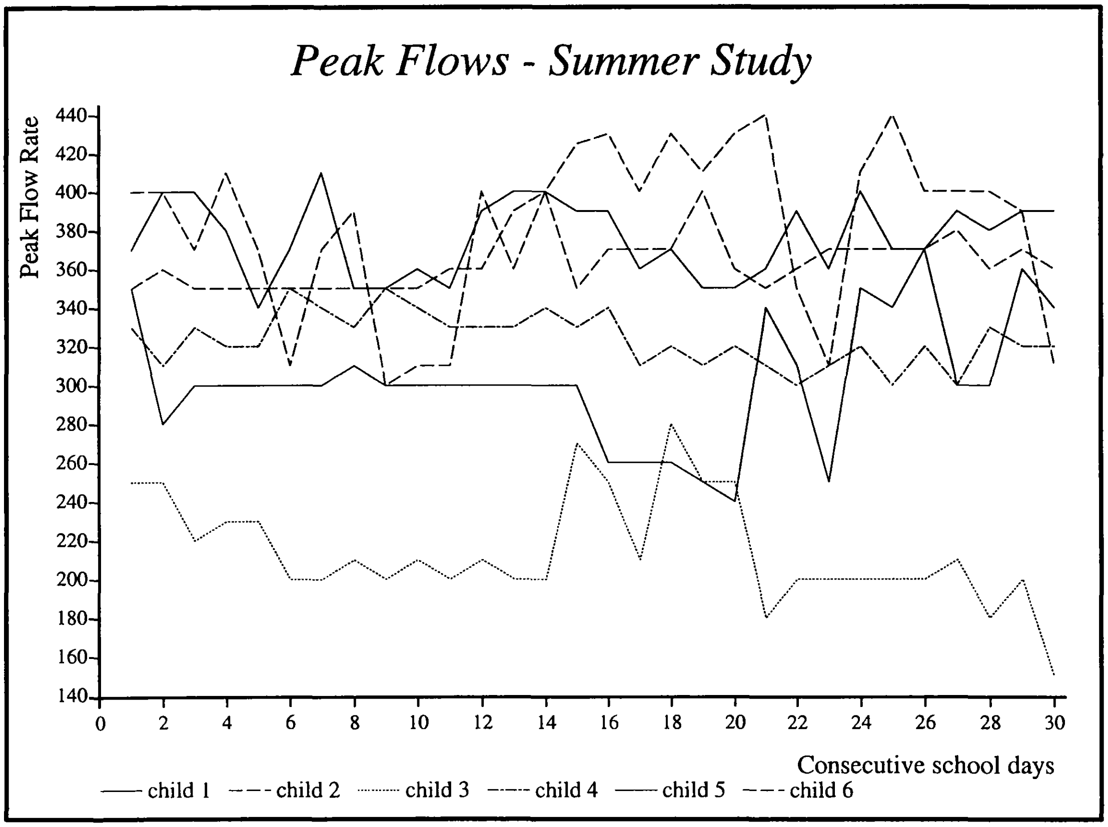
$$\left(R_i \right)_{jk} = corr\left(\varepsilon_{it_{ij}}, \varepsilon_{it_{ik}} \right) = \rho^{\left| t_{ij} - t_{ik} \right|}$$

Hence it follows that $\underline{z}_i \sim MVN\left(D_i \underline{\beta}, \sigma^2 A_i R_i A_i^T \right)$.

For a group of m children, the vector $\underline{z}^T = \left(\underline{z}_1, \underline{z}_2, \cdots, \underline{z}_m \right)^T$ consists of all $\underline{z}$ vectors stacked on top of each other. Similarly, the overall design matrix D is an $\left(n_1 + n_2 + \cdots + n_m \right) \times p$ matrix whose first n_1 rows equal D_1, the next n_2 rows are D_2 and so on.

Figure 6.1 Some Peak Flow Processes





Since observations on different children are statistically independent, define a diagonally partitioned correlation matrix V by

$$V = \begin{pmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & V_m \end{pmatrix} \qquad \text{where } V_i = A_i R_i {A_i}^T$$

Hence it follows that

$$\underline{z} \sim MVN\left(D\underline{\beta}, \sigma^2 V\right) \tag{4}$$

The theory of generalised linear models then provides an estimate $\underline{\hat{\beta}}$ of $\underline{\beta}$ where

$$\begin{aligned} \underline{\hat{\beta}}(\rho) &= \left(D^T V^{-1} D\right)^{-1} D^T V^{-1} \underline{z} \\ &= \left(\sum_{i=1}^{m} {D_i}^T V_i^{-1} D_i\right)^{-1} \sum_{i=1}^{m} {D_i}^T V_i^{-1} \underline{z} \end{aligned}$$

with associated residual sum of squares

$$RSS(\rho) = \sum_{i=1}^{m} \left(\underline{z}_i - D_i \underline{\hat{\beta}}\right)^T V_i^{-1} \left(\underline{z}_i - D_i \underline{\hat{\beta}}\right)$$

Note however that $\underline{\hat{\beta}}$ depends on the unknown AR(1) parameter ρ of the error term $\underline{\varepsilon}_{ij}$.

6.4 REML: Restricted Maximum Likelihood Estimation

Equation (4) implies that the log-likelihood function, maximised over $\underline{\beta}$ is given by

$$L(\rho, \sigma) = -\frac{1}{2} \log\left|\sigma^2 V\right| - \frac{1}{2}\left(\underline{z} - D\hat{\beta}\right)^T \left(\sigma^2 V\right)^{-1} \left(\underline{z} - D\hat{\beta}\right)$$

In Diggle [86, Chapters 4 and 5], the technique of Restricted Maximum Likelihood Estimation is discussed, and in general, endorsed for application in longitudinal studies, particularly from the point of view of reducing the bias in estimating the parameters in the covariance matrix (here, in particular, ρ).

The REML methodology proceeds by replacing $L(\rho,\sigma)$ by

$$L^*(\rho,\sigma)=-\frac{1}{2}\log\left|\sigma^2 V\right|-\frac{1}{2}\log\left|D^T\left(\sigma^2 V\right)^{-1}D\right|-\frac{1}{2}\left(\underline{z}-D\underline{\hat{\beta}}\right)^T\left(\sigma^2 V\right)^{-1}\left(\underline{z}-D\underline{\hat{\beta}}\right)$$

$$=-\frac{1}{2}\log\sigma^{2N}-\frac{1}{2}\log|V|+p\log\sigma-\frac{1}{2}\log\left|D^T V^{-1}D\right|-\frac{1}{2\sigma^2}\left(\underline{z}-D\underline{\hat{\beta}}\right)^T V^{-1}\left(\underline{z}-D\underline{\hat{\beta}}\right)$$

$$=-(N-p)\log\sigma-\frac{1}{2}\log|V|-\frac{1}{2}\log\left|D^T V^{-1}D\right|-\frac{1}{2\sigma^2}RSS(\rho)$$

Note that the matrix $D^T\left(\sigma^2 V\right)^{-1}D$ is of size $p\times p$, and this is why the coefficient of $\log\ \sigma$ is $(N-p)$ and not N as in the reference - private communication with Professor Diggle confirms this.

Differentiating $L^*(\rho,\sigma)$ with respect to σ and equating to zero gives

$$\hat{\sigma}^2(\rho)=\frac{RSS(\rho)}{N-p}$$

and hence $L^*(\rho,\hat{\sigma}(\rho))$ can be maximised with respect to ρ with direct search methods.

SUMMARY:

With this approach, model fitting proceeds by the following stages;

- Choose an assumed "homogeneous" group of m children and compute peak flow differences $\underline{z}_i$ and associated design matrices D_i
- For a given ρ, calculate $\underline{\hat{\beta}}(\rho)$ from (2) and hence evaluate $L^*(\rho,\hat{\sigma}(\rho))$
- Choose ρ to maximise L^* by direct search or other methods and hence calculate $\underline{\hat{\beta}}$ together with $RSS(\hat{\rho})$ and the estimated covariance matrix of $\underline{\hat{\beta}}$

Consequently, different competing models may then be compared by reference to their associated residual sums of squares and the significance of parameter estimates evaluated by reference to their standard errors.

6.5 Choice of Prediction Variables

The choice of the form of linear predictor remains a problem because of (a) the number of pollutant variables, (b) possible lags of these variables, and (c) the possibility of interaction terms between the pollutant variables.

For any given set of variables, the likelihood technique recommended (ordinary or REML) yields standard errors for their coefficients which play a crucial role in evaluating which variables contribute significantly to the model. The significance of sets of variables can also be investigated by asymptotic likelihood methods.

To determine appropriate lags for each variable, assuming only one lag is relevant, residual sum of squares or restricted maximised log-likelihood values may be computed for different lags and the lag giving the smallest residual sum of squares (largest log-likelihood) may be discovered.

7. Final Model

7.1 Investigating the Different Lag Effects of the Pollutants

The final model involved using the methodology described in Chapter 6, where the dependent and independent variables are differenced and the restricted maximum likelihood methodology employed. The peak flow differences for one child constitute a general linear model involving only the parameters $\beta_1, \beta_2, \cdots, \beta_p$ and a covariance matrix that depends on the unknown autocorrelation ρ. All such individual models may be stacked together to form one large model involving the unknown parameters $\beta_1, \beta_2, \cdots, \beta_p, \rho$ and a variance σ^2. The statistical analysis of such a model involved the use of programs written in the statistical programming language APL. These programs are shown in Appendix 1.2.

The methodology involved selecting a group of children of a particular sex from a particular site and then estimating beta coefficients for a given ρ value. Various values of ρ are fitted until a value of ρ is found which maximises the restricted maximum likelihood. Once this ρ value has been estimated, the beta coefficients and their associated standard errors can be found, and residuals can be estimated to validate the model.

As previously mentioned, (Chapter 6), there is a basic problem in that there are many potential models to consider, as the pollution variables selected may enter into the model more than once at different lags. For this longitudinal modelling, the predictor variables used were relative humidity, temperature, NOx, SO2, CO, Ozone and Dust all lagged simultaneously by 1 to 5 days. The restricted maximum likelihood values could then be used to assess the effect of changing the lag.

Initially, modelling proceeded by first of all fitting just relative humidity and temperature and finding the best lag times which minimised the residual sum of squares. Each pollutant was added one at a time, trying every possible lag until the lowest residual sum of squares was found. It should be noted however, that the residual sum of squares were lower still when the same lag time was used for all the pollutants, possibly suggesting some interaction occurring between the predictor variables. For this reason the following analyses use the same lag times for all the predictor variables and differences between the lag times could be compared. Note that for the winter study the residual sum of squares were always minimised for lags 1 to 3 days, but for the summer study the best lags were often 4 or 5 days. Table (7.1) summarises the results of fitting the various models. Lags of 4 and 5 days are only presented for the summer study.

Table 7. 1 Restricted Maximum Likelihood Values for Different Lags of Variables

ONE DAY LAGS

WINTER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-25194.989	0.45	937.232	3383
Glyn (2)	-19293.685	0.575	1108.227	2632
Glyn (3)	-10704.491	0.45	705.965	1547
Glyn Neath	-31428.812	0.55	971.087	4185
Swansea	-20910.627	0.525	1021.514	2840

SUMMER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-19120.999	0.375	565.714	2685
Glyn (2)	-16557.162	0.35	689.535	2315
Glyn (3)	-8693.629	0.45	587.068	1288
Glyn Neath	-26474.156	0.4	665.582	3609
Swansea	-15612.415	0.475	804.352	2184

TWO DAY LAGS

WINTER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-25194.131	0.45	936.776	3383
Glyn (2)	-19290.723	0.575	1106.146	2632
Glyn (3)	-10702.510	0.45	704.613	1547
Glyn Neath	-31427.233	0.55	970.583	4185
Swansea	-20910.686	0.525	1021.775	2840

SUMMER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-19121.432	0.375	565.918	2685
Glyn (2)	-16564.013	0.375	707.010	2315
Glyn (3)	-8693.359	0.45	587.074	1288
Glyn Neath	-26475.810	0.4	666.326	3609
Swansea	-15609.413	0.475	802.312	2184

THREE DAY LAGS

WINTER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-25195.540	0.45	937.953	3383
Glyn (2)	-19292.377	0.575	1107.600	2632
Glyn (3)	-10703.863	0.45	705.918	1547
Glyn Neath	-31431.050	0.55	972.389	4185
Swansea	-20908.746	0.525	1020.247	2840

SUMMER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-19121.503	0.375	565.657	2685
Glyn (2)	-16565.532	0.375	707.688	2315
Glyn (3)	-8690.346	0.475	598.750	1288
Glyn Neath	-26474.173	0.4	665.570	3609
Swansea	-15610.715	0.475	802.790	2184

FOUR DAY LAGS

SUMMER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-19119.774	0.375	565.033	2685
Glyn (2)	-16561.659	0.375	705.298	2315
Glyn (3)	-8698.433	0.45	591.280	1288
Glyn Neath	-26477.027	0.4	666.605	3609
Swansea	-15610.583	0.475	802.873	2184

FIVE DAY LAGS

SUMMER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-19120.630	0.375	565.137	2685
Glyn (2)	-16533.349	0.375	700.533	2315
Glyn (3)	-8689.622	0.475	598.405	1288
Glyn Neath	-26458.858	0.4	660.093	3609
Swansea	-15609.868	0.475	802.503	2184

Inspection of the table shows that the estimated value of ρ varies from 0.45 to 0.575 in the winter and 0.375 to 0.475 in the summer. Since the standard deviation of these coefficients is approximately equal to the reciprocal of the square root of the degrees of freedom, all such estimates are highly significantly different from zero, which gives partial justification of the assumption of the AR(1) model of the error process.

7.2 Modelling the different lag times for the pollutants

Tables 7.3 to 7.6 present the restricted maximum likelihood (REML) estimates of the beta coefficients and their standard errors. Each table uses all variables at a specific lag. Recall, for the winter studies, the restricted maximum likelihood values are maximised for a lag of 2 or 3, whilst for the summer studies, lags 4 or 5 were required.

The beta coefficients in bold indicate that they are statistically significantly different from zero at the 5% level (two-sided test), or approximately so (any coefficient whose absolute value is greater than 1.96 standard deviations).

7.3 Appraising the Results of the Longitudinal Analysis

By comparing the restricted maximum likelihood values for different lags, the lag which produces the largest restricted maximum likelihood value can be found.

Table 7.2 presents the results. For the winter studies, a lag of 2 was indicated for all the sites except Swansea. Note that the variation in these restricted maximum likelihood values, as the lag varied, was not great - at most 2 units. For the summer studies however, lags of 4 or 5 were required, and in one case (Glyn Neath (2)) the variation was quite substantial - about 34 units. (To understand the significance of these comments, note that for a coefficient of a new variable to be statistically significant at the 5% level, we would anticipate an increase in log restricted maximum likelihood of approximately 2).

Table 7. 2 Best lag times for the pollutant variables

WINTER	Swansea	Lag of **3** days
	All other sites	Lag of **2** days
SUMMER	Bishopston	Lag of **4** days
	Swansea and other Glyn Neath sites	Lag of **5** days

SITE - ONE DAY LAGS

Table 7. 3 Beta Coefficient Estimates for the Longitudinal Model

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.00431	0.0985	-0.0106	0.0671
Temp	-0.47040	0.3725	-0.3823	0.3368
NOx	0.01363	0.2140	0.0995	0.2968
SO2	-1.53223	1.0910	0.2561	0.3035
CO			**-19.2450**	7.7580
Ozone	-0.04984	0.0779	0.1116	0.1091
Dust	-0.21432	0.2999	**-0.2079**	0.1124

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.08909	0.1166	-0.0081	0.0681
Temp	-0.99653	0.5285	**-0.6474**	0.2612
NOx	0.00229	0.0975	**1.1378**	0.2481
SO2	-0.01538	0.0802	0.0535	0.0881
Ozone	0.22131	0.1797	**0.4283**	0.1232
Dust	0.01764	0.0370	**-0.5417**	0.1363

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.04450	0.1324	**-0.2257**	0.0796
Temp	-0.32834	0.5720	-0.2846	0.3177
NOx	**0.29888**	0.1139	-0.4199	0.3235
SO2	0.09602	0.0921	**0.3099**	0.1034
Ozone	**0.58199**	0.1955	-0.2394	0.1483
Dust	0.01997	0.0437	0.0253	0.1643

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0602	0.0885	-0.0854	0.0524
Temp	-0.7020	0.3976	**-0.5514**	0.2042
NOx	0.1180	0.0744	**0.6392**	0.1974
SO2	0.0260	0.0610	**0.1560**	0.0677
Ozone	**0.3550**	0.1350	**0.1989**	0.0959
Dust	0.0150	0.0282	**-0.3438**	0.1063

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0213	0.1014	-0.0499	0.0804
Temp	-0.1146	0.5650	0.5372	0.3995
NOx	-0.0528	0.0383	0.0041	0.0776
SO2	**-0.5901**	0.2357	-0.3779	0.4454
CO	2.8114	4.3514	-3.9295	7.2622
Ozone	-0.2651	0.2101	0.0209	0.1240
Dust	0.1403	0.1362	0.0762	0.1389

SITE - TWO DAY LAGS

Table 7. 4 Beta Coefficient Estimates for the Longitudinal Model

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.1060	0.082	-0.0720	0.058
Temp	0.3824	0.364	0.4830	0.375
NOx	0.0753	0.189	-0.0709	0.307
SO2	**-2.5778**	1.206	-0.0558	0.291
CO			-5.1584	7.122
Ozone	-0.1010	0.085	-0.0879	0.096
Dust	0.3265	0.288	-0.0578	0.116

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.1933**	0.101	**-0.1181**	0.067
Temp	-0.1444	0.557	**-0.8169**	0.297
NOx	0.0871	0.152	**0.6452**	0.259
SO2	0.0785	0.088	0.0570	0.121
Ozone	0.1664	0.189	**0.2340**	0.122
Dust	-0.0402	0.039	-0.1887	0.146

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1352	0.119	0.0795	0.081
Temp	**-1.4631**	0.605	-0.2276	0.366
NOx	0.0861	0.175	**0.7654**	0.312
SO2	-0.0212	0.100	-0.1557	0.142
Ozone	0.1952	0.206	0.0519	0.145
Dust	0.0615	0.046	-0.2658	0.175

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0701	0.077	-0.0520	0.052
Temp	-0.6980	0.419	**-0.6216**	0.232
NOx	0.0615	0.116	**0.6879**	0.201
SO2	0.0344	0.067	-0.0144	0.093
Ozone	0.1565	0.142	0.1634	0.094
Dust	-0.0030	0.030	**-0.2109**	0.113

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0296	0.092	-0.0985	0.069
Temp	-0.0162	0.538	**0.8937**	0.431
NOx	0.0227	0.043	-0.0321	0.083
SO2	0.3437	0.248	-0.2367	0.499
CO	7.0467	4.418	3.5576	6.791
Ozone	**0.5224**	0.224	-0.1479	0.110
Dust	-0.1664	0.133	0.0552	0.135

SITE - THREE DAY LAGS

Table 7. 5 Beta Coefficient Estimates for the Longitudinal Model

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0434	0.094	0.0634	0.056
Temp	0.0835	0.373	-0.1317	0.364
NOx	0.1303	0.164	0.0226	0.271
SO2	-1.7199	1.200	-0.1220	0.279
CO			-5.9189	6.442
Ozone	-0.1349	0.106	-0.0263	0.081
Dust	0.1085	0.513	0.1700	0.097

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0093	0.109	0.0287	0.067
Temp	-0.3796	0.542	-0.3943	0.281
NOx	-0.1033	0.106	-0.0098	0.197
SO2	**-0.2892**	0.126	0.0491	0.110
Ozone	-0.1971	0.180	0.0932	0.092
Dust	-0.0113	0.044	-0.1999	0.121

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1186	0.128	-0.0096	0.079
Temp	-0.4179	0.590	**-1.3300**	0.346
NOx	0.1347	0.125	-0.3169	0.233
SO2	**0.2731**	0.147	**0.3951**	0.130
Ozone	**0.5488**	0.205	0.0083	0.104
Dust	0.0343	0.049	0.0997	0.142

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0367	0.083	0.0155	0.052
Temp	-0.4351	0.408	**-0.7223**	0.218
NOx	-0.0148	0.081	-0.1183	0.152
SO2	-0.0956	0.096	**0.1687**	0.085
Ozone	0.0806	0.137	0.0602	0.070
Dust	-0.0022	0.033	-0.0865	0.093

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1059	0.107	0.0275	0.067
Temp	0.1639	0.544	0.1088	0.372
NOx	-0.0766	0.044	-0.0042	0.077
SO2	0.0903	0.241	0.1789	0.377
CO	**7.3750**	3.979	-9.1950	7.065
Ozone	0.1277	0.205	0.1049	0.114
Dust	0.0718	0.122	0.0427	0.110

SITE - FOUR AND FIVE DAY LAGS: SUMMER ONLY

FOUR DAY LAGS FIVE DAY LAGS

Table 7. 6 Beta Coefficient Estimates for the Longitudinal Model

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0310	0.055	-0.0319	0.061
Temp	-0.6077	0.438	0.4793	0.343
NOx	0.4440	0.317	-0.3948	0.517
SO2	-0.2833	0.261	0.3695	0.288
CO	**-17.6871**	7.882	3.9046	2.638
Ozone	0.1602	0.090	-0.0807	0.082
Dust	0.0068	0.092	0.0095	0.086

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0191	0.073	0.1112	0.084
Temp	**-0.5415**	0.276	**-0.7468**	0.281
NOx	**0.4412**	0.211	0.2153	0.217
SO2	-0.1195	0.090	**0.2408**	0.111
Ozone	0.0929	0.093	0.0007	0.094
Dust	-0.2207	0.121	-0.1874	0.131

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.1790**	0.091	**0.2036**	0.102
Temp	**-0.9349**	0.328	-0.5587	0.351
NOx	-0.1374	0.241	0.1220	0.254
SO2	0.0161	0.108	0.0523	0.131
Ozone	0.1036	0.109	**0.2107**	0.109
Dust	0.0427	0.140	**-0.4709**	0.149

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0651	0.057	**0.1333**	0.065
Temp	**-0.6847**	0.212	**-0.7021**	0.219
NOx	0.2338	0.161	0.1721	0.167
SO2	-0.0661	0.070	**0.1741**	0.086
Ozone	0.1034	0.072	0.0709	0.072
Dust	-0.1327	0.092	**-0.2784**	0.099

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0688	0.072	0.0269	0.078
Temp	0.3707	0.419	0.4099	0.422
NOx	-0.0660	0.082	-0.0061	0.080
SO2	-0.2577	0.348	0.2040	0.352
CO	**15.0970**	7.027	-8.3217	6.458
Ozone	-0.0708	0.120	-0.0086	0.125
Dust	0.1117	0.113	0.0266	0.127

Table 7.7 attempts to extract the essential features of Tables 7.3 to 7.6. From the definition of the model, a negative coefficient of a pollution variable indicates a detrimental effect on peak flow.

Table 7. 7 Essential features of Tables 7.3 to 7.6

BISHOPSTON:

WINTER				
Lag	REML	Source	Sign	p-value
1	-25195.0	SO2	neg	0.1600
2	-25194.1	SO2	neg	0.0330
3	-25195.5	SO2	neg	0.1500
SUMMER				
1	-19121.0	CO	neg	0.0130
		Dust	neg	0.0640
2	-19121.4			
3	-19121.5			
4	-19119.8	CO	neg	0.0250
5	-19120.6			

GLYN NEATH (2):

WINTER				
Lag	REML	Source	Sign	p-value
1	-19293.7			
2	-19290.7			
3	-19292.4	SO2	neg	0.0220
SUMMER				
1	-16557.2	NOx	pos	0.00001
		Dust	neg	0.0007
2	-16564.0	NOx	pos	0.0127
		Ozone	pos	0.0550
		Dust	neg	0.1900
3	-16565.5	Dust	neg	0.0980
4	-16561.6	NOx	pos	0.0360
		Dust	neg	0.0680
5	-16533.3	SO2	pos	0.0300

GLYN NEATH (3):

WINTER				
Lag	REML	Source	Sign	p-value
1	-10704.5	NOx	pos	0.0090
		Ozone	pos	0.0030
2	-10702.5			
3	-10703.9	SO2	pos	0.0600
		Ozone	pos	0.0070

SUMMER				
Lag	REML	Source	Sign	p-value
1	-8693.6	SO2	pos	0.0030
		Ozone	neg	0.1070
2	-8693.3	NOx	pos	0.0140
3	-8690.3	SO2	pos	0.0024
4	-8698.4			
5	-8689.6	Ozone	pos	0.0530
		Dust	neg	0.0020

GLYN NEATH COMBINED:

WINTER				
Lag	REML	Source	Sign	p-value
1	-31428.8	Ozone	pos	0.0085
2	-31427.2			
3	-31431.1			
SUMMER				
1	-26474.2	NOx	pos	0.0012
		SO2	pos	0.0212
		Ozone	pos	0.0381
		Dust	neg	0.0012
2	-26475.8	NOx	pos	0.0006
		Dust	neg	0.0620
3	-26474.2	SO2	pos	0.0472
4	-26477.0			
5	-26458.9	SO2	pos	0.0430
		Dust	neg	0.0050

SWANSEA:

WINTER				
Lag	REML	Source	Sign	p-value
1	-20910.6	SO2	neg	0.0123
2	-20910.7	O3	pos	0.0197
3	-20908.7	CO	pos	0.0638
SUMMER				
1	-15612.4			
2	-15609.4			
3	-15610.7			
4	-15610.6	CO	pos	0.0317
5	-15609.9			

In common with other investigations of a similar nature we see there are some positive and significant coefficients which are difficult to explain or comprehend. Results tend to vary across the sites and season, in that different variables appear to play significant roles. However, across lags within sites, although levels of significance vary, similar variables seem to be actively contributing to the model.

The least polluted site Bishopston gives the clearest message, (SO2 in the winter and CO and Dust in the summer). Whereas, the most polluted site, Swansea, has no significant variable coefficients in the summer and in the winter, the coefficients of SO2 and CO are statistically significant but of opposite sign.

Given that CO readings were unavailable during the winter periods at sites other than Swansea, (and were also not available in Glyn Neath for the summer), we investigated the most successful models at Swansea (winter period) and Swansea and Bishopston (summer period), with the CO variable removed to see what the effect on the other coefficients were. Table 7.8 gives the results, which indicate little change in the coefficients, except that the coefficients of SO2 tended to become larger (negatively) and hence more significant.

Table 7. 8 Results of Longitudinal Models without the CO variable

Bishopston	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0228	0.055
Temp	0.1032	0.303
NOx	0.2464	0.305
SO2	-0.3944	0.257
Ozone	0.0396	0.071
Dust	0.0736	0.087

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1491	0.105	0.0258	0.078
Temp	0.2812	0.540	0.3358	0.418
NOx	-0.0217	0.033	-0.0627	0.067
SO2	0.1764	0.236	0.1208	0.346
Ozone	0.0446	0.200	-0.0390	0.122
Dust	0.1503	0.114	0.0614	0.124

8. Application and Analysis of Final Model

8.1 Investigating Subgroups of Children.

Using the suggested lags (see Table 7.2), the methodology of the longitudinal analysis described in Chapter 7, illustrates how a group of children can be selected and the log-likelihood maximised using REML to obtain estimates of the coefficients and their standard errors. Initially children at each site can be compared and then subgroups of children within each site. This enabled contrasts to be made between:

Asthmatics v Non-Asthmatics,
Wheezers v Non-Wheezers,
Passive-Smokers v Non-Passive-Smokers,
Males v Females.

Table 8.1 reports the estimated ρ values for the subgroups. In general, these estimates are consistent with the overall estimate of ρ for each site or season.

Table 8. 1 Values of rho which maximise the REML

Site	Winter	Summer
Bishopston	0.45	0.375
Glyn Neath (2)	0.575	0.375
Glyn Neath (3)	0.45	0.45
Glyn Neath Combined	0.55	0.4
Swansea	0.525	0.475

	Asthmatics		Non-asthmatics	
Site	Winter	Summer	Winter	Summer
Bishopston	0.525	0.35	0.45	0.375
Glyn Neath (2)	0.45	0.475	0.625	0.3
Glyn Neath (3)	0.5	0.475	0.425	0.45
Glyn combined	0.475	0.475	0.575	0.375
Swansea	0.475	0.725	0.55	0.375

	Wheezers		Non-wheezers	
Site	**Winter**	**Summer**	**Winter**	**Summer**
Bishopston	0.475	0.35	0.45	0.375
Glyn Neath (2)	0.55	0.425	0.6	0.3
Glyn Neath (3)	0.475	0.45	0.425	0.475
Glyn combined	0.525	0.425	0.55	0.375
Swansea	0.55	0.625	0.5	0.325

	Passive Smokers		Non Passive Smokers	
Site	**Winter**	**Summer**	**Winter**	**Summer**
Bishopston	0.475	0.4	0.45	0.35
Glyn Neath (2)	0.5	0.4	0.675	0.35
Glyn Neath (3)	0.4	0.5	0.5	0.35
Glyn combined	0.475	0.425	0.625	0.35
Swansea	0.5	0.55	0.55	0.325

	Males		Females	
Site	**Winter**	**Summer**	**Winter**	**Summer**
Bishopston	0.45	0.425	0.475	0.3
Glyn Neath (2)	0.625	0.375	0.525	0.35
Glyn Neath (3)	0.45	0.45	0.45	0.475
Glyn combined	0.575	0.4	0.5	0.4
Swansea	0.65	0.425	0.45	0.475

8.2 The Results of the Longitudinal Analysis

Tables 8.2 to 8.10 present the coefficients and their standard errors, first for an overall site, and then for subgroups in that site. Some interesting comparisons can be made, the most important observations are summarised below.

Asthmatics.

For Bishopston the coefficients of SO2 and CO for the winter study more than double for asthmatics, indicating that the effect of pollution on asthmatics is greater than for non-asthmatics.

Wheezers.

A similar but less pronounced affect occurs for the wheezers.

Passive-Smokers.

Children who are passive-smokers in both sites in Glyn Neath had a statistically significantly negative coefficient of dust (p=0.044 for Glyn Neath (2) and p=0.012 for Glyn Neath (3)).

Males/Females.

Perhaps the most surprising result of the subgroup analysis is the appearance of more statistically significant coefficients for the subgroups of female children in contrast to the overall results or those for the male children.

The results for the Glyn Neath Sites (2 and 3) showed that the girls had a negative dust coefficient (p=0.098 at Glyn Neath (2), p=0.002 at Glyn Neath (3)) during the winter study.

Overall, the subgroup analysis showed a consistency of results that was encouraging.

Table 8. 2 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.106	0.082	-0.031	0.055
Temp	0.382	0.364	-0.608	0.438
NOx	0.075	0.189	0.444	0.317
SO2	**-2.578**	1.206	-0.283	0.261
CO			**-17.687**	7.882
Ozone	-0.101	0.085	0.160	0.090
Dust	0.326	0.288	0.007	0.092

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.193**	0.101	0.111	0.084
Temp	-0.144	0.557	**-0.747**	0.281
NOx	0.087	0.152	0.215	0.217
SO2	0.079	0.088	**0.241**	0.111
Ozone	0.166	0.189	0.001	0.094
Dust	-0.040	0.039	-0.187	0.131

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.135	0.119	**0.204**	0.102
Temp	**-1.463**	0.605	-0.559	0.351
NOx	0.086	0.175	0.122	0.254
SO2	-0.021	0.100	0.052	0.131
Ozone	0.195	0.206	**0.211**	0.109
Dust	0.061	0.046	**-0.471**	0.149

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.070	0.077	**0.133**	0.065
Temp	-0.698	0.419	**-0.702**	0.219
NOx	0.062	0.116	0.172	0.167
SO2	0.034	0.067	**0.174**	0.086
Ozone	0.157	0.142	0.071	0.072
Dust	-0.003	0.030	**-0.278**	0.099

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1059	0.107	0.0269	0.078
Temp	0.1639	0.544	0.4099	0.422
NOx	-0.0766	0.044	-0.0061	0.080
SO2	0.0903	0.241	0.2040	0.352
CO	**7.3748**	3.979	-8.3217	6.458
Ozone	0.1277	0.205	-0.0086	0.125
Dust	0.0718	0.122	0.0266	0.127

ASTHMATICS

Table 8. 3 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.493**	0.192	0.007	0.119
Temp	0.103	0.869	**-1.948**	0.942
NOx	0.816	0.460	0.711	0.685
SO2	**-6.724**	2.992	-0.745	0.564
CO			**-35.517**	16.801
Ozone	-0.088	0.212	**0.392**	0.191
Dust	-0.570	0.702	0.193	0.197

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.586**	0.253	0.246	0.213
Temp	0.209	1.259	**-1.921**	0.758
NOx	0.116	0.363	-0.254	0.561
SO2	0.240	0.217	**0.776**	0.271
Ozone	0.006	0.431	0.053	0.241
Dust	-0.022	0.096	-0.187	0.334

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.358	0.247	0.247	0.226
Temp	**-3.398**	1.296	-0.107	0.774
NOx	-0.159	0.363	-0.636	0.559
SO2	0.184	0.206	0.171	0.301
Ozone	0.294	0.436	-0.143	0.241
Dust	0.175	0.094	-0.285	0.327

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.207	0.181	0.240	0.158
Temp	-1.174	0.923	**-1.234**	0.553
NOx	0.015	0.262	-0.388	0.405
SO2	0.234	0.154	**0.559**	0.204
Ozone	0.129	0.314	-0.022	0.174
Dust	0.057	0.069	-0.234	0.240

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.2864	0.304	-0.3168	0.258
Temp	0.2472	1.500	-1.2762	1.510
NOx	-0.1294	0.120	0.0684	0.255
SO2	0.2974	0.669	0.2523	1.113
CO	16.6942	10.939	-10.2079	20.205
Ozone	0.1048	0.561	0.0336	0.389
Dust	-0.1032	0.341	0.1805	0.398

NON-ASTHMATICS

Table 8. 4 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.008	0.089	-0.033	0.062
Temp	0.480	0.392	-0.545	0.494
NOx	-0.094	0.205	0.440	0.360
SO2	-1.781	1.299	-0.062	0.295
CO			**-16.758**	8.888
Ozone	-0.120	0.091	0.143	0.101
Dust	0.544	0.310	-0.041	0.104

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.138	0.109	0.098	0.088
Temp	-0.022	0.623	-0.388	0.283
NOx	0.132	0.166	0.353	0.223
SO2	0.065	0.095	0.110	0.119
Ozone	0.229	0.211	-0.025	0.097
Dust	-0.056	0.042	-0.106	0.137

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.069	0.139	0.192	0.116
Temp	-0.954	0.695	**-0.757**	0.400
NOx	0.166	0.203	0.389	0.290
SO2	-0.096	0.117	0.027	0.148
Ozone	0.156	0.237	**0.352**	0.124
Dust	0.027	0.054	**-0.568**	0.170

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.050	0.086	0.116	0.070
Temp	-0.527	0.474	**-0.540**	0.233
NOx	0.091	0.130	**0.339**	0.179
SO2	-0.005	0.074	0.076	0.093
Ozone	0.166	0.161	0.094	0.077
Dust	-0.025	0.033	**-0.252**	0.107

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0727	0.113	0.0794	0.077
Temp	0.0521	0.577	0.7699	0.414
NOx	-0.0551	0.047	-0.0272	0.080
SO2	0.0331	0.257	0.3018	0.352
CO	4.3225	4.231	-6.8079	6.490
Ozone	0.1206	0.220	-0.0680	0.126
Dust	0.1141	0.129	-0.0334	0.128

WHEEZERS

Table 8. 5 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.178	0.131	-0.009	0.082
Temp	0.067	0.580	**-1.304**	0.650
NOx	0.239	0.302	0.528	0.472
SO2	**-4.401**	1.937	-0.229	0.387
CO			**-25.298**	11.624
Ozone	-0.213	0.137	**0.278**	0.132
Dust	0.405	0.456	-0.006	0.136

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.124	0.161	0.220	0.137
Temp	-1.222	0.870	**-0.993**	0.472
NOx	-0.007	0.240	0.203	0.359
SO2	0.100	0.140	**0.573**	0.177
Ozone	0.216	0.295	0.082	0.154
Dust	-0.044	0.062	-0.143	0.214

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.130	0.181	0.179	0.160
Temp	-1.293	0.933	-0.820	0.544
NOx	-0.021	0.266	-0.515	0.397
SO2	0.140	0.152	0.087	0.208
Ozone	0.022	0.316	0.084	0.170
Dust	0.122	0.070	-0.308	0.233

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.041	0.123	**0.205**	0.106
Temp	**-1.271**	0.653	**-0.945**	0.361
NOx	-0.026	0.182	-0.063	0.272
SO2	0.113	0.105	**0.412**	0.137
Ozone	0.137	0.221	0.073	0.116
Dust	0.016	0.047	-0.197	0.161

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.2504	0.175	-0.0618	0.134
Temp	-0.6537	0.886	0.9990	0.750
NOx	**-0.1603**	0.072	0.0253	0.136
SO2	0.0972	0.392	0.0568	0.592
CO	**14.1049**	6.495	-18.8995	10.697
Ozone	0.0361	0.338	-0.0936	0.210
Dust	0.0032	0.199	-0.0487	0.213

NON-WHEEZERS

Table 8. 6 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.045	0.104	-0.030	0.074
Temp	0.619	0.459	-0.380	0.593
NOx	-0.085	0.241	0.489	0.435
SO2	-1.010	1.529	-0.214	0.356
CO			-15.573	10.679
Ozone	-0.007	0.107	0.135	0.121
Dust	0.277	0.367	-0.003	0.125

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.274**	0.129	0.023	0.103
Temp	0.762	0.722	-0.580	0.330
NOx	0.142	0.196	0.231	0.259
SO2	0.068	0.111	-0.027	0.140
Ozone	0.078	0.245	-0.052	0.113
Dust	-0.038	0.050	-0.205	0.160

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.122	0.158	0.213	0.132
Temp	**-1.575**	0.789	-0.367	0.455
NOx	0.173	0.230	0.610	0.330
SO2	-0.139	0.132	0.026	0.170
Ozone	0.287	0.268	**0.309**	0.141
Dust	0.016	0.062	**-0.597**	0.193

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.110	0.100	0.073	0.082
Temp	-0.266	0.542	**-0.527**	0.269
NOx	0.107	0.150	0.352	0.206
SO2	-0.027	0.086	-0.018	0.108
Ozone	0.116	0.184	0.068	0.089
Dust	-0.018	0.039	**-0.333**	0.124

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0135	0.135	0.0913	0.092
Temp	0.5320	0.682	0.0698	0.488
NOx	-0.0137	0.055	-0.0188	0.094
SO2	0.0293	0.303	0.4803	0.417
CO	1.9021	4.985	-0.0750	7.731
Ozone	0.1531	0.255	-0.0269	0.149
Dust	0.1224	0.153	0.0325	0.151

PASSIVE SMOKERS

Table 8. 7 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.055	0.160	-0.004	0.118
Temp	-0.360	0.709	-0.355	0.952
NOx	0.226	0.376	0.601	0.680
SO2	-1.496	2.388	-0.586	0.562
CO			-26.394	17.180
Ozone	0.108	0.168	0.196	0.192
Dust	0.409	0.575	-0.065	0.199

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.079	0.154	0.147	0.119
Temp	-0.579	0.801	-0.315	0.401
NOx	0.127	0.228	0.370	0.303
SO2	0.087	0.133	0.233	0.155
Ozone	0.341	0.274	0.080	0.132
Dust	-0.029	0.059	**-0.368**	0.183

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.089	0.175	0.203	0.162
Temp	-0.892	0.861	0.038	0.563
NOx	-0.009	0.253	-0.196	0.402
SO2	0.126	0.145	0.026	0.208
Ozone	-0.031	0.292	0.037	0.173
Dust	0.087	0.068	**-0.594**	0.236

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.018	0.116	0.156	0.096
Temp	-0.790	0.596	-0.240	0.325
NOx	0.046	0.171	0.152	0.243
SO2	0.095	0.099	0.162	0.125
Ozone	0.172	0.203	0.057	0.105
Dust	0.016	0.045	**-0.436**	0.145

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.2557	0.160	0.0563	0.118
Temp	-0.8912	0.799	0.6907	0.648
NOx	-0.0599	0.064	-0.1374	0.121
SO2	0.1940	0.350	-0.2939	0.522
CO	6.8352	5.841	0.8437	9.646
Ozone	0.3413	0.298	-0.2632	0.185
Dust	0.0903	0.179	0.1212	0.188

NON-PASSIVE SMOKERS

Table 8. 8 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.164	0.095	-0.022	0.061
Temp	0.643	0.422	**-0.965**	0.483
NOx	0.009	0.219	0.468	0.356
SO2	**-2.965**	1.399	-0.081	0.290
CO			**-17.405**	8.654
Ozone	**-0.181**	0.098	**0.199**	0.099
Dust	0.311	0.332	0.023	0.102

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.327**	0.129	0.076	0.120
Temp	0.185	0.763	**-1.212**	0.394
NOx	0.040	0.197	0.006	0.309
SO2	0.060	0.112	0.243	0.160
Ozone	-0.027	0.257	-0.102	0.133
Dust	-0.052	0.050	0.019	0.188

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.189	0.158	0.166	0.122
Temp	**-2.173**	0.832	**-1.106**	0.403
NOx	0.224	0.235	0.443	0.306
SO2	-0.222	0.135	0.054	0.162
Ozone	0.446	0.282	**0.363**	0.129
Dust	0.020	0.062	**-0.338**	0.179

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.139	0.101	0.105	0.088
Temp	-0.674	0.577	**-1.171**	0.289
NOx	0.080	0.153	0.182	0.225
SO2	-0.047	0.087	0.178	0.118
Ozone	0.117	0.195	0.078	0.096
Dust	-0.026	0.039	-0.108	0.135

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0451	0.142	-0.0056	0.099
Temp	1.1183	0.729	0.3981	0.524
NOx	-0.0812	0.060	0.0644	0.100
SO2	-0.0637	0.328	0.6474	0.449
CO	6.5630	5.340	-13.1015	8.228
Ozone	-0.1127	0.279	0.1360	0.161
Dust	0.0663	0.163	-0.0903	0.163

MALES

Table 8. 9 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.098	0.118	0.002	0.080
Temp	0.263	0.523	-0.517	0.644
NOx	0.208	0.274	0.403	0.456
SO2	**-3.996**	1.740	-0.240	0.381
CO			-16.927	11.728
Ozone	-0.042	0.122	0.102	0.130
Dust	0.240	0.417	0.066	0.134

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.150	0.137	0.198	0.124
Temp	-0.332	0.781	-0.495	0.411
NOx	0.126	0.208	-0.157	0.312
SO2	0.034	0.119	0.270	0.165
Ozone	0.346	0.264	-0.077	0.137
Dust	0.008	0.053	-0.106	0.189

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.031	0.192	**0.455**	0.161
Temp	**-1.764**	0.970	0.035	0.549
NOx	-0.232	0.281	-0.616	0.396
SO2	0.139	0.161	0.007	0.211
Ozone	0.098	0.329	0.051	0.170
Dust	0.136	0.075	-0.340	0.234

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.093	0.112	**0.268**	0.099
Temp	-0.889	0.616	-0.345	0.330
NOx	-0.018	0.168	-0.313	0.248
SO2	0.058	0.097	0.184	0.131
Ozone	0.239	0.208	-0.045	0.108
Dust	0.051	0.043	-0.173	0.149

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0679	0.143	0.0451	0.108
Temp	0.8758	0.749	0.9884	0.574
NOx	-0.0270	0.062	0.0671	0.110
SO2	0.2755	0.337	0.3438	0.489
CO	-0.8713	5.493	-12.8153	8.803
Ozone	-0.0975	0.291	-0.1312	0.173
Dust	-0.0236	0.165	-0.0186	0.176

FEMALES

Table 8. 10 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.115	0.112	-0.039	0.075
Temp	0.506	0.501	**-1.036**	0.590
NOx	-0.076	0.261	0.559	0.445
SO2	-1.020	1.676	-0.224	0.357
CO			**-21.213**	10.452
Ozone	-0.168	0.119	**0.280**	0.121
Dust	0.437	0.396	-0.048	0.126

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.275**	0.149	0.010	0.113
Temp	0.210	0.787	**-1.015**	0.375
NOx	0.029	0.222	**0.697**	0.296
SO2	0.116	0.128	0.213	0.149
Ozone	-0.111	0.269	0.106	0.126
Dust	**-0.096**	0.058	-0.293	0.180

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.216	0.149	0.003	0.131
Temp	-1.313	0.759	**-1.043**	0.453
NOx	0.340	0.219	**0.730**	0.328
SO2	-0.168	0.125	0.091	0.166
Ozone	0.238	0.258	**0.345**	0.140
Dust	-0.005	0.058	**-0.578**	0.191

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.077	0.108	-0.002	0.086
Temp	-0.446	0.562	**-1.047**	0.289
NOx	0.133	0.159	**0.691**	0.221
SO2	-0.005	0.092	0.169	0.111
Ozone	0.011	0.191	**0.197**	0.094
Dust	-0.059	0.042	**-0.392**	0.132

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1380	0.154	0.0082	0.109
Temp	-0.6792	0.762	0.0410	0.595
NOx	-0.0782	0.061	-0.0566	0.112
SO2	-0.1243	0.337	0.2040	0.488
CO	10.5087	5.554	-5.1161	9.128
Ozone	0.3097	0.283	0.0436	0.174
Dust	0.1746	0.172	0.0247	0.176

8.3 The Analysis of the Trends in the Individual Peak Flow Rates

The technique of longitudinal data analysis should be applied to a homogeneous group of individuals. We have attempted to achieve this in the last Chapter by looking at children with similar personal or environmental backgrounds. Recall site of residence, asthmatic tendency, wheezing tendency, whether the children were passive smokers, males and females were grouped together and previously considered.

Alternatively we can use a plot of an individual's Peak Expiratory Flow Rate over time to assess whether any trend is present and use these trends to identify homogeneous groups of children. Fig. 6.1 on page 81 shows an example of the different Peak Expiratory Flow Rate trends exhibited by the children in this study.

Each individual's Peak Expiratory Flow Rate was plotted over time, visually inspected and grouped into one of four new categories exhibiting the following apparent trends:

1 = No trend apparent (a relatively flat line)
2 = A distinct downward trend over time
3 = A distinct upward trend over time
4 = A variable trend over time

The same models were fitted as used in the subgroup analysis above (section 8.2) using the best lag times for the pollutant variables as defined in Table 7.2. Note that the Glyn Neath sites were considered together so as to avoid small numbers in some groups. The values of rho which maximised the models are shown in Table 8.11.

Table 8.11 Values of rho which maximise the models

	Trend 1 (No trend)		**Trend 2 (Downward)**	
Site	**Winter**	**Summer**	**Winter**	**Summer**
Bishopston	0.25	0.275	0.25	0.25
Glyn combined	0.275	0.25	0.675	0.4
Swansea	0.275	0.225	0.625	0.5

	Trend 3 (Upward)		Trend 4 (Variable)	
Site	**Winter**	**Summer**	**Winter**	**Summer**
Bishopston	0.325	0.475	0.45	0.325
Glyn combined	0.725	0.325	0.55	0.35
Swansea	0.575	0.5	0.5	0.425

The values of rho were generally higher in the winter models than the summer models. The results are shown in tables 8.12 - 8.15. The residuals were plotted and appear to be normally distributed with no trend. Recall that a value printed in bold font in the following tables indicates a statistically significant result at the 5% level.

TREND 1 - No Trend in the Individual Peak Flow Rates

Table 8. 12 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0792	0.1068	-0.0427	0.0576
Temp	0.6024	0.4495	0.0932	0.4517
NOx	-0.0244	0.2068	-0.2200	0.3398
SO2	-0.6677	1.3016	-0.1209	0.2729
CO			2.0183	8.0490
Ozone	0.1454	0.0905	-0.1052	0.0935
Dust	-0.0799	0.3183	**0.2255**	0.0968

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0121	0.0935	**0.1916**	0.0647
Temp	-0.1464	0.4118	-0.2333	0.2044
NOx	-0.0310	0.1280	0.2108	0.1644
SO2	0.0470	0.0726	0.1132	0.0896
Ozone	-0.0442	0.1477	0.0511	0.0697
Dust	0.0042	0.0357	0.0444	0.1002

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.2736	0.1482	0.1196	0.0685
Temp	0.7830	0.6734	0.2738	0.3614
NOx	0.0276	0.0585	-0.0027	0.0694
SO2	0.3069	0.3240	-0.0267	0.3163
CO	-1.0063	4.8010	3.0814	5.7566
Ozone	0.2512	0.2464	-0.0400	0.1128
Dust	-0.0054	0.1576	0.0788	0.1145

The coefficient for Relative Humidity was nearly significant in Swansea for both seasons and significant in Glyn Neath during the summer study.

TREND 2 - Downward Trend in the Individual Peak Flow Rates

Table 8. 13 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1542	0.2572	**-0.2491**	0.1291
Temp	-1.7802	1.0735	-1.5786	0.9524
NOx	**-1.6803**	0.4870	-0.6322	0.7487
SO2	-0.1008	3.0855	1.0452	0.5908
CO			28.3196	17.7626
Ozone	**-1.0715**	0.2108	0.0296	0.2013
Dust	0.4291	0.7318	0.0265	0.2093

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.2621	0.2300	-0.1969	0.1101
Temp	**-2.5942**	1.3084	**-2.9836**	0.3731
NOx	**-0.8515**	0.3428	**0.6812**	0.2828
SO2	0.0577	0.1937	0.0566	0.1437
Ozone	-0.6176	0.4571	**0.4237**	0.1215
Dust	0.1135	0.0878	**-0.8311**	0.1684

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.9726**	0.2924	-0.5059	0.3159
Temp	**-4.0079**	1.4166	**-3.3516**	1.6693
NOx	**-0.4468**	0.1284	0.0085	0.3250
SO2	**-1.6777**	0.7481	2.4573	1.4228
CO	**44.9922**	10.2694	9.4891	25.1943
Ozone	-0.3719	0.5826	0.3766	0.5070
Dust	**-0.6550**	0.3244	**-1.1192**	0.5039

For this group of children, who had experienced a downward trend in their individual peak flow rates over time, many more coefficients were statistically significant.

There were significant negative effects of NOx during the winter study at all sites. The Dust coefficient was significant and negative at Glyn Neath during the summer study and also at Swansea for both periods. Glyn Neath had more significant effects during the summer study. Swansea, however, found nearly all of the coefficients significant during the winter study.

TREND 3 - Upward Trend in the Individual Peak Flow Rates

Table 8. 14 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0793	0.2318	0.3001	0.1809
Temp	1.2174	0.9915	1.1113	1.5458
NOx	**1.2566**	0.4698	1.7618	1.0415
SO2	0.3742	2.9988	**-1.9612**	0.8707
CO			**-113.2958**	26.9948
Ozone	**0.7266**	0.2071	0.3070	0.3018
Dust	-0.0309	0.7135	0.0390	0.2988

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1318	0.2007	0.2598	0.2011
Temp	-1.1508	1.1846	**3.9357**	0.6493
NOx	0.0989	0.3019	0.3575	0.5014
SO2	-0.0292	0.1701	0.0162	0.2802
Ozone	0.1884	0.4164	**-0.4206**	0.2147
Dust	0.0198	0.0766	0.0769	0.3036

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**1.3612**	0.3295	0.2367	0.1775
Temp	0.8263	1.6632	**3.9118**	0.9480
NOx	**0.4672**	0.1461	-0.1951	0.1810
SO2	**2.1799**	0.8553	-0.1905	0.7910
CO	**-54.8250**	11.8084	-19.0588	14.3728
Ozone	0.6511	0.6495	-0.3779	0.2807
Dust	**1.4764**	0.3759	0.2008	0.2896

Swansea again had many coefficients significant during the winter study, however these were small and positive with the exception of CO whose coefficient was negative and significant - as it was at Bishopston during the summer study. The temperature coefficient was positive and significant during the summer study at both Glyn Neath and Swansea.

TREND 4 - Variable Trend in the Individual Peak Flow Rates

Table 8. 15 Results of using various lag effects.

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.3393**	0.1593	-0.0144	0.1126
Temp	1.0430	0.7050	-1.1328	0.8989
NOx	0.2691	0.3475	1.1101	0.6696
SO2	-3.6196	2.2704	-0.4506	0.5417
CO			-23.2331	15.8799
Ozone	-0.1745	0.1556	**0.3883**	0.1830
Dust	0.6491	0.5217	-0.2373	0.1923

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0933	0.1518	**0.4850**	0.2534
Temp	-0.7132	0.7857	0.6575	0.8189
NOx	0.2521	0.2164	-0.7997	0.6392
SO2	0.0069	0.1272	0.5571	0.3419
Ozone	0.3517	0.2727	-0.3578	0.2755
Dust	-0.0387	0.0554	-0.2311	0.3871

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0635	0.1826	-0.0612	0.1579
Temp	0.9505	0.8634	-0.1962	0.8506
NOx	-0.1335	0.0774	0.0531	0.1617
SO2	0.1259	0.4426	0.2164	0.7105
CO	**15.0413**	6.2896	-17.3732	13.2999
Ozone	0.1415	0.3413	0.0980	0.2540
Dust	-0.0075	0.2008	0.0748	0.2577

This was the subgroup of children who had very variable trends in their individual peak flow rates over time. Surprisingly, few coefficients were significant.

8.4 Modelling the Peak Flow Trend Subgroups without CO

Since CO was not available for collection at all sites, the models for Swansea (both seasons) and Bishopston (summer only) were refitted without the CO variable for comparison. The values of rho which maximise the models, see table 8.16, are similar to those for the models with CO included, compare with table 8.11.

Table 8.16 Values of rho which maximise the models without CO

No CO	**Trend 1 (No trend)**		**Trend 2 (Downward)**	
Site	**Winter**	**Summer**	**Winter**	**Summer**
Bishopston	-----	0.275	-----	0.25
Swansea	0.275	0.225	0.7	0.5

No CO	**Trend 3 (Upward)**		**Trend 2 (Variable)**	
Site	**Winter**	**Summer**	**Winter**	**Summer**
Bishopston	-----	0.55	-----	0.325
Swansea	0.675	0.5	0.5	0.425

Results are shown in tables 8.17 - 8.20.

TREND 1 - No Trend in the Individual Peak Flow Rates

Table 8. 17 Results of using various lag effects without CO.

Bishopston	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0433	0.0576
Temp	0.0092	0.3030
NOx	-0.1955	0.3253
SO2	-0.1077	0.2677
Ozone	-0.0903	0.0723
Dust	**0.2170**	0.0907

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.2665	0.1442	0.1185	0.0685
Temp	0.7590	0.6631	0.2966	0.3587
NOx	0.0193	0.0428	0.0169	0.0589
SO2	0.2889	0.3122	-0.0053	0.3136
Ozone	0.2599	0.2427	-0.0296	0.1111
Dust	-0.0108	0.1554	0.0702	0.1133

These results are very similar to those obtained previously, see table 8.12 for a comparison.

TREND 2 - Downward Trend in the Individual Peak Flow Rates

Table 8. 18 Results of using various lag effects without CO.

Bishopston	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.2607**	0.1292
Temp	**-2.7195**	0.6301
NOx	-0.3234	0.7251
SO2	**1.2308**	0.5808
Ozone	0.2234	0.1609
Dust	-0.0755	0.1998

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.4920	0.2753	-0.5069	0.3147
Temp	**-3.1791**	1.4316	**-3.2853**	1.6538
NOx	-0.0681	0.0954	0.0740	0.2734
SO2	-0.6952	0.7346	2.5706	1.3854
Ozone	-0.4490	0.6061	0.4145	0.4950
Dust	-0.1762	0.3165	**-1.1628**	0.4886

Without CO in the model, the temperature and SO2 coefficients now become significant in Bishopston, whereas coefficients for the winter study at Swansea are no longer significant, except for temperature.

TREND 3 - Upward Trend in the Individual Peak Flow Rates

Table 8. 19 Results of using various lag effects without CO.

Bishopston	Mean (Summer)	Sdev (Summer)
Rel Hum	**0.3741**	0.1834
Temp	**5.1740**	1.1713
NOx	0.5408	0.9972
SO2	**-2.5210**	0.8786
Ozone	-0.4136	0.2538
Dust	0.4459	0.2876

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**0.8093**	0.3180	0.2365	0.1777
Temp	-0.2602	1.7045	**3.7641**	0.9426
NOx	0.0152	0.1104	**-0.3239**	0.1529
SO2	1.0053	0.8593	-0.4011	0.7759
Ozone	0.7793	0.6862	-0.4508	0.2755
Dust	**0.8964**	0.3711	0.2852	0.2828

The SO2 coefficient remains negative and significant at Bishopston and now relative humidity and temperature have significant and positive coefficients. At Swansea only Dust and relative humidity have positive and significant coefficients during the winter study. The NOx coefficient becomes significant and negative during the summer study.

TREND 4 - Variable Trend in the Individual Peak Flow Rates

Table 8. 20 Results of using various lag effects without CO.

Bishopston	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0048	0.1125
Temp	-0.1657	0.6095
NOx	0.8309	0.6421
SO2	-0.5964	0.5327
Ozone	0.2203	0.1426
Dust	-0.1431	0.1813

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1880	0.1753	-0.0608	0.1580
Temp	1.3224	0.8506	-0.3520	0.8426
NOx	-0.0060	0.0562	-0.0631	0.1351
SO2	0.4314	0.4244	0.0632	0.7011
Ozone	0.0578	0.3400	0.0399	0.2502
Dust	0.0983	0.1962	0.1395	0.2530

No coefficients are significant for those children whose individual peak flow rates showed a variable trend over time.

Each of these four subgroups of children had similar results to the models with CO included. Without CO in the models some coefficients loose their significance, in particular at Swansea during the winter study.

9. Investigating The Robustness Of The Final Model

9.1 Zero Residuals

Examination of the peak flow graphs showed that some children had fairly stable peak flow rates and so resulted in some zero differences, which seemed to be responsible for a plethora of small estimated residuals when the model was fitted. A scatterplot of these residuals over time, revealed a random scattering and a fairly normal distribution. However frequency histograms of the residuals, with an overlaid normal distribution, revealed that there were too many zero differences. Children who had ten or more zero differences were removed and the model was refitted using the best lags as found previously, no great changes occurred in the results. Appendix 2.6 shows the residuals before and after the removal of these zero residuals. Table 9.1 shows the actual number of children who were used in each modelling procedure, for both the full and reduced models respectively.

Table 9. 1 Number of children in each modelling procedure

	Full database		Reduced database	
Site	**Winter**	**Summer**	**Winter**	**Summer**
Bishopston	127	117	111	103
Glyn Neath (2)	98	93	81	78
Glyn Neath (3)	55	52	47	38
Glyn Neath Combined	153	145	128	116
Swansea	108	91	94	78

The values of rho which maximised the log-likelihoods for both databases are shown in Table 9.2 they are very similar.

Table 9. 2 New values of rho which maximise the likelihood

	Full database		Reduced database	
Site	**Winter**	**Summer**	**Winter**	**Summer**
Bishopston	0.45	0.375	0.45	0.375
Glyn Neath (2)	0.575	0.375	0.6	0.325
Glyn Neath (3)	0.45	0.475	0.4	0.475
Glyn Neath Combined	0.55	0.4	0.55	0.375
Swansea	0.525	0.475	0.5	0.45

The results of the longitudinal modelling follow, very little has changed in the results which is encouraging.

SITE

Table 9. 3 Results of using various lags on the reduced database

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.1356	0.0925	0.0019	0.0590
Temp	0.4157	0.4090	-0.2074	0.4701
NOx	0.0587	0.2126	0.4618	0.3401
SO2	**-2.6573**	1.3536	-0.4199	0.2799
CO			-13.4535	8.4123
Ozone	-0.1499	0.0953	0.1186	0.0956
Dust	0.3937	0.3222	0.0147	0.0985

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.2693**	0.1156	0.0700	0.0844
Temp	-0.1956	0.6470	**-0.7655**	0.2759
NOx	0.0801	0.1743	0.2804	0.2164
SO2	0.0501	0.1006	**0.2506**	0.1135
Ozone	0.1893	0.2194	-0.0286	0.0936
Dust	-0.0221	0.0445	-0.1309	0.1321

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0588	0.1343	**0.2447**	0.1260
Temp	-1.0845	0.6601	-0.6523	0.4307
NOx	0.1538	0.1945	0.1562	0.3128
SO2	0.0650	0.1113	0.0772	0.1625
Ozone	0.2071	0.2244	**0.3023**	0.1343
Dust	0.0831	0.0525	**-0.5286**	0.1836

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.1437	0.0883	0.1049	0.0704
Temp	-0.6857	0.4777	**-0.7696**	0.2334
NOx	0.0548	0.1319	0.2327	0.1792
SO2	0.0436	0.0759	**0.1918**	0.0933
Ozone	0.1578	0.1619	0.0736	0.0771
Dust	0.0183	0.0340	**-0.2503**	0.1077

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0679	0.1200	0.0377	0.0874
Temp	0.4022	0.6036	0.2872	0.4675
NOx	**-0.1019**	0.0487	0.0035	0.0894
SO2	0.0233	0.2654	0.1137	0.3939
CO	**10.9514**	4.4106	-6.8618	7.2251
Ozone	0.0805	0.2255	0.0068	0.1394
Dust	-0.0700	0.1352	0.0829	0.1422

Very few changes emerged in the final models compared to those in Chapter 8. Where some coefficients have changed sign, these coefficients remain small and non significant and have not altered any of the other coefficients in the model.

1. Glyn (3) winter;
    - SO2 coefficient is now positive (but not significant)
    - Temperature is no longer significant
2. Glyn (combined) winter;
    - Dust coefficient is now positive (but not significant)
    - Rhum coefficient has halved
3. Swansea winter;
    - Dust coefficient is now negative (but not significant)
    - Temp coefficient has doubled
    - Nox is now negative and significant
4. Bishopston summer;
    - Rhum coefficient is now positive (but not significant)
    - CO is not significant
5. Glyn (2) summer;
    - Ozone coefficient is now negative (but not significant)
6. Swansea summer;
    - Ozone and NOx coefficients are now positive (but not significant)

Table 9. 4 Restricted Maximum Likelihood Values for the Reduced Database

WINTER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-21695.5972	0.45	1022.64109	2924
Glyn (2)	-15578.7979	0.6	1247.39846	2143
Glyn (3)	-8926.5286	0.4	693.30579	1303
Glyn Neath	-25713.4353	0.55	1038.07523	3452
Swansea	-17859.8230	0.5	1046.27267	2441

SUMMER	Max. REML	$\hat{\rho}$	$\hat{\sigma}^2$	DF
Bishopston	-16586.1274	0.375	567.70449	2350
Glyn (2)	-13556.3677	0.325	571.69529	1942
Glyn (3)	-6148.9065	0.475	653.09360	927
Glyn Neath	-20641.3248	0.375	597.75072	2875
Swansea	-13283.5986	0.45	833.51857	1870

9.2 Comments About These Reduced Models

For all the models defined, the R^2 values in the context of the longitudinal model can be approximately calculated as $\left[1-\left(\frac{RSS}{TCSS}\right)\right]\%$ where RSS = Residual Sum of Squares, TCSS = Total Corrected Sum of Squares for the model. In this case the R^2 values are small and not very helpful in aiding a comparison of any of these models presented to the full models. However in the reduced models the results remain relatively unchanged, with similar coefficients for the variables. This shows a reassuring robustness of the models used in the final analysis.

9.3 Models Without Ozone

Ozone was found to be highly correlated with all other pollutants at each site. To check for any possible confounding and interaction of the pollutants, which may be obscuring some of the results, Ozone was removed from the models. The differences Z_{ij} were related to relative humidity, temperature, NOx, SO2, CO and Dust using the same lags as those given in Table 7.2, so that a direct comparison could be made with the existing results. Table 9.5 presents the details for each model fitted such as the value of rho and the maximum likelihood estimate and Table 9.6 presents the results.

Table 9. 5 Details of final models with Ozone removed

WINTER	Details of model				
Site	REML:	RHO	RSS	DF	VAR
Bishopston	-25198.84	0.45	3170443.35	3384	936.892
Glyn Neath (2)	-19295.66	0.575	2912234.81	2633	1106.052
Glyn Neath (3)	-10707.05	0.45	1090670.71	1548	704.568
Glyn Neath	-31432.56	0.55	4063064.53	4186	970.632
Swansea	-20918.19	0.525	2907391.74	2841	1023.369

SUMMER	Details of model				
Site	REML:	RHO	RSS	DF	VAR
Bishopston	-19125.24	0.375	1518930.71	2686	565.499
Glyn Neath (2)	-16557.09	0.375	1621734.56	2316	700.231
Glyn Neath (3)	-8694.81	0.475	772985.06	1289	599.678
Glyn Neath	-26463.26	0.4	2382922.59	3610	660.089
Swansea	-15613.83	0.475	1752670.27	2185	802.137

Table 9. 6 Results of modelling without Ozone

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0944	0.0816	-0.0513	0.0539
Temp	0.3180	0.3597	-0.0872	0.3288
NOx	0.0468	0.1879	0.1882	0.2837
SO2	**-2.3682**	1.1931	-0.3619	0.2578
CO			-9.1769	6.2970
Dust	0.1868	0.2627	0.0785	0.0831

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.2103**	0.0994	0.1110	0.0815
Temp	-0.0005	0.5326	**-0.7461**	0.2627
NOx	-0.0101	0.1048	0.2149	0.2068
SO2	0.0690	0.0870	**0.2407**	0.1088
Dust	-0.0329	0.0383	-0.1869	0.1158

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1184	0.1181	0.1558	0.0991
Temp	**-1.3524**	0.5934	-0.2919	0.3233
NOx	-0.0358	0.1185	-0.0112	0.2448
SO2	-0.0371	0.0991	-0.0016	0.1286
Dust	0.0714	0.0452	**-0.3246**	0.1282

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0855	0.0762	**0.1165**	0.0632
Temp	-0.5709	0.4032	**-0.6223**	0.2041
NOx	-0.0311	0.0796	0.1245	0.1595
SO2	0.0246	0.0661	**0.1564**	0.0840
Dust	0.0040	0.0292	**-0.2307**	0.0872

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0694	0.0901	0.0274	0.0781
Temp	-0.0421	0.5382	0.3976	0.3823
NOx	-0.0389	0.0337	-0.0031	0.0669
SO2	-0.0340	0.1875	0.2155	0.3098
CO	2.7079	4.0093	-8.4059	6.3399
Dust	-0.0179	0.1170	0.0223	0.1104

The coefficients are similar to before, with very few changes. At Bishopston the CO coefficient is still negative but no longer significant. At Glyn Neath (3) relative humidity is nearly significant, NOx and SO2 are now negative, but not significant. The Dust coefficient is still negative and significant at the combined Glyn Neath site.

9.4 Models Without SO2, CO and Ozone

Ozone and SO2 were significantly correlated to other variables at each site, and CO data was only available at Swansea and during the summer study in Bishopston. So, the models were refitted relating the differences Z_{ij} to relative humidity, temperature, NOx and Dust to see the effects on the remaining variables. Table 9.7 shows the details of each model and Table 9.8 shows the results.

Table 9. 7 Details of models without SO2, CO and Ozone

WINTER	Details of model				
Site	REML:	RHO	RSS	DF	VAR
Bishopston	-25207.46	0.45	3174134.54	3385	937.706
Glyn Neath (2)	-19299.74	0.575	2912930.64	2634	1105.896
Glyn Neath (3)	-10710.48	0.45	1090769.45	1549	704.177
Glyn Neath	-31436.58	0.55	4063199.41	4187	970.432
Swansea	-20927.72	0.525	2901711.56	2843	1020.651

SUMMER	Details of model				
Site	REML:	RHO	RSS	DF	VAR
Bishopston	-19140.60	0.375	1521539.49	2688	566.049
Glyn Neath (2)	-16563.44	0.375	1625161.31	2317	701.408
Glyn Neath (3)	-8698.28	0.475	772985.15	1290	599.213
Glyn Neath	-26469.08	0.4	2385207.42	3611	660.539
Swansea	-15627.66	0.475	1754327.61	2187	802.162

Table 9. 8 Results of modelling without SO2, CO and Ozone

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.1015	0.0816	-0.0171	0.0514
Temp	0.4440	0.3542	0.0387	0.2633
NOx	-0.1681	0.1536	-0.1245	0.2056
Dust	0.2893	0.2577	0.1013	0.0824

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.1978**	0.0981	0.0375	0.0744
Temp	-0.1850	0.4791	**-0.6186**	0.2565
NOx	-0.0128	0.1047	0.2719	0.2052
Dust	-0.0347	0.0382	**-0.2346**	0.1139

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1115	0.1166	**0.1563**	0.0894
Temp	**-1.2558**	0.5342	-0.2926	0.3178
NOx	-0.0347	0.1185	-0.0117	0.2418
Dust	0.0725	0.0451	**-0.3243**	0.1255

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0810	0.0752	0.0676	0.0575
Temp	**-0.6366**	0.3627	**-0.5432**	0.1997
NOx	-0.0320	0.0795	0.1663	0.1579
Dust	0.0033	0.0292	**-0.2615**	0.0856

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1543	0.0968	0.0210	0.0770
Temp	0.0534	0.4211	0.3169	0.3628
NOx	-0.0264	0.0194	-0.0569	0.0553
Dust	0.1845	0.1028	0.0865	0.0759

For the winter study, the NOx coefficient has again changed sign, and is now negative but still not-significant. Dust is nearly significant but positive at Swansea. For the summer study, little has changed. The temperature coefficient at Bishopston is now positive (but not significant), NOx is again negative (but not significant). At Glyn Neath (combined) relative humidity is no longer significant, however the Dust coefficient remains negative and significant.

9.5 Modelling an SO2 and Dust Interaction

In the following models the Z_{ij} were related to relative humidity, temperature, NOx, SO2, Ozone, Dust and SO2*Dust - investigating a possible SO2 and Dust interaction. This was modelled for the Glyn Neath sites only, as it was thought that this could explain the significant Dust effects which were occurring.

Table 9. 9 Details of models with an SO2 and Dust interaction at Glyn Neath

WINTER	Details of model				
Site	REML:	RHO	RSS	DF	VAR
Glyn Neath (2)	-19287.07	0.575	2904546.38	2631	1103.970
Glyn Neath (3)	-10698.68	0.45	1084653.25	1546	701.587
Glyn Neath	-31426.15	0.55	4061170.14	4184	970.643

SUMMER	Details of model				
Site	REML:	RHO	RSS	DF	VAR
Glyn Neath (2)	-16550.66	0.375	1620420.34	2314	700.268
Glyn Neath (3)	-8685.77	0.475	767765.67	1287	596.555
Glyn Neath	-26453.96	0.4	2378365.83	3608	659.192

Table 9. 10 Results of modelling an SO2 and Dust interaction at Glyn Neath

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.2409**	0.1029	0.1072	0.0845
Temp	0.1518	0.5691	**-0.9966**	0.3352
NOx	0.2138	0.1603	0.1551	0.2211
SO2	0.0757	0.0876	**0.2857**	0.1160
Ozone	0.2468	0.1914	0.0100	0.0941
Dust	**-0.0815**	0.0425	**-0.2365**	0.1360
SO2* Dust	**0.0090**	0.0036	-0.0173	0.0126

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.2073	0.1219	**0.2028**	0.1019
Temp	**-1.9756**	0.6312	**-0.9020**	0.3828
NOx	-0.1050	0.1875	0.0200	0.2578
SO2	-0.0216	0.1003	0.1239	0.1351
Ozone	0.0794	0.2095	**0.2246**	0.1089
Dust	**0.1092**	0.0494	**-0.5386**	0.1516
SO2*Dust	**-0.0096**	0.0035	**-0.0315**	0.0141

Glyn Neath Combined	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0832	0.0789	**0.1290**	0.0654
Temp	-0.6134	0.4308	**-1.0049**	0.2521
NOx	0.0960	0.1225	0.0951	0.1694
SO2	0.0339	0.0667	**0.2305**	0.0889
Ozone	0.1779	0.1443	0.0829	0.0718
Dust	-0.0137	0.0324	**-0.3378**	0.1025
SO2*Dust	0.0023	0.0026	**-0.0228**	0.0094

There is little change in the coefficients, for Dust it is negative and significant particularly during the summer, as is the coefficient for the interaction between SO2 and Dust.

9.6 Investigating the Effects of Ozone and Dust in the Models

Initial models had relative humidity and temperature, then Dust and Ozone were added both separately and then together. Table 9.11 below summarises how the models were built up for the Glyn Neath sites separately and combined.

Table 9. 11 Details of adding Ozone and Dust to the model

GLYN NEATH (2):

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Rhum, Temp	-19307.14	0.575	2914313.33	2636	1105.582
With Dust	-19303.70	0.575	2912947.12	2635	1105.483
With Ozone	-19302.62	0.575	2913499.28	2635	1105.692
With both	-19299.33	0.575	2912513.95	2634	1105.738

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Rhum, Temp	-16574.12	0.375	1628395.96	2319	702.197
With Dust	-16568.85	0.375	1626419.60	2318	701.648
With Ozone	-16569.45	0.375	1626870.70	2318	701.842
With both	-16564.81	0.375	1625904.30	2317	701.728

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.1976**	0.0853	0.0767	0.0698
Temp	-0.1749	0.2994	**-0.8392**	0.2238

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.2037**	0.0855	0.0720	0.0699
Temp	-0.1395	0.3011	**-0.6287**	0.2565
Dust	-0.0370	0.0333	-0.1770	0.1054

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.1574	0.0973	0.0451	0.0730
Temp	-0.5041	0.4867	**-0.6498**	0.2580
Ozone	0.1086	0.1266	-0.1170	0.0793

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	**-0.1729**	0.0987	0.0530	0.0733
Temp	-0.3904	0.5014	**-0.5566**	0.2699
Ozone	0.0813	0.1299	-0.0747	0.0871
Dust	-0.0322	0.0342	-0.1359	0.1158

GLYN NEATH (3):

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Rhum, Temp	-10717.91	0.45	1092814.49	1551	704.587
With Dust	-10714.06	0.45	1090829.74	1550	703.761
With Ozone	-10714.10	0.45	1092631.40	1550	704.923
With both	-10709.96	0.45	1090262.91	1549	703.850

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Rhum, Temp	-8709.60	0.45	758352.30	1292	586.960
With Dust	-8702.40	0.475	772986.54	1291	598.750
With Ozone	-8706.41	0.45	758287.64	1291	587.365
With both	-8697.49	0.475	771001.05	1290	597.675

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0854	0.1007	0.1470	0.0836
Temp	**-1.0684**	0.3410	**-0.7307**	0.2766

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.0944	0.1008	**0.1547**	0.0834
Temp	**-1.1360**	0.3432	-0.2935	0.3172
Dust	0.0660	0.0393	**-0.3265**	0.1170

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1129	0.1144	**0.1544**	0.0865
Temp	**-1.2662**	0.5166	**-0.7857**	0.3226
Ozone	0.0691	0.1356	0.0299	0.0900

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	0.1451	0.1156	**0.1984**	0.0867
Temp	**-1.5006**	0.5318	-0.4966	0.3359
Ozone	0.1246	0.1388	0.1851	0.1016
Dust	0.0739	0.0403	**-0.4400**	0.1324

GLYN NEATH (combined):

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Rhum, Temp	-31443.82	0.55	4063365.61	4189	970.009
With Dust	-31400.81	0.55	4063356.84	4188	970.238
With Ozone	-31438.99	0.55	4062380.67	4188	970.005
With both	-31435.95	0.55	4062364.07	4187	970.233

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Rhum, Temp	-26482.48	0.4	2391371.69	3613	661.880
With Dust	-26474.35	0.4	2385939.99	3612	660.559
With Ozone	-26478.19	0.4	2390656.29	3612	661.865
With both	-26470.47	0.4	2385907.07	3611	660.733

Glyn	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0955	0.0653	0.0933	0.0539
Temp	**-0.5253**	0.2266	**-0.8217**	0.1746

Glyn	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0959	0.0654	0.0888	0.0539
Temp	**-0.5230**	0.2280	**-0.5427**	0.1997
Dust	-0.0024	0.0255	**-0.2281**	0.0795

Glyn	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0596	0.0744	0.0768	0.0562
Temp	**-0.8137**	0.3650	**-0.7156**	0.2022
Ozone	0.0958	0.0950	-0.0625	0.0601

Glyn	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rel Hum	-0.0580	0.0754	0.0925	0.0565
Temp	**-0.8254**	0.3760	**-0.5575**	0.2105
Ozone	0.0986	0.0975	0.0149	0.0666
Dust	0.0034	0.0261	**-0.2366**	0.0883

Once again the coefficients remain stable and Dust remains negative and significant with or without Ozone present in the model.

9.7 Investigating the Effects of Temperature and Dust in the Model

In order to investigate the potential confounding effect of having temperature and Dust in the model, temperature was first modelled by itself and Dust by itself and then they were compared to the model with them both in together. Table 9.12 presents the details and results for each site.

Table 9. 12 Details of modelling Temperature and Dust

BISHOPSTON:

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Temp only	-25222.70	0.45	3177438.26	3388	937.851
Dust only	-25223.23	0.45	3177938.80	3388	937.998
With both	-25217.70	0.45	3177297.84	3387	938.086

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Temp only	-19153.22	0.375	1522481.29	2691	565.768
Dust only	-19153.75	0.375	1521786.61	2691	565.510
With both	-19148.85	0.375	1521783.39	2690	565.719

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	0.2249	0.2749	0.2147	0.1618

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Dust	0.0782	0.2122	0.0896	0.0518

Bishopston	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	0.2273	0.2750	0.0181	0.2398
Dust	0.0821	0.2123	0.0853	0.0768

GLYN NEATH (2):

WINTER Model	Details of model				
	REML:	RHO	RSS	DF	VAR
Temp only	-19313.58	0.575	2920241.96	2637	1107.411
Dust only	-19315.71	0.575	2920111.95	2637	1107.361
With both	-19310.30	0.575	2919220.94	2636	1107.443

SUMMER Model	Details of model				
	REML:	RHO	RSS	DF	VAR
Temp only	-16578.19	0.375	1629240.64	2320	702.259
Dust only	-16582.58	0.375	1634313.57	2320	704.446
With both	-16572.85	0.375	1627165.70	2319	701.667

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	-0.2935	0.2953	**-0.9642**	0.1926

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Dust	-0.0348	0.0331	**-0.3692**	0.0875

Glyn (2)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	-0.2661	0.2966	**-0.7410**	0.2322
Dust	-0.0319	0.0333	**-0.1812**	0.1054

GLYN NEATH (3):

WINTER Model	Details of model				
	REML:	RHO	RSS	DF	VAR
Temp only	-10721.65	0.45	1093320.43	1552	704.459
Dust only	-10727.53	0.45	1098596.51	1552	707.859
With both	-10717.88	0.45	1091446.45	1551	703.705

SUMMER Model	Details of model				
	REML:	RHO	RSS	DF	VAR
Temp only	-8714.22	0.45	760169.040	1293	587.911
Dust only	-8713.54	0.45	758370.274	1293	586.520
With both	-8707.09	0.45	755808.911	1292	584.991

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	**-1.0120**	0.3344	**-0.9955**	0.2323

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Dust	0.0504	0.0391	**-0.4575**	0.0987

Glyn (3)	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	**-1.0718**	0.3363	**-0.5800**	0.2772
Dust	0.0640	0.0392	**-0.3220**	0.1179

GLYN NEATH (Combined):

WINTER Model	Details of model				
	REML:	RHO	RSS	DF	VAR
Temp only	-31448.84	0.55	4065440.94	4190	970.272
Dust only	-31454.43	0.55	4072065.69	4190	971.853
With both	-31445.83	0.55	4065440.94	4189	970.504

SUMMER Model	Details of model				
	REML:	RHO	RSS	DF	VAR
Temp only	-26487.63	0.4	2393354.37	3614	662.245
Dust only	-26491.54	0.4	2397456.56	3614	663.380
With both	-26479.35	0.4	2387732.67	3613	660.873

Glyn	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	**-0.5853**	0.2229	**-0.9793**	0.1490

Glyn	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Dust	-0.0066	0.0253	**-0.4021**	0.0662

Glyn	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	**-0.5853**	0.2240	**-0.6877**	0.1793
Dust	-0.0001	0.0254	**-0.2320**	0.0795

SWANSEA:

WINTER Model	Details of model				
	REML:	RHO	RSS	DF	VAR
Temp only	-20940.74	0.525	2907364.66	2846	1021.562
Dust only	-20942.69	0.525	2908673.03	2846	1022.021
With both	-20936.18	0.525	2906064.90	2845	1021.464

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Temp only	-15638.91	0.475	1756171.51	2190	801.905
Dust only	-15640.62	0.475	1756566.10	2190	802.085
With both	-15635.02	0.475	1755505.27	2189	801.967

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	0.5072	0.3417	**0.5674**	0.2299

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Dust	0.0880	0.0917	**0.1263**	0.0534

Swansea	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Temp	0.5493	0.3438	0.3664	0.3186
Dust	0.1040	0.0922	0.0674	0.0740

The Dust coefficient remains negative and significant at the Glyn Neath sites, but not at Swansea or Bishopston.

9.8 Comparing the Effects of Including Meteorological Variables in the Model

In order to evaluate the overall significance of the variables relative humidity and temperature when added to the basic model consisting of only pollutants, we follow the straight forward argument applied to $\underline{\hat{\beta}} = \begin{pmatrix} \hat{\beta}_{rel.humidity} \\ \hat{\beta}_{temperature} \end{pmatrix}$ which has a multivariate normal distribution with (estimated) covariance matrix $V_{\underline{\hat{\beta}}}$. The value $\underline{\hat{\beta}}^{T} V^{-1} \underline{\hat{\beta}}$ is computed and can be directly compared with a Chi-squared distribution on 2 degrees of freedom (due to the matrix $\underline{\hat{\beta}}$ having rank of 2) i.e. $\chi^{2}{}_{2,0.05}$ (see Diggle [87] - Page 94]. Table 9.13 shows the effects of adding the meteorological variables for each of the sites. There is a significant improvement in the Glyn Neath models when these variables are included. The improvement in the model (reduced model with log likelihood L_n) due to adding p variables (full model with log likelihood L_{n+p}) can be assessed by calculating twice the difference in these log likelihoods. Let $W_k = 2(L_{n+p} - L_n)$ which has a Chi-squared distribution on p degrees of freedom.

Table 9. 13 Effects of adding meteorological variables to the basic pollutant models

	WINTER				SUMMER			
SITE	Chi-stat	DF	p-value		Chi-stat	DF	p-value	
Bishopston	1.7628	2	0.4142		2.1222	2	0.3461	
Glyn (2)	5.7839	2	0.0555	-	9.9271	2	0.0070	*
Glyn (3)	5.8773	2	0.0529	-	8.1077	2	0.0174	*
Glyn Neath	7.1030	2	0.0287	*	16.8577	2	0.0002	*
Swansea	1.5607	2	0.4582		0.9472	2	0.6228	

* indicates significance at the 5% level.

Table 9.14 below, shows the details of each model fitting procedure, together with the results for the models with just the pollutants in. Note that the full models, (pollutants and climate), are already given in Chapter 7.

Table 9. 14 Details and results of adding meteorological variables to the basic pollutant models

BISHOPSTON:

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-25204.18	0.45	3170766.18	3385	936.711
Full Model	-25194.13	0.45	3169113.21	3383	936.776

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-19129.67	0.375	1518311.53	2687	565.058
Full Model	-19119.77	0.375	1517113.61	2685	565.033

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
NOx	0.0398	0.1875	0.3319	0.3030
SO2	**-2.6833**	1.1874	-0.3840	0.2393
CO			-9.7719	5.4486
Ozone	-0.0855	0.0838	0.0887	0.0627
Dust	0.2940	0.2866	0.0092	0.0921

GLYN NEATH (2):

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-19303.00	0.575	2917773.87	2634	1107.735
Full Model	-19290.72	0.575	2911376.27	2632	1106.146

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-16566.76	0.375	1628688.75	2317	702.930
Full Model	-16533.35	0.375	1621733.90	2315	700.533

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
NOx	0.0723	0.1467	0.2769	0.2096
SO2	0.1033	0.0809	0.0686	0.0927
Ozone	0.1725	0.1805	-0.1385	0.0827
Dust	-0.0346	0.0374	**-0.2903**	0.1272

GLYN NEATH (3):

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-10714.00	0.45	1094176.98	1549	706.376
Full Model	-10702.51	0.45	1090036.31	1547	704.613

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-8701.38	0.475	775597.54	1290	601.238
Full Model	-8689.62	0.475	770745.64	1288	598.405

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
NOx	0.2215	0.1650	0.2339	0.2427
SO2	0.0746	0.0923	-0.1609	0.1077
Ozone	0.1010	0.2018	0.0583	0.0938
Dust	0.0272	0.0441	**-0.5341**	0.1470

GLYN NEATH (COMBINED):

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-31440.54	0.55	4068781.87	4187	971.765
Full Model	-31427.23	0.55	4061889.86	4185	970.583

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-26476.12	0.4	2393404.12	3611	662.809
Full Model	-26458.86	0.4	2382275.64	3609	660.093

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
NOx	0.0973	0.1112	0.2400	0.1605
SO2	0.0875	0.0615	-0.0085	0.0712
Ozone	0.1044	0.1366	-0.0726	0.0627
Dust	-0.0118	0.0286	**-0.3682**	0.0972

SWANSEA:

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-20919.18	0.525	2899094.85	2842	1020.090
Full Model	-20928.75	0.525	2897501.48	2840	1020.247

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Pollutants only	-15619.01	0.475	1753426.58	2186	802.116
Full Model	-15609.87	0.475	1752666.55	2184	802.503

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
NOx	**-0.0997**	0.0401	-0.0110	0.0767
SO2	-0.0202	0.2170	0.3135	0.3143
CO	**8.9073**	3.7809	-7.4790	6.3933
Ozone	0.0905	0.1948	0.0409	0.1094
Dust	0.0672	0.1200	0.0277	0.1265

Again there is little change in results, compared with Chapter 7. The Dust coefficient is still negative and significant in Glyn Neath hence showing that the climatic variables are not confounding the effects of Dust. With the meteorological variables removed, the dust coefficient is still not significant at either Bishopston or Swansea.

9.9 Comparing the Effects of Including Pollutant Variables

Using the method described in section 9.7 the effects of adding the pollutants to the basic model, consisting of the meteorological variables only, was also calculated as shown in Table 9.15. There is a significant improvement in the Glyn Neath models during the summer study when the pollutant variables are included.

Table 9. 15 Effects of adding pollutants to basic meteorological models

	WINTER			SUMMER			
SITE	Chi-stat	DF	p-value	Chi-stat	DF	p-value	
Bishopston	6.9471	4	0.1387	9.3819	5	0.0948	-
Glyn (2)	2.6555	4	0.6170	9.5091	4	0.0496	*
Glyn (3)	3.9437	4	0.4137	11.5358	4	0.0212	*
Glyn Neath	1.5226	4	0.8226	13.7787	4	0.0080	*
Swansea	7.7930	5	0.1680	4.1494	5	0.5281	

* indicates significance at the 5% level.

Below, Table 9.16 shows the details of each model fitting procedure, together with the results for the models with just the meteorological variables in. These can be compared to the full models given in Chapter 7.

Table 9. 16 Details and results of adding pollutants to the basic meteorological models

BISHOPSTON:

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-25217.80	0.45	3175622.69	3387	937.592
Full Model	-25194.13	0.45	3169113.21	3383	936.776

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-19149.83	0.375	1522413.49	2690	565.953
Full Model	-19119.77	0.375	1517113.61	2685	565.033

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rhum	-0.1100	0.0791	-0.0176	0.0508
Temp	0.5087	0.3422	0.1807	0.1894

GLYN NEATH (2):

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-19307.14	0.575	2914313.33	2636	1105.582
Full Model	-19290.72	0.575	2911376.27	2632	1106.146

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-16574.12	0.375	1628395.96	2319	702.197
Full Model	-16533.35	0.375	1621733.90	2315	700.533

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rhum	**-0.1976**	0.0853	0.0766	0.0698
Temp	-0.1749	0.2994	**-0.8392**	0.2238

GLYN NEATH (3):

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-10717.91	0.45	1092814.49	1551	704.587
Full Model	-10702.51	0.45	1090036.31	1547	704.613

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-8709.60	0.45	758352.30	1292	586.960
Full Model	-8689.62	0.475	770745.64	1288	598.405

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rhum	0.0854	0.1007	0.1470	0.0836
Temp	**-1.0684**	0.3410	**-0.7307**	0.2766

GLYN NEATH (COMBINED):

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-31443.82	0.55	4063365.61	4189	970.009
Full Model	-31427.23	0.55	4061889.86	4185	970.583

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-26482.48	0.4	2391371.69	3613	661.880
Full Model	-26458.86	0.4	2382275.64	3609	660.093

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rhum	-0.0955	0.0653	**0.0933**	0.0539
Temp	**-0.5253**	0.2266	**-0.8217**	0.1746

SWANSEA:

WINTER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-20935.84	0.525	2905453.30	2845	1021.249
Full Model	-20928.75	0.525	2897501.48	2840	1020.247

SUMMER	Details of model				
Model	REML:	RHO	RSS	DF	VAR
Climate only	-15635.35	0.475	1755996.36	2189	802.191
Full Model	-15609.87	0.475	1752666.55	2184	802.503

	Mean (Winter)	Sdev (Winter)	Mean (Summer)	Sdev (Summer)
Rhum	0.1309	0.0957	0.0335	0.0717
Temp	0.2901	0.3767	**0.6417**	0.2795

Very little has changed from the results presented in Chapter 7. Some coefficients have even become significant - temperature at Glyn Neath and Swansea during the summer.

10. Conclusions

The final methodology used in this research is recommended for analysing longitudinal data, such as the type presented here looking at the health effects of varying pollution levels.

Future studies being considered, should certainly seek to minimise the amount of missing data where possible. The less information available, the harder it is to find any links between the independent and dependent variables. In particular, it limits the amount of work which can be done when considering lag effects. In this study, it was fortunate that there was at least one week's overlap of pollution data gathered.

In this research the biggest problems with lack of pollution data arose where:

1. There was no CO (Carbon Monoxide) recorded at Bishopston during the winter study and none at all in Glyn Neath. In order to compare sites, consistency is essential, however the final models were also fitted without CO and little change occurred in the results. But it would have been interesting to see the effects, if any, in Glyn Neath, (recall CO was negative and significant at Bishopston during the summer and positive and significant at Swansea during the winter months).

2. There was a large chunk of pollution data missing, namely the first 16 days of the winter study in Bishopston, due to a mechanical failure.

3. The mechanical failures in the pollution monitoring equipment led to a loss of information which could potentially affect the calculated daily averages, if enough was missing. In this study this was not a problem, but it should be borne in mind for similar future research. Convention requires a substantial amount of the daily measurements to be present for a daily average to be calculated, otherwise a missing value is recorded.

With regards to the children's daily peak flow readings, the way in which suspicious readings were identified and coded was a powerful way of excluding readings which would otherwise have had a detrimental impact on the results. The irregular absences from school, did not lead to the loss of too much peak flow data. Some children did record their own peak flow rates at home with parental supervision. Not all were used in the analysis however. It could, therefore, have been plausible to also record the weekend peak flow readings.

It is debatable as to whether any peak flow measurements taken on the weekends, without supervision, should be considered as good quality data and included in the analysis. Including weekend readings could have been risky. Parental supervision could have been one possibility in an attempt to ensure that an accurate peak flow reading was recorded and at the correct time of day. Some children would quite likely cheat and make up some results, but in this study, it was generally well known who the cheaters were and their results could have been ignored.

For the longitudinal modelling used in this research, the missing data would have been useful, particularly since the peak flow differences were considered. The missing weekend data lead to differences, taken from Friday to Monday, being considered with no idea of what trend happened to the peak flow rates over the weekend period. This is a theoretical issue, which could be considered in future research. Some form of linear interpolation or Kalman Filter could be employed to fill in the odd gap of missing data.

As previously mentioned, some children had very stable peak flow rates with two or three identical consecutive readings on most weeks, which lead to zero differences. These readings could be genuine or it could be due to laziness on behalf of the child. The differences were either zero or close to zero, and were included in the longitudinal modelling. As discussed in Chapter 9, removing those children with many zero differences, even though it lead to a loss of information, it did not change the overall results.

Not differencing the data is another option, but since each child had their own random intercept, which varied greatly from child to child, it was not plausible to consider this methodology in this instance.

Using the method of restricted maximum likelihood to maximise the log likelihoods for a homogeneous group of children allowed subgroups to be investigated. Due to the large number of possible subgroups not all could be investigated here. Others which could be looked at are, for example, children:

- who have a furry pet in their home
- who have a home environment where gas is used for cooking
- who have suffered from a specific symptom in the past, e.g. a dry cough at night in the last 12 months

It would also have been useful if the children had noted in their daily symptom diaries whether asthma medication had been used. If the air pollution had caused a decreased peak flow rate in some asthmatic children, the medication would rebalance their peak flow level and hence any effects of the air pollution would go relatively undetected.

There was such a wealth of information available in this dataset that not all aspects of it could be fully analysed. Due to time constraints, the daily symptom diaries could not be fully exploited. Logistic regression of each symptom against the various pollutants, might determine whether a specific pollutant might cause coughs or sneezes amongst the children.

It may be possible to analyse this data using a multiple input - single output transfer function model. It would be interesting to compare results. However all the gaps of missing data in this dataset, make it very difficult to apply the transfer function methodology.

The final modelling strategy demonstrated in this thesis, was the Cochran-Orcutt regression modelling of each individual's peak flow differences, using restricted maximum likelihood techniques for a homogeneous group of children. I believe this to be the most powerful and flexible tool for analysing the longitudinal nature of this data.

The final models were robust, in that, it did not matter which combination of variables were modelled, the estimates of the coefficients and their associated standard errors did not change. Results were similar across the different subgroups of children within each site.

To summarise, females in Glyn Neath had significantly lower average peak flow rates and were found to be significantly affected by many of the pollutants. At all sites, where a significant effect was found, for a particular pollutant, its magnitude doubled for certain subgroups of children - namely Wheezers, Asthmatics and non-passive smokers.

A negative effect of dust was found to exist in Glyn Neath, but not at either of the other sites, even when possible confounding was considered and modelled. It is not possible to say exactly where this dust had come from. Plotting levels of dust by wind direction, (using a rose plot), could be one way of pinpointing where higher dust levels come from. However, it should be noted that due to the situation of Glyn Neath in a long valley which runs from north-east to south-west, the hills could obscure the actual wind directions, by funnelling the wind down the valley.

Wind directional analysis is not so straight forward either [91]. Calculating an average wind direction can not be done by calculating a mathematical average. This is due to the fact that two wind directions of 1 degree and 359 degrees, (both from the North), should be averaged to give a direction of 360 = 0 degrees (North). However the numerical average would be calculated incorrectly as 180 degrees which is a Southerly direction. The correct calculation involves transforming the wind direction into a function of sine and cosine, then using the method described in [91], an average can be calculated. Daily averages could obscure any short-term trends. For instance, on one particular day, the wind may blow from the East for 80% of the day, and dust levels may be quite low. If, for the other 20% of that day, the wind blows from the North and dust levels are really high. This would lead to the incorrect assumption that high dust levels come from the East. It is important to take into account how often the wind blows from a particular direction. All these things must be borne in mind when analysing such rose diagrams. It may be better, for days with high average dust levels, to look directly at the raw data and attempt to find the parts of the day with the high dust levels and see where the wind direction was from at those particular times.

There are three possible sources for the dust in Glyn Neath:

i. the traffic passing through and along the A465 dual carriageway
ii. the open-cast coal mines
iii. Llandarcy BP oil refinery which burns coal

However the lack of a dust effect in Swansea, were the traffic volume is greater, suggests that the dust in Glyn Neath comes from another source. An electron microscope failed to pinpoint the source of the dust exactly, but was able to tell that the dust samples gathered contained dust from both traffic and coal burning. It has been suggested that there may be a different combination of pollutants in the air in Glyn Neath, possibly from industry. This mixture may well cause the dust particles to become more acidic, hence the reason why it is having a more harmful effect on lung function. What is important however is that this dust is having a detrimental effect on the health of the children living in the Glyn Neath area, and any reduction of dust levels would be beneficial.

APPENDIX 1

COMPUTER PROGRAMS

Appendix 1.1 Programs for Korn and Whittemore Methodology

All these programs were used for the winter study, slight amendments were made when modelling the summer study.

PREPAREDATA:

```
Y_NY_T←PREPAREDATA S;site;Y;M;C;DIM;NOMISS;NY;DAY;DOFW;ADJX;YADJ
⍝ prepares y data for Cochran Orcutt Longitudinal Analysis
⍝S is the site to be used
⍝ this version assumes to find the peakflow days in vector pfdays
⍝ THIS VERSION PRODUCES A LONG Y VECTOR, A VECTOR GIVING ELEMENTS FOR
EACH INDIVIDUAL Y VECTOR
⍝ AND THE TIME VECTOR TO GO WITH IT
 site←PEAK[;2]
 Y←PEAK[;24+⍳30]
 Y←(site∊S)⌿Y ⍝ GET SITE DATA
 Y←(15<+/Y≠¯1)⌿Y   ⍝ DROP OUT THOSE CASES WITH FEWER THAN HALF
 M←+/Y×Y≠¯1
 C←+/Y≠¯1
 M←(⍴Y)⍴(¯1  ⍴Y)/M←M÷C
 Y←Y-M×Y≠¯1   ⍝ TAKE OFF PERSONAL MEANS
 DIM←⍴Y
 NOMISS←Y≠¯1
 NY←+/NOMISS
 Y←,Y
 NOMISS←,NOMISS
 Y←NOMISS/Y
 DAY←,DIM⍴pfdays
 DAY←NOMISS/DAY
⍝ THESE NEXT LINES ARE USED IF YOU WISH TO ADJUST FOR DAYS
⍝ DOFW←,DIM⍴⍳5
⍝ DOFW←NOMISS/DOFW
⍝ ADJX←1,(DOFW∘.=2 3 4 5)
 YADJ←Y
⍝*****-ADJX+.×B←Y⌹ADJX ⍝ YADJ IS NOW ADJUSTED FOR DOW AND TEMP
⍝NY GIVES THE NO OF VALUES IN EACH Y
⍝ NOW CAN START TO LOOK AT EACH INDIVIDUAL'S RESULTS
 Y_NY_T←YADJ NY DAY
```

LONGANAL:

```
RES←site LONGANAL Y_NY_T
;Y;NY;T;C;I;POLLDATA;lut;YI;TI;XI;NOX;RHU;SO2;CO;O3;DUST;RI
⍝ produces lognitudinal analyses for each individual
⍝ using Cochrane Orcutt Regression Analysis
⍝ RES contains the regression estimates and their variances
 Y NY T←Y_NY_T
 lut←4 7⍴2 3 16 21 13 19 20 25 22 26 31 13 29 30 25 22 26 31 13 29 30
2 3 9 12 13 14 15
⍝  THIS IS A LOOK-UP TABLE  TELLING US WHERE TO FIND VARS IN
INTWINTDAY
 C←1 ⋄ RES←((⍴NY),(7+site=4),2)⍴0
lp:
 I←(+/(C-1) NY)+⍳NY[C]
 TI←(¯1+INTWINTDAY[;1]⍳T[I])⍝ finds the appropriate pollution days -
note ¯1 offset
 RHUM←INTWINTDAY[TI;lut[site;1]]
 TEMP←INTWINTDAY[TI;lut[site;2]]
 NOX←INTWINTDAY[TI;lut[site;3]]
 SO2←INTWINTDAY[TI;lut[site;4]]
 ⍎(site=4)/'CO←INTWINTDAY[TI;lut[site;5]]'
 O3←INTWINTDAY[TI;lut[site;6]]
 DUST←INTWINTDAY[TI;lut[site;7]]
 XI←1,RHUM,TEMP,NOX,[1.5]SO2
 ⍎(site=4)/'XI←XI,CO'
 XI←XI,O3,[1.5]DUST
 YI←Y[I]
⍝NOW DO THE CRUNCHY BIT
 RI←(((XI YI TI)∘LLIKECOCHORC)DERIV)EQUALS(,0)(¯0.999 0.999)
 ⍎(1<|RI)/'RI←0.999××RI'
 RES[C;;]←XI YI TI COCHORC RI
 →((⍴NY)≥C←C+1)/lp
```

LLIKECOCHORC:

```
RES←X_Y_T LLIKECOCHORC RHO;TRHO;B;TEMP;IR;R;RSS;RS;X;Y;T;C;D
⍝calculates REML log-likelihood for cochran orcutt
⍝ X=design matrix and Y=y data
⍝ T=times associated with y data
⍝ If transformation required, generateRHO from  RHOX by folloing line
⍝ RHO←(2×÷○1)×¯3○RHOX
 X Y T←X_Y_T
```

```
 D←ρRHO ⋄ TRHO←,RHO ⋄ C←1 ⋄ RES←(ρTRHO)ρ0
lp:RHO←TRHO[C]
 IR←⌹R←RHO*|T∘.-T
 B←(⌹TEMP←TEMP+.×X)+.×(TEMP←(⍉X)+.×IR)+.×Y
 RSS←RS+.×IR+.×RS←Y-X+.×B
 P←¯1 ρX
 RES[C]←-0.5×((ρY-P)×⍟RSS)+(⍟DET R)+⍟DET TEMP
 →((ρTRHO)≥C←C+1)/lp
 RES←DρRES
```

EQUALS:

```
r←(f EQUALS)a_x;eps;f;x;y;a;i;fx;nx;nfx
⍝ re-codes ILLINOIS method see VECTOR vol 12 No 2
 a x←a_x
 ⍎(2≠ρρx)/'x←1 2ρx'
 x←((ρa),2)ρx
 eps←0.001
 fx←((f x[;1])-a),[1.5](f x[;2])-a
lp:nx←x[;1]-fx[;1]×(-/x)÷-/fx
 nfx←(f nx)-a
 i←0>nfx×fx[;2]
 x←((i×x[;2])+x[;1]×~i),[1.5]nx
 fx←(fx[;2]×i)+0.5×fx[;1]×~i
 fx←fx,[1.5]nfx
 →(0=∧/(|fx[;2])<(eps⌊eps×|a)+(a=0)×eps)/lp
 r←nx
```

COCHORC:

```
RES←X_Y_T COCHORC RHO;TRHO;B;TEMP;IR;R;RSS;RS;X;Y;T;C;D
⍝calculates REML log-likelihood estimates for cochran orcutt
⍝ given a RHO value
⍝ returns estimates and their VARIANCES
⍝ X=design matrix
⍝ Y=y data
⍝ T=times associated with y data
 X Y T←X_Y_T
 IR←⌹R←RHO*|T∘.-T
 B←(TEMP←⌹TEMP+.×X)+.×(TEMP←(⍉X)+.×IR)+.×Y
 RSS←RS+.×IR+.×RS←Y-X+.×B
 RES←B,[1.5](RSS÷(ρY)-¯1 ρX)×1 1⍉TEMP
```

DERIV:

```
R←(F DERIV)X;DELTA
 DELTA←0.0001
 R←(F X+DELTA)-F X-DELTA
 R←R÷2×DELTA
```

WB:

```
R←WB DATA;ETA
⍝'ETA='
⍝⎕←ETA←DATA∘ETAEQN EQUALS(,0)(0.01 20)
⍝'MEAN AND STD DEV'
⍝DATA WBETAM_SD ETA
⍝'MEAN AND SDEV USING 1/SE SQUARED'
 R←DATA WBETAM_SD 0
```

WBETA:

```
R←B_VARB WBETA ETA;B;VARB;WT
⍝ calculates a weighted regression estimate given eta -see Korn and
Whittemore
 B←B_VARB[;1]
 VARB←B_VARB[;2]
 WT←÷VARB+ETA×ETA
 R←(WT+.×B)÷+/WT
```

WBETA_SD:

```
R←B_VARB WBETAM_SD ETA;B;VARB;WT
⍝ calculates a weighted regression estimate given eta -see Korn and
Whittemore
 B←B_VARB[;1]
 VARB←B_VARB[;2]
 WT←÷VARB+ETA×ETA
 R←(WT+.×B)÷+/WT
 R←R,(÷+/WT)*0.5
```

Appendix 1.2 Programs for Longitudinal Data Analysis

All these programs were used for the winter study, slight amendments were made when modelling the summer study.

REMLOGLIKE:

```
RESRHO←site_YNYT REMLOGLIKE
RHO;Y;NY;YNYT;T;C;P;I;site;TI_1;POLLDATA;lut;ZI;YI_1;Z;TI;DI;NOX;RHU;
SO2;CO;O3;DUST;RI;TEMP;AI;DETDVD;DVD;DVZ;IVI;LDETVI;VI
⍝ Calculates the L*(RHO) for Y_NY_T data from site and specified RHO
 site YNYT←site_YNYT
 Y NY T←YNYT
 lut←4 7⍴2 3 16 21 13 19 20 25 22 26 31 13 29 30 25 22 26 31 13 29 30
2 3 9 12 13 14 15
⍝ THIS IS A LOOK-UP TABLE  TELLING US WHERE TO FIND VARS IN
INTWINTDAY
 C←1 ⋄ DVD←(2/P←6+site=4)⍴0 ⋄ DVZ←P⍴0⍝ used for accumulating DtV-1D
and DtV-1z
 LDETVI←0 ⋄ DETDVD←0 ⋄ TIMES←⍳0 ⍝ required for REMLOGLIKE
lp:
 I←(+/(C-1) NY)+⍳NY[C]
 TI_1←¯2+INTWINTDAY[;1]⍳¯1↓T[I]⍝ finds the appropriate pollution days
- note ¯1 offset
 TI←¯2+INTWINTDAY[;1]⍳1↓T[I]
 TIMES←TIMES,TI
 RHUM←INTWINTDAY[TI;lut[site;1]]-INTWINTDAY[TI_1;lut[site;1]]
 TEMP←INTWINTDAY[TI;lut[site;2]]-INTWINTDAY[TI_1;lut[site;2]]
 NOX←INTWINTDAY[TI;lut[site;3]]-INTWINTDAY[TI_1;lut[site;3]]
 SO2←INTWINTDAY[TI;lut[site;4]]-INTWINTDAY[TI_1;lut[site;4]]
 ⍎(site=4)/'CO←INTWINTDAY[TI;lut[site;5]]-
INTWINTDAY[TI_1;lut[site;5]]'
 O3←INTWINTDAY[TI;lut[site;6]]-INTWINTDAY[TI_1;lut[site;6]]
 DUST←INTWINTDAY[TI;lut[site;7]]-INTWINTDAY[TI_1;lut[site;7]]
 DI←RHUM,TEMP,NOX,[1.5]SO2
 ⍎(site=4)/'DI←DI,CO'
 DI←DI,O3,[1.5]DUST
 ZI←(1↓Y[I])-¯1↓Y[I]
 RI←RHO*|T[I]∘.-T[I]
 AI←(¯1 0+2/NY[C])⍴¯1 1,(¯1+NY[C])⍴0
 IVI←⌹VI←AI+.×RI+.×⍉AI
 LDETVI←LDETVI+⍟DET VI
```

```
⍝Now start to assemble all the different terms for the REM log-
likelihood
 DVD←DVD+TEMP←(⍉DI)+.×IVI+.×DI
 DETDVD←DETDVD+DET TEMP
 DVZ←DVZ+(⍉DI)+.×IVI+.×ZI ⍝ note ordering here is so as to save
transposing unnecessarily
 →((⍴NY)≥C←C+1)/lp
 BETA←(COV←⌹DVD)+.×DVZ ⍝ at last we have the beta estimates
 RSS←0 ⋄ C←1 ⋄ TCSS←0 ⋄ RESID←⍳0
lp2:⍝ need to re-create DI and ZI unfortunately (could save large
array if there is space)
 I←(+/(C-1) NY)+⍳NY[C]
 TI_1←¯2+INTWINTDAY[;1]⍳¯1↓T[I]⍝ finds the appropriate pollution days
- note ¯1 offset
 TI←¯2+INTWINTDAY[;1]⍳1↓T[I]
 RHUM←INTWINTDAY[TI;lut[site;1]]-INTWINTDAY[TI_1;lut[site;1]]
 TEMP←INTWINTDAY[TI;lut[site;2]]-INTWINTDAY[TI_1;lut[site;2]]
 NOX←INTWINTDAY[TI;lut[site;3]]-INTWINTDAY[TI_1;lut[site;3]]
 SO2←INTWINTDAY[TI;lut[site;4]]-INTWINTDAY[TI_1;lut[site;4]]
 ⍎(site=4)/'CO←INTWINTDAY[TI;lut[site;5]]-
INTWINTDAY[TI_1;lut[site;5]]'
 O3←INTWINTDAY[TI;lut[site;6]]-INTWINTDAY[TI_1;lut[site;6]]
 DUST←INTWINTDAY[TI;lut[site;7]]-INTWINTDAY[TI_1;lut[site;7]]
 DI←RHUM,TEMP,NOX,[1.5]SO2
 ⍎(site=4)/'DI←DI,CO'
 DI←DI,O3,[1.5]DUST
 ZI←(1↓Y[I])-¯1↓Y[I]
 RI←RHO*|T[I]∘.-T[I]
 AI←(¯1 0+2/NY[C])⍴¯1 1,(¯1+NY[C])⍴0
 IVI←⌹VI←AI+.×RI+.×⍉AI
 RSS←RSS+(⍉RESD)+.×IVI+.×RESD←ZI-DI+.×BETA
 RESID←RESID,RESD
 TCSS←TCSS++/(ZI-(+/ZI)÷⍴ZI)*2
 →((⍴NY)≥C←C+1)/lp2
 RESRHO←(DF←(+/NY-1)-P)×⍟RSS
 RESRHO←RESRHO+LDETVI
 RESRHO←-0.5×RESRHO+⍟DETDVD
 VAR←RSS÷DF
 COV←COV×VAR
```

PREPARE:

```
Y_NY_T←PREPARE S;site;Y;M;C;DIM;NOMISS;NY;DAY;DOFW;ADJX;YADJ
⍝ prepares y data for Cochran Orcutt Longitudinal Analysis
⍝S is the site to be used
⍝ this version assumes to find the peakflow days in vector pfdays
⍝ THIS VERSION PRODUCES A LONG Y VECTOR, A VECTOR GIVING ELEMENTS FOR
EACH INDIVIDUAL Y VECTOR
⍝ AND THE TIME VECTOR TO GO WITH IT
⍝ *****Need to add any further selection of cases e.g.
asthmatic/nonasthmatic****
 site←PEAK[;2]
 Y←PEAK[;24+⍳30]
 Y←(site∊S)⌿Y ⍝ GET SITE DATA
 Y←(15<+/Y≠¯1)⌿Y    ⍝ DROP OUT THOSE CASES WITH FEWER THAN HALF
 DIM←⍴Y
 NOMISS←Y≠¯1
 NY←+/NOMISS
 Y←,Y
 NOMISS←,NOMISS
 Y←NOMISS/Y
 DAY←,DIM⍴pfdays
 DAY←NOMISS/DAY
 Y_NY_T←Y NY DAY
```

DET:

```
D←DET M;C;Z
 C←1
lp:M←C ΔQ M ⋄ →((1 ⍴M)≥C←C+1)/lp
 D←|×/1 1⍉M
```

ΔQ:

```
V←K ΔQ D;S;A;T;B;U
 S←(+/T×T←(K-1)↓D[;K])*0.5
 B←÷S×S+|T←D[K;K]
 U←((1 ⍴D),1)⍴U←((⍳1 ⍴D)≥K)×D[;K]
 U[K;1]←((×T)+T=0)×(S+|T)
 V←D-B×U+.×(⍉U)+.×D
```

APPENDIX 2

MAPS AND GRAPHS

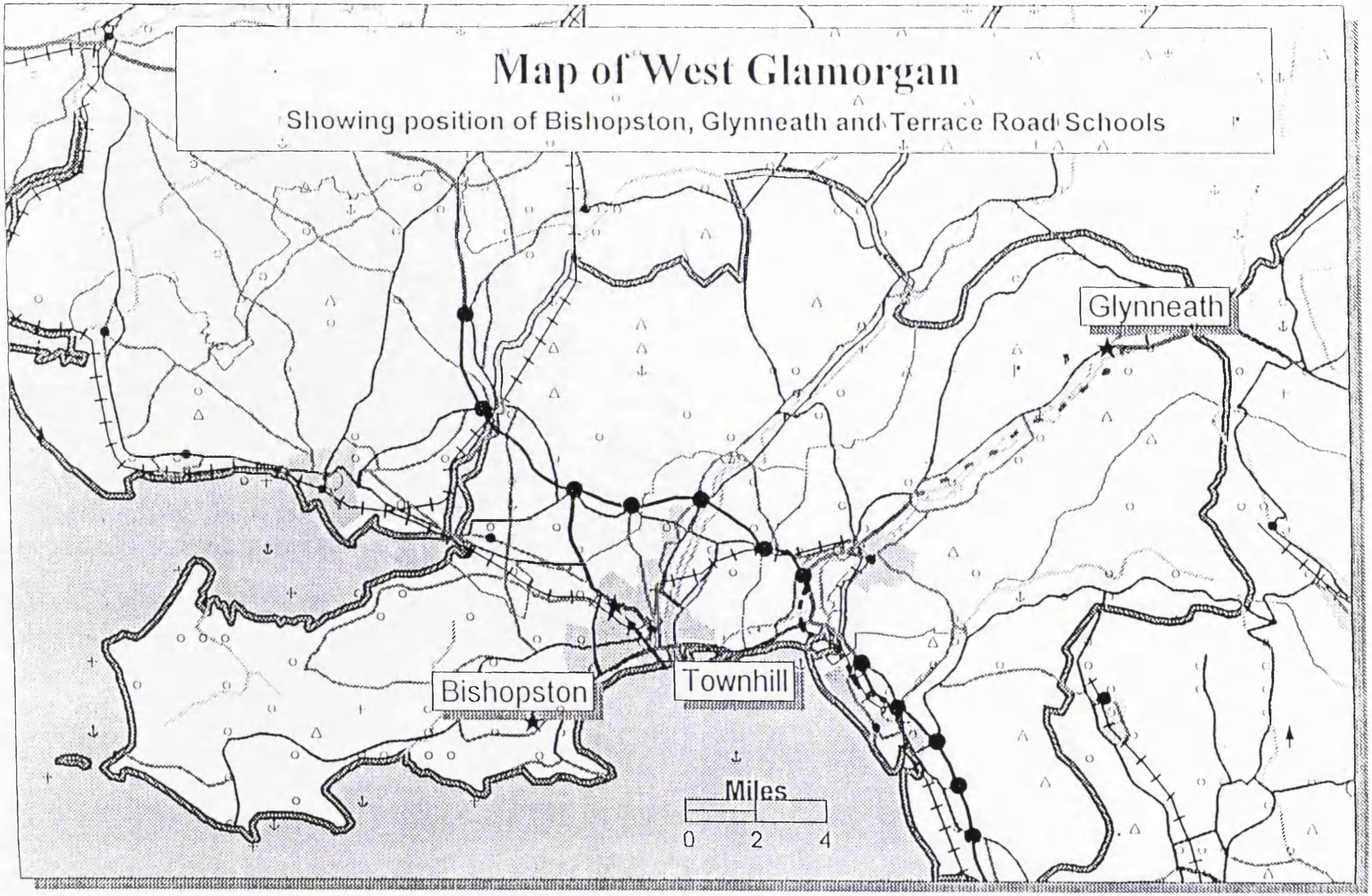
Appendix 2 .1

The Maps showing the locations used in the study

Map 1 shows an Ordnance Survey map of Swansea and the locations of Bishopston and Glyn Neath in relation to Swansea. The direction and location of the roads can be clearly seen.

Map 2 is a detailed breakdown of the enumeration districts and shows the locations of the four schools.

Map 3 is a street map of central Swansea showing the exact location of Terrace Road Primary School.

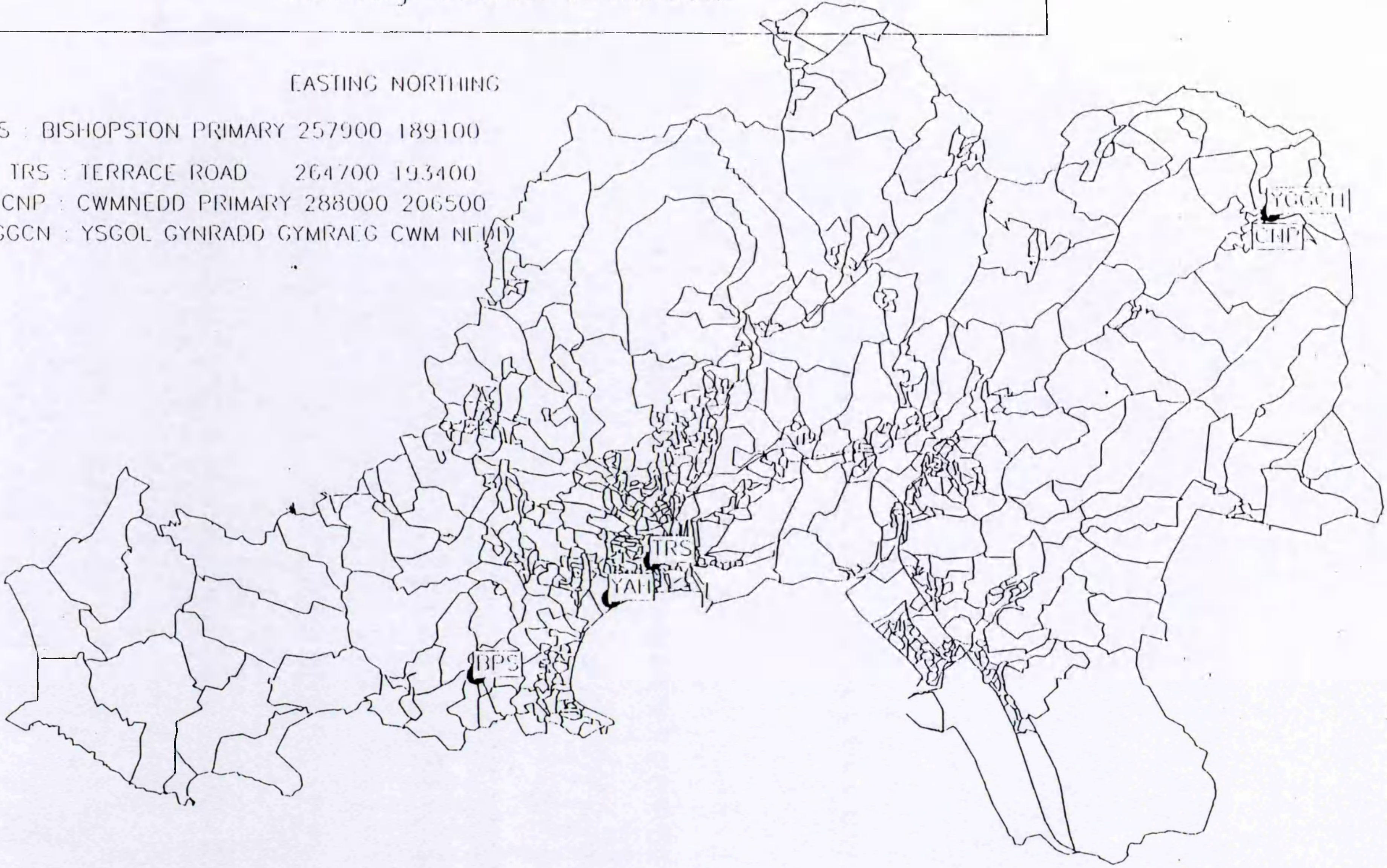


Map of West Glamorgan

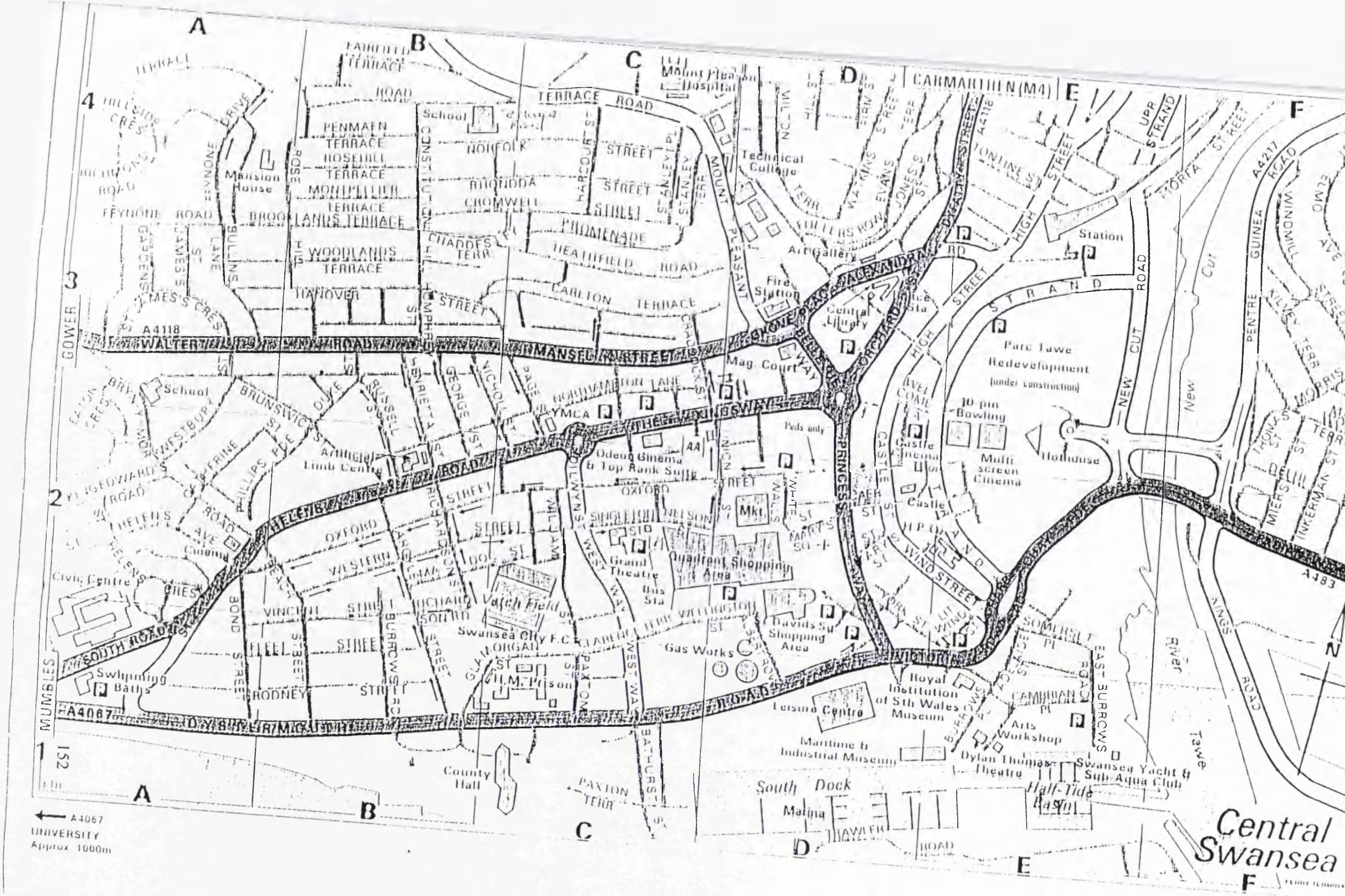
Showing position of Bishopston, Glynneath and Terrace Road Schools

Showing location of Schools

		EASTING	NORTHING
BPS	BISHOPSTON PRIMARY	257900	189100
TRS	TERRACE ROAD	264700	193400
CNP	CWMNEDD PRIMARY	288000	206500
YGGCN	YSGOL GYNRADD GYMRAEG CWM NEDD		



YAH : YOU ARE HERE


Central Swansea
152
A4067 UNIVERSITY Approx 1000m
CARMARTHEN (M4)
MUMBLES
GOWER
Mount Pleasant Hospital
Technical College
Art Gallery
Fire Station
Central Library
Station
Parc Tawe Redevelopment (under construction)
10 pin Bowling
Multi screen Cinema
Hothouse
Castle Cinema
Castle
Mag. Court
YMCA
Odeon Cinema & Top Rank Suite
Mkt
Quadrant Shopping Area
Grand Theatre
Bus Sta
St Davids Sq Shopping Area
Gas Works
Vetch Field
Swansea City F.C.
H.M. Prison
County Hall
Civic Centre
Swimming Baths
Mansion House
School
Artificial Limb Centre
Cinema
Leisure Centre
Royal Institution of Sth Wales Museum
Maritime & Industrial Museum
Arts Workshop
Dylan Thomas Theatre
Swansea Yacht & Sub-Aqua Club
South Dock
Marina
Half-Tide Basin
River Tawe
New Cut
WALTER ROAD
MANSEL STREET
THE KINGSWAY
ST HELENS ROAD
OYSTERMOUTH ROAD
VICTORIA ROAD
PRINCESS WAY
HIGH STREET
STRAND
NEW CUT ROAD
KINGS ROAD
MOUNT PLEASANT
A4118
A4217
A483
A4067

Appendix 2 .2

Original Questionnaire

These questions are to do with your child's chest.

Q1: Has your child ever had wheezing or whistling in the chest at any time in the past?

Answer	Coding
YES	1
NO	0

IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 6

Q2: Has your child had wheezing or whistling in the chest in the last 12 months?

Answer	Coding
YES	1
NO	0

IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 6

Q3: How many attacks of wheezing has your child had in the last 12 months?

Answer	Coding
NONE	0
1 to 3	1
4 to 12	2
More than 12	3

Q4: In the last 12 months how often, on average, has your child's sleep been disturbed due to wheezing?

Answer	Coding
Never woken with wheezing	0
Less than one night per week	1
One or more nights per week	2

Q5: In the last 12 months has wheezing ever been severe enough to limit your child's speech to only one or two words at a time between breaths?

Answer	Coding
YES	1
NO	0

Q6: Has your child ever had asthma?

Answer	Coding
YES	1
NO	0

IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 8

Q7: Has your child received treatment for asthma in the last 4 weeks?

Answer	Coding
YES	1
NO	0

Q8: In the last 12 months, has your child's chest sounded wheezy during or after exercise?

Answer	Coding
YES	1
NO	0

Q9: In the last 12 months, has your child had a dry cough at night, apart from a cough associated with a cold or chest infection?

Answer	Coding
YES	1
NO	0

The next few questions have to do with your home.

Q10: Do you use gas for cooking?

Answer	Coding
YES	1
NO	0

Q11: Do you use a calor type free standing bottle gas heater in your home?

Answer	Coding
YES	1
NO	0

Q12: Do you use paraffin heaters in your home?

Answer	Coding
YES	1
NO	0

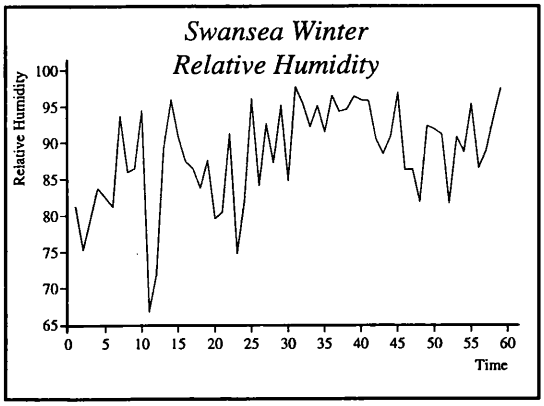
Q13: Does anybody smoke cigarettes, cigars, or a pipe in your home on most days of the week?

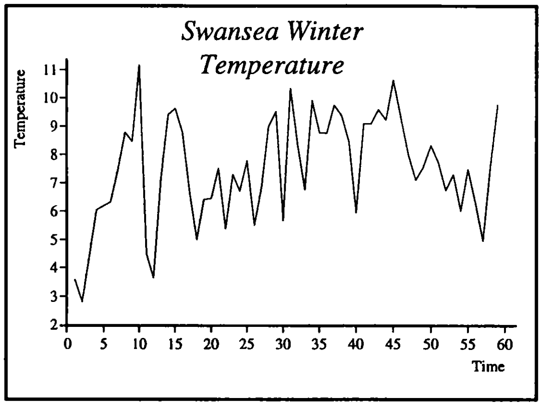
Answer	Coding
YES	1
NO	0

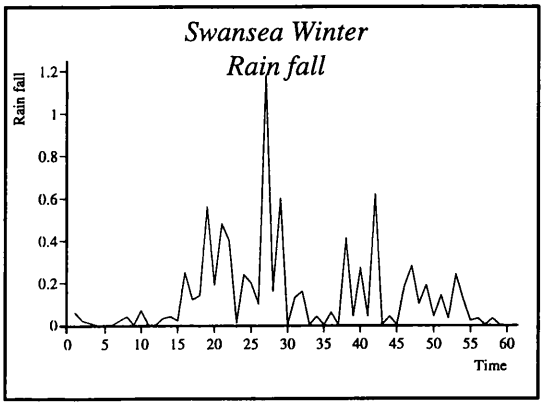
Q14: Do you have a pet bird, dog, cat or other furry animal in the home?

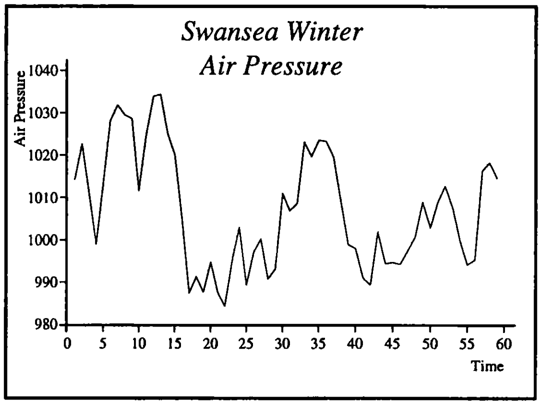
Answer	Coding
YES	1
NO	0

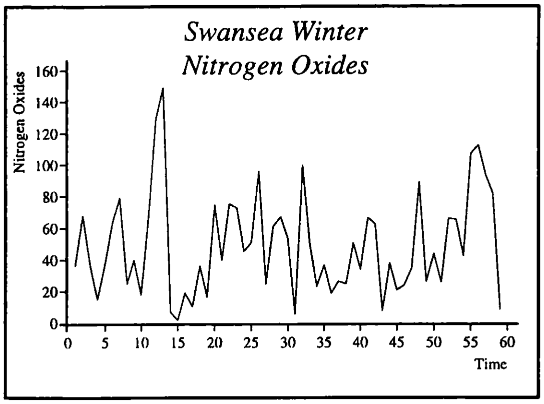
Appendix 2 .3

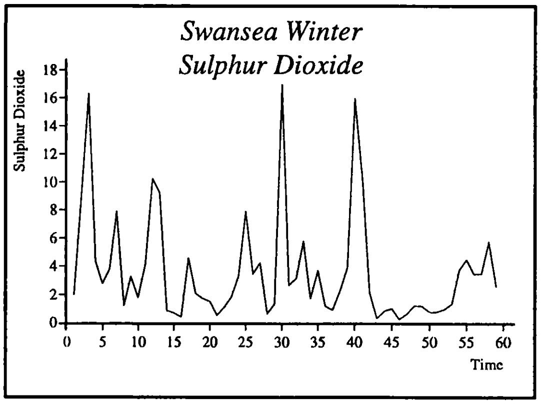

Swansea Winter
Relative Humidity
Relative Humidity
100 95 90 85 80 75 70 65
0 5 10 15 20 25 30 35 40 45 50 55 60
Time

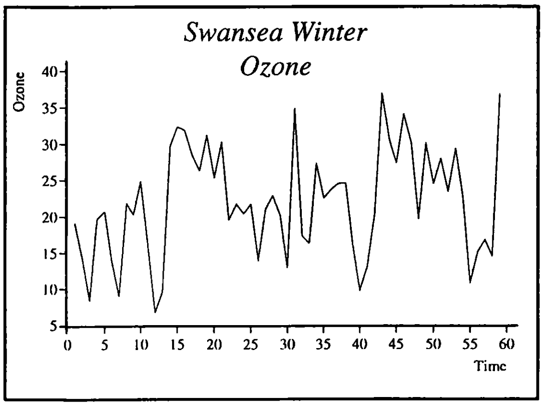

Swansea Winter
Temperature
Temperature
11 10 9 8 7 6 5 4 3 2
0 5 10 15 20 25 30 35 40 45 50 55 60
Time

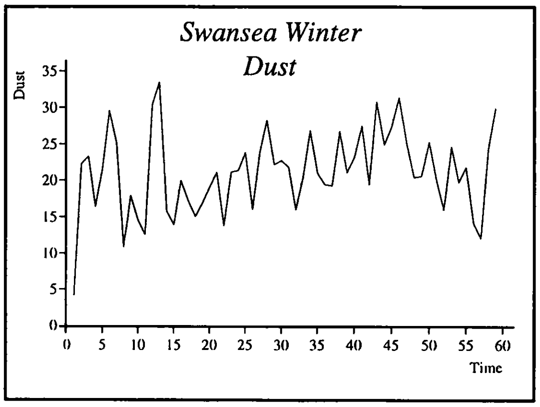

Swansea Winter
Rain fall
Rain fall
1.2 1 0.8 0.6 0.4 0.2 0
0 5 10 15 20 25 30 35 40 45 50 55 60
Time

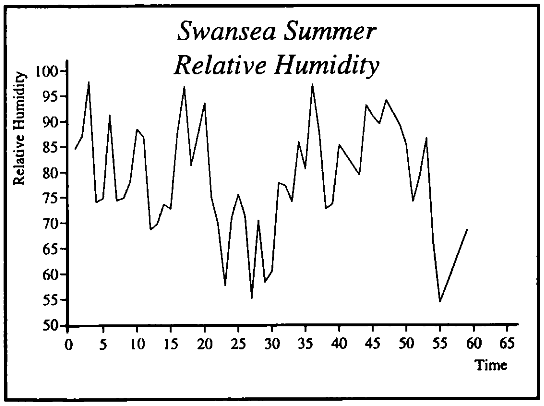

Swansea Winter
Air Pressure
Air Pressure
1040 1030 1020 1010 1000 990 980
0 5 10 15 20 25 30 35 40 45 50 55 60
Time

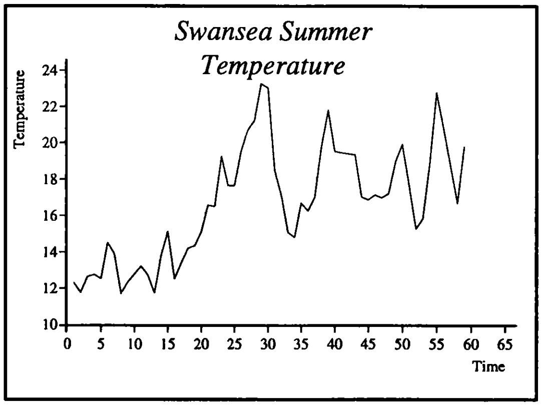

Swansea Winter
Nitrogen Oxides
Nitrogen Oxides
160 140 120 100 80 60 40 20 0
0 5 10 15 20 25 30 35 40 45 50 55 60
Time

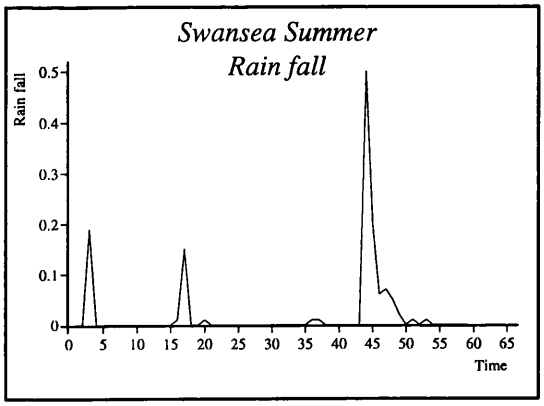

Swansea Winter
Sulphur Dioxide
Sulphur Dioxide
18 16 14 12 10 8 6 4 2 0
0 5 10 15 20 25 30 35 40 45 50 55 60
Time

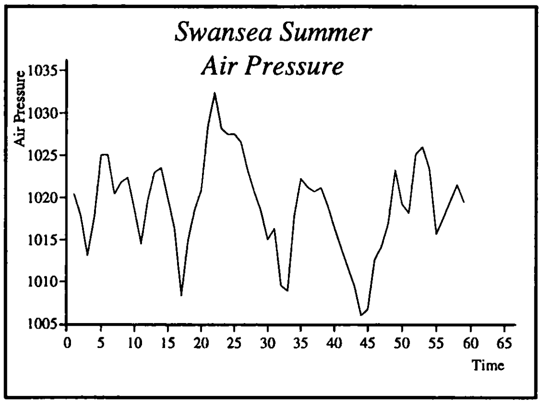

Swansea Winter
Ozone
Ozone
40 35 30 25 20 15 10 5
0 5 10 15 20 25 30 35 40 45 50 55 60
Time

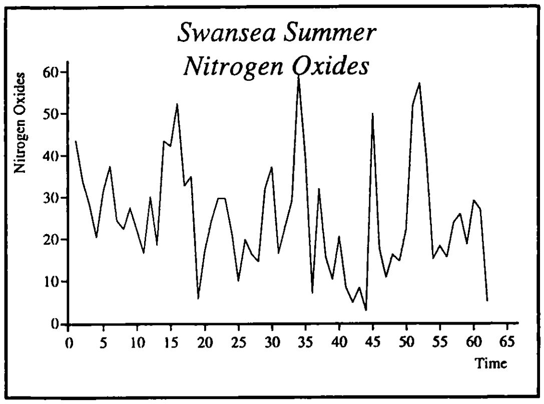

Swansea Winter
Dust
Dust
35 30 25 20 15 10 5 0
0 5 10 15 20 25 30 35 40 45 50 55 60
Time

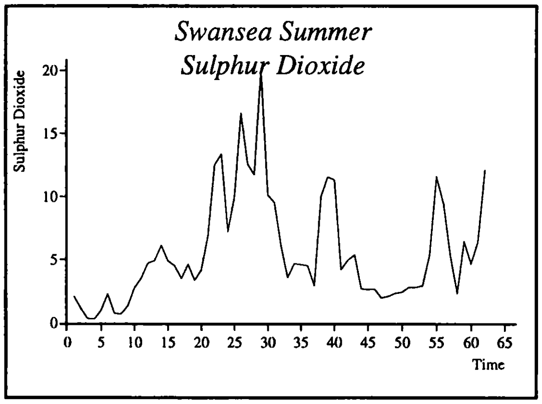

Swansea Summer
Relative Humidity
Relative Humidity
100
95
90
85
80
75
70
65
60
55
50
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

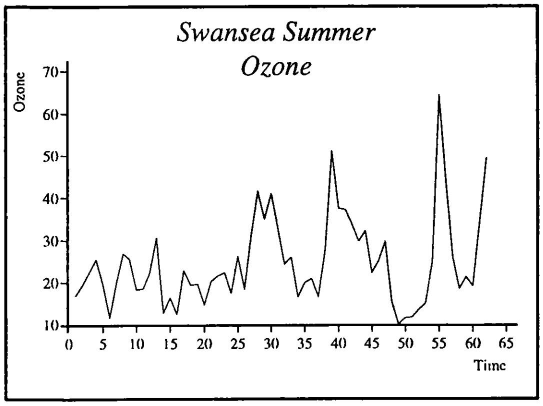

Swansea Summer
Temperature
Temperature
24
22
20
18
16
14
12
10
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

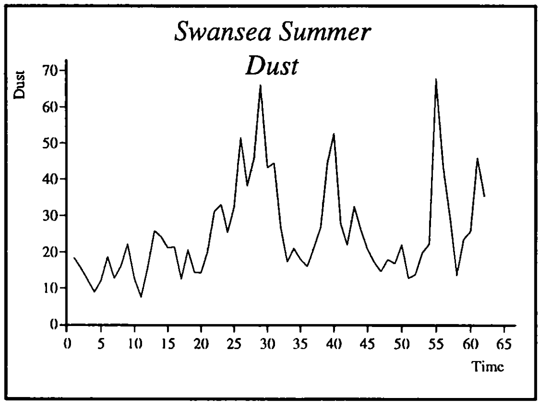

Swansea Summer
Rain fall
Rain fall
0.5
0.4
0.3
0.2
0.1
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

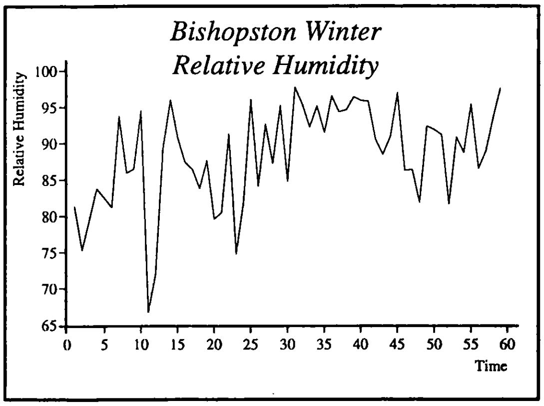

Swansea Summer
Air Pressure
Air Pressure
1035
1030
1025
1020
1015
1010
1005
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

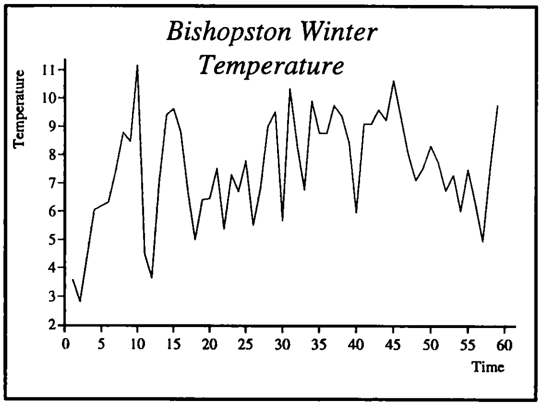

Swansea Summer
Nitrogen Oxides
Nitrogen Oxides
60
50
40
30
20
10
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

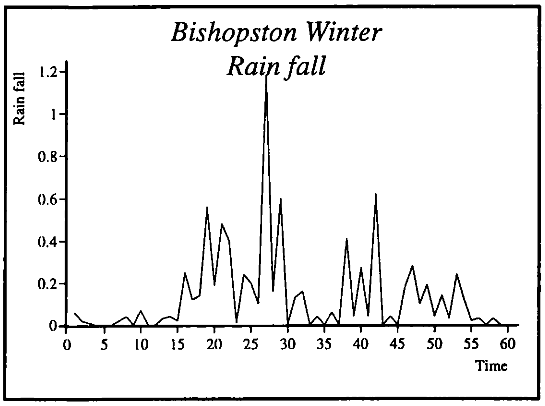

Swansea Summer
Sulphur Dioxide
Sulphur Dioxide
20
15
10
5
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

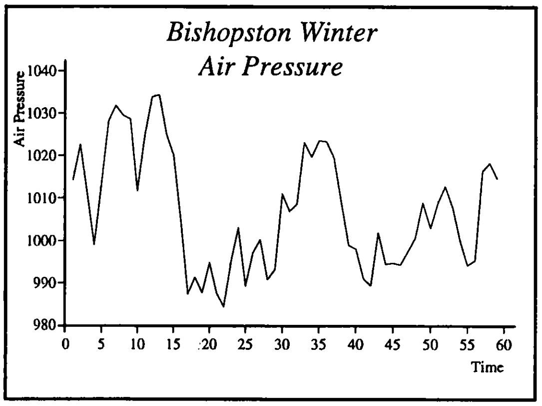

Swansea Summer
Ozone
Ozone
70
60
50
40
30
20
10
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

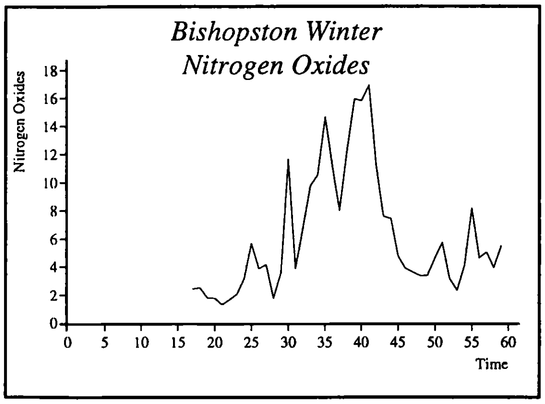

Swansea Summer
Dust
Dust
70
60
50
40
30
20
10
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

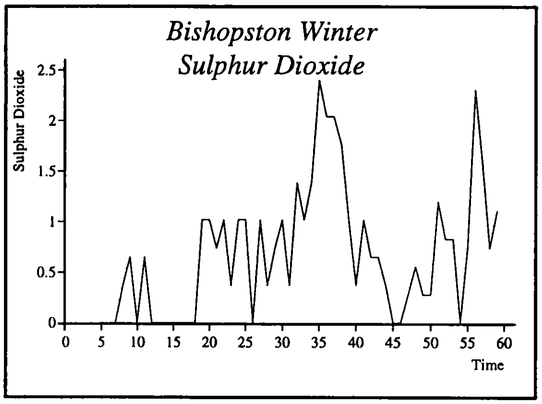

Bishopston Winter
Relative Humidity
Relative Humidity
100
95
90
85
80
75
70
65
0
5
10
15
20
25
30
35
40
45
50
55
60
Time

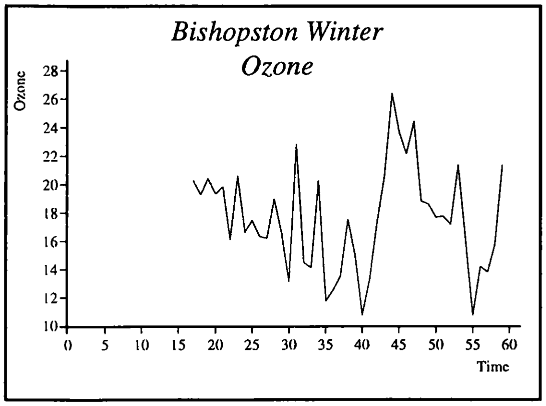

Bishopston Winter
Temperature
Temperature
11
10
9
8
7
6
5
4
3
2
0
5
10
15
20
25
30
35
40
45
50
55
60
Time

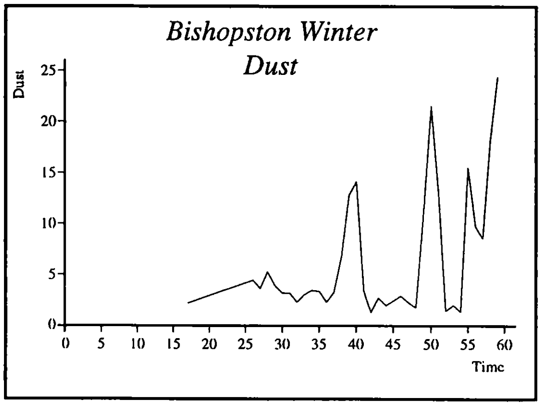

Bishopston Winter
Rain fall
Rain fall
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
35
40
45
50
55
60
Time

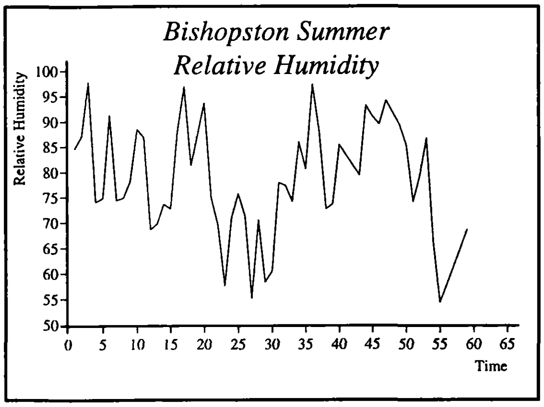

Bishopston Winter
Air Pressure
Air Pressure
1040
1030
1020
1010
1000
990
980
0
5
10
15
20
25
30
35
40
45
50
55
60
Time

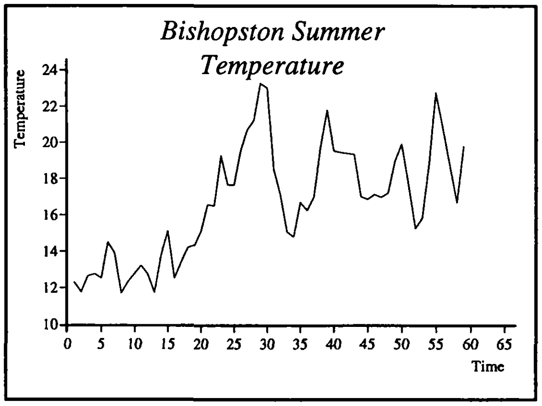

Bishopston Winter
Nitrogen Oxides
Nitrogen Oxides
18
16
14
12
10
8
6
4
2
0
0
5
10
15
20
25
30
35
40
45
50
55
60
Time

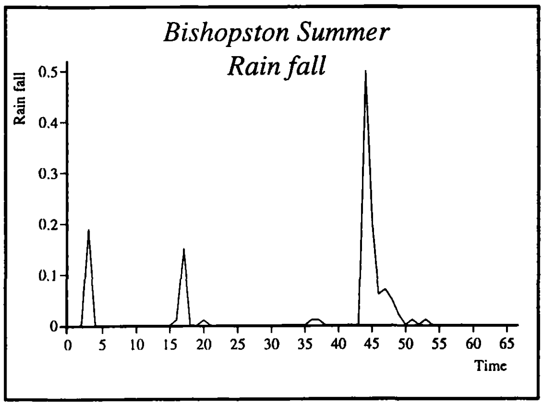

Bishopston Winter
Sulphur Dioxide
Sulphur Dioxide
2.5
2
1.5
1
0.5
0
0
5
10
15
20
25
30
35
40
45
50
55
60
Time

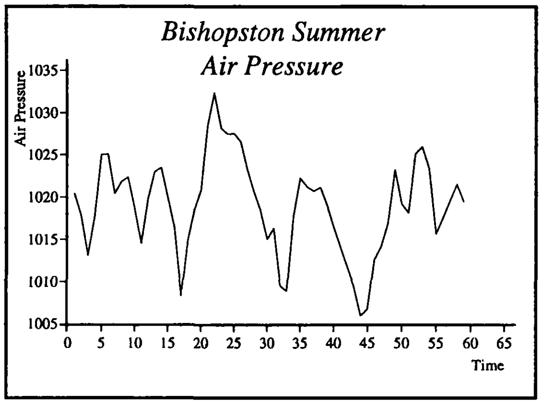

Bishopston Winter
Ozone
Ozone
28
26
24
22
20
18
16
14
12
10
0
5
10
15
20
25
30
35
40
45
50
55
60
Time

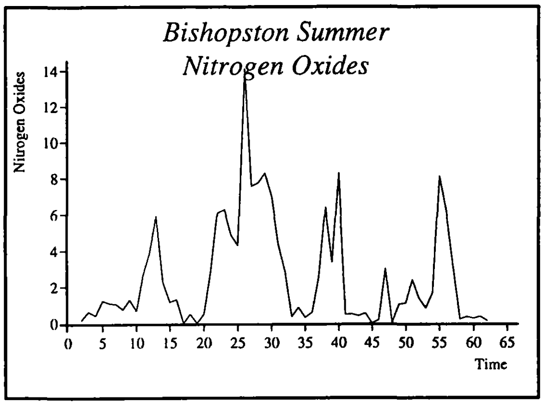

Bishopston Winter
Dust
Dust
25
20
15
10
5
0
0
5
10
15
20
25
30
35
40
45
50
55
60
Time

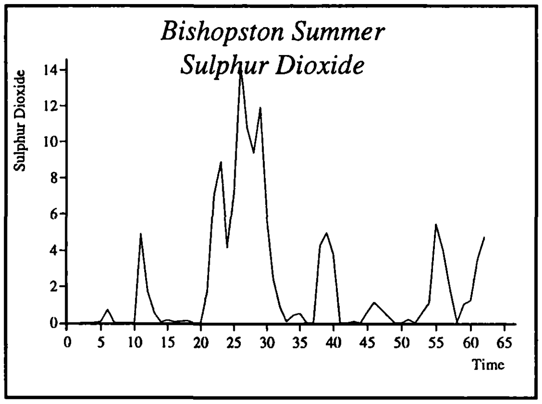

Bishopston Summer
Relative Humidity

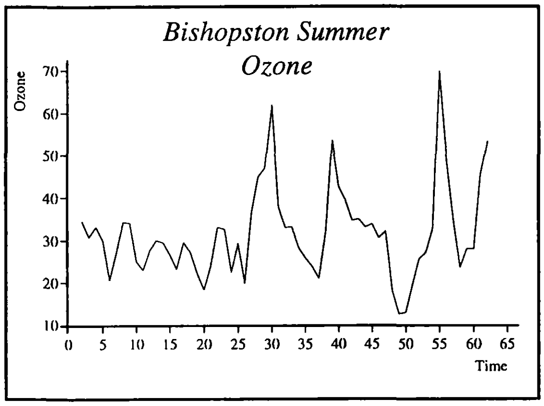

Bishopston Summer
Temperature

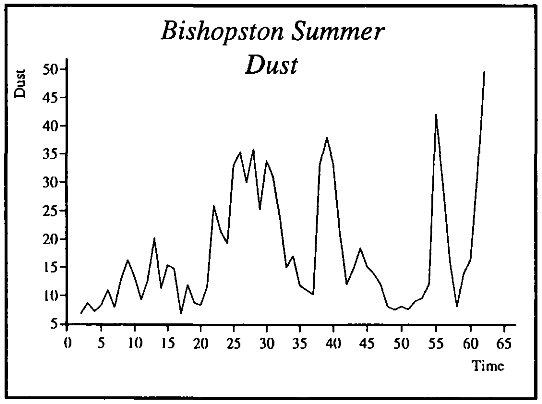

Bishopston Summer
Rain fall

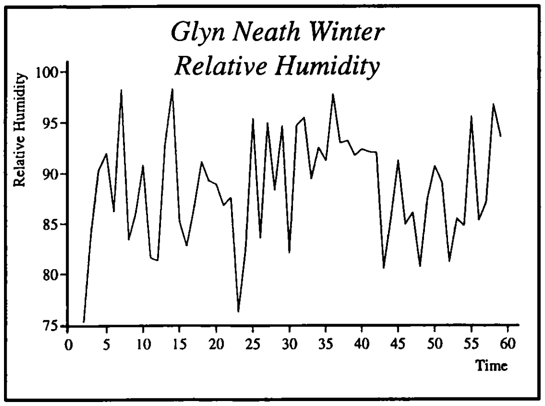

Bishopston Summer
Air Pressure

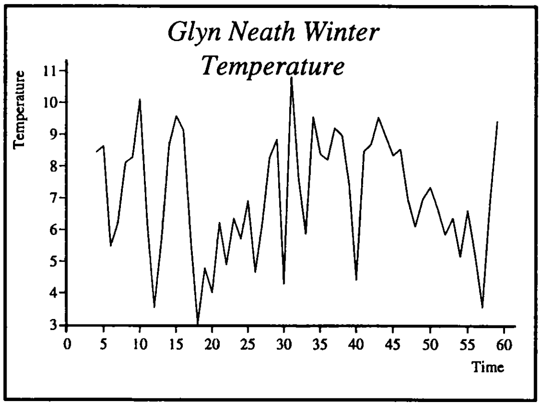

Bishopston Summer
Nitrogen Oxides

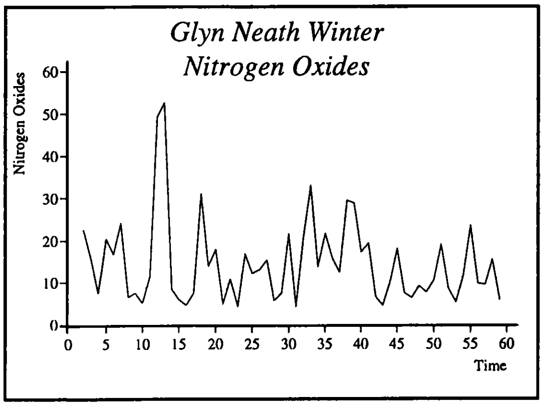

Bishopston Summer
Sulphur Dioxide

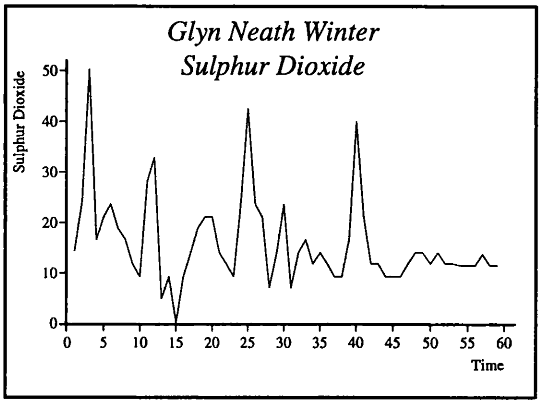

Bishopston Summer
Ozone

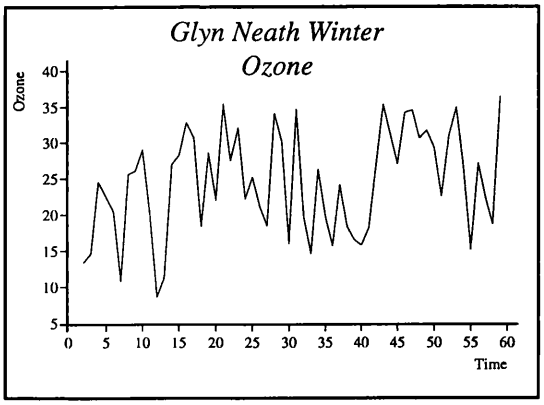

Bishopston Summer
Dust

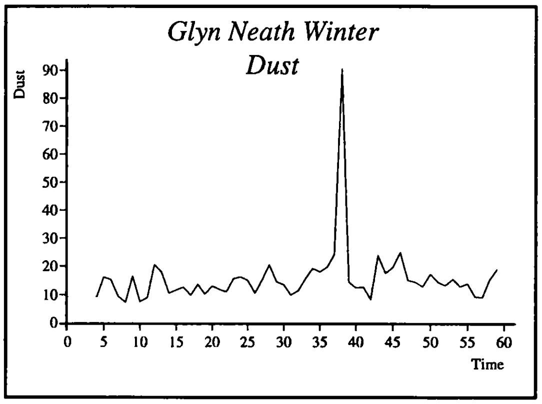

Glyn Neath Winter
Relative Humidity

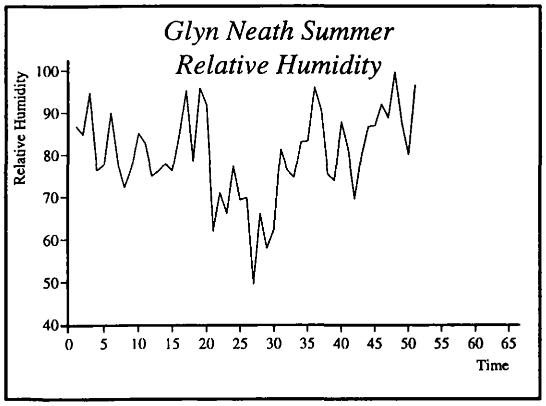

Glyn Neath Winter
Temperature

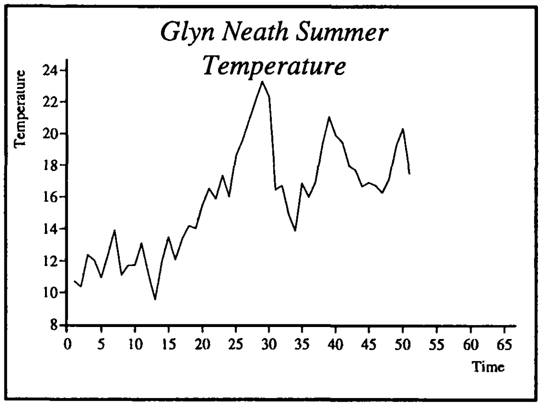

Glyn Neath Winter
Nitrogen Oxides

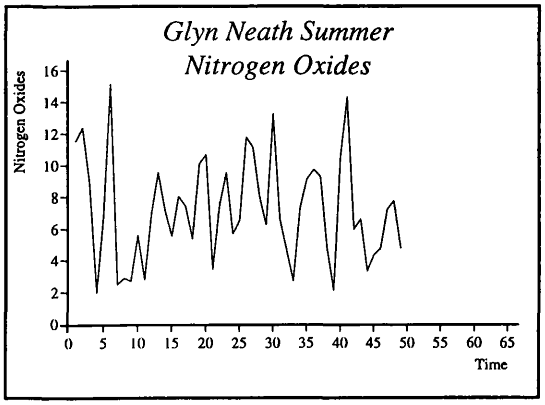

Glyn Neath Winter
Sulphur Dioxide

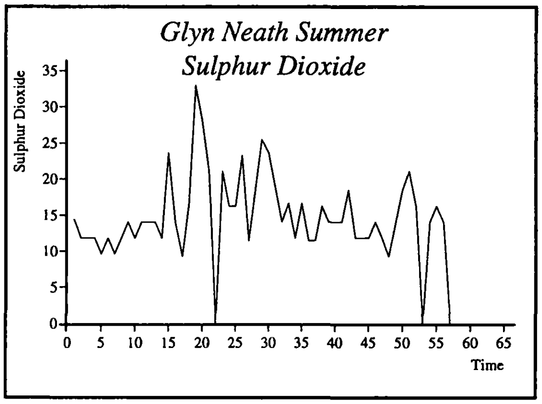

Glyn Neath Winter
Ozone

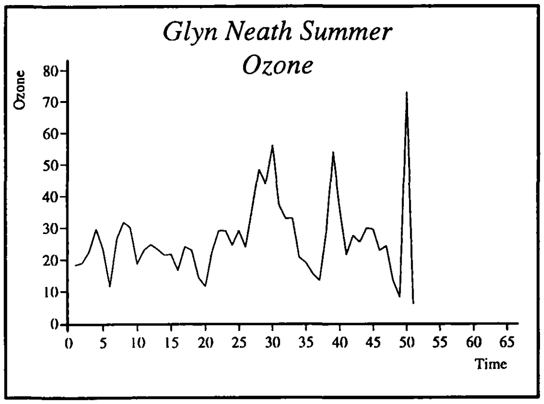

Glyn Neath Winter
Dust

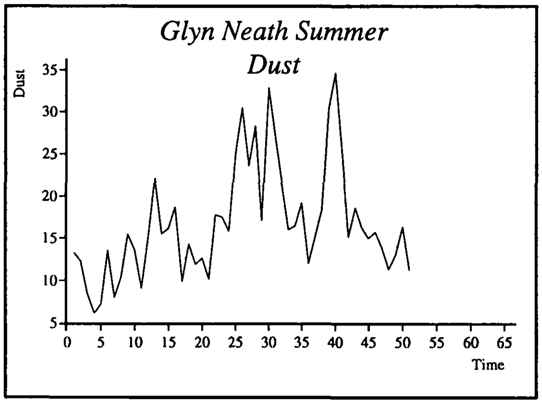

Glyn Neath Summer
Relative Humidity
Relative Humidity
100
90
80
70
60
50
40
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

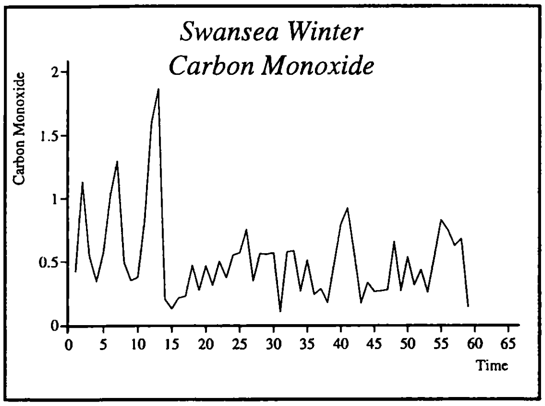

Glyn Neath Summer
Temperature
Temperature
24
22
20
18
16
14
12
10
8
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

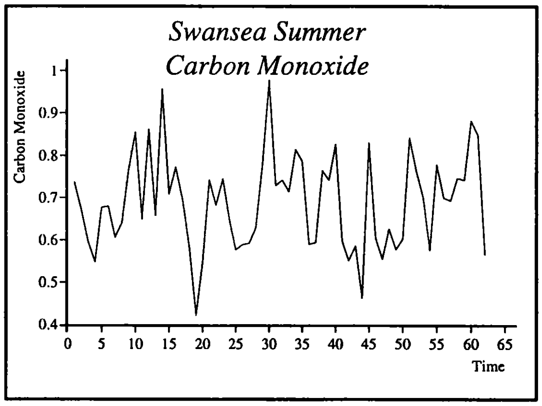

Glyn Neath Summer
Nitrogen Oxides
Nitrogen Oxides
16
14
12
10
8
6
4
2
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

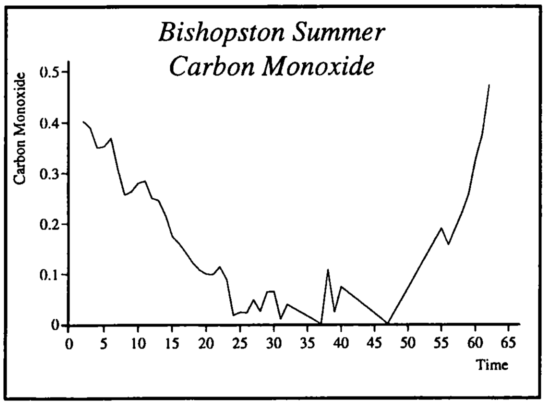

Glyn Neath Summer
Sulphur Dioxide
Sulphur Dioxide
35
30
25
20
15
10
5
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time


Glyn Neath Summer
Ozone
Ozone
80
70
60
50
40
30
20
10
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time


Glyn Neath Summer
Dust
Dust
35
30
25
20
15
10
5
0 5 10 15 20 25 30 35 40 45 50 55 60 65
Time

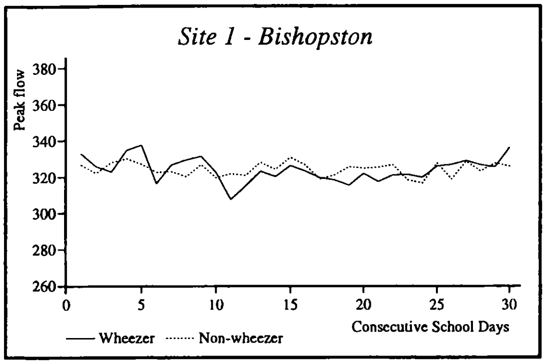

Swansea Winter
Carbon Monoxide
Carbon Monoxide
0
0.5
1
1.5
2
0
5
10
15
20
25
30
35
40
45
50
55
60
65
Time


Swansea Summer
Carbon Monoxide
Carbon Monoxide
0.4
0.5
0.6
0.7
0.8
0.9
1
0
5
10
15
20
25
30
35
40
45
50
55
60
65
Time

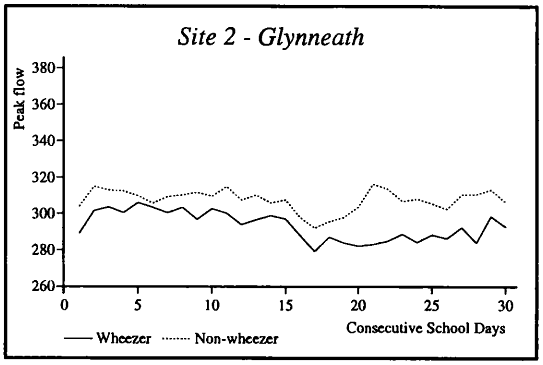

Bishopston Summer
Carbon Monoxide
Carbon Monoxide
0
0.1
0.2
0.3
0.4
0.5
0
5
10
15
20
25
30
35
40
45
50
55
60
65
Time

Appendix 2 .4: Graphs of daily variation in average peak flow rates of wheezers and non-wheezers.

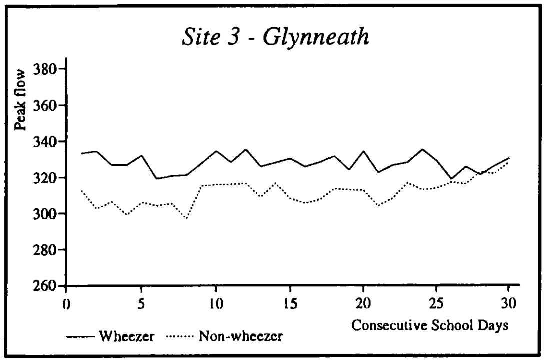
Winter



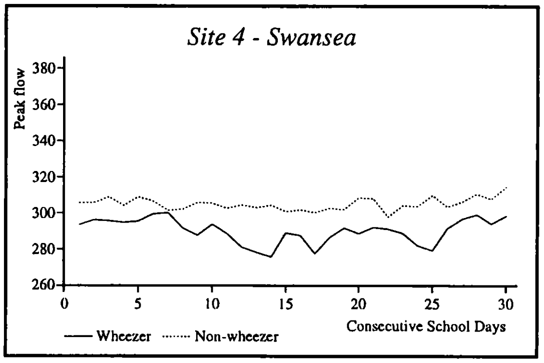
Site 1 - Bishopston



Site 2 - Glynneath

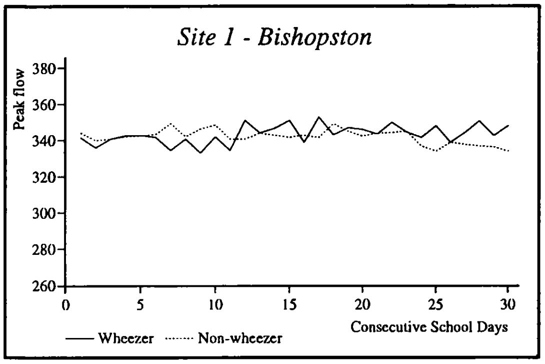


Site 3 - Glynneath

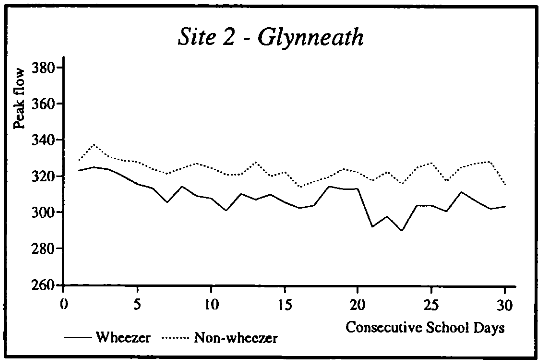


Site 4 - Swansea

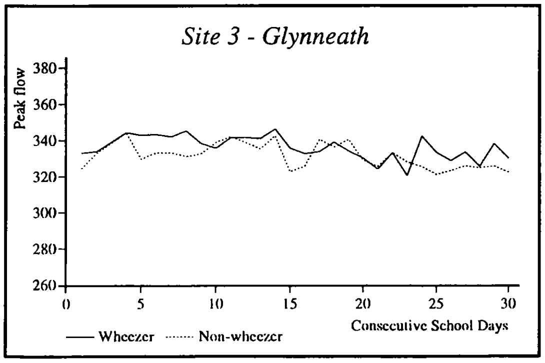
Summer



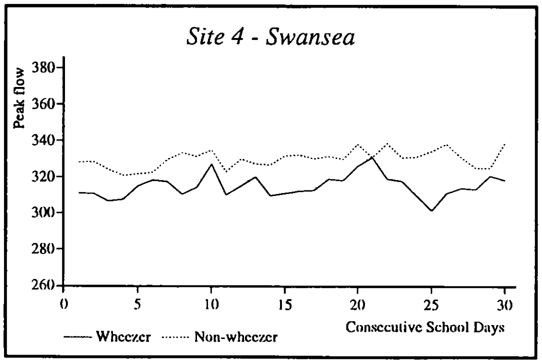
Site 1 - Bishopston



Site 2 - Glynneath

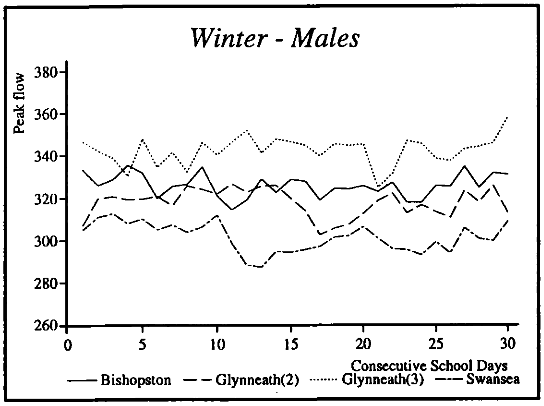


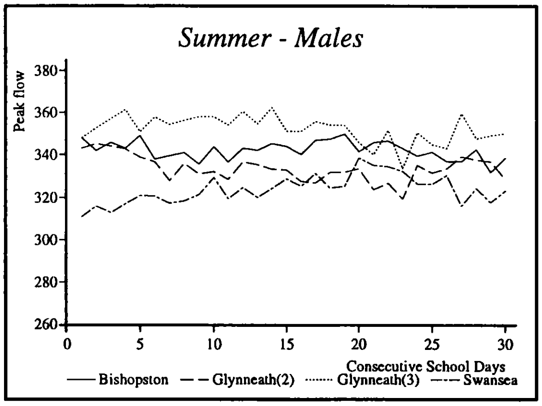
Site 3 - Glynneath

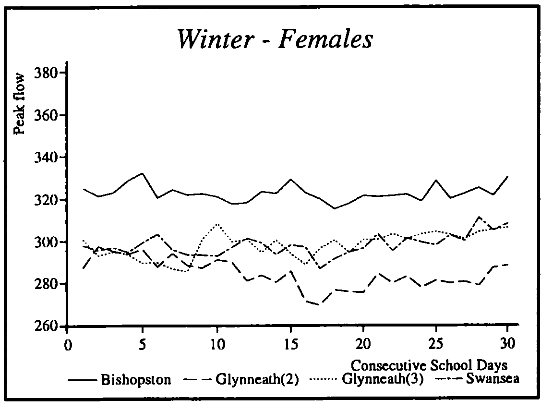


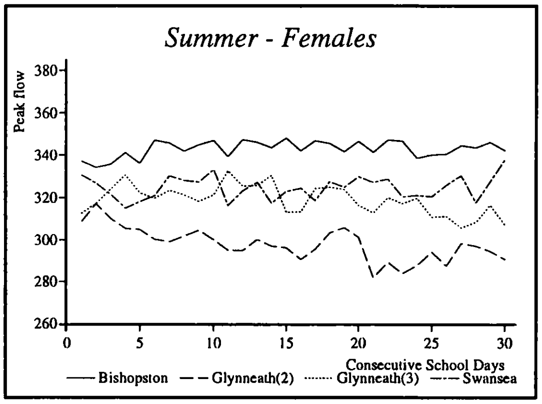
Site 4 - Swansea

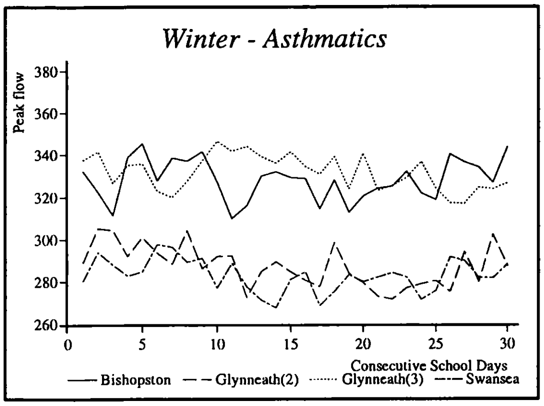
Graphs of daily variation in average peak flow rates by sex and by asthmatic tendency.

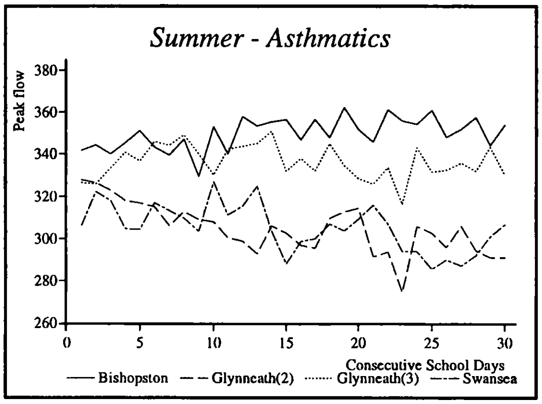


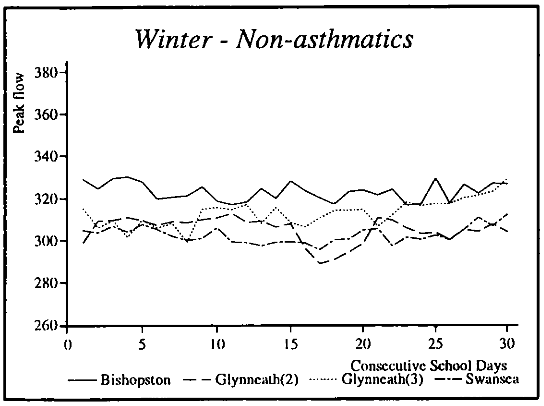


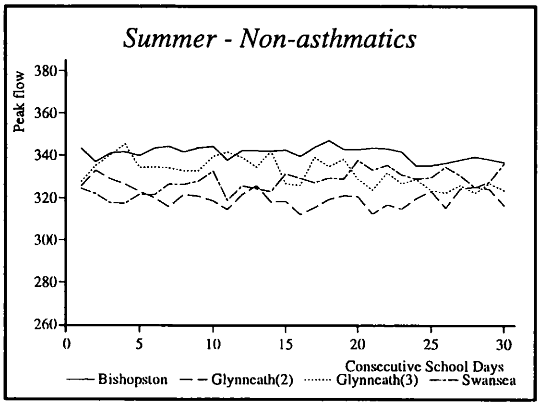




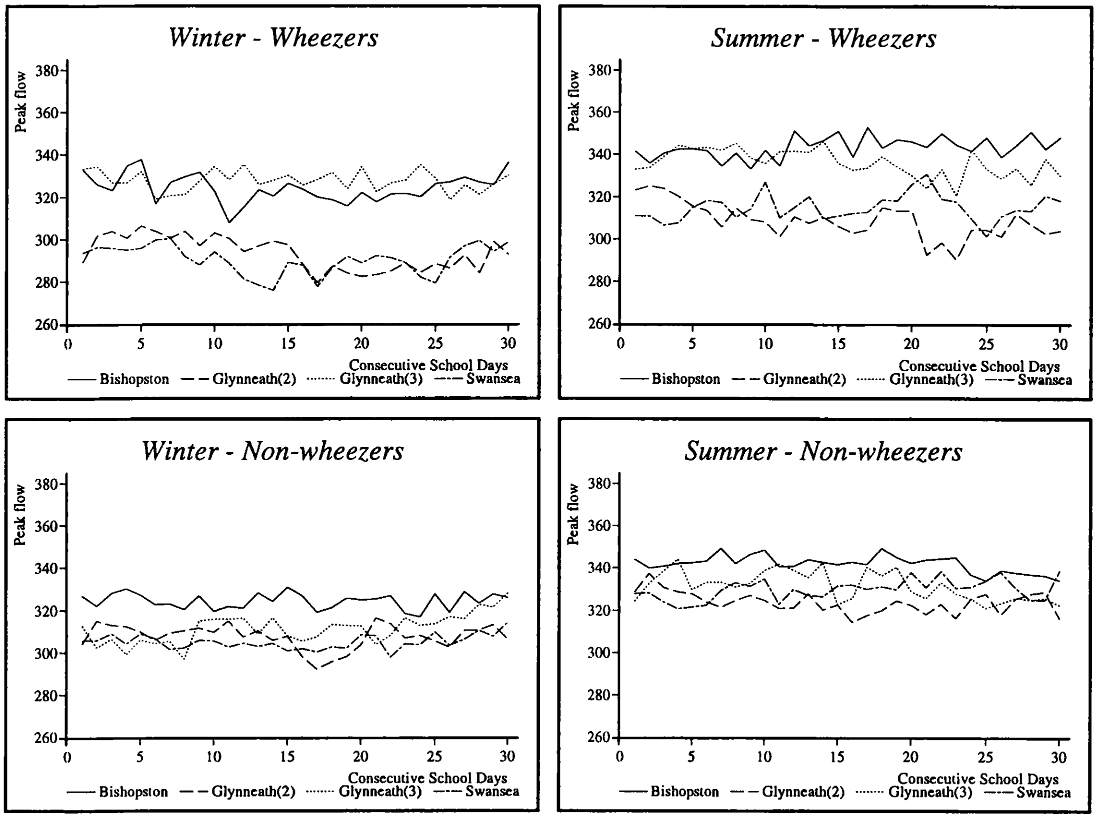






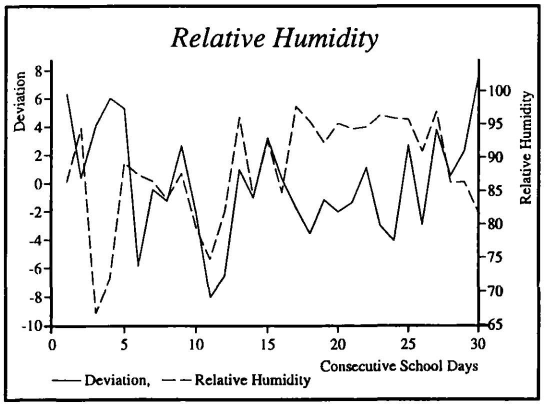


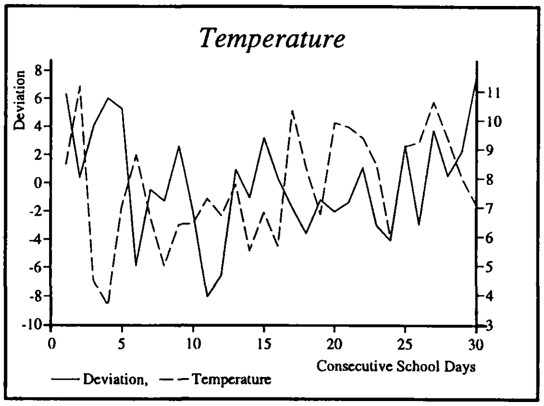
Graph of daily variation in average peak flow rates of wheezers.

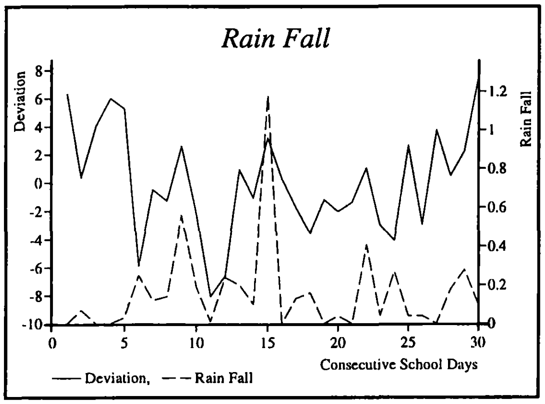


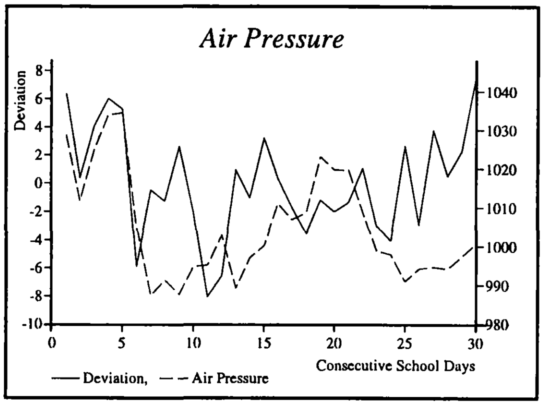
Appendix 2 .5: Average daily deviations and pollution variables

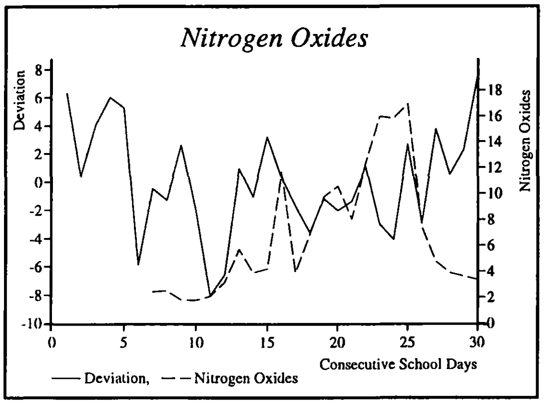
BISHOPSTON - WINTER

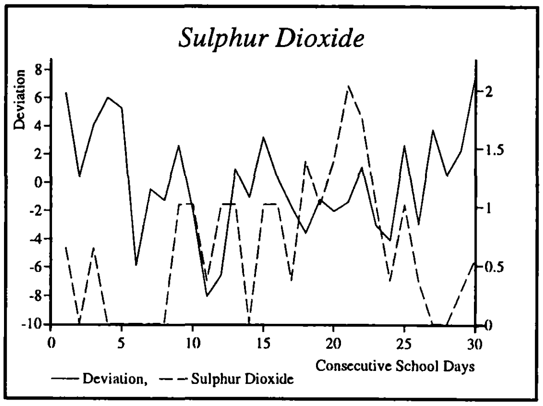


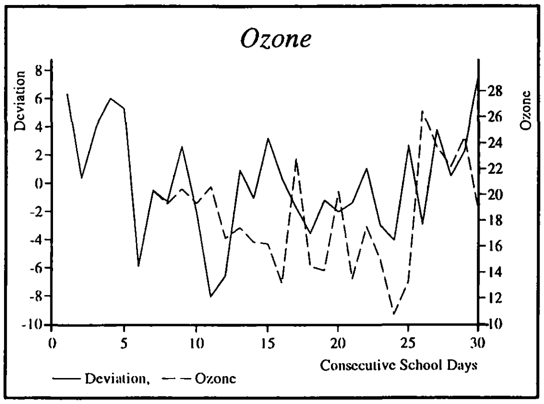


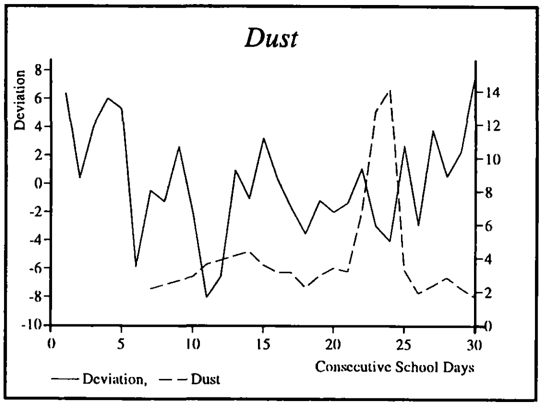




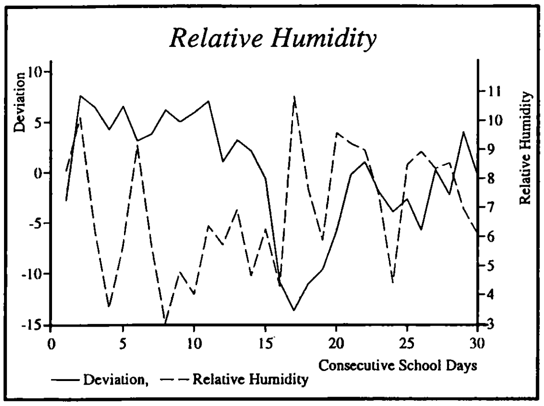


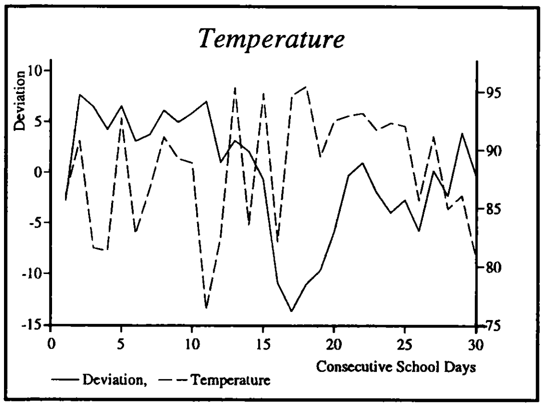


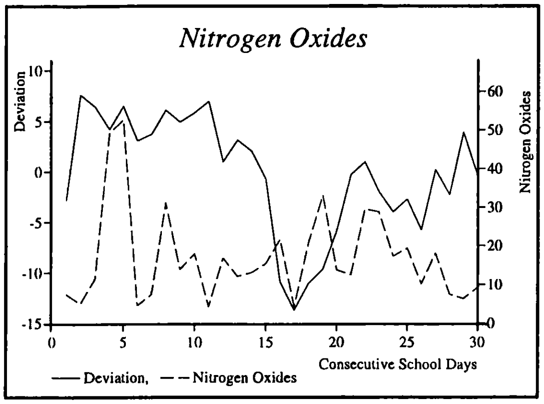


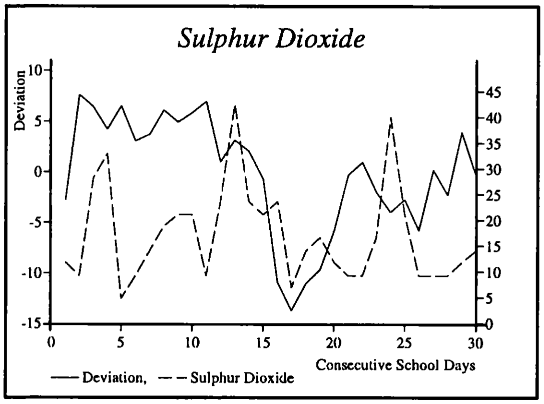


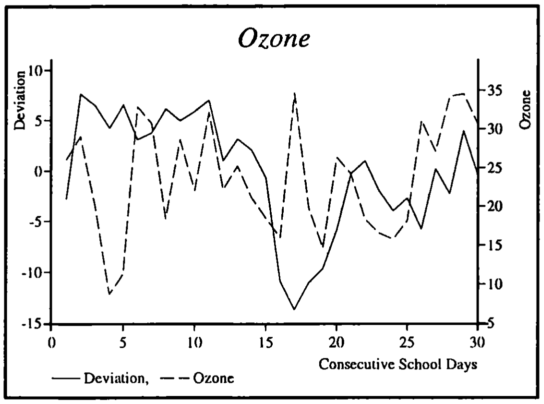
GLYN NEATH (2) - WINTER

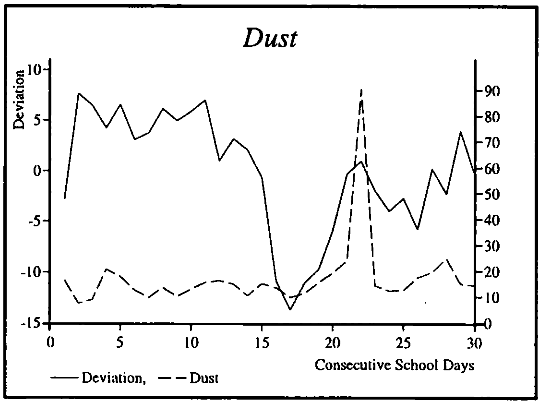

Relative Humidity


Temperature

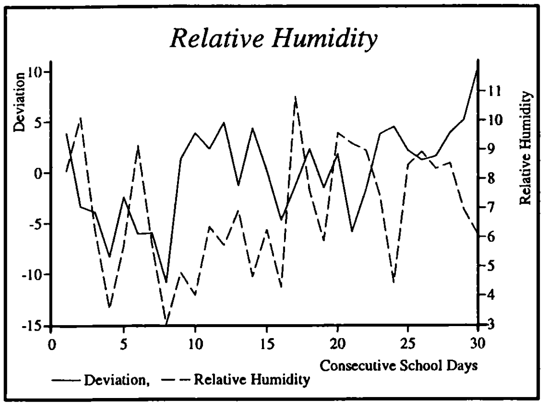

Nitrogen Oxides

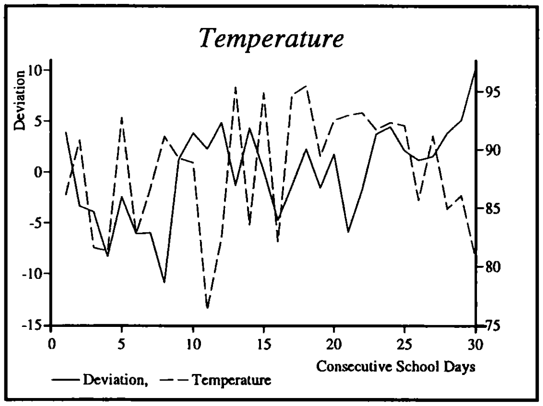

Sulphur Dioxide

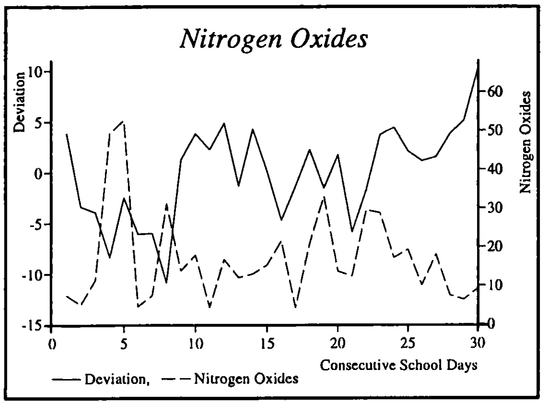

Ozone

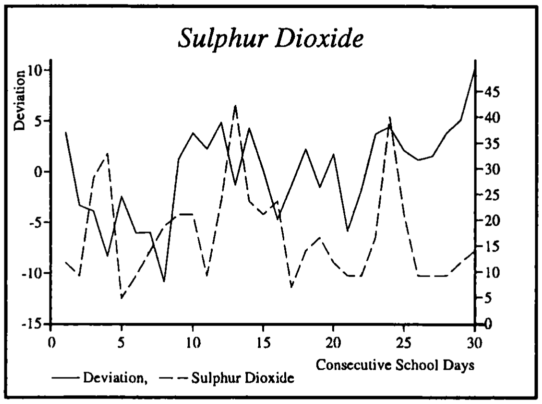

Dust

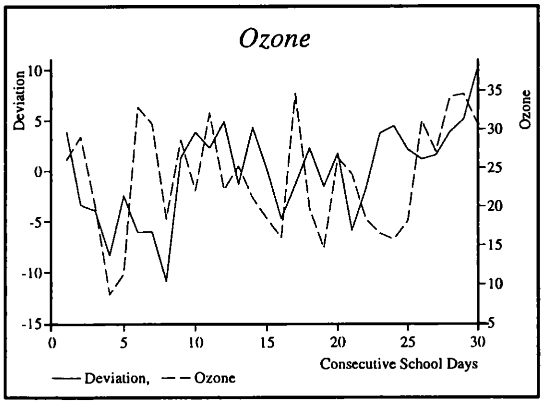
GLYN NEATH (3) - WINTER

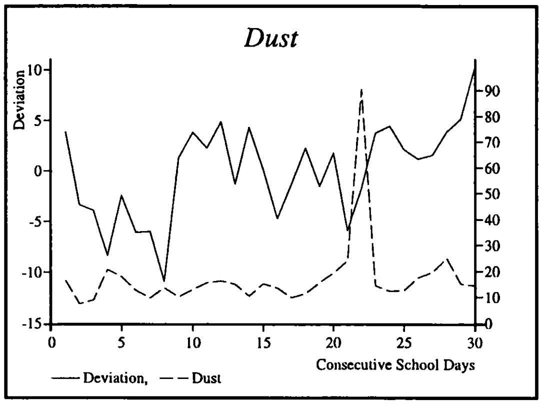




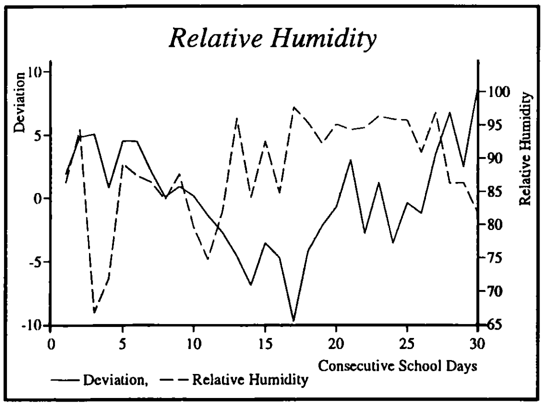


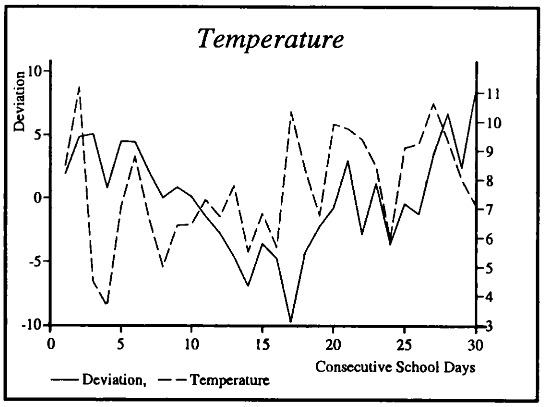


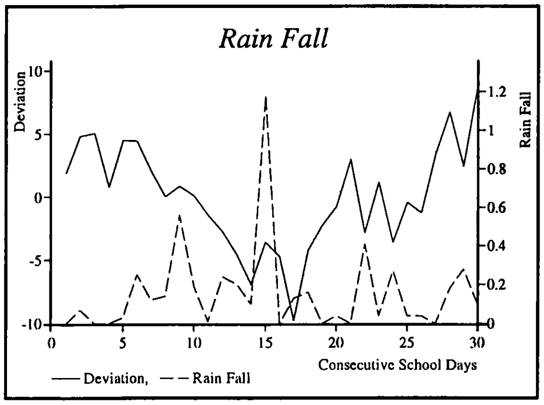


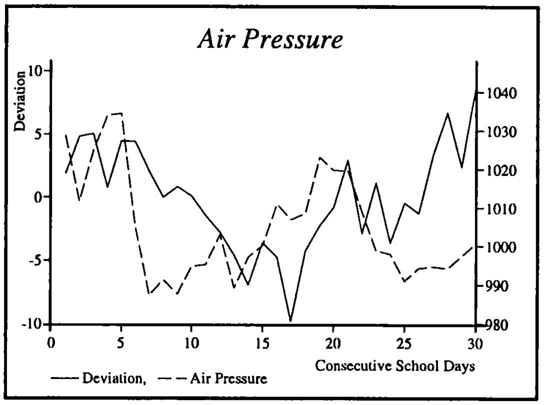


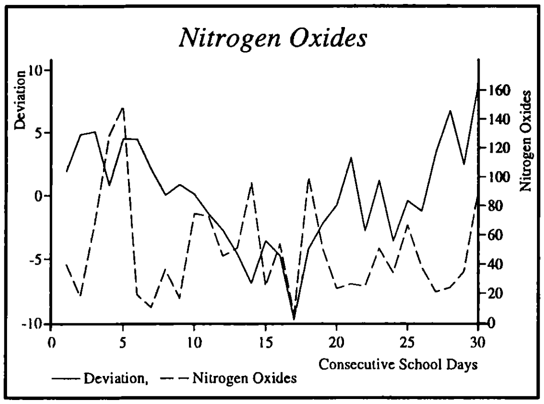
SWANSEA - WINTER

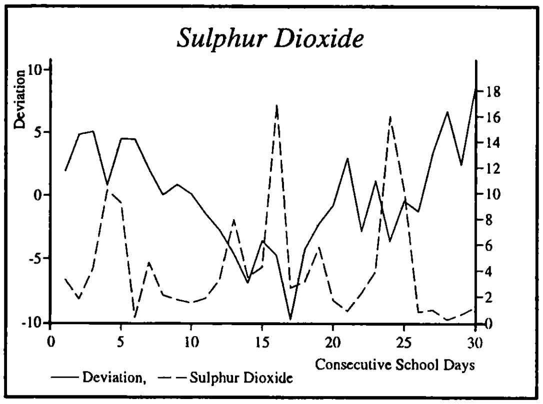

Relative Humidity

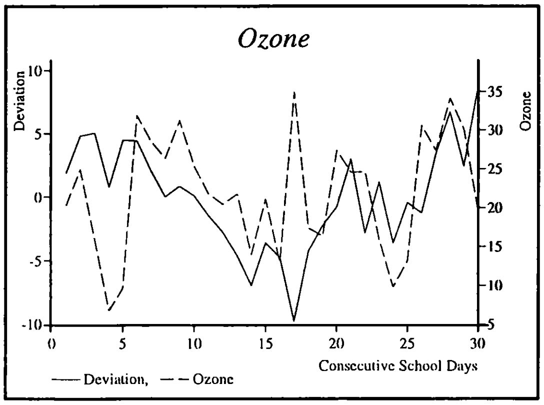

Temperature

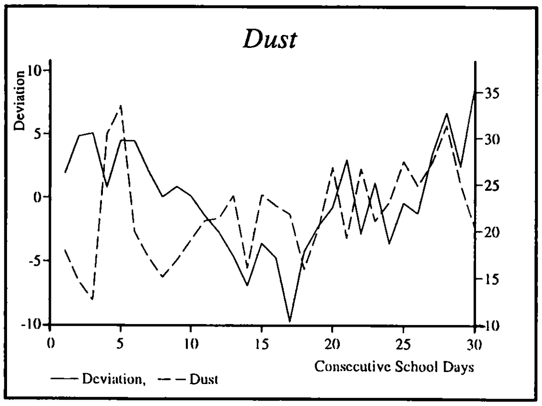

Rain Fall


Air Pressure

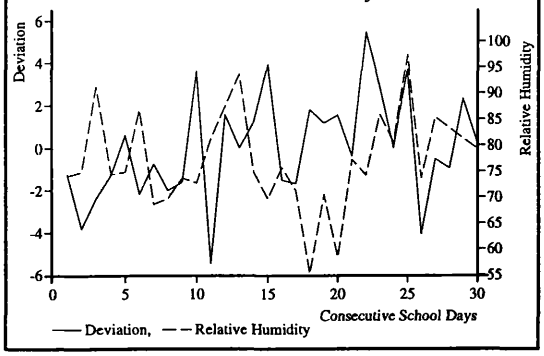

Nitrogen Oxides


Sulphur Dioxide

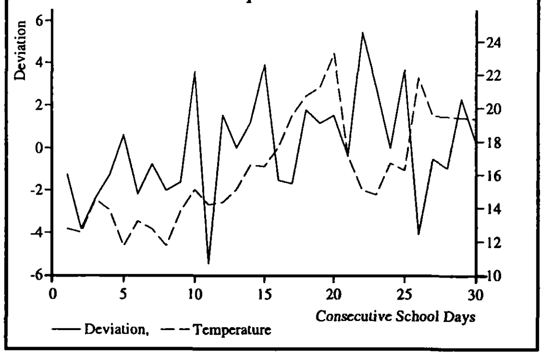

Ozone


Dust

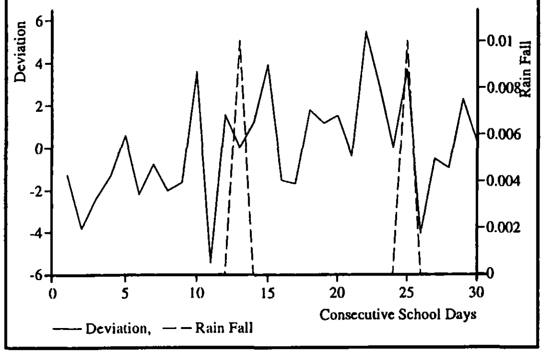
BISHOPSTON - SUMMER



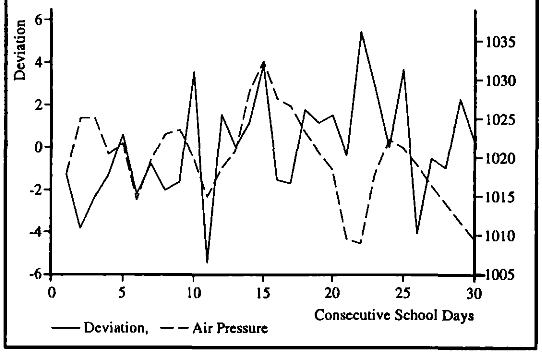
Relative Humidity



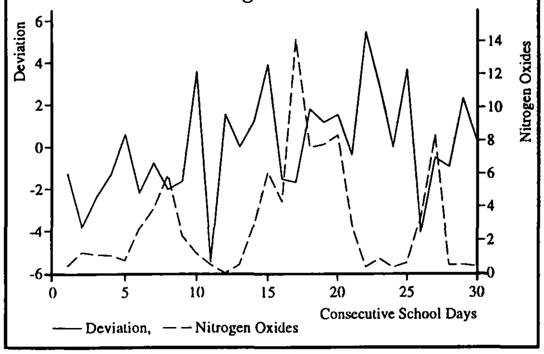
Temperature



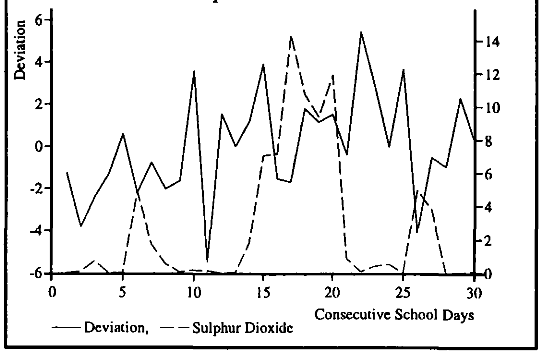
Rain Fall



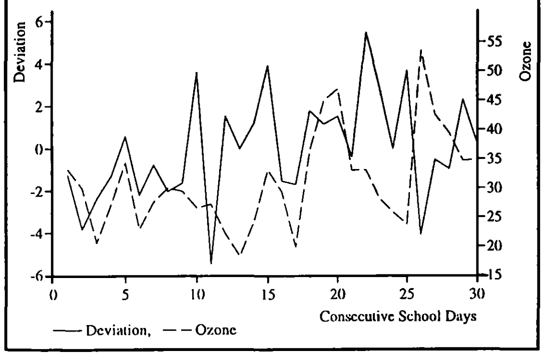
Air Pressure



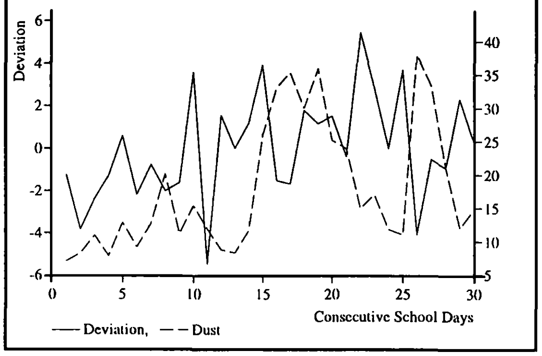
Nitrogen Oxides



Sulphur Dioxide

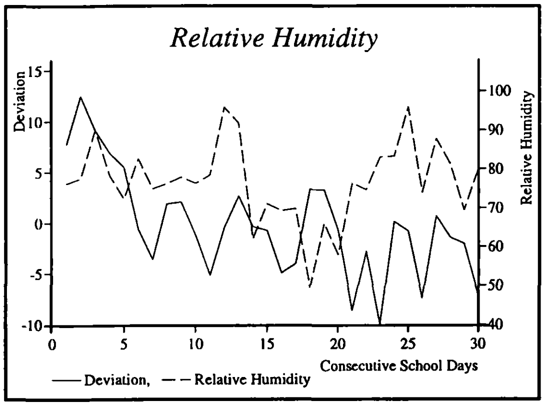


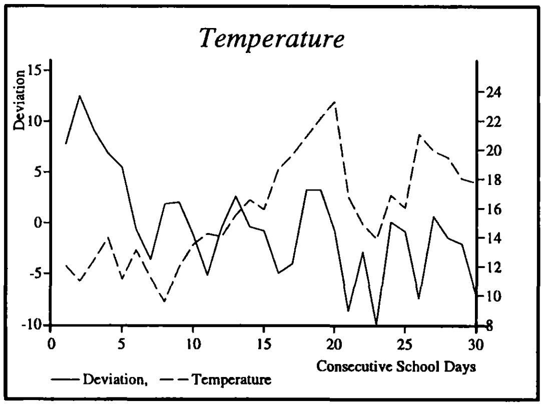
Ozone

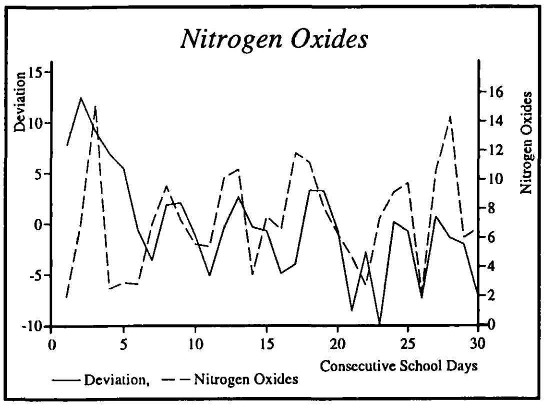


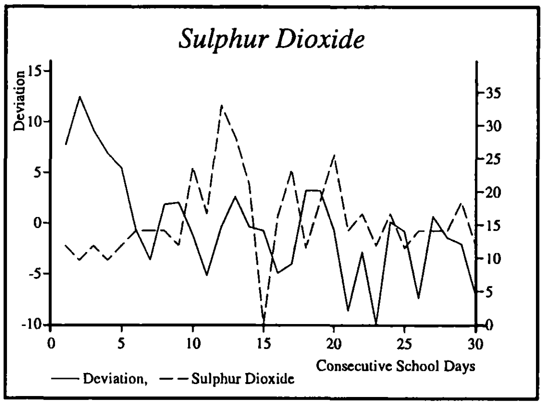
Dust

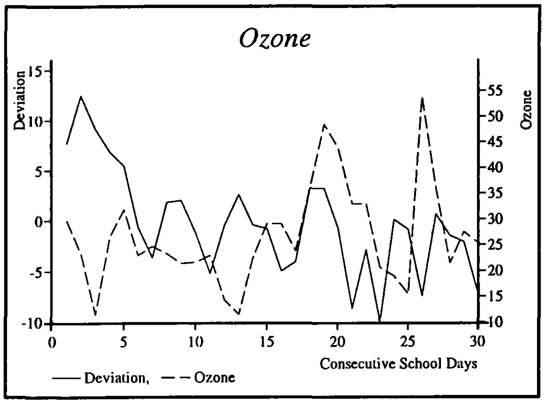
GLYN NEATH (2) - SUMMER

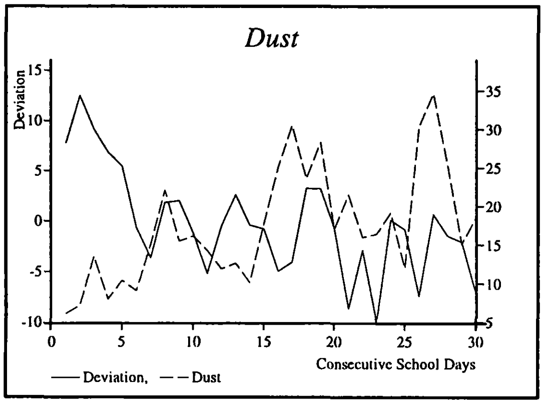

Relative Humidity


Temperature

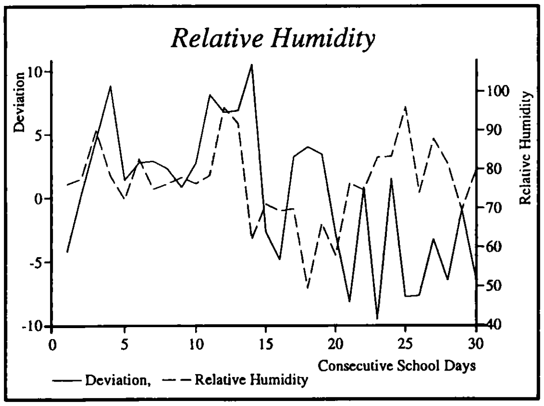

Nitrogen Oxides

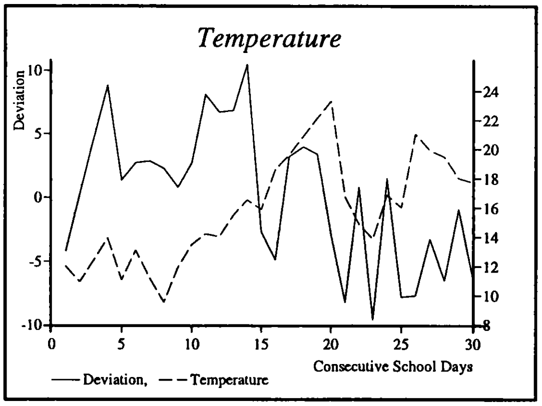

Sulphur Dioxide

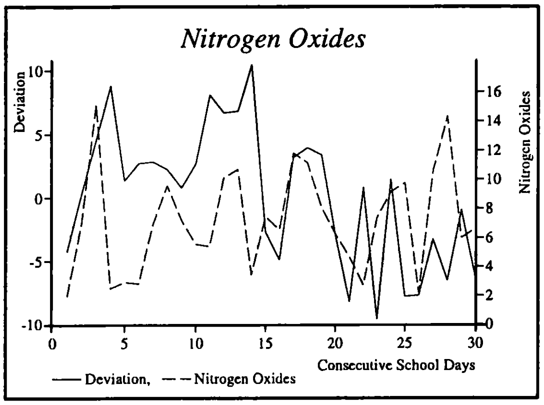

Ozone

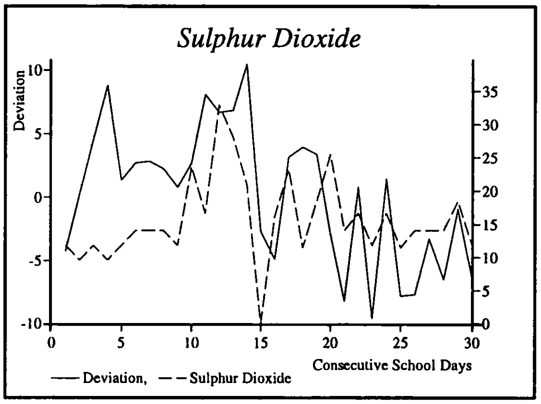

Dust

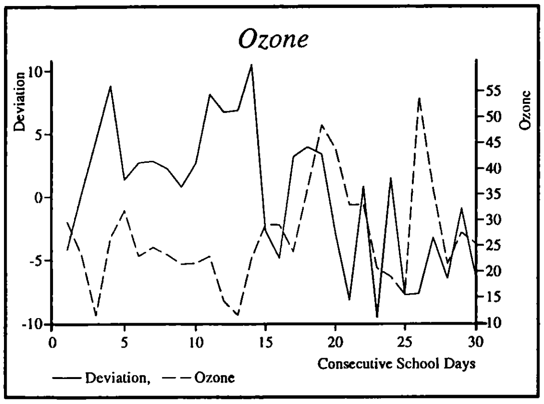
GLYN NEATH (3) - SUMMER

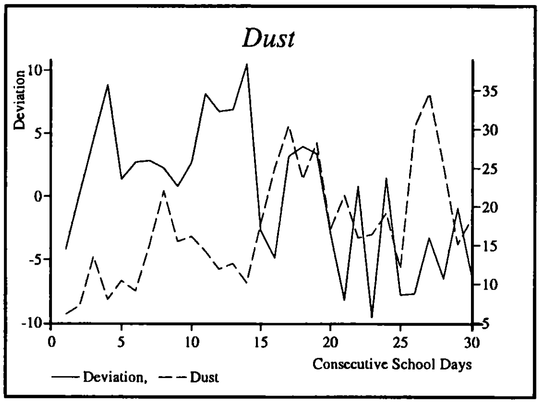

Relative Humidity


Temperature

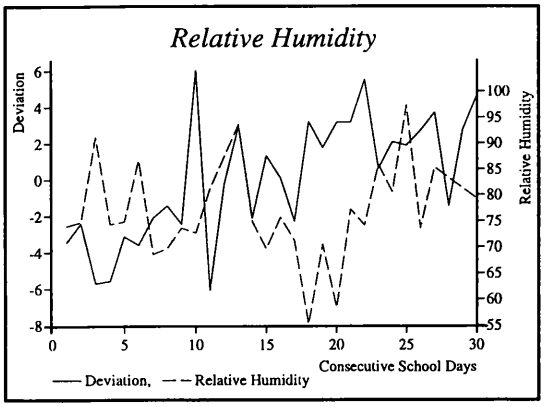

Nitrogen Oxides

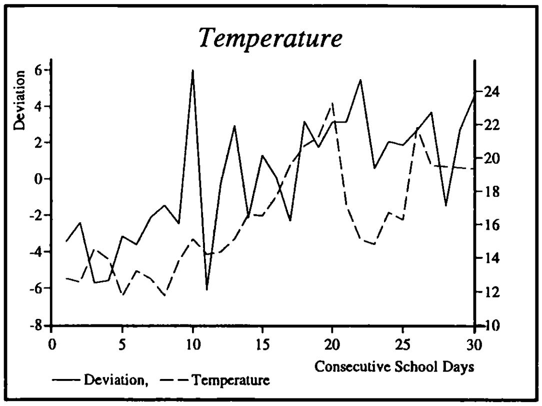

Sulphur Dioxide

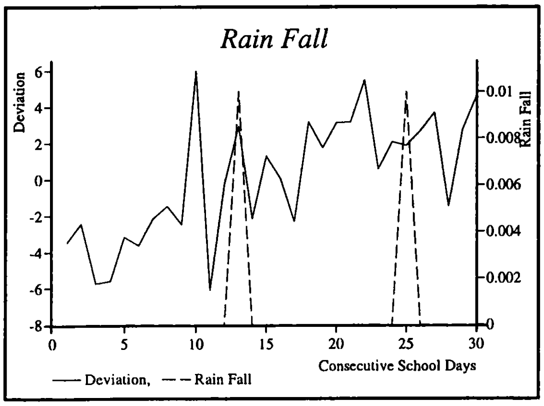

Ozone

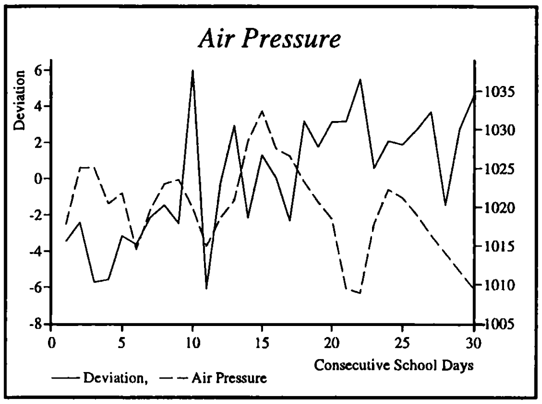

Dust

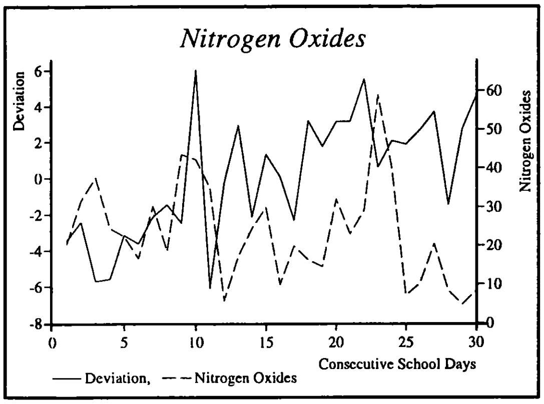
SWANSEA - SUMMER

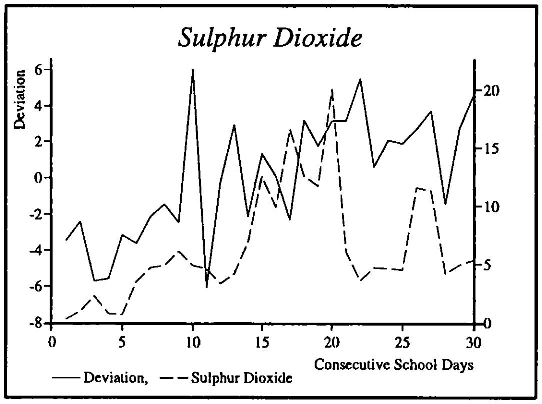


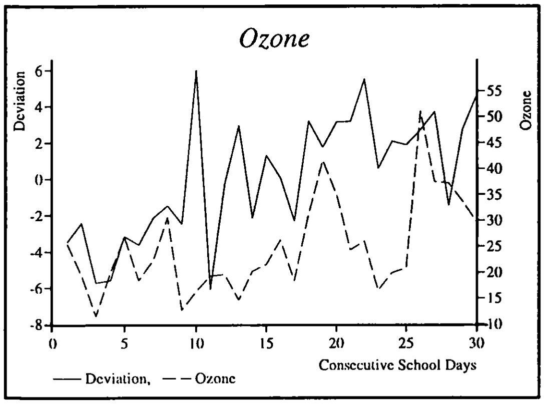


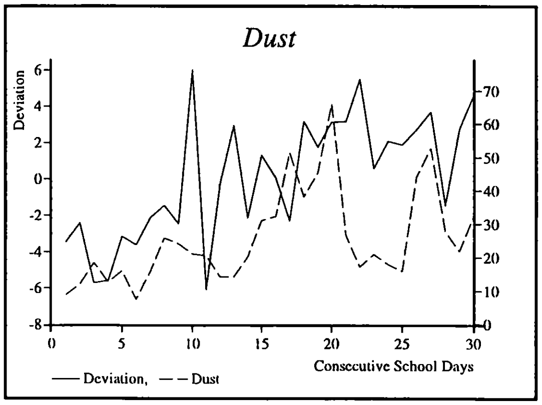


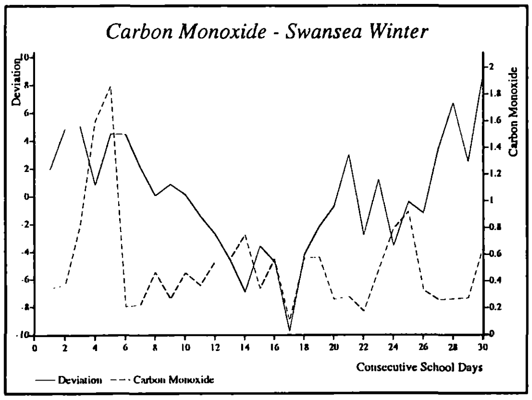


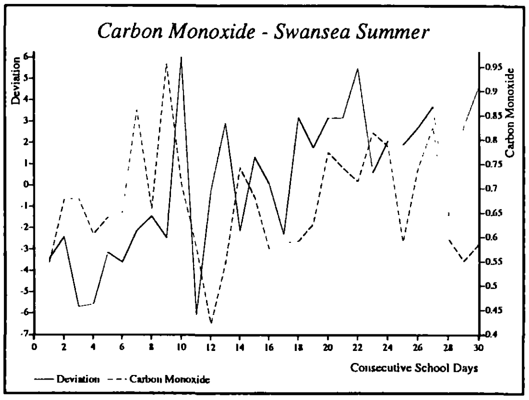


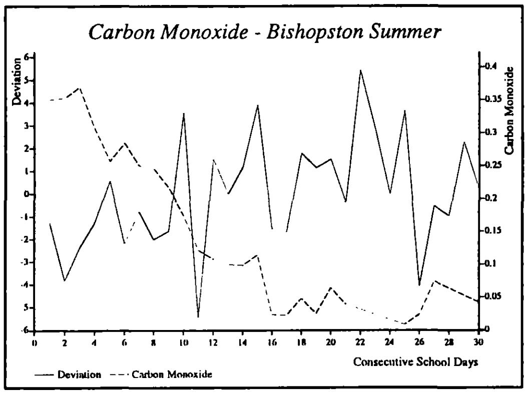




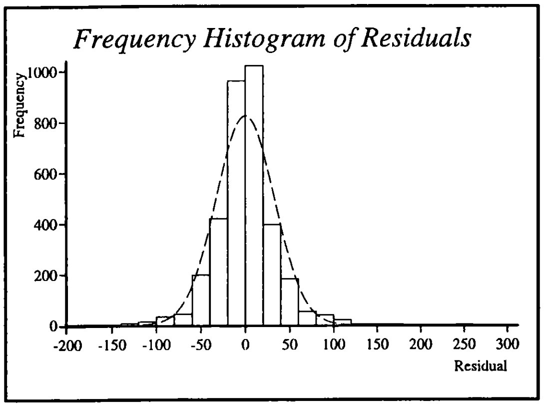


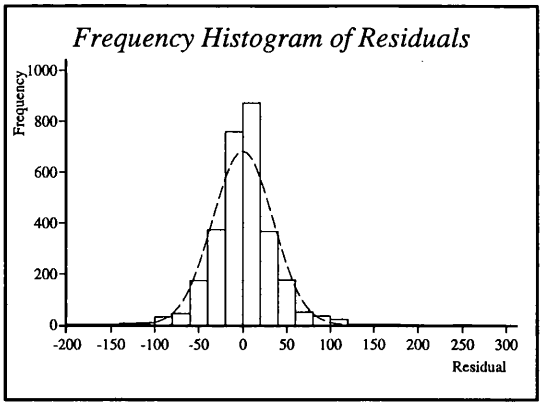

Carbon Monoxide - Swansea Winter

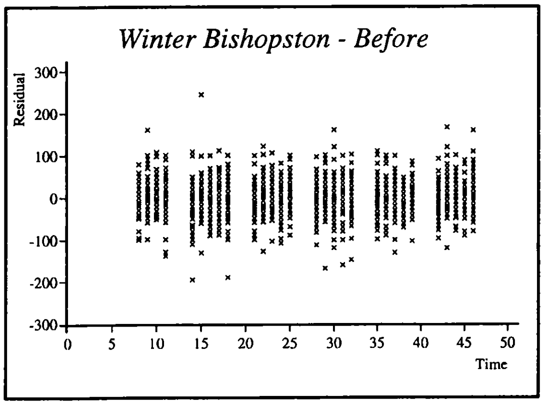

Carbon Monoxide - Swansea Summer

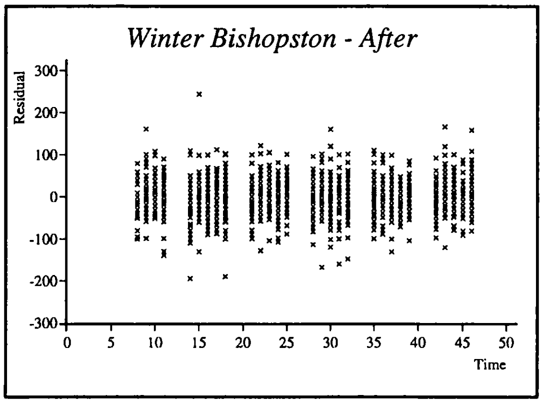

Carbon Monoxide - Bishopston Summer

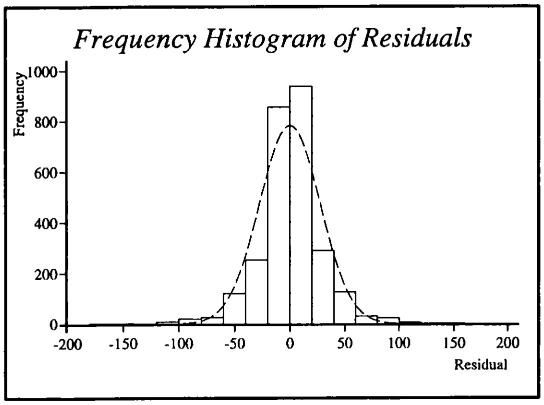
Appendix 2.6: Before and After the Removal of the Zero Residuals

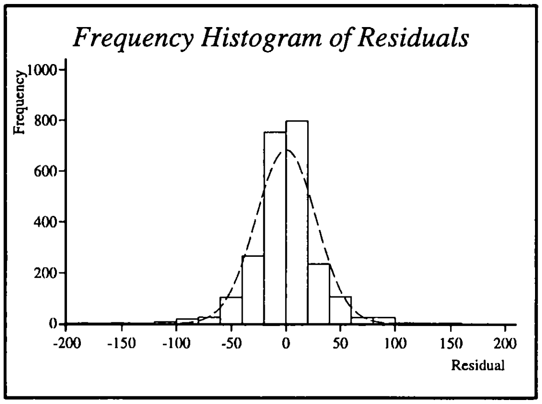

Frequency Histogram of Residuals

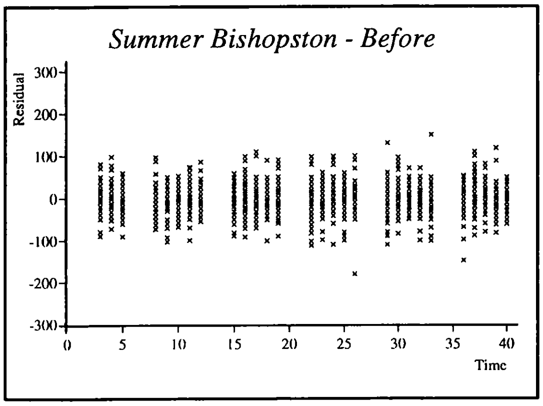

Frequency Histogram of Residuals

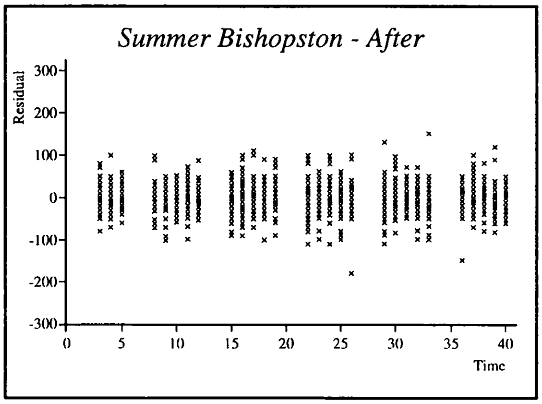

Winter Bishopston - Before

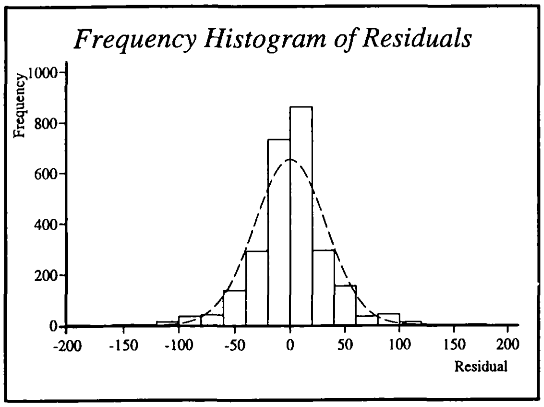

Winter Bishopston - After

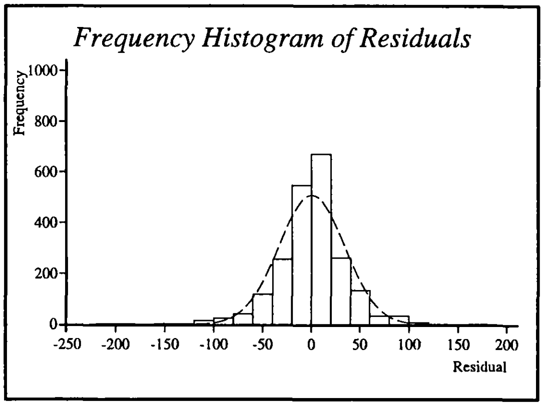

Frequency Histogram of Residuals

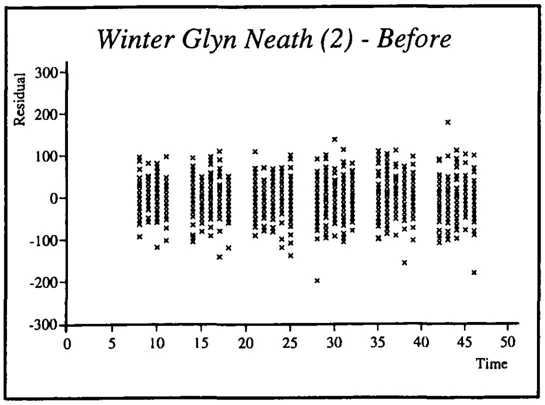

Frequency Histogram of Residuals

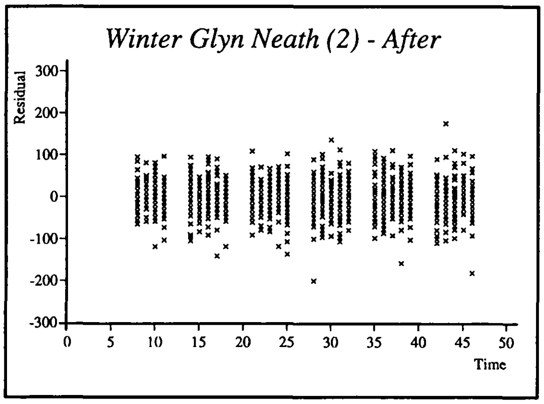

Summer Bishopston - Before

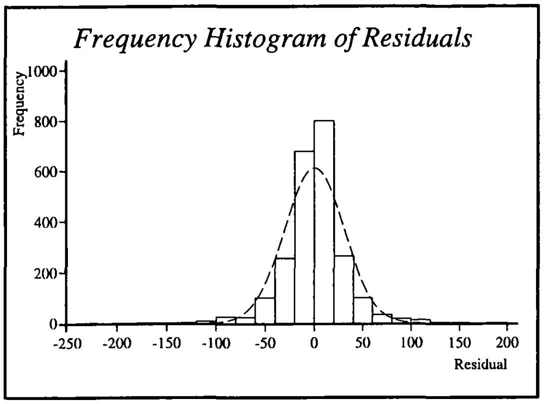

Summer Bishopston - After

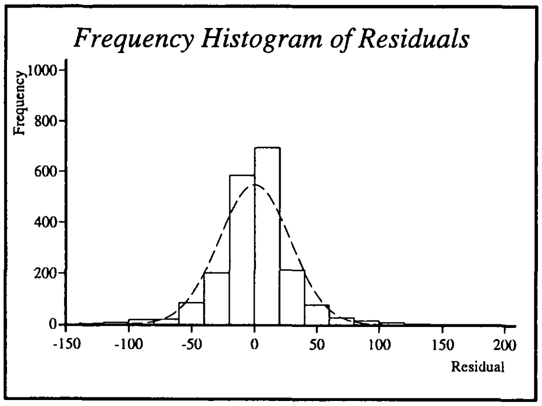

Frequency Histogram of Residuals

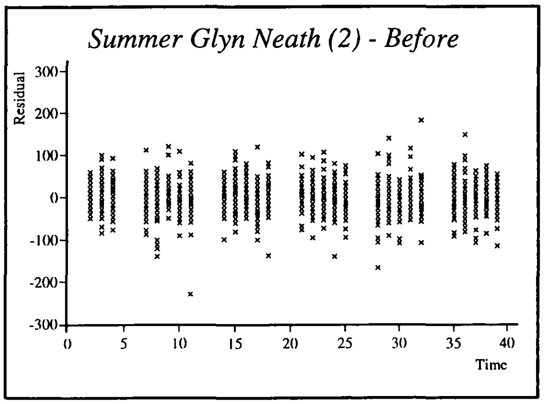

Frequency Histogram of Residuals

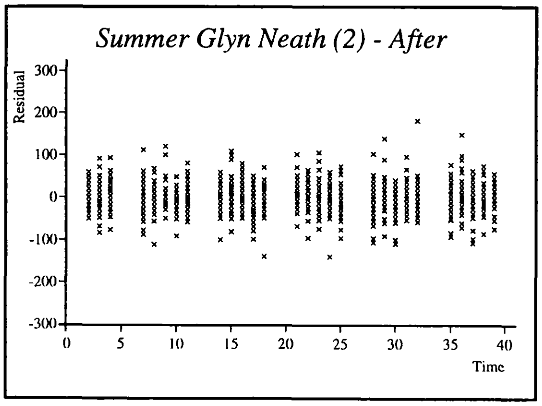

Winter Glyn Neath (2) - Before

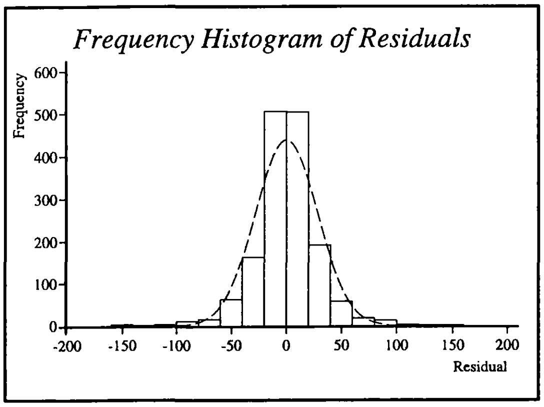

Winter Glyn Neath (2) - After

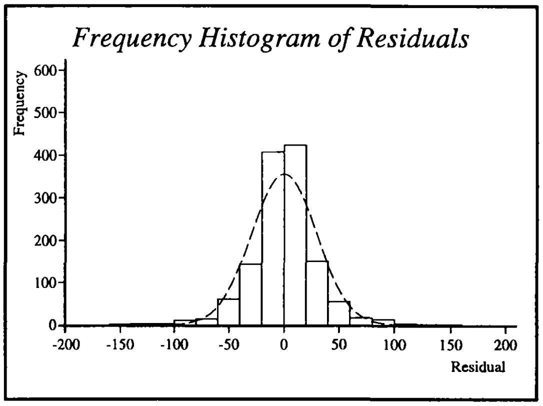

Frequency Histogram of Residuals

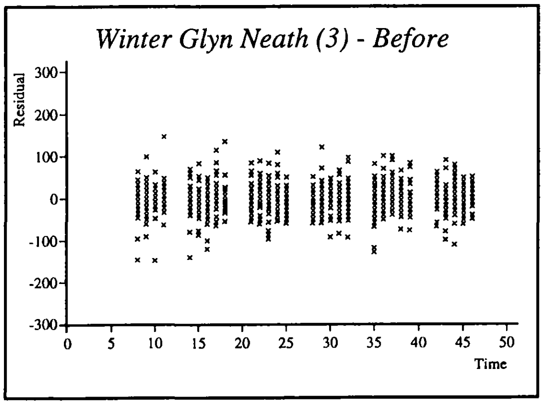

Frequency Histogram of Residuals

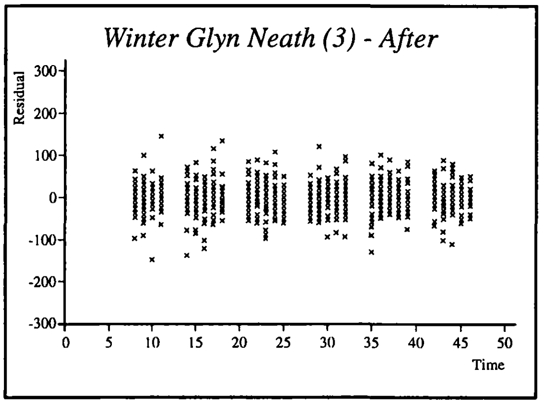

Summer Glyn Neath (2) - Before

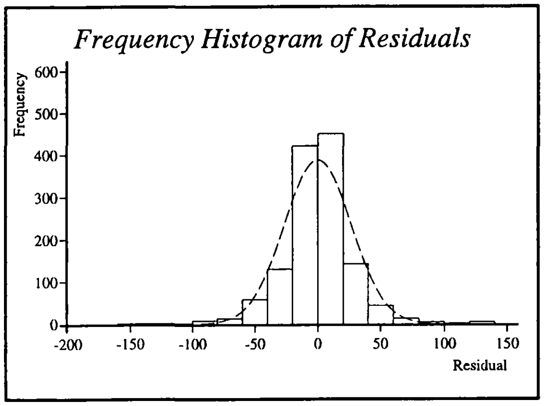

Summer Glyn Neath (2) - After

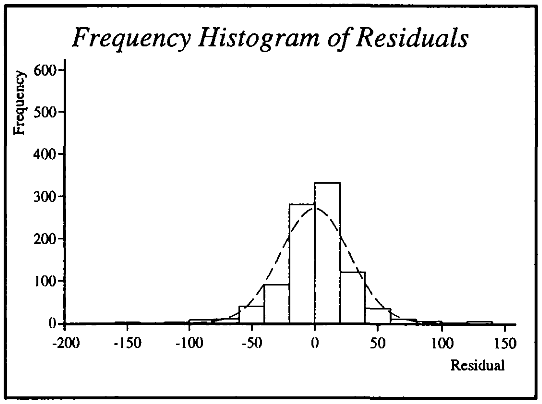

Frequency Histogram of Residuals

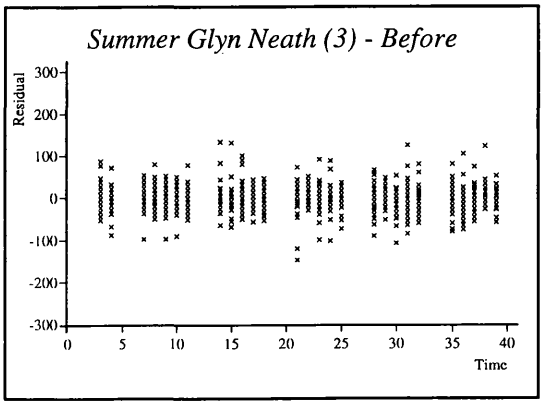

Frequency Histogram of Residuals

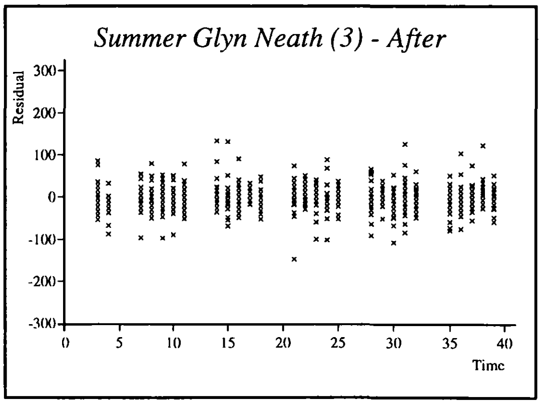

Winter Glyn Neath (3) - Before

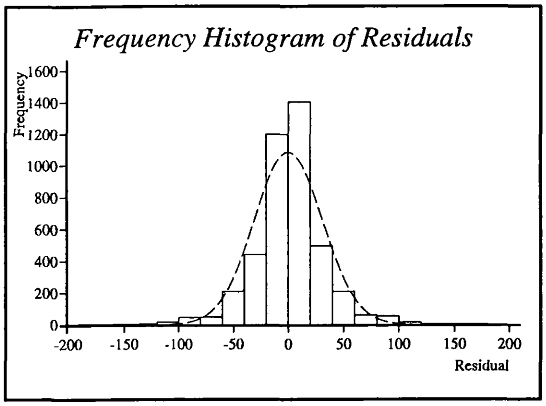

Winter Glyn Neath (3) - After

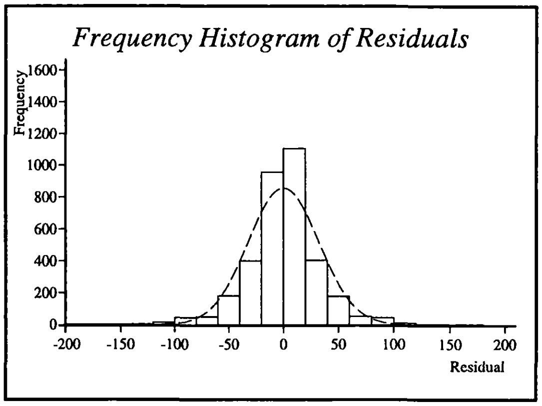

Frequency Histogram of Residuals

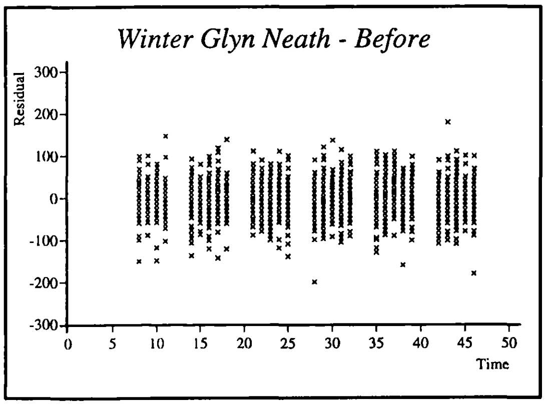

Frequency Histogram of Residuals

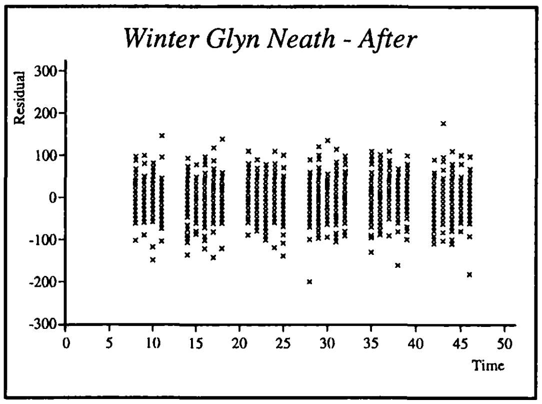

Summer Glyn Neath (3) - Before

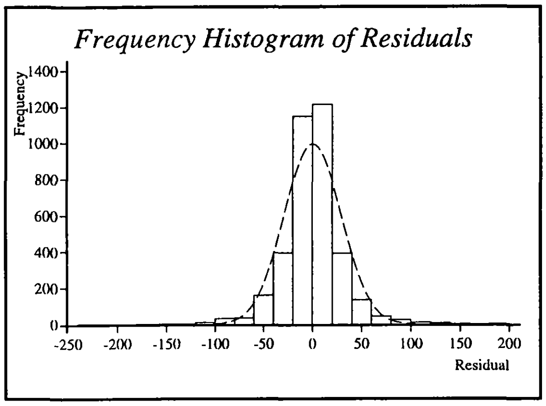

Summer Glyn Neath (3) - After

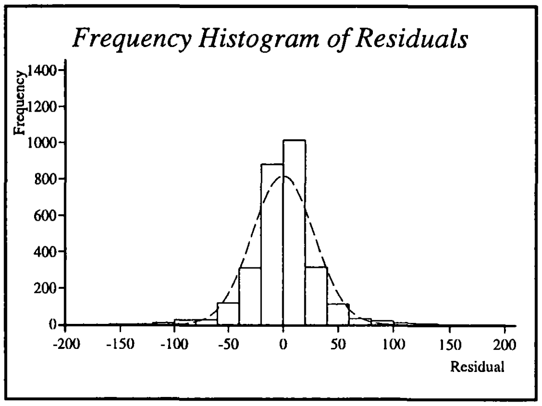

Frequency Histogram of Residuals

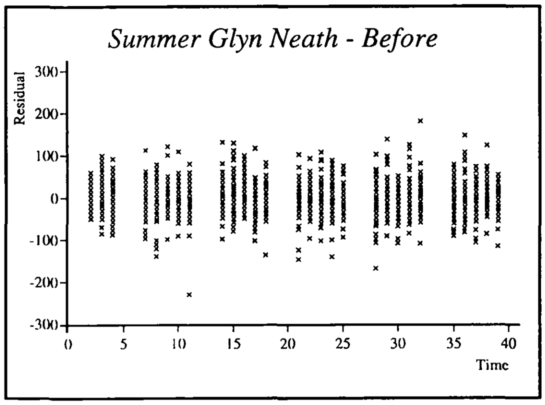

Frequency Histogram of Residuals

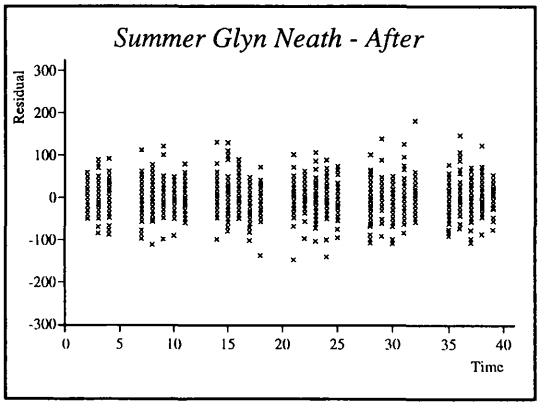

Winter Glyn Neath - Before

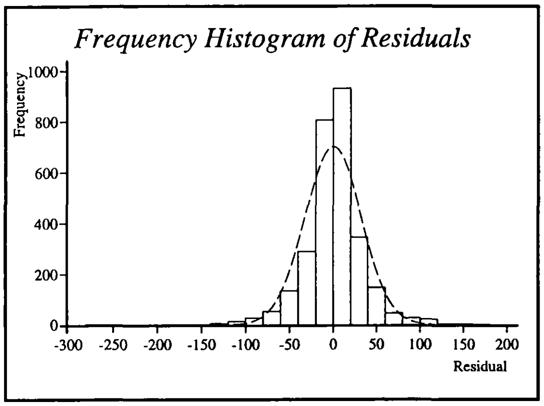

Winter Glyn Neath - After

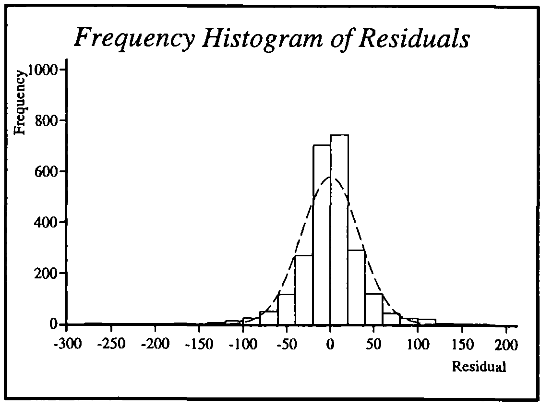

Frequency Histogram of Residuals

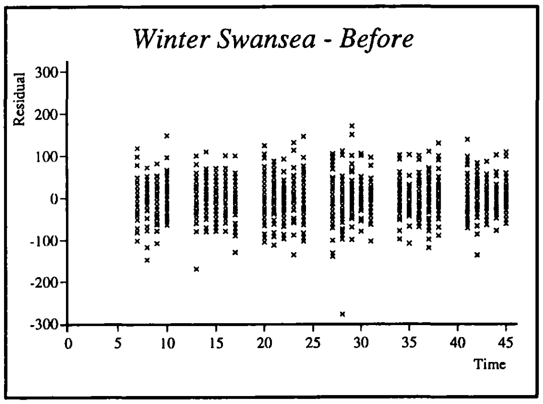

Frequency Histogram of Residuals

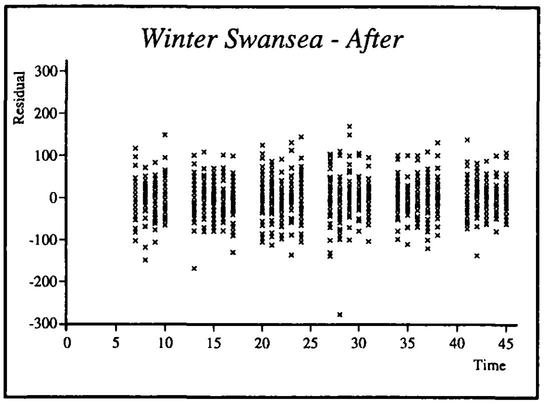

Summer Glyn Neath - Before

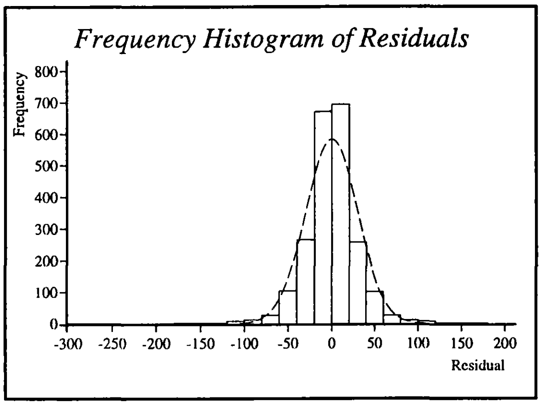

Summer Glyn Neath - After

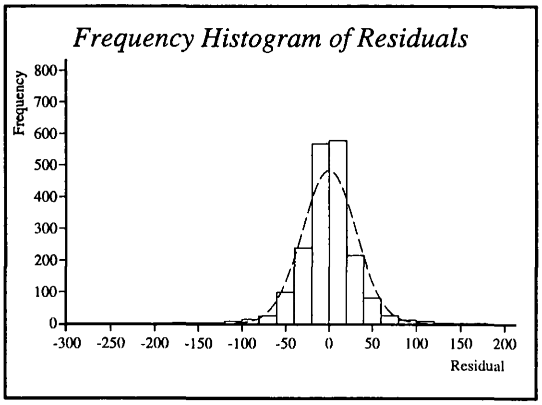

Frequency Histogram of Residuals

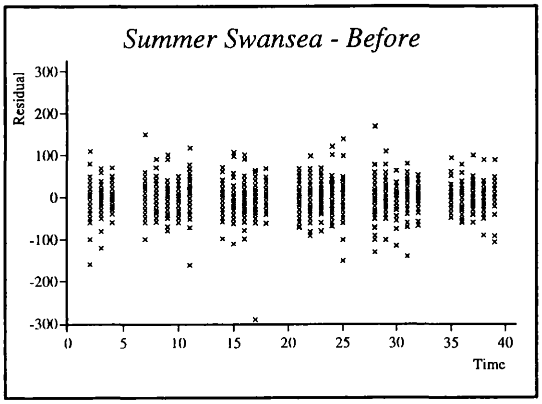

Frequency Histogram of Residuals

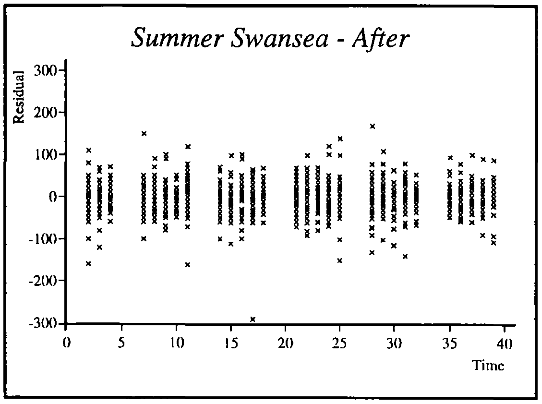

Winter Swansea - Before


Winter Swansea - After


Frequency Histogram of Residuals


Frequency Histogram of Residuals


Summer Swansea - Before


Summer Swansea - After

Bibliography

1. Tim Radford (8/5/95). Air of Mystery. *Guardian*, Page 2.

2. Celia Hall (7/3/96). Epidemic of lung disease is forecast. *Electronic Telegraph.*

3. *Air Health Strategy* (January 1996). Page 8.

4. *Cutting Edge* (10/10/95) - Fighting for breath. S4C, 10.30pm - 11.30pm.

5. Cassandra Marks (July 1995). Breathe easy with homeopathy. *Therapies in focus HH.* pp 31-33.

6. Kate Ardern (June 1996). Air Pollution and Health. *The atmospheric research and information centre.*

7. *Public Service and Local Government* (September 1995). Page 28.

8. *New Scientist* (4/11/95). Page 5.

9. *New Scientist* (18/11/95). Page 6.

10. Mick Hamer and Debora Mackenzie (18/11/95). Brussels blocks Britain's clean air plan. *New Scientist.*

11. JMF Temple and AM Sykes (15/08/92). Asthma and open cast mining. *British Medical Journal;* **Vol. 305**: pp 396-397.

12. JN Tsanakas, RDG Milner, OM Bannister, AW Boon (1988). Free running asthma screening test. *Arch Dis Child*; **63**: pp 261-265.

13. Burr ML, Eldridge BA, Borysiewicz LK (1974). Peak expiratory flow rates before and after exercise in school children. *Arch Dis Child*; **49**: pp 923-926.

14. S. Godfrey, P.L. Kamburoff, J.R. Nairn (1970). Spirometry, lung volumes and airway resistance in normal children aged 5 to 18 years. *British Journal of Dis. Chest*; **64**: pp 15-24.

15. S. Walters, M. Phupinyokul, J. Ayres (1995). Hospital admission rates for asthma and respiratory disease in the west midlands - their relationship to air pollution levels. *Thorax;* **Vol.50, No 9**: pp 948-954.

16. Joel Schwarz (1992). Air pollution and the duration of acute respiratory symptoms. *Archives of Environmental Health;* March/ April. **Vol. 47, No 2**: pp 116-122.

17. Ostro, Lipsett, Mann, Weiner, Selner (1994). Indoor air pollution and asthma - results from a panel study. *American Journal of Respiratory and Critical Care Medicine;* **Vol.149, No 6**: pp 1400-1406.

18. J .J. Stoddard, T. Miller (1995). Impact of parental smoking on the prevalence of wheezing respiratory illness in children. *American Journal of Epidemiology;* **Vol. 141, No 2**: pp 96-102.

19. The effect of maternal smoking during pregnancy on early infant lung function (1992). *American Review of Respiratory Disease;* **145**: pp 1129-1135.

20. A.J. Gray, J.K. Peat, C.M. Mellis, J. Harrington, A.J. Woolcock (1994). Asthma severity and morbidity in a population sample of sydney school children. *Australian and New Zealand Journal of medicine;* **Vol. 24, No.2**: pp 168-175.

21. First studies of indoor air pollution (1996). *Air Quality Issues;* July, **Vol. 1, Issue 6**: page 4.

22. Debora Jackson (11/10/95). Could one quick jab make your child gasp for breath? *Independent Section 2.*

23. B. Brabin, M. Smith, P. Milligan, C. Benjamin, E. Dunne, M. Pearson (1994). Respiratory morbidity in merseyside schoolchildren exposed to coal-dust and air-pollution. *Archives of disease in childhood;* **Vol. 70**, **No 4**: pp 305-312.

24. R.S. Bhopal, P. Phillimore, S. Moffatt, C. Foy (1994). Is living near a coking works harmful to health - a study of industrial air-pollution. *Journal of epidemiology and community health;* **Vol. 48**, **No 3**: pp 237-247.

25. E. Rebstockbourgkard, et al. (1994). Respiratory symptoms and pulmonary function in coal workers with abnormal chest x-rays. *Revue de epidemiologie et de sante publique;* **Vol. 42**, **No 6**: pp 533-541.

26. N. Chau et al. (1992). Lung function in retired coke oven plant workers. *British journal of industrial medicine;* **Vol. 49**, **No 5**: pp 316-325.

27. C. Infanterivard (1993). Childhood asthma and indoor environmental risk factors. *American journal of epidemiology;* **Vol. 137**, **No 8**: pp 834-844.

28. H.H. Dayal, S. Khuder, R. Sharrar, N. Trieff (1994). Passive smoking in obstructive respiratory diseases in an industrialized urban population. *Environmental research;* **Vol. 65**, **No 2**: pp 161-171.

29. National Society for clean air and environmental protection (August 1994). *Air pollution and human health.*

30. A brief guide to the key urban air pollutants (1996). *Atmospheric Research and Information Centre.*

31. National Asthma Campaign Factsheet.

32. No tolerance in Mexico City kids (January 1996). *Air Health Strategy*: Page 11.

33. Health effects of smog (1996). *CCME: Canadian Council of Ministers of the Environment.*

34. P.J. Barnes (1994). Air pollution and asthma. *Postgraduate medical journal;* **Vol. 70, No 823:** pp 319-325.

35. I.S. McCrae, I.D. Williams (1994). Road traffic pollution and public nuisance. *Science of the total environment;* **Vol. 147**: pp 81-91.

36. Budget '96 promotes use of cleaner fuels (January 1997). *Air Quality Issues;* **Vol. 2**.

37. Exhaust emissions study gives IMO food for thought (1995). *Mer-Marine engineers review*: pp 15-16.

38. Tim Brown (29/6/95). In a position to tackle air pollution. *Surveyor*.

39. Robert Bond (26/10/95). Asthma - air pollution link diminshed by research. *Surveyor*: Page 8.

40. Studies fail to link asthma symptoms with traffic pollution (January 1996). *Air Health Strategy*: Page 7.

41. *Air Quality Issues* (November 1996); **Vol. 2, Issue 2**.

42. A. Dusseldorp et al. (November 1995). Association of PM10 and air-borne iron with respiratory health of adults living near a steel factory. *American Journal Respiratory and Critical Care Medicine;* **152 (5)** pp 1932.

43. The APHEA project (April 1996). Short term effects of air pollution on health: a European approach using epidemiological time series data. *Journal of epidemiology and Community Health;* **Vol. 50**, **Supplement 1**, pp S1-S80.

44. L.J. Folinsbee (1992). Does nitrogen dioxide exposure increase airways responsiveness. *Toxicology and industrial health;* **Vol. 8, No 5**: pp 273-283.

45. L.Moseholm, E. Taudorf, A. Frosig (1993). Pulmonary-function changes in asthmatics associated with low level SO2 and NO2 air pollution, weather and medicine intake - an 8 month prospective study analyzed by neural networks. *Allergy;* **Vol. 48, No. 5**: pp 334-344.

46. O.V.J. Rossi, V.L. Kinnula, J. Tienari, E. Huhti (1993). Association of severe asthma attacks with weather, pollen and air polltants. *Thorax;* **Vol. 48, No 3**: pp 244-248.

47. L.M. Kieding, A.K. Rindel, D. Kronborg (1995). Respiratory illnesses in children and air pollution in Copenhagen. *Archives of environmental health;* **Vol. 50, No 3**: pp 200-206.

48. W.L. Eschrnbacher, T.B. Moore, T.J. Lorenzen, J.G. Weg, K.B. Gross (1992). Pulmonary responses of asthmatic and normal subjects to different temperature and humidity conditions in an environmental chamber. *Lung;* **Vol. 170, No 1**: pp 51-62.

49. W. Roemer, G. Hoek, B. Brunekreef (1993). Effect of ambient winter air pollution on respiratory health of children with chronic respiratory symptoms. *American review of respiratory disease;* **Vol. 147, No 1**: pp 118-124.

50. B.G. Higgins, H.C. Francis, C.J. Yates, C.J. Warburton, A.M. Fletcher, J.A. Reid, C.A.C. Pickering, A.A. Woodcock (1995). Effects of air pollution on symptoms and PEF measurements in subjects with obstructive airways disease. *Thorax;* **Vol. 50, No 2**: pp 149-155.

51. C. Rusznack, J.L. Devalia, R.J. Davies (1994). The impact of pollution on allergic disease. *Allergy;* **Vol. 49, No 18**: pp 21-27.

52. Krupnick, Harrington and Ostro (1990). Ambient ozone and acute health effects: Evidence from daily data. *Journal of Environmental Economics and Management;* **18**: pp 1-18.

53. C.A. Pope, R.E. Kanner (1993). Acute effects of PM10 pollution on pulmonary function of smokers with mild to moderate chronic obstructive pulmonary disease. *American review of respiratory disease;* **Vol. 147, No 6**: pp 1336-1340.

54. C. Arden Pope, III and Douglas W. Dockery (1992). Acute health effects of PM10 pollution on symptomatic and asymptomatic children. *American Review Respiratory Disease;* **145**: pp 1123-1128.

55. G. Hock, W. Roemer, B. Brunekreef (1995). THE PEACE STUDY: Pollution effects on asthmatic children in Europe. Acute effects of air pollution on peak flow. Department of Epidemiology and Public Health. Wageningen Agricultural University.

56. B. Forsberg, N. Stjernberg, M. Falk, B. Lundback, S. Wall (1993). Air pollution levels, meterological conditions and asthma symptoms. *European repiratory journal;* **Vol. 6, No 8**: pp 1109-1115.

57. E. Vonmutius, D.L. Sherrill, C. Fritzsch, F.D. Martinez, M.D. Lebowitz (1995). Air pollution and upper respiratory symptoms in children from East Germany. *European respiratory journal;* **Vol. 8, No 5**: pp 723-728.

58. A.S. Whittemore, E.L. Korn (July 1980). Asthma and air pollution in the Los Angeles Area. *American Journal of Public Health;* **Vol. 70, No 7**.

59. A.H. Holguin. The effects of ozone on asthmatics in the Houston area.

60. Daily diaries of respiratory symptoms and air pollution: Methodological issues and results (1991). *Environmental Health Perspectives;* **Vol. 90**: pp 181-187.

61. Acute effects of summer air pollution on respiratory reporting in children (1994). *American Journal Respiratory Crit. Care Med;* **Vol 150**: pp 1234-1242.

62. A. Goren, S. Hellmann, S. Brenner (1993). Respiratory problems among schoolchildren living in the vicinity of an industrialized zone. *Water science and technology;* **Vol. 27, No 7-8**: pp 339-344.

63. C. Braunfahrlander, U. Ackermannliebrich, J. Schwarz, H.P. Gnehm, M. Rutishauser, H.U. Wanner (1992). Air pollution and respiratory symptoms in preschool children. *American review of respiratory disease;* **Vol. 145, No 1**: pp 42-47.

64. S. Walters, R.K. Griffiths, J.G. Ayres (1994). Temporal association between hospital admissions for asthma in Birmingham and ambient levels of sulfur dioxide and smoke. *Thorax;* **Vol. 49, No 2**: pp 133-140.

65. R.Y.M. Tseng, C.K. Li, J.A. Spinks (1992). Particulate air pollution and hospitalization for asthma. *Annals of allergy;* **Vol. 68, No 5**: pp 425-432.

66. J. Schwarz (1995). Short term fluctuations in air pollution and hospital admissions of the elderly for respiratory disease. *Thorax;* **Vol. 50, No 5**: pp 531-538.

67. X.P. Xu, B.L. Li, H.Y. Huang (1995). Air pollution and unscheduled hospital outpatient and emergency room visits. *Environmental health perspectives;* **Vol. 103, No. 3**: pp 286-289.

68. X.P. Xu, J. Gao, D.W. Dockery, Y. Chen (1994). Air pollution and daily mortality in residential areas of Beijing, China. *Archives of environmental health;* **Vol. 49, No 4**: pp 216-222.

69. G. Touloumi, S.J. Pocock, K. Katsouyanni, D. Trichopoulos (1994). Short term effects of air pollution on daily mortality in Athens - A time series analysis. *International journal of epidemiology;* **Vol. 23, No 5**: pp 957-967.

70. J. Schwarz, D.W. Dockery (1992). Particulate air pollution and daily mortality in Steubenville, Ohio. *American journal of epidemiology;* **Vol. 145, No 5**: pp 1123-1128.

71. K. Katsouyanni, et al (1995). Short term effects of air pollution on health - a European approach using epidemiologic time series data - the APHEA project - background, objectives, design. *European respiratory journal;* **Vol. 8, No 6**: pp 1030-1038.

72. B.D. Ostro, M.J. Lipsett, J.K. Mann, H. Braxtonowens, M.C. White (1995). Air pollution and asthma exacerbations among african-american children in L.A. *Inhalation technology;* **Vol. 7, No 5**: pp 711-722.

73. M. Wust, P. Reitmeir, S. Dold, A. Wulff, T. Nicolai, E.F. Vonloeffelholzcolberg, E. Vonmutius (1993). Road traffic and adverse effects on respiratory health in children. *British medical journal;* **Vol. 307, No 6904**: pp 596-600.

74. J.R. Greer, D.E. Abbey, R.J. Burchette (1993). Asthma related to occupational and ambient air pollutants in non-smokers. *Journal of occupational medicine;* **Vol. 35, No 9**: pp 909-915.

75. National Air Quality Strategy (May 1997). *Air Quality Issues;* **Vol. 2, Issue 5**: Page 1.

76. Acid Rain and Smog (May 1997). *Air Quality Issues;* **Vol. 2, Issue 5**.

77. Particulate studies in Mexico (November 1996). *Air Quality Issues;* **Vol. 2, Issue 2**: Page 3.