

**Atmospheric Dispersion Modelling
Liaison Committee**

Annual Report 2000/2001

INCLUDING

**Options for the Most Appropriate Meteorological
Data for Use in Short Range Dispersion Modelling**

AND

Methods for Undertaking Uncertainty Analyses

This study was funded by the Atmospheric Dispersion Modelling Liaison Committee.

Annex A © The Meteorological Office
Annex B © See the annex

© National Radiological Protection Board
Chilton
Didcot
Oxon OX11 0RQ

Approval: January 2002
Publication: February 2002
£15.00
ISBN 0 85951 471 4

PREFACE

In 1977 a meeting of representatives of government departments, utilities and research organisations was held to discuss methods of calculation of atmospheric dispersion for radioactive releases. Those present agreed on the need for a review of recent developments in atmospheric dispersion modelling, and a Working Group was formed. Those present at the meeting formed an informal Steering Committee, that subsequently became the UK Atmospheric Dispersion Modelling Liaison Committee. That Committee operated for a number of years. Members of the Working Group worked voluntarily and produced a series of reports. A workshop on dispersion at low wind speeds was also held, but its proceedings were never published.

The Committee has been reorganised and has adopted terms of reference. The organisations represented on the Committee, and the terms of reference adopted, are given in this report. The organisations represented on the Committee pay a small annual subscription. The money thus raised is used to fund reviews on topics agreed by the Committee, and to support in part its secretariat, provided by NRPB. The new arrangements came into place for the start of the 1995/96 financial year. This report describes the sixth year in which the Committee has operated under the new arrangements, and during which it placed one contract, for a review of the most appropriate source of data for short range modelling calculations. The technical specification for this contract is given in this report, and the contract report is attached as an annex to this report. The Committee also organised a set of presentations on techniques for uncertainty analyses; the speakers overheads are given in an annex to this report. The Committee funded ten studies in previous years; they are described in its earlier annual reports.

The Committee intends to place further contracts in future years and would like to hear from those interested in tendering for such contracts. They should contact the Secretary:

Mr J G Smith
National Radiological Protection Board
Chilton
Didcot
Oxon OX11 0RQ

E-mail: justin.smith@nrpb.org.uk

CONTENTS

Atmospheric Dispersion Modelling Liaison Committee		1
1	Organisations represented on the Committee	1
2	Terms of reference	2
3	Reports of the Committee and its earlier Working Group on Atmospheric Dispersion	2
4	Specifications for technical annexes	4
	A Options for the most appropriate meteorological data for use in short range dispersion modelling	4
	B Methods for undertaking uncertainty analyses	5
ANNEX A	Options for the Most Appropriate Meteorological Data for Use in Short Range Dispersion Modelling	7
ANNEX B	Methods for Undertaking Uncertainty Analyses	85

ATMOSPHERIC DISPERSION MODELLING LIAISON COMMITTEE

I ORGANISATIONS REPRESENTED ON THE COMMITTEE

Atomic Weapons Establishment, Aldermaston

British Energy

British Nuclear Fuels plc

BNFL Magnox Generation

Department of the Environment Northern Ireland

Environment Agency

Health and Safety Executive

 Hazardous Installations Directorate

 Nuclear Installations Inspectorate

Ministry of Agriculture, Fisheries and Food

Meteorological Office

National Nuclear Corporation

National Radiological Protection Board

Nycomed Amersham plc

Royal Naval College, Greenwich

Rolls Royce Power Engineering plc

Scottish Environment Protection Agency

Urenco (Capenhurst)

Westlakes Research Institute

The Chairman and Secretary are provided by NRPB.

2 TERMS OF REFERENCE

1 To review current understanding of atmospheric dispersion and related phenomena and to identify suitable models for application primarily in authorisation or licensing, in the context of discharges to atmosphere resulting from nuclear industry activities.

2 The Committee shall consist of representatives of government departments, government agencies and primarily the nuclear industry. Each organisation represented on the Committee shall pay an annual membership fee of £1000.

3 The Committee will consider selected topics. These should be selected following discussion and provisional agreement at meetings of the Committee, followed by confirmation after the meeting. Where possible, it will produce reports describing suitable models for that topic. These will reflect either the views of an Expert Working Group appointed by the Committee or the outcome of a workshop organised on behalf of the Committee. The Working Group will determine who should be invited to speak at workshops, and will subsequently review their outcome and identify suitable models.

4 The money raised from membership fees and registration fees for the workshops will be used to support the Working Group, the drafting of reports, and any other matters which the Committee may decide.

3 REPORTS OF THE COMMITTEE AND ITS EARLIER WORKING GROUPS ON ATMOSPHERIC DISPERSION

Clarke, R H (1979). The first report of a Working Group on Atmospheric Dispersion: a model for short and medium range dispersion of radionuclides released to the atmosphere. Harwell, NRPB-R91

Jones, J A (1981). The second report of a Working Group on Atmospheric Dispersion: a procedure to include deposition in the model for short and medium range dispersion of radionuclides. Chilton, NRPB-R122

Jones, J A (1981). The third report of a Working Group on Atmospheric Dispersion: the estimation of long range dispersion and deposition of continuous releases of radionuclides to atmosphere. Chilton, NRPB-R123

Jones, J A (1981). The fourth report of a Working Group on Atmospheric Dispersion: a model for long range atmospheric dispersion of radionuclides released over a short period. Chilton, NRPB-R124

Jones, J A (1983). The fifth report of a Working Group on Atmospheric Dispersion: models to allow for the effects of coastal sites, plume rise and

buildings on dispersion of radionuclides and guidance on the value of deposition velocity and washout coefficients. Chilton, NRPB-R157

Jones, J A (1986). The sixth report of a Working Group on Atmospheric Dispersion: modelling wet deposition from a short release. Chilton, NRPB-R198

Jones, J A (1986). The seventh report of a Working Group on Atmospheric Dispersion: the uncertainty in dispersion estimates obtained from the Working Group models. Chilton, NRPB-R199

Atmospheric Dispersion Modelling Liaison Committee. Annual Report 1995/96. Chilton, NRPB-R292

Includes annexes

- Atmospheric Dispersion at Low Wind Speed
- Application of Computational Fluid Dynamics Codes to Near-field Atmospheric Dispersion
- Rise of a Buoyant Plume from a Building Wake

Atmospheric Dispersion Modelling Liaison Committee. Annual Report 1996/97. Chilton, NRPB-R302

Includes annexes

- Atmospheric Dispersion at Low Wind Speed
- Review of Models for Calculating Air Concentrations when Plumes Impinge on Buildings or the Ground

Atmospheric Dispersion Modelling Liaison Committee. Annual Report 1997/98. Chilton, NRPB-R316

Includes annex

- Portability of Weather Data for Dispersion Calculations

Atmospheric Dispersion Modelling Liaison Committee. Annual Report 1998/98. Chilton, NRPB-R322

Includes annexes

- Review of Deposition Velocity and Washout Coefficient
- Review of Flow and Dispersion in the Vicinity of Groups of Buildings

Atmospheric Dispersion Modelling Liaison Committee. Annual Report 1999/2000. Chilton, NRPB-W2

Includes annexes

- Review of Dispersion of Bodies of Water
- Best Practice for Binning Meteorological Data

4 SPECIFICATIONS FOR TECHNICAL ANNEXES

A Options for the most appropriate meteorological data for use in short range dispersion modelling

This proposal is a result of the Committee's discussion at the ADMLC meeting held at NRPB on the 2 June 2000. These suggested number of areas which the Committee might wish to fund, all falling within the broad topic of the options available when choosing the most appropriate meteorological data for use in short-range dispersion models such as ADMS (see www.CERC.co.uk).

This topic is especially of concern in data-sparse areas of the UK, and indeed the globe, where recent data from a suitably equipped observing station is not always available in reasonable proximity of the study area.

This topic is by its very nature wide-ranging and as such will include qualitative comment accompanied by limited modelling to support the conclusions reached. It was apparent from the meeting that many of the Committee members were unaware of some of the options available and for this reason this proposed project would aim to provide a useful reference for the selection of meteorological data.

The main options to address are,

The use of less recent datasets which are close to the modelling site compared to datasets which although more recent are more distant from the study area;

The use of high spatial resolution data derived from analysis fields of the Meteorological Office's Numerical Weather Prediction (NWP) model;

The use of further refinement to the NWP data by application of the Site Specific Forecast Model (SSFM) which accounts for detailed topography and terrain.

Work Packages

This project aims to tackle 3 main areas in what is only a moderate time and as such the detail must be limited:

- 1 There is often a choice as to which meteorological dataset is more appropriate, one which is less recent but closer to the study area or a more recent one that is more distant. This work will investigate the 'lifetime' of a dataset by considering London Heathrow as an example station, for which data exist back until 1949. ADMS will be run for the years '50, '55, '65, '70, '80, '90, '93, '96 and '99 (or a similar set of years dependent on precise availability) and 'headline' statistics generated for each year and compared to investigate any trend. This will be for a small industrial emission scenario. A short qualitative discussion of the impact of observing practices and urbanisation will also be included.

2 The use of NWP data will be considered by comparing ADMS output, both longterm averages and percentile values, for the same scenarios as used in the recent study concerning the binning of statistical data. These runs will be repeated for 2 sites using 3 types of the met input file ("traditional" observed data. NWP-derived data extracted to look like the observed data files, and NWP-derived data where secondary variables, such as boundary layer height, are included where available from the NWP source). The output will be compared and differences discussed, as will any other consideration for the use of NWP data.

3 A brief investigation will be made of meteorological data derived from the Site Specific Forecast Model (SSFM). A description of the SSFM model will be presented and the results of an annual ADMS run compared for each of the 3 scenarios using SSFM data and observed data from a site would have been used in the absence of the SSFM meteorological data.

4 All components will be presented in a single report and a presentation given to ADMLC following completion.

B Methods for undertaking uncertainty analyses

ADMLC expressed an interest in methods for undertaking uncertainty analyses. The Committee did not place a contract to review this topic, but arranged for a series of presentations during one of the Committee's normal meetings. The overheads used in the presentations are presented in Annex B.

ANNEX A

OPTIONS FOR THE MOST APPROPRIATE METEOROLOGICAL DATA FOR USE IN SHORT RANGE DISPERSION MODELLING

J Smith
THE METEOROLOGICAL OFFICE

A1 Introduction

The aim of this project is to investigate the options that are available when choosing the most appropriate meteorological dataset for use in short-range dispersion models, such as ADMS (Carruthers et al., 2000) or AERMOD (Cimorelli et al., 1998). This topic is especially of concern in data-sparse regions of the UK and indeed anywhere else in the world where recent data from an observing station is not always available in reasonable proximity to the study area. Indeed, several studies (Davies and Thompson 1997) have investigated the sensitivity of dispersion predictions to the meteorological data used.

With only a limited number of observational Met Office sites currently open, finding the most representative site can on occasions be a task requiring much thought. Often the site nearest to the area of study is no longer open and although old observations are available there is a question of how applicable these are today. Indeed many consultants purchasing data are often not interested in less recent data because of the issue of climate change. However, it is sometimes the case that there is a site some distance from the area of study that is open, but this raises the question of representivity again (Hough and Nelson 2000). In addition to this, a number of stations have become fully automated, and this presents a problem as ADMS requires manual cloud observations and cannot use those recorded by a cloud base recorder (CBR).

The main aim of the first part of the project is to investigate whether older observations can be used in dispersion modelling, without detriment to the output. In short, whether a meteorological dataset has a lifetime beyond which it produces inaccuracies and should therefore not be used. This is an important area of study as the change in the data licensing terms of the Met Office has meant that modellers can now use the datasets they purchase for any project rather than just one. The situation may then arise where a modeller wishes to use a dataset purchased some time ago for a recent project but is unsure whether that dataset will represent the current meteorology well.

The following section investigates the use of data from the Met Office's Numerical Weather Prediction (NWP) model in preference to observed meteorological data. This is probably the best option when trying to find data for a site in a data sparse region, where perhaps the nearest observational station is too far away to be representative. The objective being to investigate if NWP is a viable and good option by comparing the output from ADMS for two sites in the UK using NWP and observed data.

The last part of this study takes the previous section further by examining the use of data from the Site Specific Model (SSM) for one site, and comparing it with observational data for the same site. The aim of this is to investigate if NWP data that has been refined to be more site specific offers increased representivity when no observational data is available.

It is hoped that this project will make modellers aware of the number of options open to them when selecting meteorological data, and serve as a useful reference for them.

A2 Study of historical data sets

Heathrow Airport (51.5 N, -0.45 W) has had Met Office staff observing the weather there since 1947 and can therefore provide over 50 years worth of meteorological data. To investigate the issue of a dataset having a 'lifetime' ADMS3 was run for the years '50, '55, '60, '65, '70, '75, '80, '85, '90, '93, '96 and '99 using data from Heathrow.

The model runs were carried out using a small industrial scenario (see Table A1 for details). An inert tracer was used as the pollutant with an emission rate of 1000 grams per second. The concentration output was as 15 minute averages.

TABLE A1 Source characteristics

Stack height (m)	Stack diameter (m)	Emission velocity (ms^{-1})	Temperature of emission ($^{\circ}\text{C}$)
10	1	15	150

The 99.9th percentile was studied as this is the objective for sulphur dioxide emissions in the National Air Quality Strategy for the year 2000 and this will capture the extreme events. In addition the long term average concentration was looked at, and both the spatial maxima and dispersion patterns were examined for this and the 99.9th percentile.

The size of the modelling grid was 50 by 50 kilometres with 31 grid points spaced regularly and the stack was located at the centre. No topographical features or buildings were featured in the runs to keep the scenario basic with less influencing factors.

This part of the project discusses the impact of both observing practices and urbanisation on the parameters recorded, as these factors are largely responsible for determining the 'lifetime' of a dataset.

A2.1 Results

A2.1.1 Wind roses

As wind speed and direction are largely responsible for the concentration contours that are produced by ADMS 3, the wind roses (see Appendix A) were examined to make a brief study of how these two parameters have changed since 1950. The wind roses show the number of occurrences of winds from different directions, banded into 30⁰ sectors. Due to the limitations of the software used in generating these plots, the scale for the number of occurrences

(increasing in the vertical on the 0° line) cannot be altered making some comparisons difficult.

By comparing the shape of each wind rose some interesting differences between years can be seen. In 1950 and 1955 the main wind direction is from the west, after 1955 the winds are predominantly from the southwest. There are more low winds, of speeds between 0 and 1.5 ms⁻¹, recorded from 1985 to 1999 than previously. 1965 and 1970 appear to have been rather windy years with higher instances of winds above 8.2 ms⁻¹ (16 knots) than other years. It is also noticeable that in 1975 and the years following, there are more winds in the 3.1 to 5.1 ms⁻¹ category. 1996 also stands out with a higher percentage of winds from the northeast than any of the other wind roses. Generally, Heathrow does not often experience winds from the southeast, but this is a common feature of the weather in the UK.

A2.1.2 Long term concentrations

When the ADMS output for the long-term average concentration was plotted it was seen that the area covered by the inert tracer was approximately the same for each year. This is largely to be expected considering the same site characteristics are used in each model run.

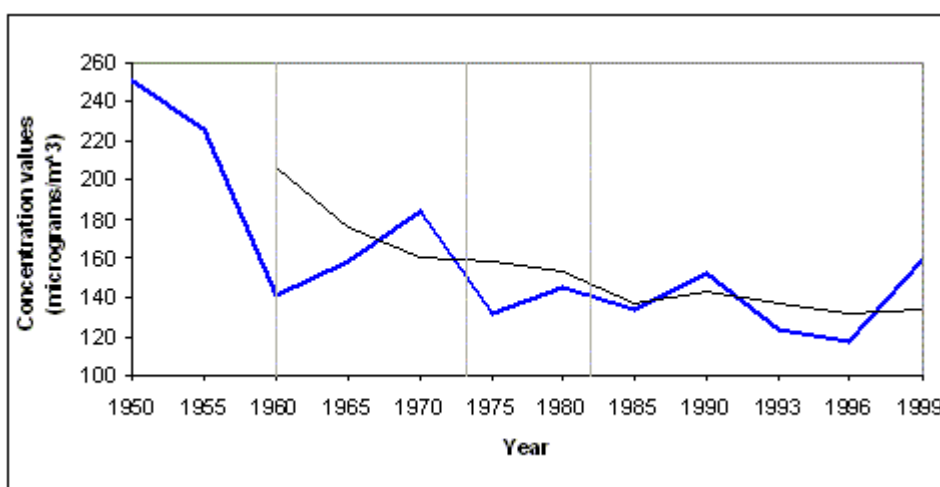


FIGURE A1 Graph to show how the spatial maximum of the long term concentration value has changed from 1950 (blue line) to 1999, and the 10 year rolling average (black line). The grey lines indicate the years when the position of the anemometer was changed

The most notable feature was the maximum concentration value shown in Figure A1. From 1950 to 1985 there was a gradual decrease in the size of the maximum concentration value, and after 1985 this decrease slowed down and there was even a slight increase from 1996 to 1999. This can be seen more clearly from the trend line drawn on figure A1. Figure A1 also illustrates some inter-annual variability with 1950, 1955, 1970, 1990 and 1999 producing noticeable peaks. These are probably caused by the weather pattern for that year being different to

the other years in some way and will be discussed fully later. (The reader should note that all concentration values have the units of $\mu\text{g}/\text{m}^3$.)

The position of the peaks in concentration also varied between years (Appendix A). For the years 1950 to 1970 the maximum value is to be found to the east of the stack, which is what is expected as the wind is predominantly from the west in those years. After 1975, the location of the maximum value in concentration was located on the northeast side of the stack. Again this can be expected with a mainly southwesterly wind in those years. In some years (1960, 1965, 1985, 1993 and 1996) there was seen to be a small maximum to the south or southwest of the stack. This corresponds to a large percentage of winds from the north to northeast in these years.

A2.1.3 99.9th percentile values

The contour plots for the 99.9th percentile values showed very little variation in the general pattern produced by the pollutant for each year studied (Appendix A). The location of the maximum values varied from year to year but was commonly to the east or north of the stack. Directly around the stack concentration values decreased readily. Figure A2 below shows that during the period of study the maximum values have steadily decreased from 1950, and that the data for 1955 produced a noticeable peak. As in Figure A1 for the long-term concentration values, a slight rise towards the end of 1999 is seen.

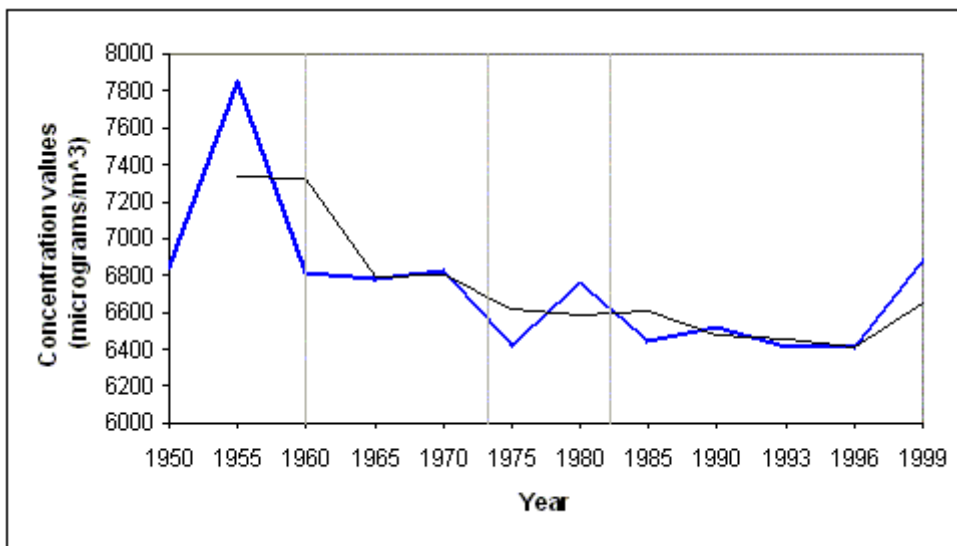


FIGURE A2 Graph to show how the maximum concentration value for the 99.9th percentile has changed from 1950 (blue line) and the trend for the period 1950 to 1999 (black line). The grey lines indicate the years when the position of the anemometer was changed

A2.2 Discussion of results

The differences described in 2.1.2 and 2.1.3 are likely to be due to changes in the weather data recorded each year rather than due to climate change. It is apparent from the individual wind roses that each year is slightly different. These differences are likely to be due to changes in observing practices and the effects of increasing urbanisation in the Heathrow area. Using a five year statistical dataset would not make these differences so apparent.

It would appear that much of the variation shown in the wind roses is largely due to the changes in observing practices since Heathrow opened. In 1973 the Met Office installed the "Mk5" wind system at the airport. This is basically a Munro Mk4 wind system with a telemetry system, which passes data across the airfield to the Met Office building, and was used extensively throughout the Met Office until recently. The anemometer that formed this "Mk5" system needs quite a high starting torque, typically 2.5 ms^{-1} or more to get the cups rotating. This may be responsible for the increased number of winds in the 0 to 1.5 ms^{-1} range from 1975 onwards, more so after 1985. Indeed, this wind system will be replaced in the near future by the "Mk6" wind system as it was found that in general the anemometer in the "Mk5" wind system does not meet the Met Office requirement for low speed wind measurement. (Unfortunately no information regarding the type of anemometer/s used at Heathrow before 1973 could be found in time to be included into this report.) Another point of note is that the anemometer position at Heathrow has moved 3 times since observations began. It was moved once in 1960, again in 1973 and in 1982 it was moved for the last time to date. Although the anemometer has been moved about, the distances moved are not significant enough to have had a significant impact on the wind speed and direction that was recorded.

As mentioned above, the wind roses show that the winds have become slightly lighter in the period from 1950 to 1999, and it is therefore reasonable to assume that there will have been more instances when the wind was calm towards the end of the study period. Indeed, this may be the reason for the overall decrease in maximum concentration values shown in Figure A1. As ADMS does not use the lines in the meteorological data set which have calm winds in its calculations, it will be missing the instances when concentration values may have been at their highest. To investigate how many lines of meteorological data were ignored, the number of calm winds for each year was counted and a graph drawn (see Figure A3 below).

It is shown on Figure A3 that the number of calm winds has risen over all since 1990 although very few were recorded in the years 1960 to 1990. There were a high number of calms recorded in 1950 and 1955 however, and this gives the graph a rather unusual shape. It is thought that maybe the winds recorded for these years produced different results because they were spot winds (averaged over the 10 minutes prior to the observation) and not mean winds (averaged over the previous hour). However, when the winds for 1960 and 1965 were replaced with spot wind values the results were almost identical to those produced by the mean winds, so it seems likely that using spot winds would

have not produced the high instances of calm winds in 1950 and '55. It might be the case that the peak may be due to the type of anemometer used at Heathrow then.

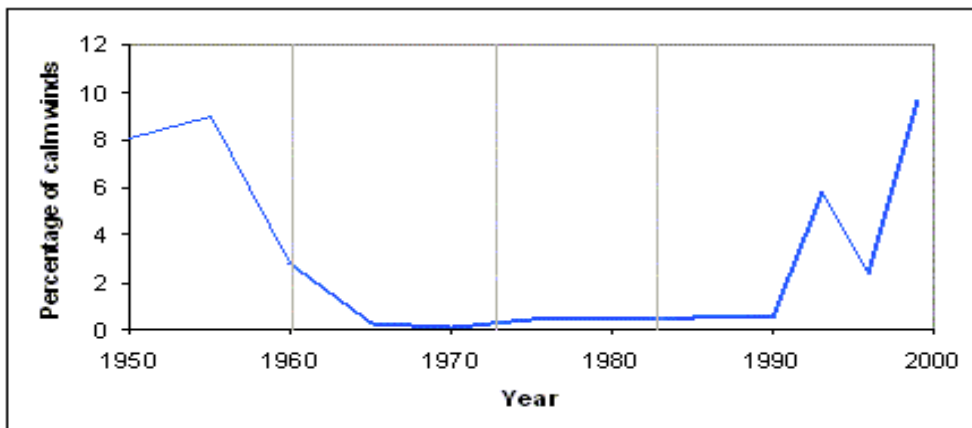


FIGURE A3 Graph to show the number of calm winds for each year studied as a percentage of total wind observations. Again, the grey lines show when the anemometer was moved

As the lines in the data set with calms will have been skipped, ADMS is not including the instances when concentrations may have been at their highest. Therefore a graph of wind speeds less than 2 ms⁻¹, excluding the calms, was plotted, as these will contribute significantly to the maxima in concentration. (See Figure A4 below.)

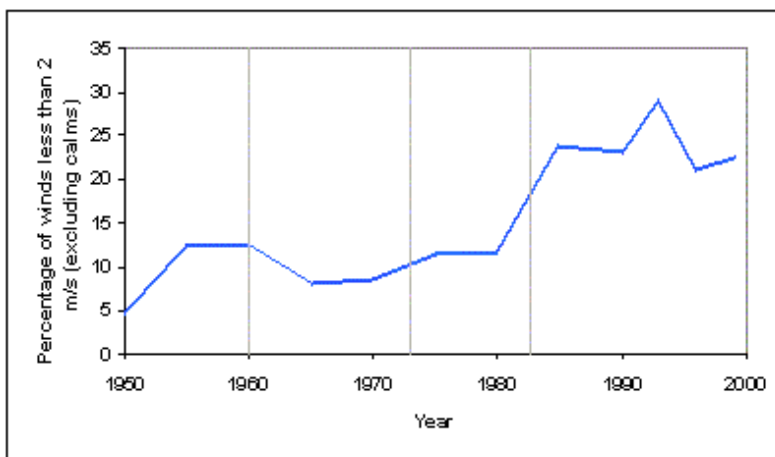


FIGURE A4 Plot to show the percentage of wind observations recording speeds of less than 2ms⁻¹, excluding the calm winds, and when the changes in anemometer position occurred (grey lines)

Figure A4 shows an important trend of the number of low wind speeds increasing towards 1999.

Although low winds generally give higher concentrations it is more complex for a buoyant source such as this, and the choice of grid resolution has probably had

some effect also on the concentrations predicted. From the graph above it is clear that the number of lower wind speeds has increased whilst the magnitude of the maximum concentrations has decreased, so there is likely to be some link between the two for this particular source. However this does not fully explain the very high maximum long term concentrations given for 1950 and 1955.

It is likely that the installation of the "Mk5" wind system anemometer sometime in 1973 is responsible for some of the changes, mainly due to the way wind speed was recorded. This is also likely to be the cause of the steep increase in the percentage of calm winds from 1980. In addition the increasing urbanisation in the Heathrow area since 1950 will have caused wind speed to decrease in general.

With the exceptionally high long-term concentration values for 1950 and 1955 in mind ADMS was run with the same settings for Waddington in Lincolnshire (53.18 N, 0.52 W), for the same years to see if this pattern was repeated at another site. The advantage of using data from Waddington is that the site has not been subject to the same increase in urbanisation and air traffic as at Heathrow. As for Heathrow, the wind roses for each year studied are shown in Appendix A. Unfortunately, the dataset for 1950 was incomplete so 1951 was used in its place. The values for both the long term average concentration and the 99.9th percentile were used to construct the graphs below.

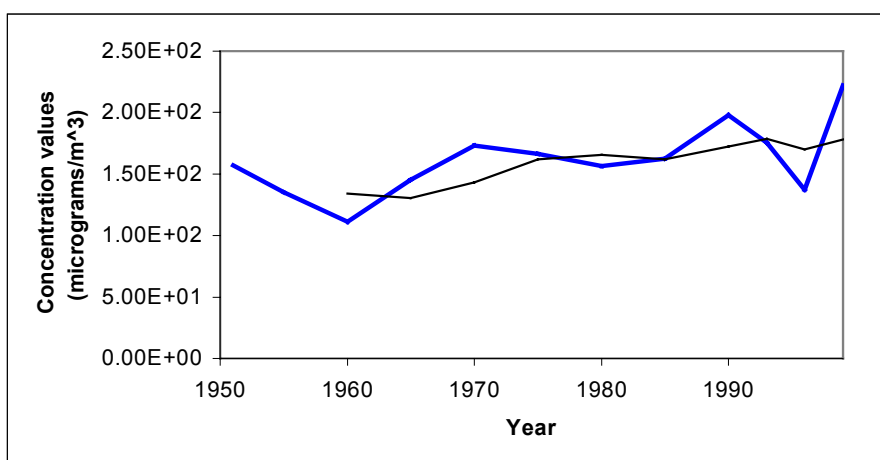


FIGURE A5 Graph to show how the maximum long term concentration values for Waddington change from 1951 to 1999 (blue line), and the trend (black line)

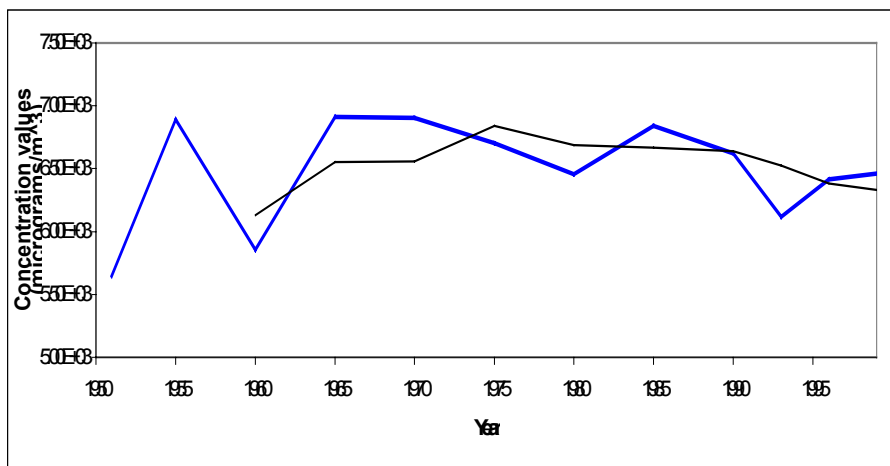


FIGURE A6 Graph to show how the maximum concentration values for the 99.9th percentile at Waddington change from 1951 to 1999 (blue line), and the trend (black line)

These graphs show that there is some degree of inter-annual variability producing slightly different results for each year and this is what is expected. However the maximum long-term concentration values increase very slightly from 1951 to 1999 and give a different pattern to the Heathrow results. The maximum concentration values for the 99.9th percentile also display a differing trend to those at Heathrow, with a less pronounced decrease over the last 50 years or so. From looking at the results from Waddington it can be concluded that 1950 and 1955 were anomalous years meteorologically at Heathrow, giving unusually high values for maximum long-term concentrations.

Although these changes in wind strength and direction are likely to have produced changes in the dispersion patterns, changes in local temperature may have also had some effect. Each different year brings generally warmer or cooler months than the last year giving annual variability. Urbanisation will have a certain effect on the overall temperature profile of the year with more tarmac, buildings etc add to the local heating affect. Also with Heathrow on the outer edge of Greater London, the heat island effect produced by the growth of London over the years may have also caused a general increase in temperature. There is likely also to have been an increase in the number of flights (especially jet aircraft) into and out of Heathrow since 1950, which may have raised local temperatures slightly.

To see if the temperatures have indeed risen over the years since 1950, the annual mean temperature of each year studied was plotted. See figure A7 below.

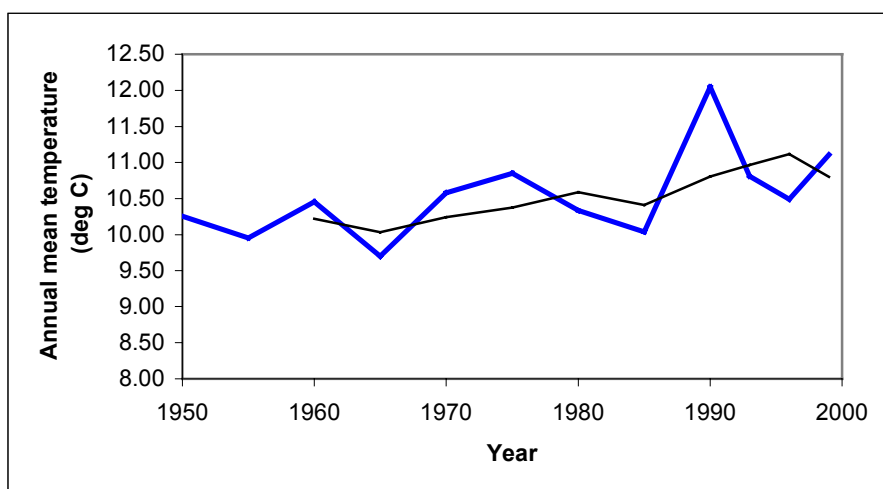


FIGURE A7 Comparison of annual mean temperature from 1950 to 1999 at Heathrow. (The black line is the trend line.)

From this graph it is clear that there has been a slow rise in the temperature over the last forty years or so as a result of urbanisation. This may have caused some slight differences in the dispersion patterns for the long-term concentrations. These differences are notably the considerably higher maximum concentration for 1950 compared with 1999.

To investigate whether this gradual rise in the annual mean temperature was due to warmer daytime temperatures, the annual mean summer daytime temperature was calculated for each year. This considered the months from the start of May till the end of August inclusively, and temperatures from the hours of 11am to 5 p.m. as between these hours the temperatures will be at their highest for the day. Figure A8 shows how this parameter varied throughout the period.

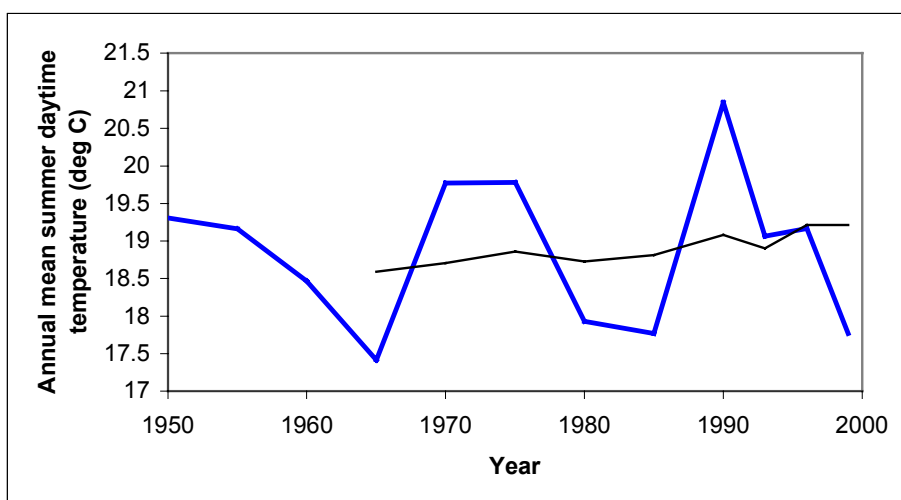


FIGURE A8 Graph to show the variation of annual mean summer daytime temperature from 1950 to 1999 at Heathrow

This plot shows that the annual mean summer daytime temperature has indeed risen since 1950, but by less than 1°C. As the pattern shown above is so variable, it can not really be concluded that the number of convective days has had a significant impact on the dispersion pattern.

From these studies of various meteorological parameters, a clear explanation for the high long-term maximum concentrations in 1950 unfortunately cannot be found.

A2.3 Conclusions

Looking at the results generated by Heathrow data for each year it is clear that even though some years are slightly different from others meteorologically, some of the datasets should not be used for modelling events today. These datasets are those of 1950 and 1955, which produced rather large differences, notably larger values for both the long term and 99.9th percentile concentrations. Therefore it can be said that for Heathrow changes in urbanisation and observing instrumentation have given some of the datasets a lifetime.

However, in comparison the Waddington data did not display such large differences in results for the same period. This is primarily due to the fact that the Waddington site is located in the countryside where there are very few buildings, and this has been the case since the site opened. There will have indeed been changes in observing practices there but these appear not to have had such an impact on the results as those at Heathrow.

Indeed, at Heathrow it is likely that changes in urbanisation will have had a larger effect on meteorological parameters than changes in instrumentation, as shown by the increase in annual mean temperature (Figure A7).

Therefore, it can be concluded that some meteorological datasets can be said to have a lifetime, but this is dependent on the changes in urbanisation, observing, and maybe even location at the observing station in question, and is not universal for all stations. So when modellers wish to use old data from stations near to their study area rather than recent data from further away, they should consider changes in urbanisation and observing. The Met Office can advise modellers on these issues before they obtain their meteorological datasets.

A3 The use of Numerical Weather Prediction (NWP) data in modelling studies

The aim of this part of the project is to investigate whether NWP data is as good as, or even better than, observed data for modelling studies. The NWP data used for this study was extracted from the 'Unified Model', which is the Met Office's NWP model for providing the basis of weather forecasts in the UK. The model is run on a Cray T3E supercomputer in two main configurations: a global and a mesoscale version, and uses a wide range of observed weather data to produce

its forecasts. The data for NWP models is organised on a three-dimensional grid, and values for various weather elements, for example winds and temperatures, are stored at the model grid points. The size of the grid varies between the two versions of the 'Unified Model'. The global version covers the globe and has a horizontal resolution of 60km across Europe, with 30 levels in the vertical. The mesoscale model covers the UK and Europe with a grid resolution of 12km, and it has 38 vertical levels. It can therefore provide more detailed forecasts than the global model, and it is data from this model that will be used for the study. Some further details on the mesoscale model can be found in the papers by Golding(1990) and Norquist(1999).

Data for 1999 was extracted from the mesoscale model for Ringway and Waddington, as these two sites are the most requested by customers. The data was extracted for every three hours (i.e. 00Z, 03Z, 06Z etc), and then run through a purpose made program to interpolate, linearly, between the three hours and create a meteorological dataset for the sites. The nearest grid point to the actual location of Ringway and Waddington was extracted. The position of these grid points, and the actual location of the stations are shown below in Table A2. Once this NWP data was extracted it was then converted into ADMS format so that it looked like the observed data files.

TABLE A2 Latitude and longitude of the actual location of Ringway and Waddington, and the corresponding location of the nearest NWP grid points, with the distance between them

Location	Actual Co-ordinates	NWP Co-ordinates	Distance between actual and NWP co-ordinates
Ringway	53.36 N, 2.28 W	53.38 N, 2.24 W	3.5 km
Waddington	53.18 N, 0.52 W	53.15 N, 0.59 W	5.8 km

It should be noted that the NWP data is not data for one point, as the observations are, but data for an average over an area. Indeed, the winds in the model represent the horizontal average over some scale that can simply be regarded as the grid box size. The data will also have been averaged over time, especially the winds, which are averaged over a period of less than an hour, centred on the output time.

All the datasets were run through ADMS, using the 3 scenarios listed in Table A3 below.

TABLE A3 Table of source characteristics for ADMS runs

Source	Height (m)	Diameter (m)	Velocity (ms ⁻¹)	Temperature (°C)
1	150	7	15	150
2	10	1	1	40
3	10	1	15	150

The 99.9th percentile was studied as this is the objective for sulphur dioxide emissions in the National Air Quality Strategy for the year 2000, and this will capture the extreme events. In addition both the long term average concentrations and the 99th percentile were examined.

As in the previous part of the project the pollutant used was an inert tracer which was emitted at 1000 grams per second, and this output was averaged over 15 minutes. The size of the modelling grid was kept the same at 50 by 50 kilometres with 31 grid points spaced regularly and the stacks were located at the centre in each scenario. No topographical features or buildings were featured in the runs once again to keep the scenario basic with less influencing factors.

A3.1 Ringway Comparisons

The Ringway site is located at Manchester Airport, which is in the Greater Manchester area. It is situated off the M56, and 30% of the area within a 100m diameter from the site is occupied with buildings. The heat island effect from Manchester is likely to have some influence on the meteorology at this site.

The wind roses (see Appendix B) generated by the observational and NWP datasets do show some similarities, for example, both show very few winds from the north. The most noticeable feature is that the NWP derived wind rose has a larger number of winds from the southwest, and although some of this is shown on the observational wind rose, the winds are mainly from the south.

The dispersion patterns produced for source 1 were very similar using both NWP and observations. The only difference was the extent of the lowest contour for the 99th percentile plot. This contour extended further out on the plot produced with NWP data.

At source 2 the plots for the long-term concentration and the 99.9th percentile displayed the lowest contour extending slightly further out than when using observational data. The 99th percentile contour plot showed the greatest difference with the observational dataset, which produced some longer spikes to the north of the stack. Again, the contours on the plot for the NWP data spread slightly further out than those on the plot for the observational data.

There were no major variations in the dispersion patterns for source 3. The long-term concentration plot was very similar in size and spacing of contouring for both datasets. The NWP data again had the lowest contour located further away from the stack on the 99.9th percentile plot, and indeed on the plot for the 99th percentile. In addition the 99th percentile pattern for the NWP data showed the

500 $\mu\text{g}/\text{m}^3$ (lowest) contour extending slightly further round to the south of the source, than when using observational data.

Although these dispersion patterns appeared similar in general it looked from the contouring levels that on occasions the NWP datasets produced larger values of maximum concentration than the observational ones. (These plots can be seen in Appendix B.) The tables below (Tables A4 to A6) show the maximum concentrations for each percentile/concentration at each of the three sources.

TABLE A4 Maximum concentration values for source 1 – Tall and buoyant source

Concentration/Percentile	Obs.	NWP	Percentage difference (NWP-OBS)
Long term concentration	2.88E+00	3.23E+00	+11
99.9 th percentile	1.79E+02	1.91E+02	+6
99 th percentile	8.86E+01	1.12E+02	+21

TABLE A5 Maximum concentration values for source 2 – Near surface, non buoyant source

Concentration/Percentile	Obs.	NWP	Percentage difference
Long term concentration	5.39E+02	5.02E+02	-8
99.9 th percentile	2.68E+04	2.90E+04	8
99 th percentile	9.37E+03	1.08E+04	12

TABLE A6 Maximum concentration values for source 3 – Near surface, buoyant source

Concentration/Percentile	Obs.	NWP	Percentage difference
Long term concentration	2.41E+02	1.89E+02	-28
99.9 th percentile	6.80E+03	7.01E+03	3
99 th percentile	4.40E+03	3.90E+03	-13

Overall the maximum values produced by the NWP data were not too dissimilar to those of the observational data. The NWP data produced slightly higher values of maximum concentration for all outputs of source 1 and all of source 2 apart from the long-term concentration where it was slightly lower. At source 3, the NWP data tended to produce smaller concentration values, but overall the NWP data performed well.

With these results in mind the differences in meteorology between the two datasets were examined to see how close the NWP meteorology was to what was actually observed. Table A7 below shows some averages for the year as well as the root mean square error for each dataset for a quick comparison;

TABLE A7 Summary of general statistics for each dataset

	Observed data	NWP data	RMS error
Average temperature	10.6 °C	9.7 °C	1.5 °C
Average wind speed	3.7 ms ⁻¹	3.1 ms ⁻¹	1.4 ms ⁻¹
Total rainfall	837.6 mm	848.4 mm	-
Average cloud	5.7 oktas	4.3 oktas	2.6 oktas
Total cloud	50158 oktas	37847 oktas	-

Just from the results in the table above it looks like the model is doing well at this site with the averages for the year very close to those observed. The root mean square error agrees with this with low values for each of the weather elements. The total rainfall predicted by the model agrees well with what was observed, however, the mesoscale model has under predicted the cloud amount. Overall though, from this it would appear that the model provided a good approximation to the real weather data at Ringway for 1999.

Earlier it was shown that when NWP data was used, more often than not, it produced higher values of maximum concentration. Although the dispersion plots produced using NWP data looked similar to those produced by the actual observations, how different actually were the concentration values over the whole grid? Did the position of the highest concentrations from the numerical data match those from the observations? This is an important issue if NWP data is to be used for dispersion modelling over an inhabited area.

By using the ADMS output files for the two runs, the percentage differences between concentration values for the NWP and actual datasets were calculated at each grid point. These differences were then plotted over the whole of the modelling domain. It would seem from these plots that the NWP concentration values were often larger than those for the observed data.

Table A8 gives an overall summary of how the concentration values using NWP data compare in magnitude with those from observed data.

This table shows that at source 1 there was an even split as to which dataset gave the highest concentration values. At sources 2 and 3 the NWP data gave larger concentrations on most of the modelling grid, perhaps indicating that for near surface scenarios the NWP over predicts concentrations. This would imply a small degree of caution is needed when using this type of data for these types of sources.

However, as source 2 is non buoyant and close to the ground, the modelling grid used in the previous runs is perhaps too large to capture the true maximum concentrations. Therefore ADMS was run again just for this source, but this time on a 0.32 by 0.32 kilometre grid.

TABLE A8 Percentage of occasions when the concentration values generated by observed data exceeded those generated by NWP data, and vice versa, for each source

Source	Concentration/ Percentile	% of Obs concentrations >NWP concentrations	% of Obs concentrations < NWP concentrations
1	Long term	51	49
	99.9 th	32	68
	99 th	46	54
2	Long term	13	87
	99.9 th	12	88
	99 th	14	86
3	Long term	21	79
	99.9 th	13	87
	99 th	233	77

The table below shows that the maximum concentrations for this higher resolution grid are considerably larger than before, and that the NWP data still tend to give higher values than the observations.

TABLE A9 Maximum concentration values for source 2 – Near surface, non-buoyant source

Concentration/Percentile	Obs.	NWP	Percentage difference (NWP – OBS)
Long term concentration	2.27E+04	2.01E+04	-13
99.9 th percentile	5.56E+05	5.73E+05	3
99 th percentile	2.86E+05	3.25E+05	12

The dispersion plots for this smaller grid were very alike in size and shape between the two datasets, especially for the long term concentration. For the 99.9th and 99th percentiles, the NWP dataset gave slightly more dispersion to the south of the stack so the contours extended round the source location, than they did on the plot using observations. These plots are shown in Appendix B.

A3.2 Waddington Comparisons

Waddington has different site characteristics to Ringway, located in a rural area just south of Lincoln. This is also one of the reasons why Waddington was chosen for the study, as it would be a good contrast to Ringway. The site is also an airfield with again 30% of the area within a 100m diameter being occupied by buildings and man made surfaces. The land slopes slightly towards the south.

The observational wind roses for Waddington show that the predominant wind direction is from the southwest, and this is the case on the wind rose using NWP data (See Appendix B). The number of winds from the southwest is less for the

NWP data than observations, but this is again due to the winds in the NWP dataset being averaged spatially. Both wind roses also show a low number of winds from the north to the south (in a clockwise direction), with a small increase in the number of winds from the northeast reflected in both. The model has tended to give fewer winds above 8.2 ms⁻¹ than actually occurred, but this again could be due to spatial and temporal averaging in the mesoscale model.

For source 1 the NWP dataset did not provide a dispersion pattern that was very different to that of observed data. It could be seen that the lowest concentration contour extended slightly further out away from the stack when NWP data was used.

There were no great differences in the contour patterns for the long term concentration at source 2 between the two datasets. On the 99.9th percentile plot it was noticeable that the contours were closer together just to the east of the stack, indicating that the observational dataset had given a higher maximum in concentration. Also on this plot the contours extended completely around the source location when using NWP meteorological data. This did not happen when using observed data. For source 3 there were no major differences in the two sets of dispersion patterns other than the contours extending over a slightly larger area when NWP data was used.

As for Ringway the maximum values of concentration for the long term average and percentiles were examined. Tables A10 to A12 below show the differences for each of the sources.

TABLE A10 Maximum concentration values for source 1 – Tall and buoyant source

Concentration/Percentile	Obs.	NWP	Percentage difference
Long term concentration	3.90E+00	3.99E+00	+2
99.9 th percentile	1.84E+02	1.91E+02	+4
99 th percentile	1.14E+02	1.21E+02	+6

TABLE A11 Maximum concentration values for source 2 –Near surface, non-buoyant source

Concentration/Percentile	Obs.	NWP	Percentage difference
Long term concentration	4.13E+02	3.36E+02	-23
99.9 th percentile	2.42E+04	2.04E+04	-19
99 th percentile	7.48E+03	6.48E+03	-15

TABLE A12 Maximum concentration values for source 3 – Near surface, buoyant source

Concentration/Percentile	Obs.	NWP	Percentage difference
Long term concentration	2.20E+02	1.67E+02	-31
99.9 th percentile	6.46E+03	6.65E+03	+3
99 th percentile	3.37E+03	3.43E+03	+2

For source 1 the NWP data produced slightly higher maximum values for all percentiles and the long-term average, but for source 2 they were lower than those from using real observations.

At source 3 the maximum values for the percentiles using NWP were slightly larger than those using observations. However for the long term concentration maximum, the NWP data gave an under prediction.

The average value for the temperature, wind speed and cloud for the year was calculated for each dataset, as was the total rainfall and total cloud. Once again the root mean square error for each weather element was calculated to find out how close to the observations the model data was. The table below (Table A13) shows the results.

TABLE A13 Summary of general statistics for each dataset

	Observed data	NWP data	RMS error
Average temperature	10.3 °C	10.6 °C	1.4 °C
Average wind speed	4.7 ms ⁻¹	4.8 ms ⁻¹	1.9 ms ⁻¹
Total rainfall	610.8 mm	755.7 mm	-
Average cloud	4.8 oktas	4.4 oktas	2.4 oktas
Total cloud	42046 oktas	38270 oktas	-

It should be noted that the observational dataset for Waddington has a number of missing rainfall values so the total, and average rainfall is noticeably smaller than for those of the NWP dataset. Despite this missing data the root mean square error for the rainfall is similar to that for Ringway, as is the temperature value. The root mean square error for the wind speed is slightly larger than for Ringway but this is likely to be due to different site characteristics, and is still a small value.

To compare the difference between concentration values for two datasets the same method as for Ringway was used to produce plots showing where there were differences. Table A14 shows how the magnitude of the concentration values produced using observed data compares with that using mesoscale model meteorology over the whole of the modelling grid.

TABLE A14 Percentage of occasions when the concentration values generated by observed data exceeded those generated by NWP data, and vice versa, for each source

Source	Concentration/Percentile	% of Obs concentrations >NWP concentrations	% of Obs concentrations < NWP concentrations
1	Long term	37	63
	99.9 th	46	54
	99 th	34	66
2	Long term	23	77
	99.9 th	24	76
	99 th	30	70
3	Long term	41	59
	99.9 th	33	67
	99 th	42	58

Once again the concentration values produced using the NWP data are generally larger than those using observed data. However, when we look at tables A10 to A12 the maximum concentrations for each dataset were not actually too dissimilar, suggesting that whilst the NWP data tends to produce larger concentration values they are only fractionally larger and should not have a significant effect on the end result.

Once again, because source 2 is non buoyant it was also modelled on a much smaller grid to capture in more detail what the maximum concentrations were. Again they were much larger than previously seen, but once again the NWP data gave higher values generally for the maximum concentration than the actual observed data did, as table A15 below shows.

TABLE A15 Maximum concentration values for source 2 – Near surface, non-buoyant source

Concentration/Percentile	Obs.	NWP	Percentage difference (NWP – OBS)
Long term concentration	2.72E+04	2.26E+04	-20
99.9 th percentile	4.49E+05	5.23E+05	14
99 th percentile	2.03E+05	2.15E+05	6

Overall the dispersion plots for each source were not hugely dissimilar for either the long term concentration or the percentiles, which shows again that NWP data is giving good results here.

A3.3 Conclusions

In conclusion, it would appear that NWP data is a good substitute for observed data when and where it is not available. The NWP data produced similar wind roses to observed data although the predominant winds were generally backed

by a few degrees, and occurred more often than those in observed did. These wind roses are a good indication of how close the model winds are to those that actually occur despite the fact that they are spatially averaged.

The NWP often produced larger values of maximum concentration and in addition tended to over predict concentration values over all compared to the predictions based on the observed data, although not by much in most cases.

With improvements to model schemes every year the mesoscale model can only continue to improve and become more accurate. As further weather stations close or become automated across the UK the use of NWP in dispersion models should certainly be considered as a good and acceptable alternative to using actual meteorological data from stations which may have closed or are a large distance away from the study area.

A3.4 Comparison with secondary variables added to the NWP meteorological dataset

ADMS 3 can calculate meteorological variables such as the boundary layer height, from the data given in the meteorological file. However, it is preferable to add such variables into the meteorological file if they have been calculated reliably or measured at the site being modelled. This reduces errors in the modelling. Values for the boundary layer height and surface heat flux were extracted from the mesoscale model at Ringway and Waddington for all of 1999, and added to the original NWP meteorological dataset. Further ADMS runs were then undertaken for each of the three scenarios used previously with the new NWP dataset at both sites.

A3.4.1 Results of the comparison at Ringway

The tables (tables A16 – A18) below show how the maximum concentration values from the NWP data with secondary variables compare with those found previously from actual observed data and NWP data without added variables. The percentage difference between the maximum concentration from the NWP dataset without secondary variables and the observed dataset is shaded in grey.

TABLE A16 Maximum concentration values for source 1 – Tall and buoyant source

Concentration/Percentile	Obs.	NWP	NWP with secondary variables	% difference between Obs & NWP with secondary variables
Long term concentration	2.88E+00	3.23E+00	3.39E+00	15 (11)
99.9 th percentile	1.79E+02	1.91E+02	2.41E+02	26 (6)
99 th percentile	8.86E+01	1.12E+02	1.19E+02	26 (21)

TABLE A17 Maximum concentration values for source 2 – Near surface, non-buoyant source

Concentration/Percentile	Obs.	NWP	NWP with secondary variables	% difference between Obs & NWP with secondary variables
Long term concentration	5.39E+02	5.02E+02	7.69E+02	30 (-8)
99.9 th percentile	2.68E+04	2.90E+04	5.31E+04	50 (8)
99 th percentile	9.37E+03	1.08E+04	1.78E+04	47 (12)

TABLE A18 Maximum concentration values for source 3 – Near surface, buoyant source

Concentration/Percentile	Obs.	NWP	NWP with secondary variables	% difference between Obs & NWP with secondary variables
Long term concentration	2.41E+02	1.89E+02	1.87E+02	-29 (-28)
99.9 th percentile	6.80E+03	7.07E+03	6.43E+03	-6 (3)
99 th percentile	4.40E+03	3.90E+03	3.57E+03	-23 (-13)

Overall these tables show that adding boundary layer height and surface heat flux values to the NWP dataset at Ringway produces greater differences in maximum concentration than from using NWP data without, especially for the 99.9th and 99th percentiles. The greatest differences were again for Source 2 where the emissions are kept near to the surface.

These differences in maximum concentration have been caused by adding the mesoscale model boundary layer and surface heat flux values to the NWP meteorology file. To examine how the NWP values for the boundary layer height and surface heat flux are different to those calculated by ADMS 3, the observed meteorological dataset was run through the ADMS meteorological pre-processor to obtain the ADMS boundary layer height and surface heat flux values. The NWP generated boundary layer height and surface heat flux values were then compared with those from ADMS to see how different they were in an attempt to explain the differences in maximum concentrations. The comparison was undertaken for the summer and winter months as these would show up any differences more clearly. The graphs produced as part of this comparison are shown in Appendix C. The wind speed values from the two datasets were also compared as this is an important parameter in determining the boundary layer height.

For both the summer and winter months the ADMS boundary layer height values were close to those generated by the mesoscale model. Although the mesoscale values did tend to be slightly larger, most noticeably in the summer of 1999 when on a few occasions it produced a boundary layer height of over 2000 metres. On some occasions the ADMS boundary layer height was given as -999 when some of the required meteorological data was missing. Despite the mesoscale model generally giving higher values for the boundary layer height, it tended to give lower wind speeds than those that were observed. This is likely to

be due to the fact that the wind in the mesoscale model is averaged spatially, but it could be partly due to the land use given in the mesoscale for Ringway is possibly different to what it actually is. Each of the 12 km tiles in this NWP model is given one land use characteristic and one roughness length which is representative generally of the area being modelled. As the city of Manchester is just to the north of Ringway perhaps this site has an urban land use and roughness length. This would indeed explain the NWP wind speed values being lower and the boundary layer height being slightly larger.

The mesoscale model produced substantially larger values for the surface heat flux in the summer months, up to twice as large on a number of occasions. However, in the winter months the NWP heat flux values became less than those generated by ADMS and were often considerably lower.

ADMS tends to under predict the highest surface heat fluxes and this could be why in the summer months the ADMS value is so much lower than that for the mesoscale. The land use characteristics given for Ringway in the mesoscale will be largely responsible for the heat flux values produced by it. Indeed the fact that the NWP heat flux values were so much larger than the ADMS ones for the summer months may point to an urban land use tile being used for Ringway in the mesoscale. However, the NWP heat flux is the surface heat flux value for the whole of a 12km tile and not at one site which may also have had an effect on these results. Some investigation was made to establish the land use and roughness length characteristics given for Ringway in the mesoscale model, and it was found that the roughness length in the model was 0.73m. This roughness length is known in the model as the effective roughness length and takes into account the effect of sub-grid scale orography. For this reason, and the fact that the grid square in which Ringway lies does include the outskirts of Manchester, the value given is higher than would be expected at this site.

Despite the differences in maximum concentrations that were seen earlier, the dispersion plots produced by the NWP dataset with secondary variables were not significantly different to those produced using the original NWP dataset. So although the maximum concentration values were higher no other significant differences over the whole of the modelling grid could be seen.

A3.4.2 Results from the comparison at Waddington

After the new NWP meteorological dataset, containing the secondary variables, was run through ADMS the maximum concentrations values were noted and compared with those found previously at Waddington. The results are shown below in tables A19 to A21. The percentage difference between the maximum concentration from the NWP dataset without secondary variables and from the observed dataset is again shaded in grey.

TABLE A19 Maximum concentration values for source 1 – Tall and buoyant source

Concentration/Percentile	Obs.	NWP	NWP with secondary variables	% difference between Obs & NWP with secondary variables
Long term concentration	3.90E+00	3.99E+00	3.41E+00	-14 (2)
99.9 th percentile	1.84E+02	1.91E+02	2.45E+02	25 (4)
99 th percentile	1.14E+02	1.21E+02	1.01E+02	-13 (6)

TABLE A20 Maximum concentration values for source 2 – Near surface, non-buoyant source

Concentration/Percentile	Obs.	NWP	NWP with secondary variables	% difference between Obs & NWP with secondary variables
Long term concentration	4.13E+02	3.36E+02	3.33E+02	-24 (-23)
99.9 th percentile	2.42E+04	2.04E+04	4.39E+04	45 (-19)
99 th percentile	7.48E+03	6.48E+03	6.70E+03	-14 (-15)

TABLE A21 Maximum concentration values for source 3 – Near surface, buoyant source

Concentration/Percentile	Obs.	NWP	NWP with secondary variables	% difference between Obs & NWP with secondary variables
Long term concentration	2.20E+02	1.67E+02	1.25E+02	-76 (-31)
99.9 th percentile	6.46E+03	6.65E+03	5.87E+03	-10 (3)
99 th percentile	3.37E+03	3.43E+03	2.44E+03	-38 (2)

As for Ringway, the NWP dataset with boundary layer height and surface heat flux included produced greater differences in maximum concentration than without. However, the greatest differences were at source 3.

To compare the differences between the NWP boundary layer height and surface heat fluxes with those produced by ADMS, a series of graphs were drawn (Appendix C). These graphs showed that the boundary layer heights were in agreement for both of the seasons, often with the NWP values produced being slightly higher. In fact the boundary layer heights, and indeed wind speed values, at Waddington appeared to be slightly closer between the two datasets than at Ringway. This could indicate that the land use characteristics used in the mesoscale are a good representation of those at Waddington. Indeed the roughness length for the grid box containing Waddington was given as 0.06m, which is slightly lower than expected but reasonable for this site. The NWP heat flux values were again considerably larger for the summer months than those calculated by ADMS, but were in better agreement for the winter months.

The majority of dispersion plots produced by the NWP dataset including the secondary variables are a good match in size and shape to those produced using observed data and NWP without them. On a few of the plots the NWP data with secondary variables had produced larger maximums in concentration and they changed the dispersion over the modelling grid considerably e.g. plot of the 99.9th percentile at source 2.

A3.4.3 Conclusions

Overall the NWP dataset including secondary variables produced closer results to those from ADMS at Waddington than Ringway. Including boundary layer height and surface heat flux from the mesoscale model generally produced larger values in maximum concentration than NWP without. The NWP values are representative of those for an area 12 km by 12 km and are therefore not as representative of the sites as the actual observed values are, and this will produce some of the differences seen in the results. However this did not produce considerable differences at Waddington, possibly because there is little land use change in the surrounding area, compared to Ringway which has the city of Manchester just to the north and is modelled in the mesoscale as having a more urban land use. This is assuming that the ADMS schemes for calculating the boundary layer height and surface heat flux are producing the correct results. As we know ADMS tends to under predict the surface heat flux values this is proof that ADMS has limitations.

However, the aim of this modelling study was to investigate whether including secondary variables into the NWP meteorological dataset improved the closeness of the NWP results to those produced using the observed data. From these results it would appear that putting these variables into the NWP dataset makes the ADMS modelling results more dissimilar to those from the NWP dataset without, and also those produced by the observed meteorology.

Therefore, if Ringway or Waddington were to be closed tomorrow and modellers wished to have ADMS meteorological datasets for either of these sites, the results here show that while NWP data is a good and viable alternative, including secondary variables should be seriously considered.

A4 Investigation of Meteorological data from the Site Specific Model

This, the last part of the project aims to compare the use of meteorological data from the Met Office's Site Specific Model (SSM) which has been formatted to form an ADMS dataset, with observational data for the same site. The site chosen was Ringway as this is one of the locations that the SSM is set up to forecast for and data is available from last year. Also, by using Ringway a comparison can be made between the results using SSM data and those found in part 3 of the project when NWP data was used. The reason for undertaking these comparisons is to investigate how much better the SSM data, which is refined NWP data, is to the usual forecast NWP data.

Unfortunately the SSM data is stored as a one year rolling archive, and at the time this part of the project was started only the whole of November and December could be obtained and so modelling has only been undertaken for these two months.

A4.1 Description of the SSM

The SSM is used to add site-specific detail to products produced by the NWP models. This model is a 1D model in the vertical and includes a detailed treatment of all the physical processes that occur in a column of air above a location. These processes modelled include the sub-surface soil moisture, soil temperature processes, surface exchange, boundary layer mixing, cloud and precipitation processes, and radiative exchange. NWP data can be used to provide synoptic and mesoscale forcing for the SSM, as currently standard synoptic data cannot be assimilated directly into the model. The model is itself based on the Unified Model, but it is run at a higher resolution in the boundary layer so as to improve the resolution of stable layers and radiative exchange.

The SSM can produce forecasts which are more site specific than either the mesoscale or global models because it has a higher resolution in the soil and boundary layer where local factors, such as topography, can have a significant effect on the particular meteorology in that area. Indeed, this model has included in the schemes simple corrections to forcing due to local variations in the topography. In addition a sophisticated treatment of surface exchange processes in the local, upwind fetch has been included to take account of the local environment, such as land use.

The exchange of heat, momentum and moisture with the surface over the upwind fetch determines the near-surface profiles of temperature at each location. To model the inhomogeneous terrain in terms of equilibrium surface exchange, the modelled area is divided into a number of tiles, which represent homogenous characteristics. The tiles used can represent bare soil, open water, grass, deciduous trees, coniferous trees and urban land use. The proportions of each tile are derived by multiplying each element of land use in the upwind fetch by a weighting function. The weighting function is derived by estimating the changes in concentration of a tracer released from the surface as it ascends to some point above the site.

The SSM has been tested on a number of occasions. The initial version of the SSM was run routinely for 14 synoptic stations during a three month trial in autumn 1996 and it showed significant improvement over the mesoscale model. When the land use tiles were included, the model was tested again, from 1st November 1997 to 31st January 1998 using the same data used in the previous three month trial, plus 5 additional sites that were used for forecasting road temperatures. At the latter 5 sites the Met Office's open road model (which calculates road surface temperatures) was run using SSM model output, mesoscale model output and forecaster modified forcing. Overall, the runs using SSM data performed as well as the forecaster. The model has also been run routinely for a number of sites, in order to show up any problems resulting from unusual meteorological conditions.

Within the Met Office the SSM is used in a number of ways, including the provision of a first guess for forecasts for the utility companies, e.g. National Grid Company (NGC). Data from this model can be used to provide input to air

quality forecasts, and more importantly it provides better NWP information on a variety of sites which forecasters can use for general guidance.

A4.2 Use of SSM data for Ringway

The SSM data for November and December was extracted for Ringway (excluding boundary layer and surface heat flux) and then formatted into a meteorological file for use in ADMS 3. Again the three different sources and three different concentrations/percentiles were used as previously for this project. In addition to the SSM data, the observed data and NWP data for Ringway was run through ADMS for the months November and December to provide the comparisons.

First, the wind roses from the SSM and NWP datasets were compared with that from the observed data to see how the winds from these NWP sources compared with what was observed. The wind rose from the NWP dataset was the most similar to that of the observed data. The wind rose from the SSM data showed that this model produced a lower frequency of winds from the south to southwest, and a greater spread in the number of southwesterly winds. (See Appendix D.)

The maximum values of concentration were then studied. Tables A22 to A24 below show the differences between the datasets.

TABLE A22 Maximum concentration values for source 1 – Tall and buoyant source

Concentration/Percentile	Obs	NWP	SSFM
Long term concentration	2.65E+00	1.87E+00	2.16E+00
99.9 th percentile	1.33E+02	1.32E+02	1.29E+02
99 th percentile	9.11E+01	5.56E+01	5.59E+01

TABLE A23 Maximum concentration values for source 2 – Near surface, non-buoyant source

Concentration/Percentile	Obs	NWP	SSFM
Long term concentration	6.55E+02	5.69E+02	4.69E+02
99.9 th percentile	2.45E+04	2.47E+04	2.73E+04
99 th percentile	9.28E+03	9.33E+03	1.31E+04

TABLE A24 Maximum concentration values for source 3 – Near surface, buoyant source

Concentration/Percentile	Obs	NWP	SSFM
Long term concentration	3.36E+02	2.83E+02	2.26E+02
99.9 th percentile	6.68E+03	6.96E+03	5.55E+03
99 th percentile	4.64E+03	4.24E+03	3.76E+03

Table A22 shows that both the NWP and SSM data often produces maxima that are less than that produced from the actual meteorological data. Both the NWP and SSM datasets gave similar values of maximum concentration to those from the observations, but overall the computer generated meteorological data under predicted the maximum values. For source 2 the ordinary NWP data gave closer maximum concentration values than the SSM data. This was also the case for source 3. Overall the NWP dataset gave closer values of maximum concentration to the observed values, than the SSM dataset did.

From the above tables the NWP and SSM datasets tend to under predict the size of the maximum concentration values. A brief comparison of the meteorology between the datasets showed that both the NWP and SSM weather parameters were in fairly good agreement to those that were observed. Graphs comparing various weather elements is shown in Appendix (D). Table A25 below shows some averages for temperature, wind speed and cloud, as well as the root mean square errors for each, as a guide to how the NWP and SSM models have done in forecasting the meteorology.

TABLE A25 Summary of general statistics for each dataset

	Observed data	NWP data	SSM data	RMS error (NWP data)	RMS error (SSM data)
Average temperature	6.2 °C	4.8 °C	5.0 °C	1.8 °C	1.8 °C
Average wind speed	4.2 ms ⁻¹	3.7 ms ⁻¹	3.9 ms ⁻¹	1.4 ms ⁻¹	1.7 ms ⁻¹
Total rainfall	181.2 mm	109.6 mm	193.9 mm	-	-
Average cloud	5.9 oktas	4.1 oktas	5.1 oktas	2.8 oktas	2.5 oktas
Total cloud	8586 oktas	5980 oktas	7480 oktas	-	-

Table A25 above shows that for the averages the SSM tend to be closer to what was observed at Ringway, than the mesoscale model. Also, the root mean square errors for the NWP and SSM datasets are small.

The only parameter that showed noticeable differences from the observations was the wind direction. Table A26 below shows how they compare by dividing the wind directions into quadrants and counting the frequency of winds in each quadrant. The directions are measured as the direction the wind is blowing from as in the ADMS meteorological data files.

TABLE A26 Differences in wind direction for the three datasets

Wind Direction	Frequency of observed winds	Frequency of NWP winds	Frequency of SSM winds
North to East	179	101	111
East to South	31	93	98
South to West	750	813	742
West to North	504	457	513

This table shows how the mesoscale and SSM models predict too few winds from the north to east quadrant than were observed in the two months, and considerably over predicts the number of winds from the eastern to southern quadrant. As the wind direction is a critical factor in the dispersion of pollutants this is an important finding. It is also very likely to be the reason why the NWP and SSM models under predicted the maximum values in concentration on the majority of instances.

At source 1 the long term concentration plots for each dataset showed little differences. There was one noticeable difference however on the plot for the 99.9th percentile. The observational dataset gave a small maximum to the southwest of the stack that was not reflected on the plots for the numerical data. The observational dataset gave another small maximum on the plot for the 99th percentile that was not on the NWP or SSM plot, but this time it was to the southeast of the stack location.

For the plot of long term concentration at source 2 the numerical data showed more dispersion to the south of the stack. This was again shown on the plots for the 99.9th percentile. On the SSM plot for this percentile there was a small peak in concentrations to the west of the stack, this is likely to be due to the higher number of winds from the east in this dataset. The plots for the 99.9th percentile showed the greatest differences. Again, clearly present was the slight extension of the contours to the south using numerical data, and into the north western half of the modelling domain.

At source 3 there were no great differences in the plots of concentration plots for either the long term concentration or the percentile plots.

Although the dispersion patterns are largely similar between datasets there are some noticeable differences, which are almost certainly due to the differences in wind direction. These differences can be shown for the whole of the modelling grid in the table below as percentage differences.

TABLE A27 Percentage differences between the concentration values generated by observed data and SSM data for each source. The corresponding percentages for the NWP dataset are in grey

Source	Concentration/Percentile	% of Obs concentrations >SSM concentrations	% of Obs concentrations < SSM concentrations
1	Long term	46 (43)	54 (57)
	99.9 th	51 (62)	49 (38)
	99 th	52 (97)	48 (3)
2	Long term	35 (21)	65 (79)
	99.9 th	38 (38)	62 (62)
	99 th	37 (37)	63 (63)
3	Long term	21 (35)	79 (65)
	99.9 th	24 (24)	76 (76)
	99 th	18 (29)	82 (71)

The SSM data gives a fifty – fifty split with the observational dataset as to which one gives the highest concentrations at source 1. The NWP data tends to under predict at this site however, especially for the 99.9th and 99th percentiles. For the near surface source 2, the SSM data gives more values that are higher than the observations, and so does the NWP data. At source 3 the SSM produces an even greater number of concentration values that are greater than those generated by the observational dataset over almost the entire modelling grid. So, although in most cases the numerical data gave lower maximum concentration values than those using observations generally did, it gave higher concentrations over the modelling area. In summary both the NWP and SSM datasets over predicted concentration values for these two last months of 1999. This could again be related to the differences in wind direction between these datasets and the observational one.

A4.3 Conclusions

For these two months the SSM data under predicted concentration values generally across the modelling grid, and maximum concentration values. The SSM dataset gave good results for all of the sources , but it was only really at source 1 that it produced closer results to the observations than the NWP dataset did. However, the NWP dataset produced good results over the whole of 1999 and the SSM would probably also give closer results over a longer period too. Indeed the meteorology from the SSM, as shown in Appendix D, is very close to that which was observed at Ringway in these months, indicating that over a year using SSM data could provide good results.

One thing to note is about the use of SSFM data generally is that for sites that are well exposed and are on average fairly flat and representative terrain the SSM does not add significant value to the mesoscale. It is only for sites where topography and local variations in land use are important that the SSM can give better results than the mesoscale. This was found as part of the verification on

the SSM that was undertaken by the Met Office's Local Forecasting Research and Development Group.

The use of SSM in dispersion models, such as ADMS 3, would be perhaps an area for further study. However, SSM data is unlikely to be available to modellers in the foreseeable future, but is an area the Met Office is keen to develop.

A5 Summary

From the first part of the project it can be concluded that some meteorological datasets can be said to have a limited lifetime, but this is dependant on the changes in urbanisation, observing, and maybe even location of the observing station in question, and is not universal for all stations.

The remainder of the project involved the study of data from numerical weather prediction models and its use in ADMS and has found some encouraging results for its use. The NWP data did tend to produce higher maximums in concentration compared to those from the observed data, and indeed over the whole modelling grid. Adding secondary variables, such as surface heat flux and boundary layer height, from the mesoscale model to the dataset did not give closer concentrations to the ones from observational data in most cases, indicating that it adds little value.

Using ADMS datasets containing meteorology from the Site Specific model did not give concentrations that were closer to those from the observed data, than the NWP dataset did at Ringway. However only two months were studied and perhaps for a year's worth of meteorology the results would have been more encouraging.

The Environmental Consultancy Group at the Met Office is to undertake a study on the use of NWP data for dispersion modelling over a variety of sites in the near future. If the results from that study are as good as the ones found here then it would appear modellers can have confidence in using NWP data at their site of study in preference to data from an observational station which may be some distance away.

A6 References

- Carruthers et al., (2000) 'ADMS 3 Technical Specification', CERC website
<http://www.cerc.co.uk>.
- Cimorelli A.J, Perry S.G, Venkatram A, Weil J.C, Paine R.J, Lee R.S, Peters W.D (1998) 'AERMOD - Description of model formulation', Version 98315 (AERMOD and AERMET) and 98022 (AERMAP). EPA document posted on EPA website
<http://www.epa.gov/scram001>.
- Davies B.M and Thomson D.J (1997) 'Investigating the importance of pre-processing in estimating dispersion climatology', Int. J. Environment and Pollution, Vol 8, Nos. 3-6, pp 590-603.

Golding B.W (1990) 'The Meteorological Office Mesoscale Model', Meteorological Magazine (UK) Vol 119 pp 81-96.

Hough M.N and Nelson N (2000) 'The representivity of weather data for dispersion calculations', ADMLC report.

Norquist D.C (1999) 'Cloud predictions diagnosed from mesoscale weather model forecasts', Monthly Weather Review, Vol 127 pp 2465-2483.

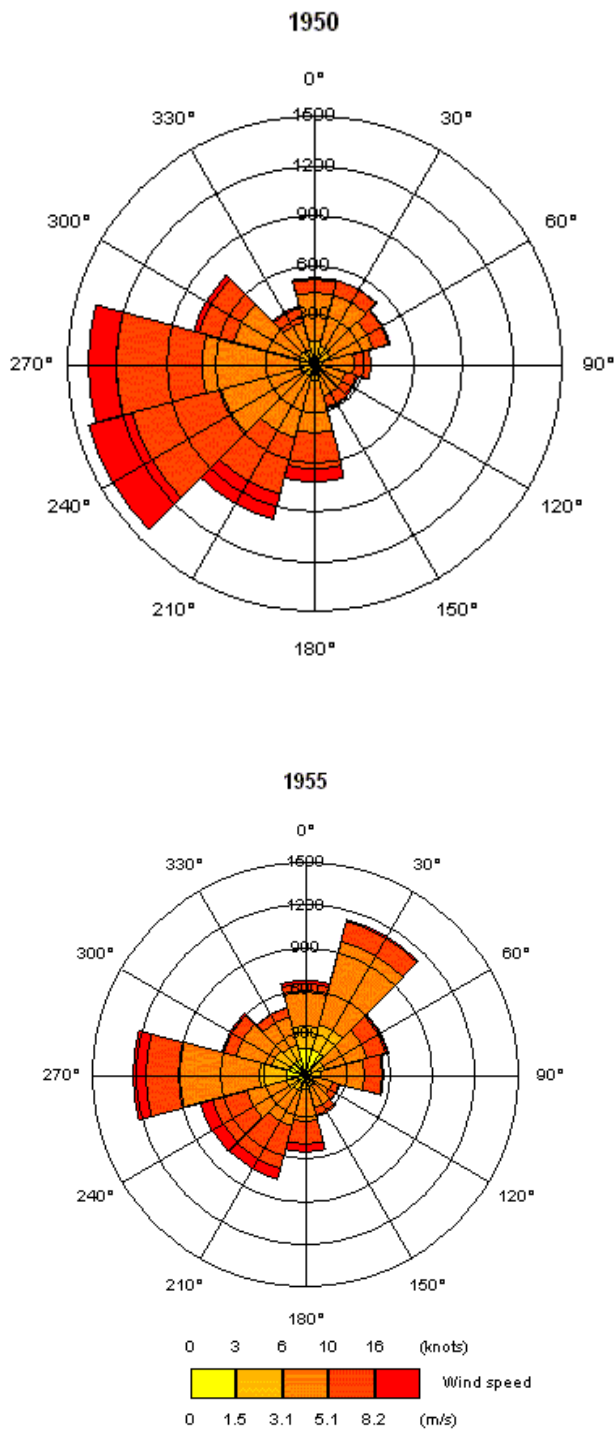
A7 Acknowledgements

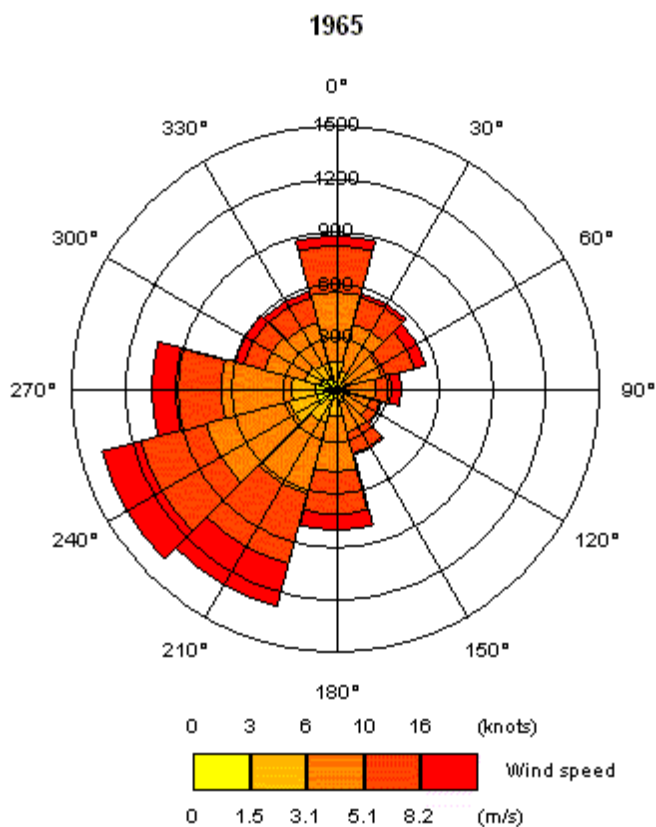
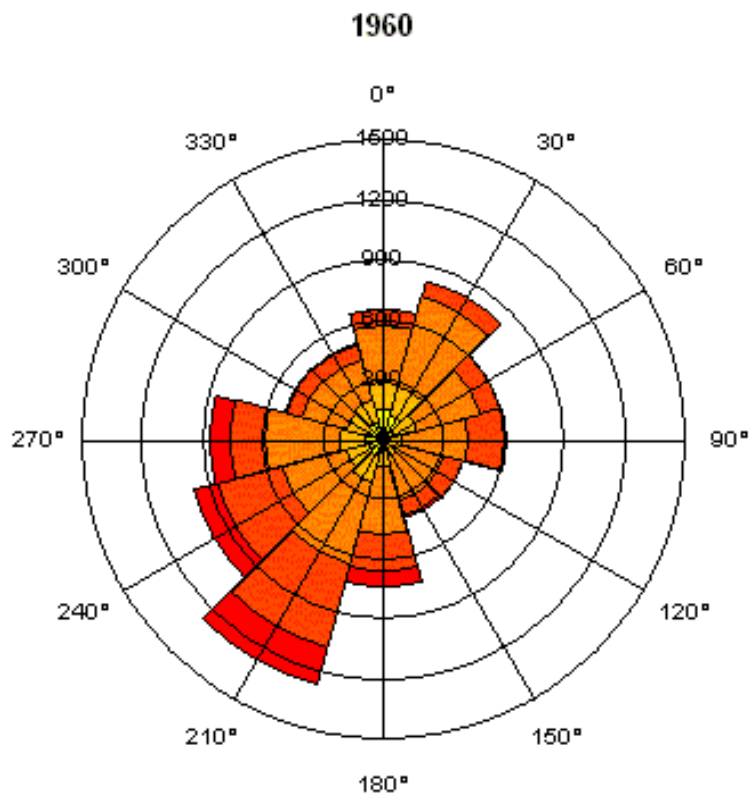
The help and advice given by Phil Hopwood and Simon Jackson (Met Office Local Forecasting Research and Development Group) in obtaining SSM data and formatting it for ADMS is acknowledged.

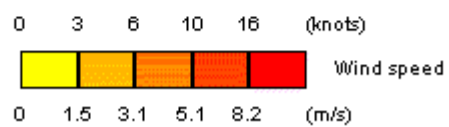
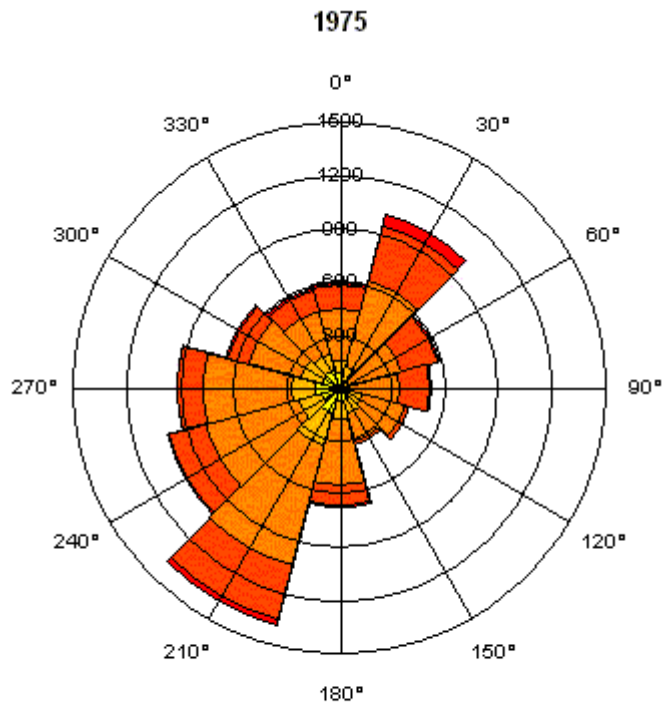
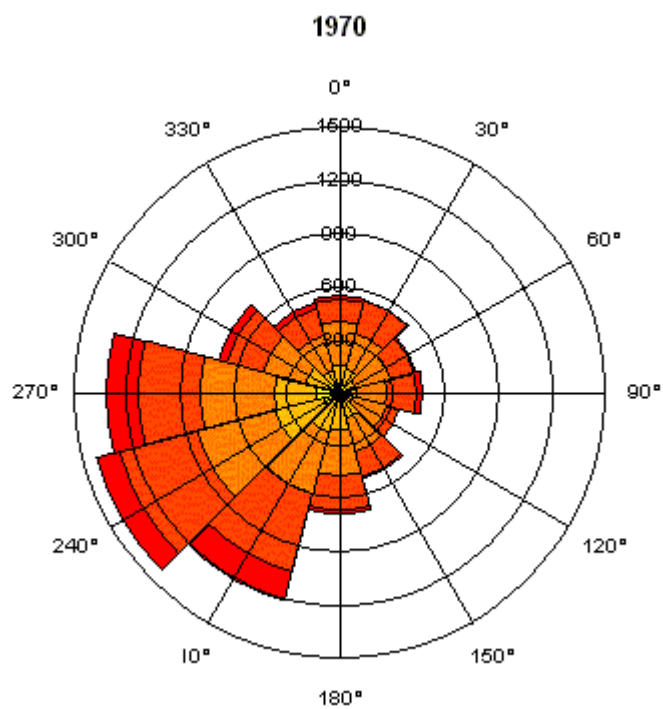
Also, the comments and suggestions provided by Ben Miners (formerly Environment Manager, Met Office Environmental Consultancy Group) and David Thompson, both during the project and regards to the final report are acknowledged.

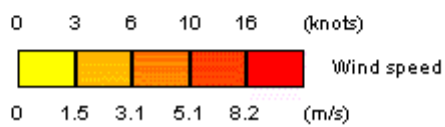
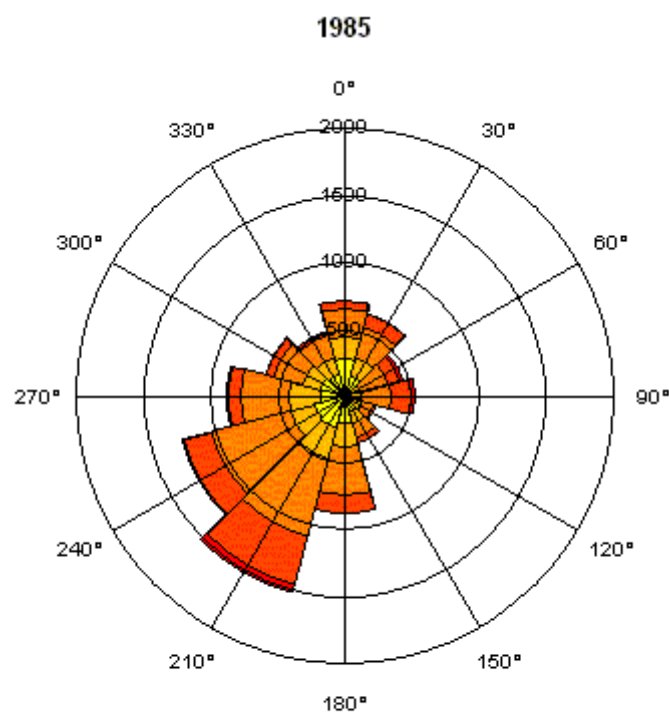
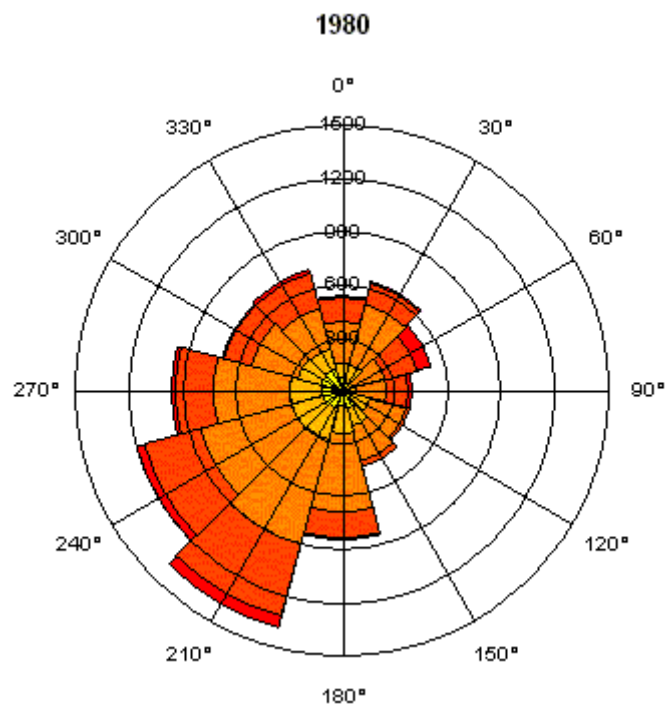
APPENDIX A Heathrow and Waddington wind roses and some dispersion plots

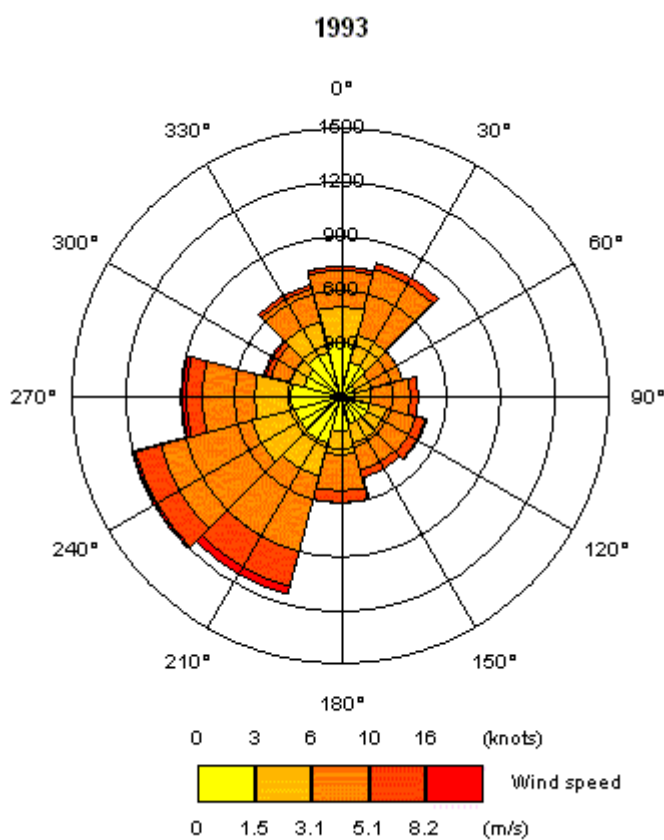
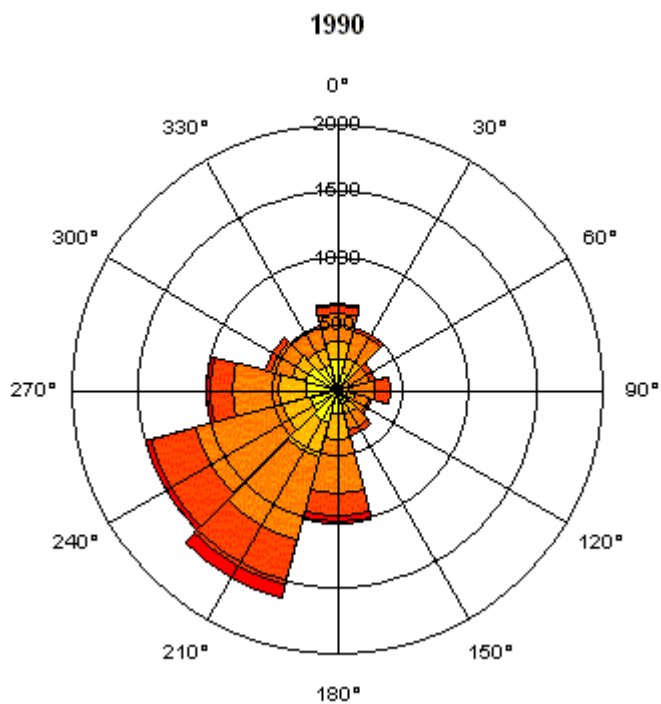
Wind roses for Heathrow 1950 – 1999

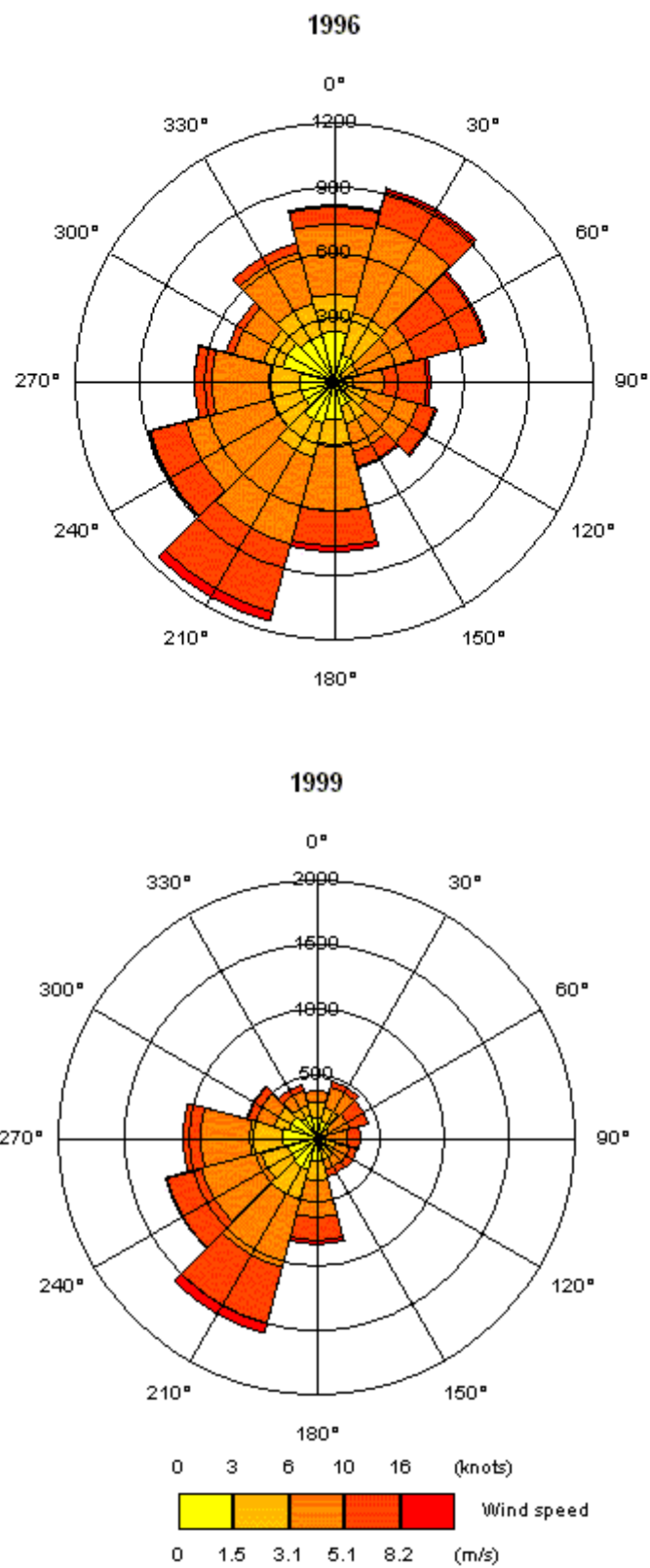




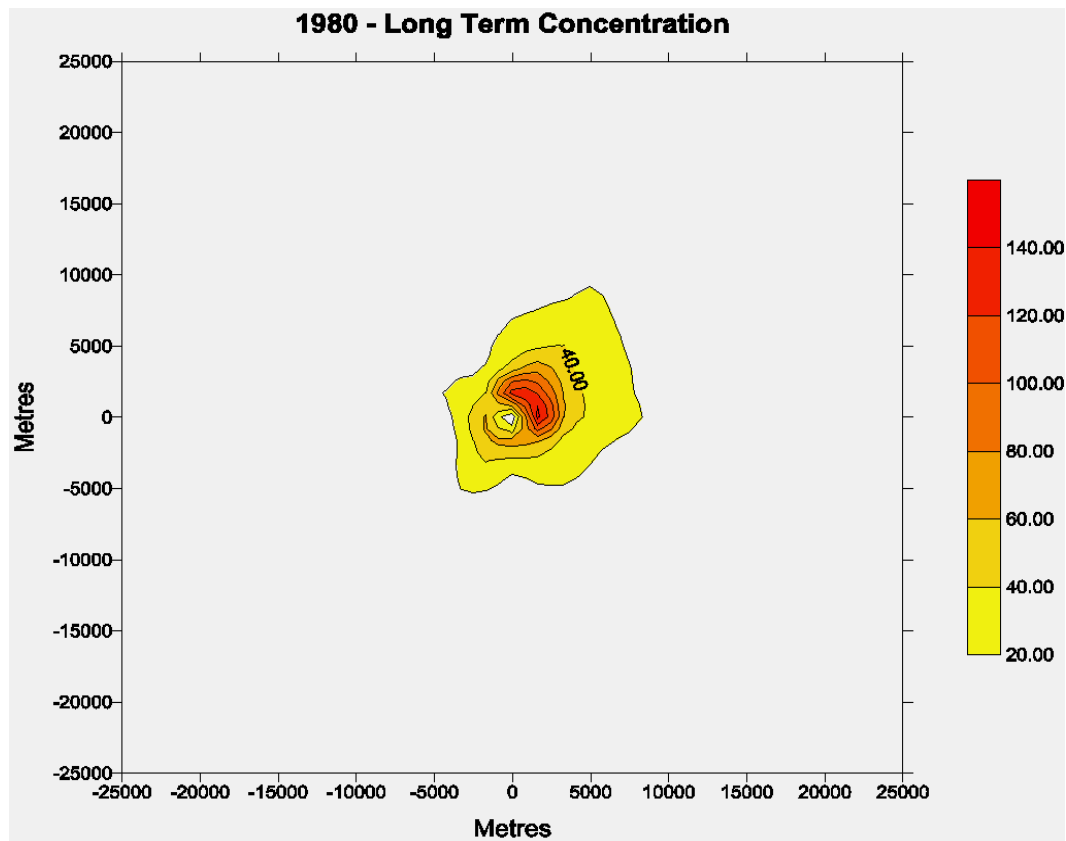
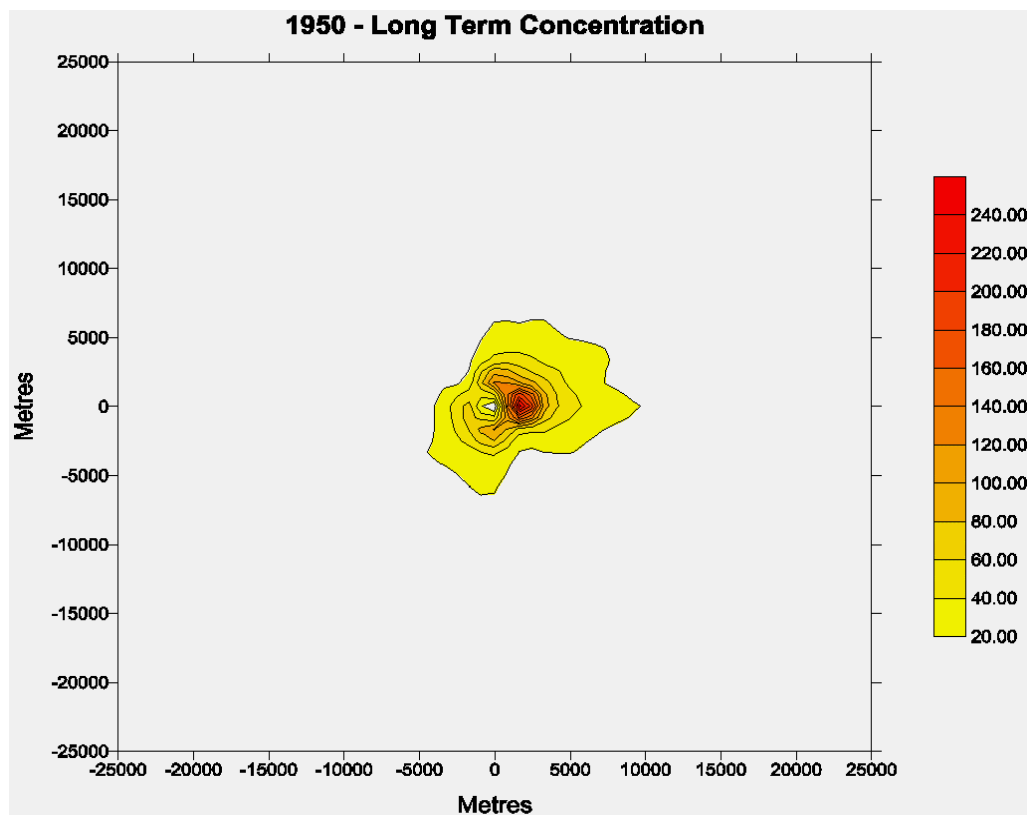


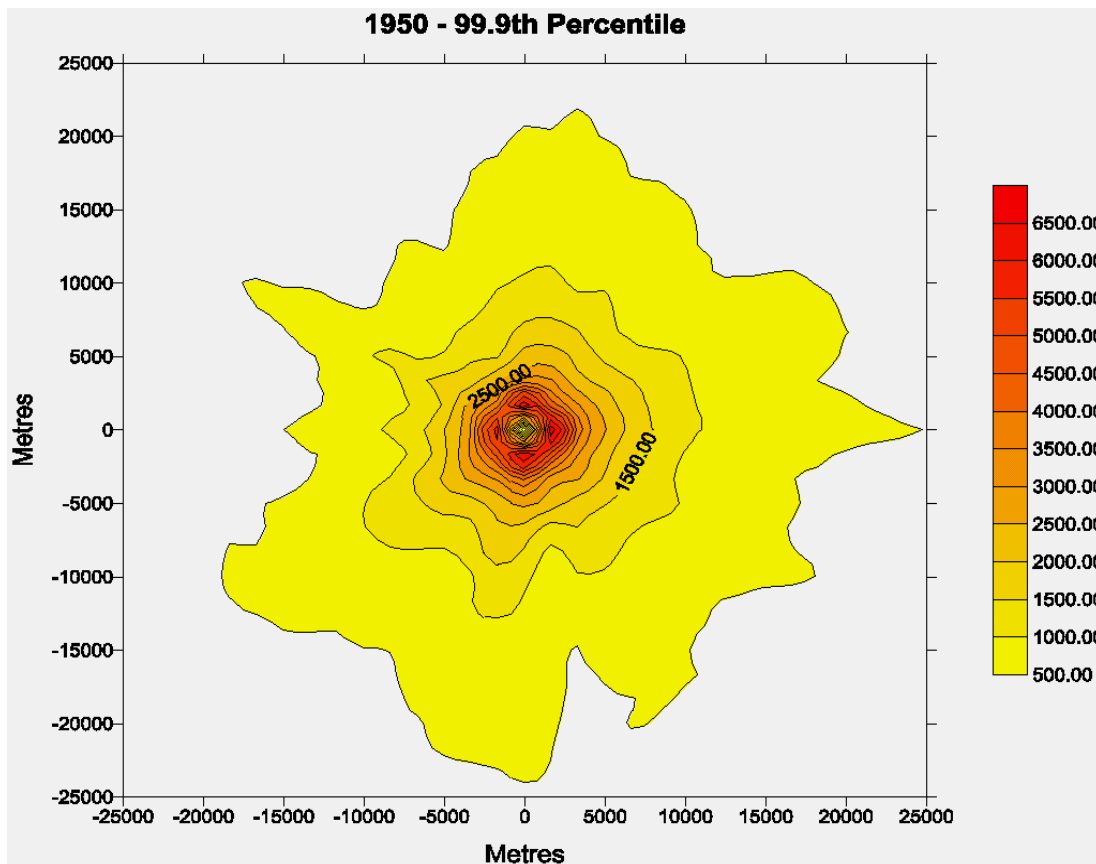
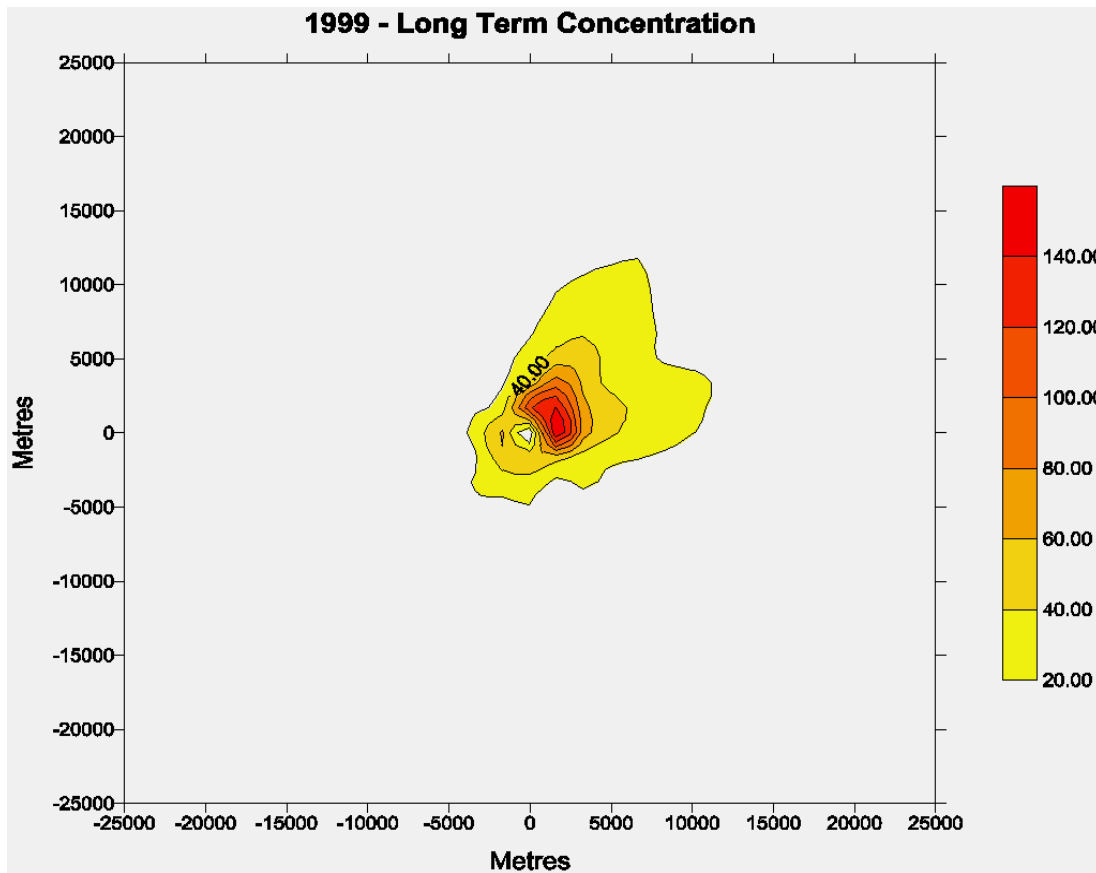


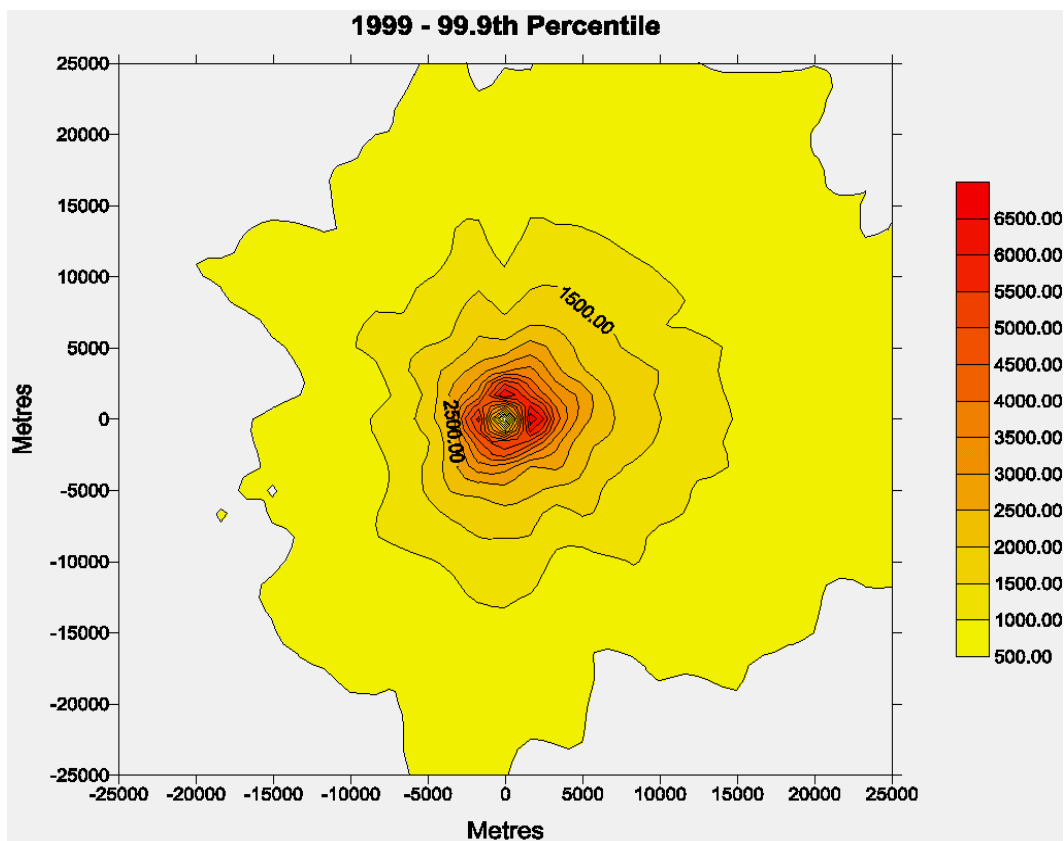
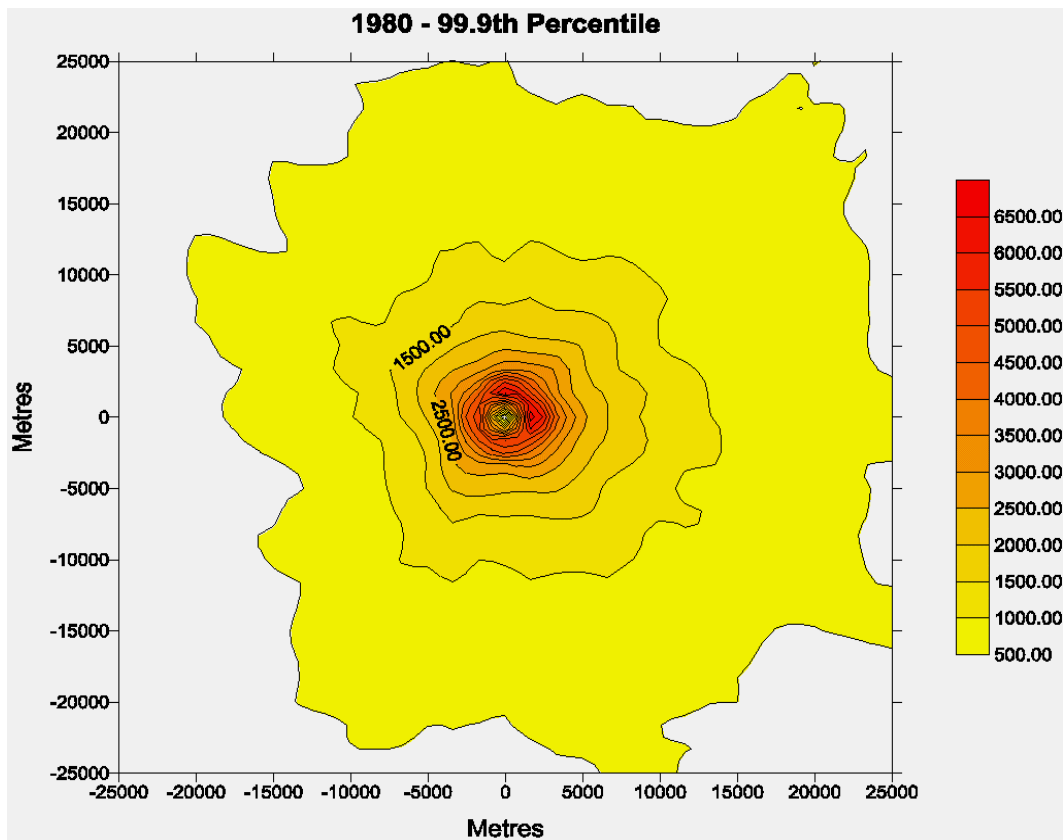




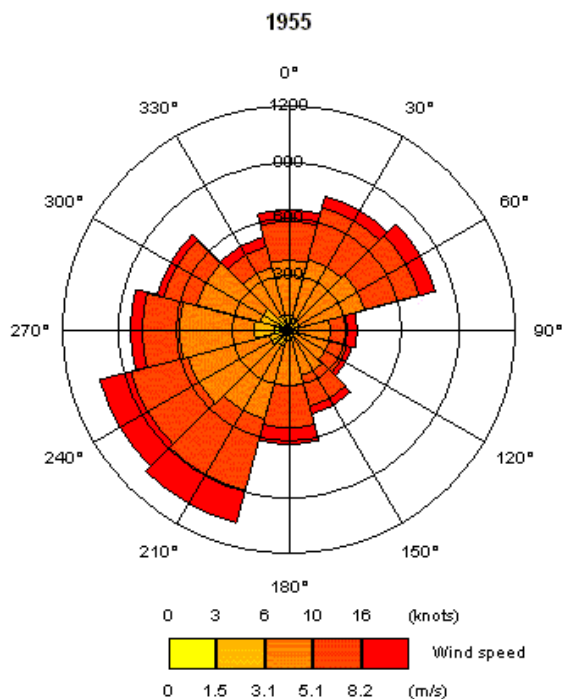
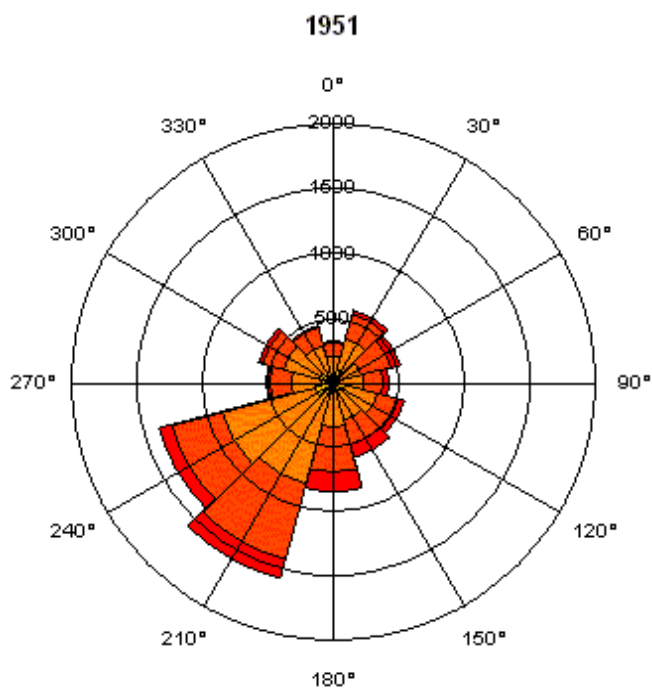
Dispersion plots for Heathrow

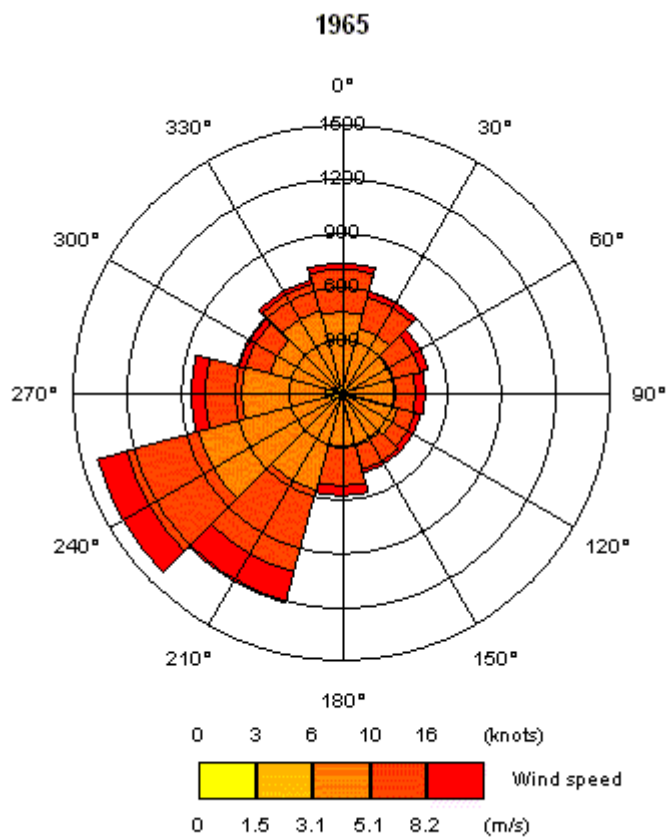
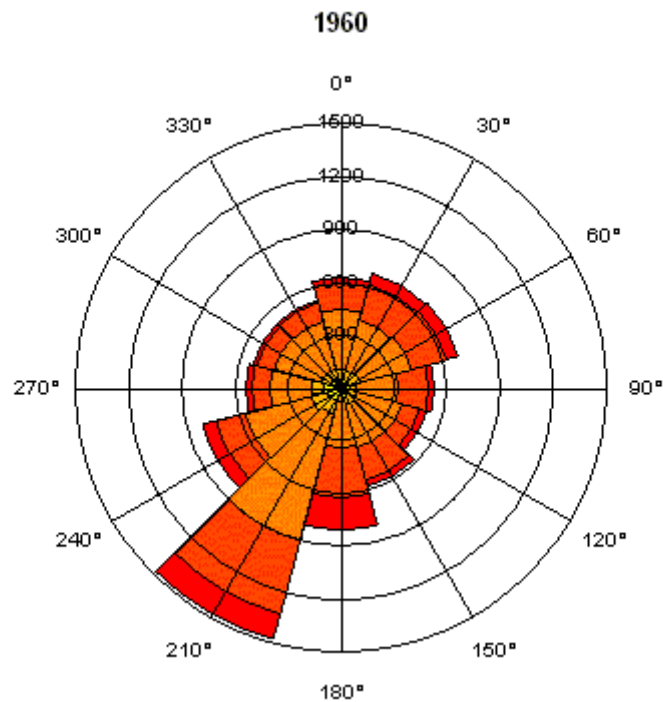


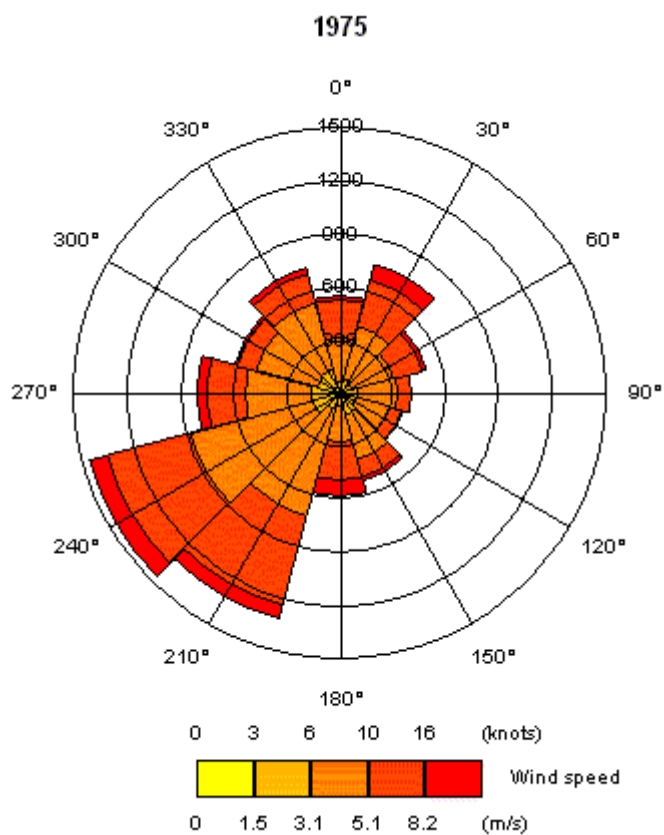
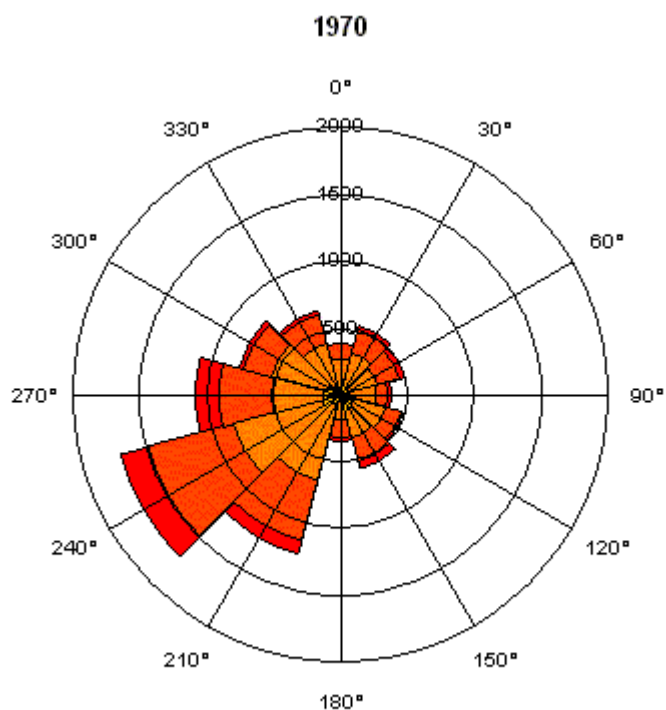


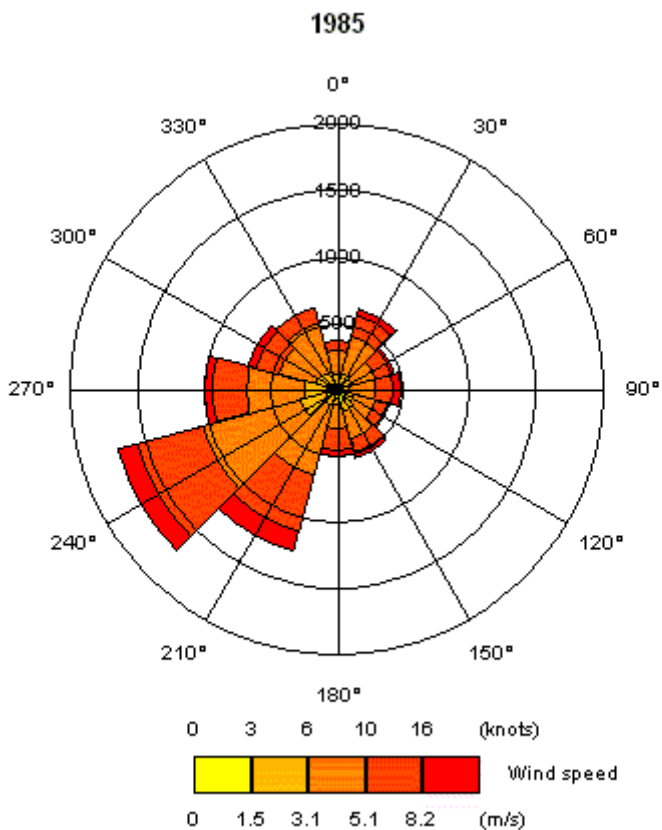
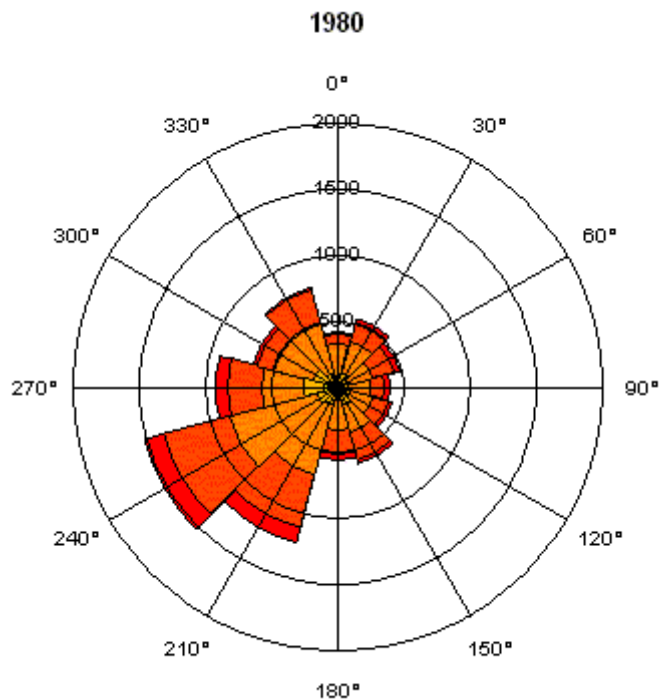


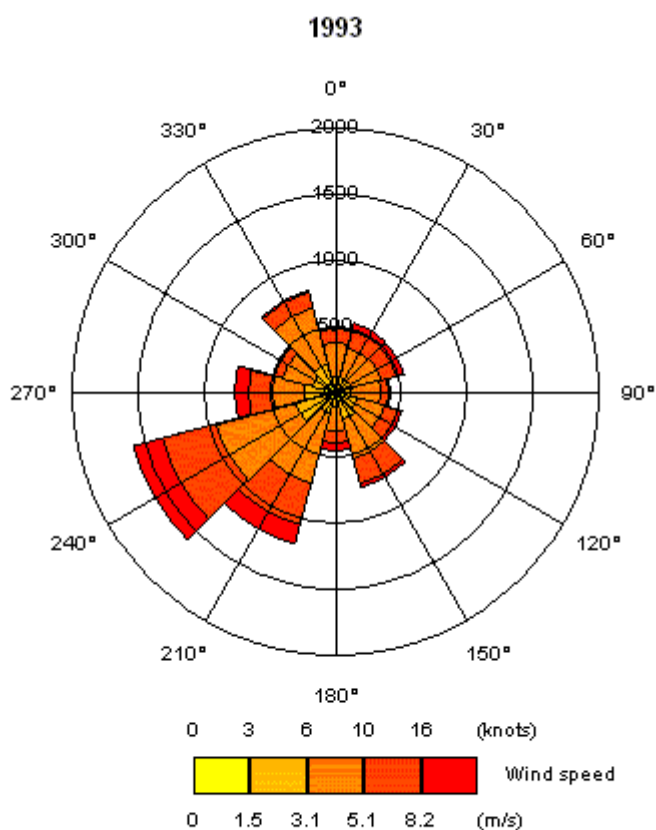
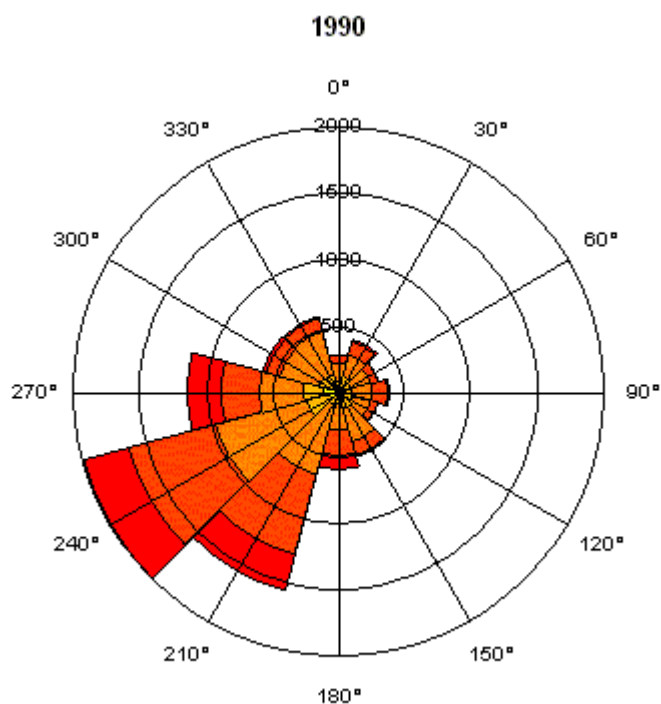
Wind roses for Waddington 1951 - 1999

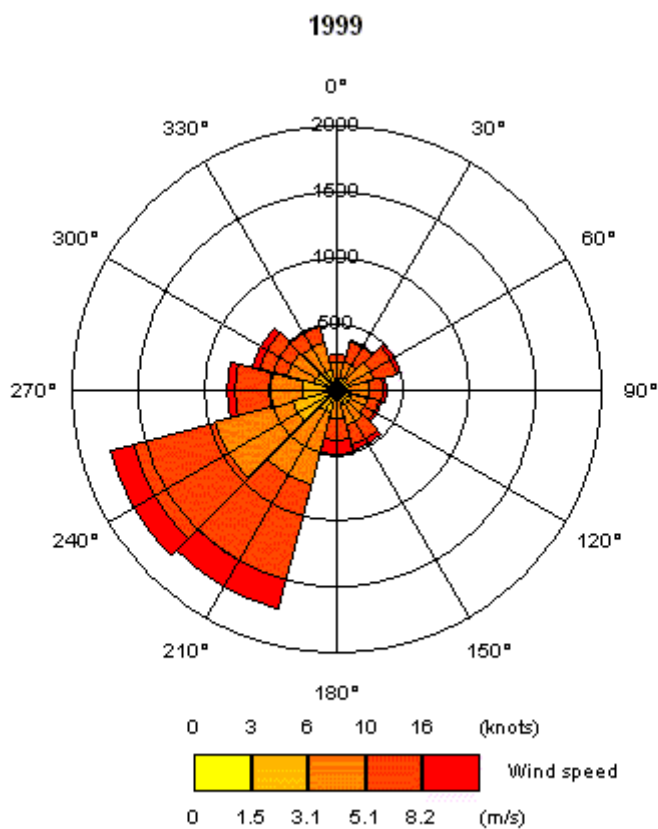
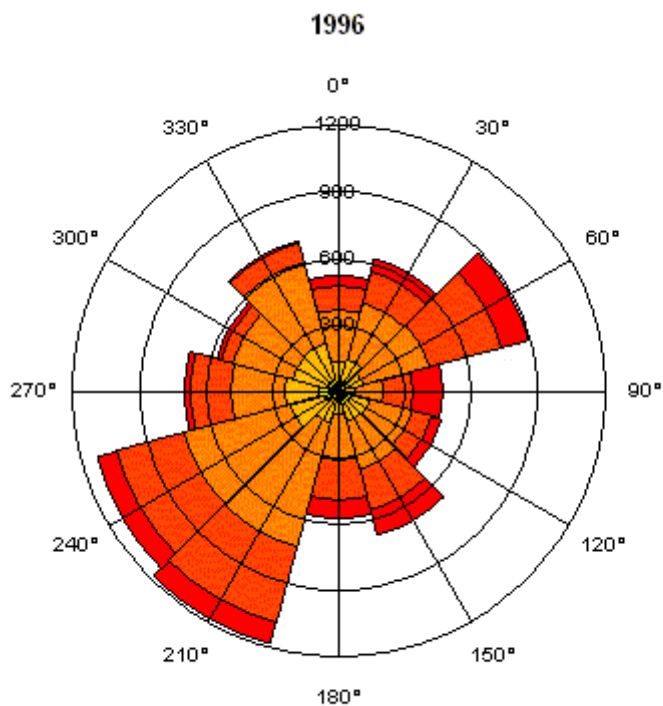






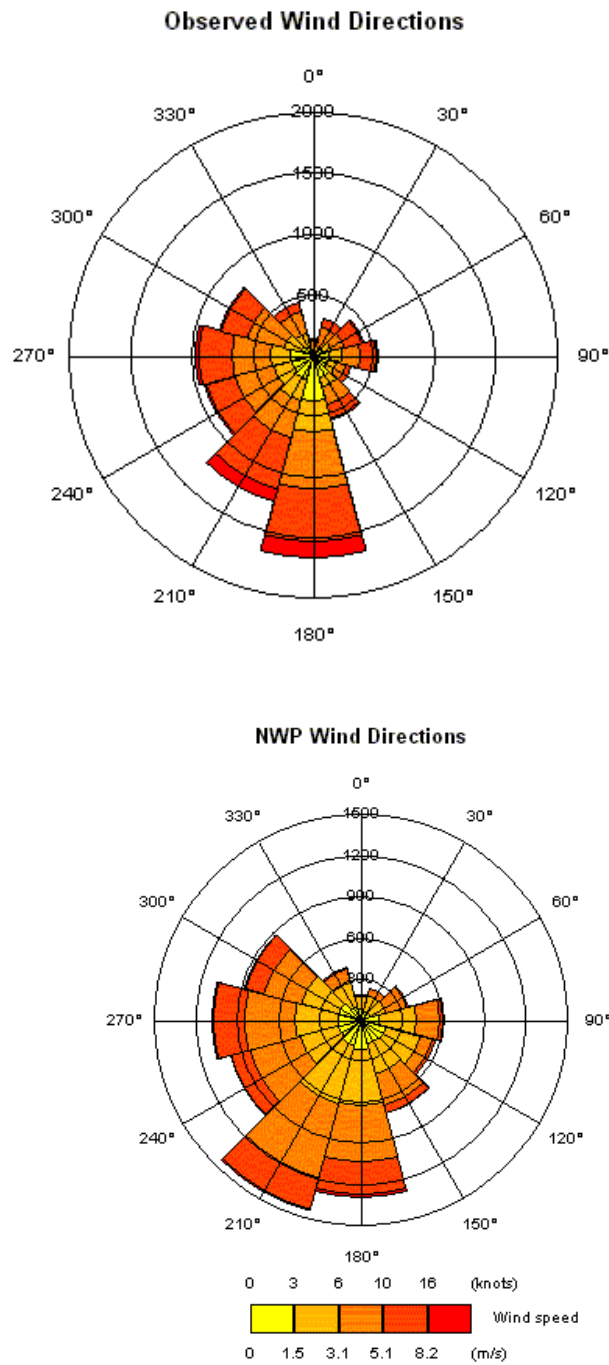




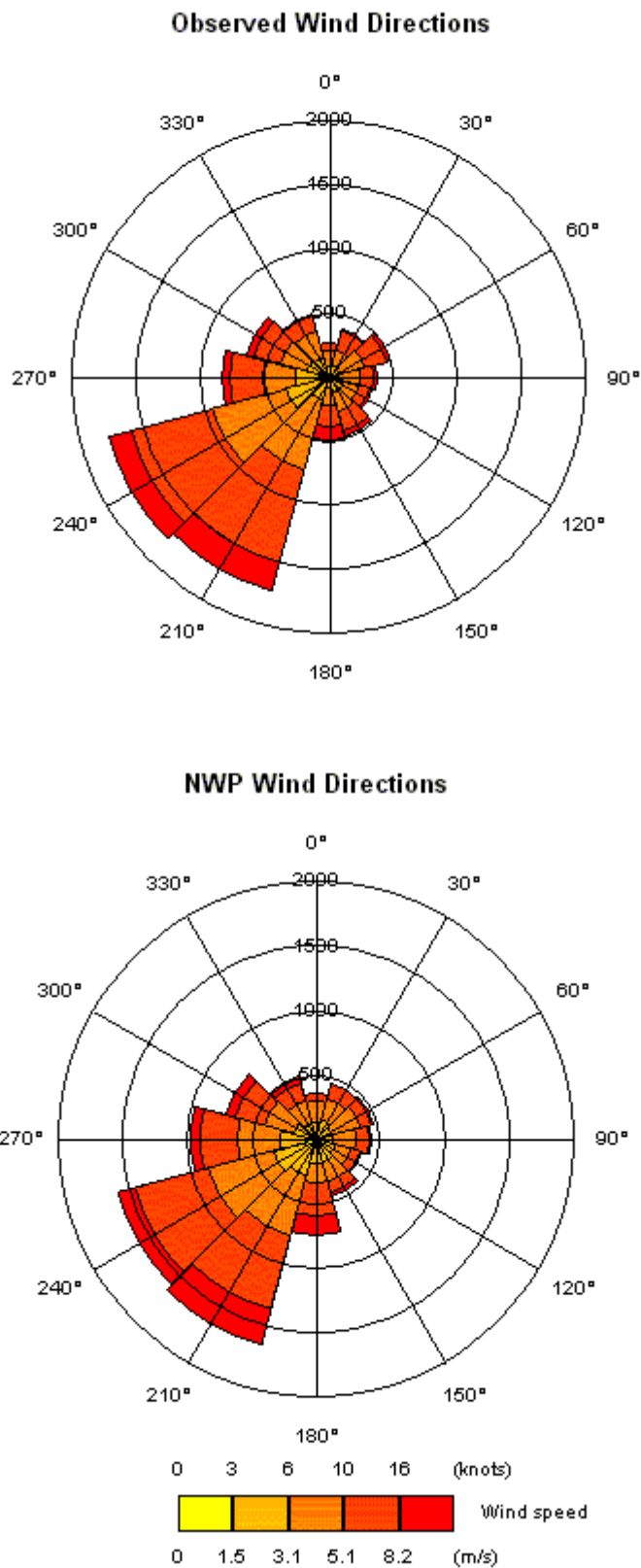


APPENDIX B Comparisons of the NWP meteorology with the observed meteorology

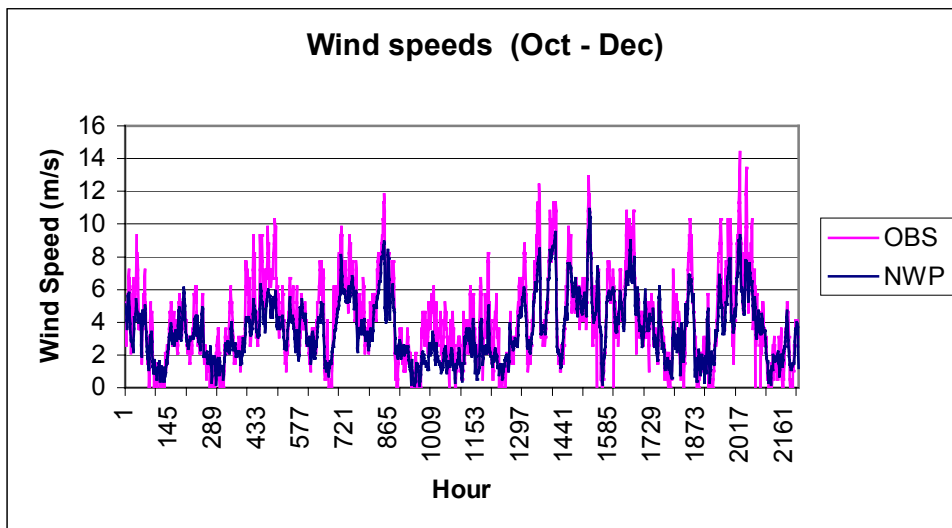
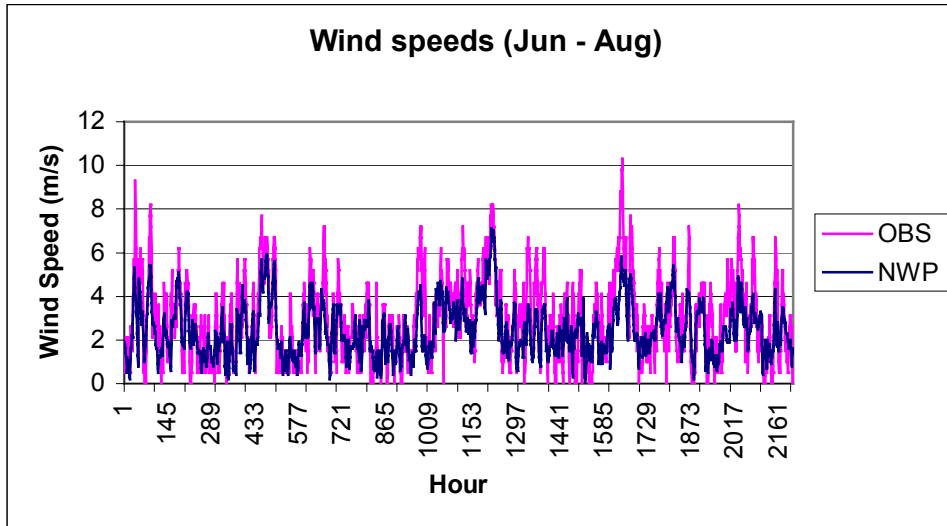
Wind roses for Ringway

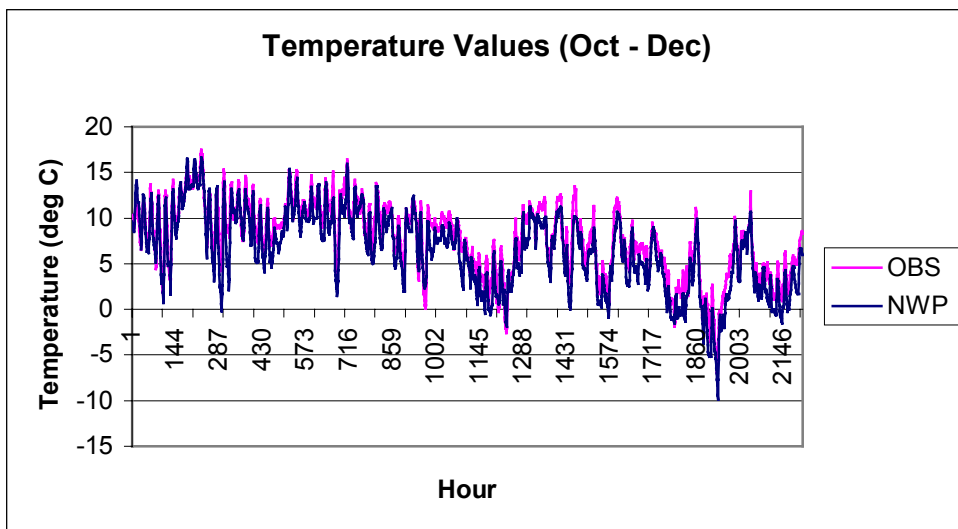
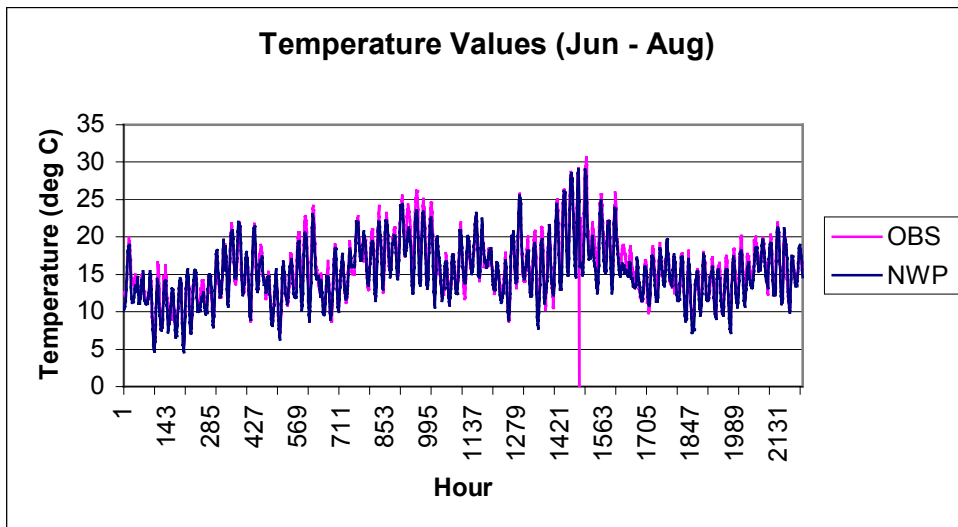


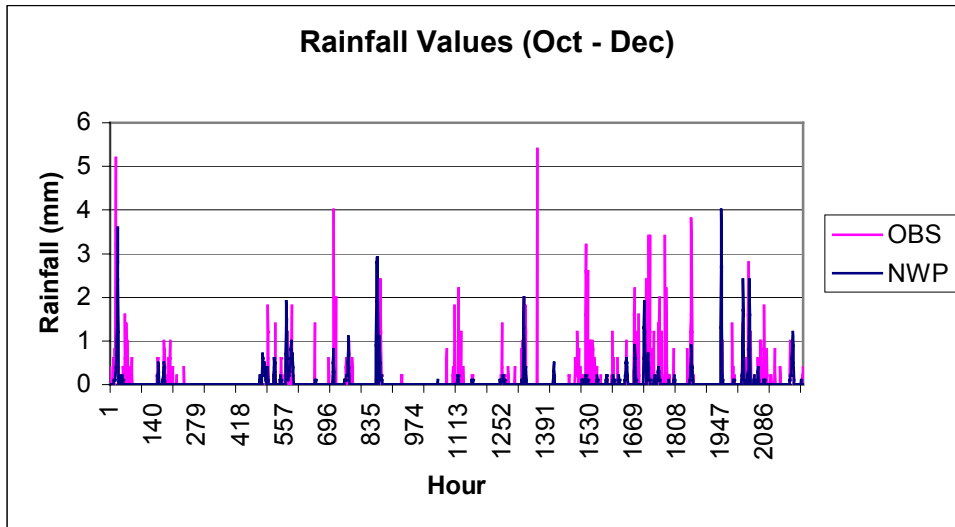
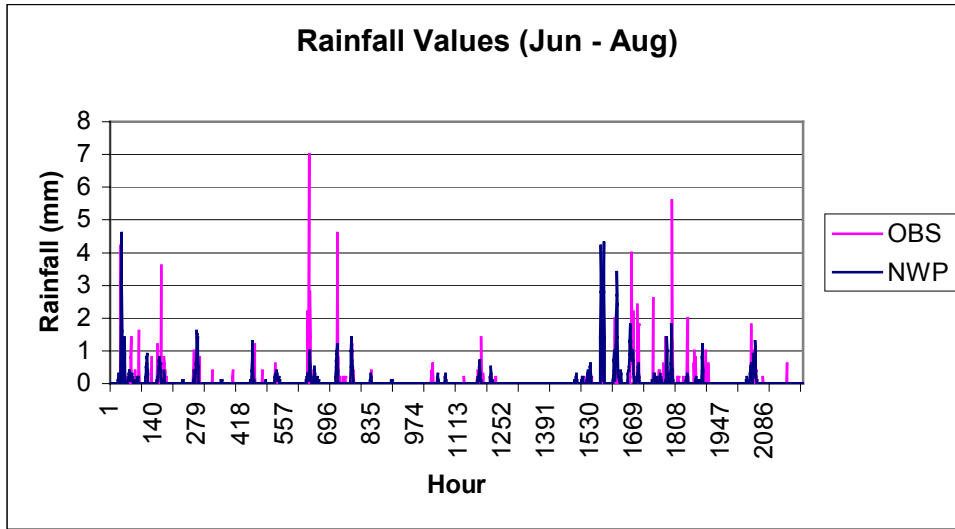
Wind roses for Waddington



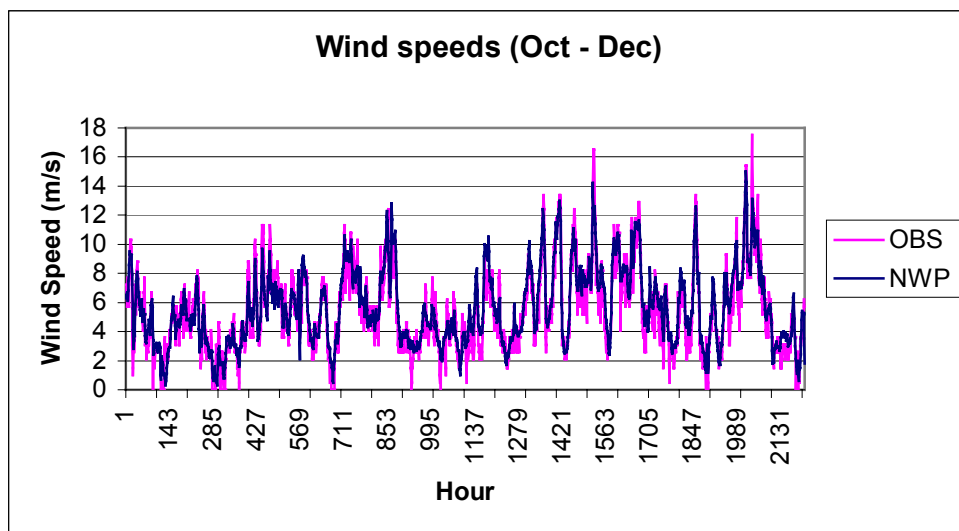
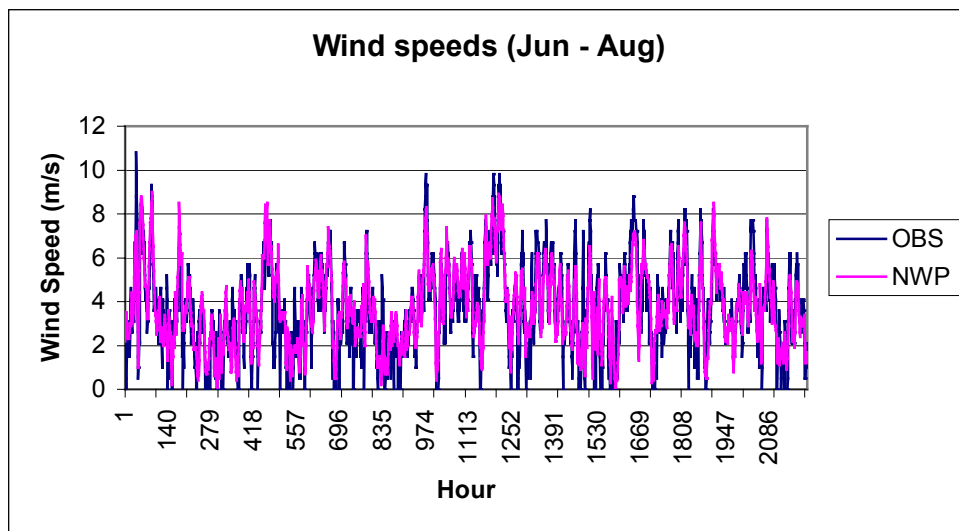
Comparison of various weather parameters for Ringway

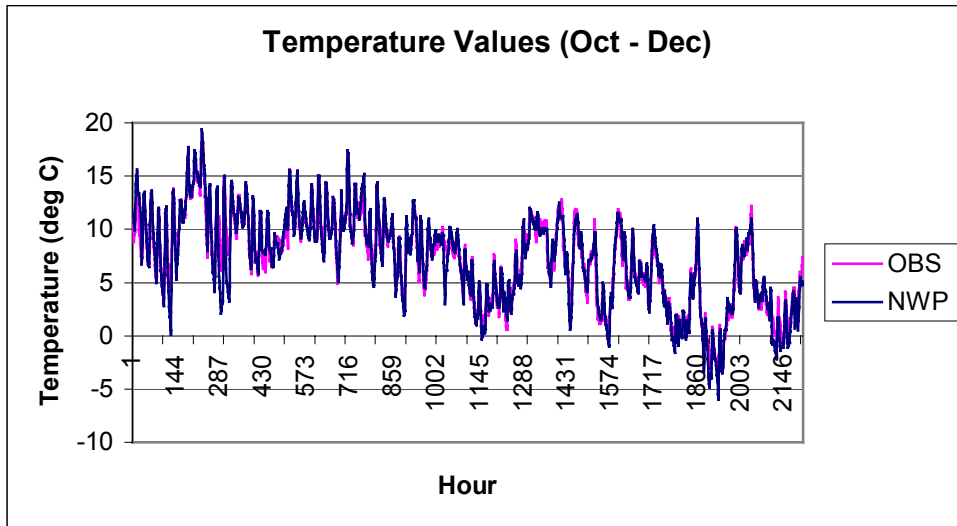
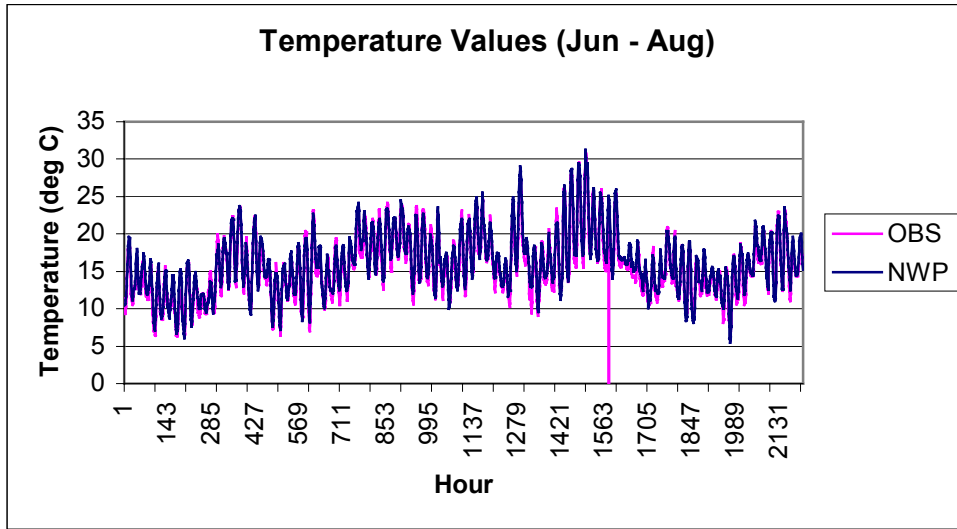


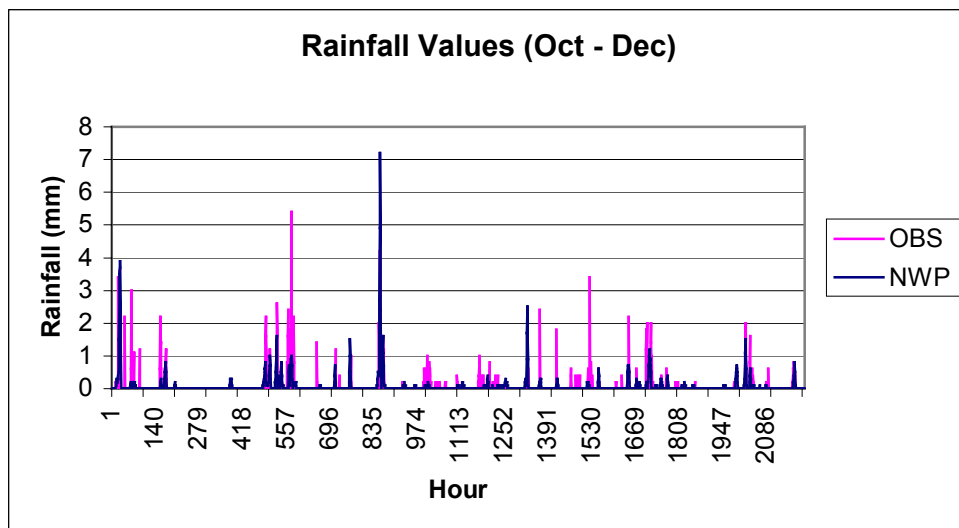
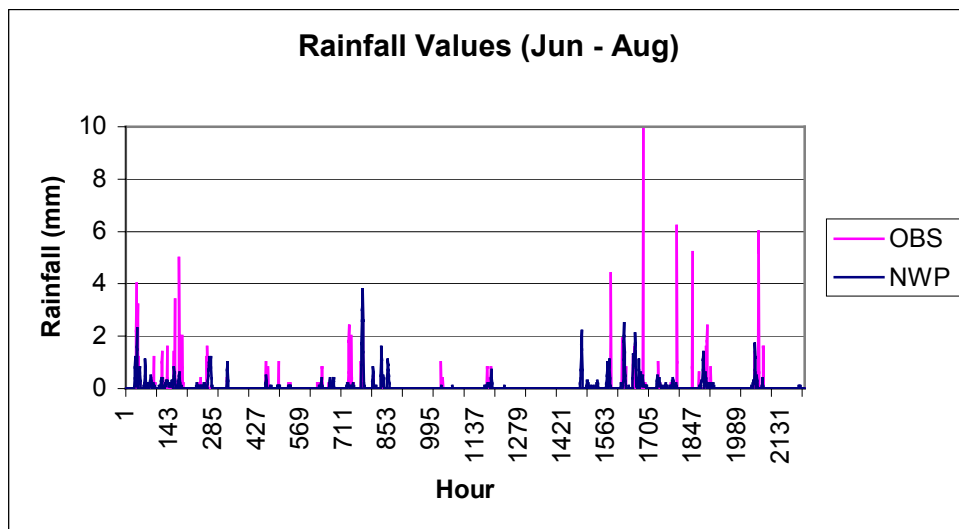




Comparison of various weather parameters for Waddington

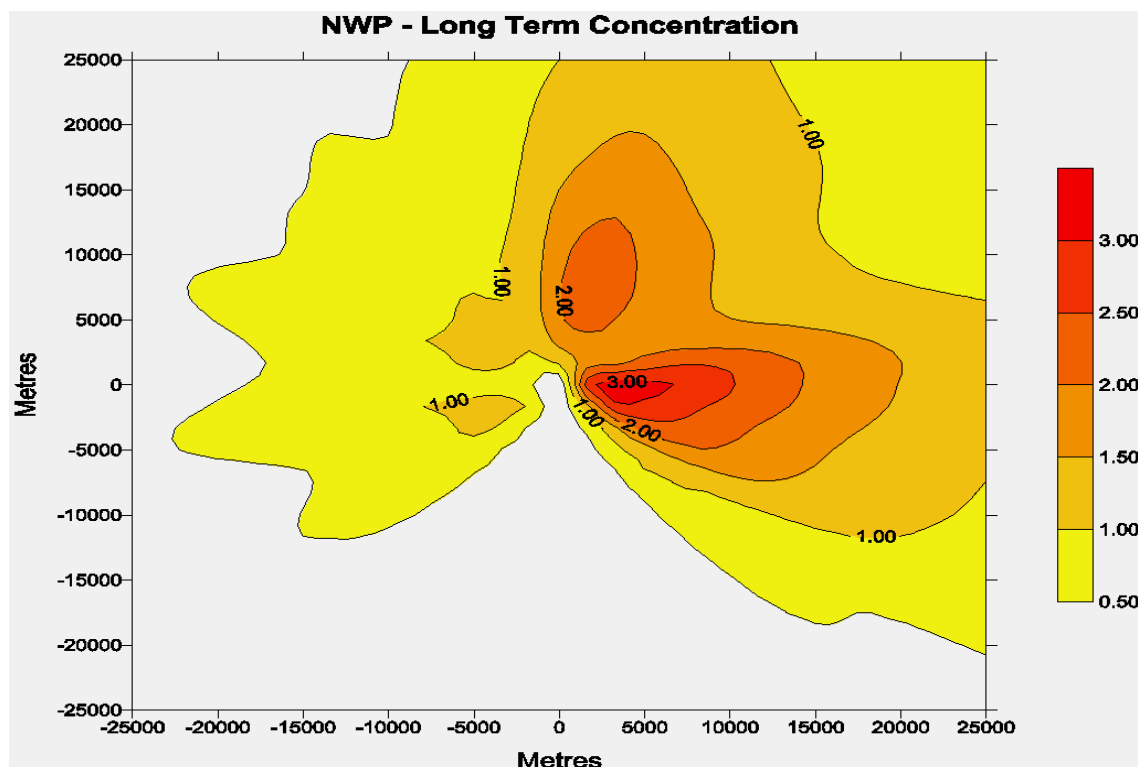
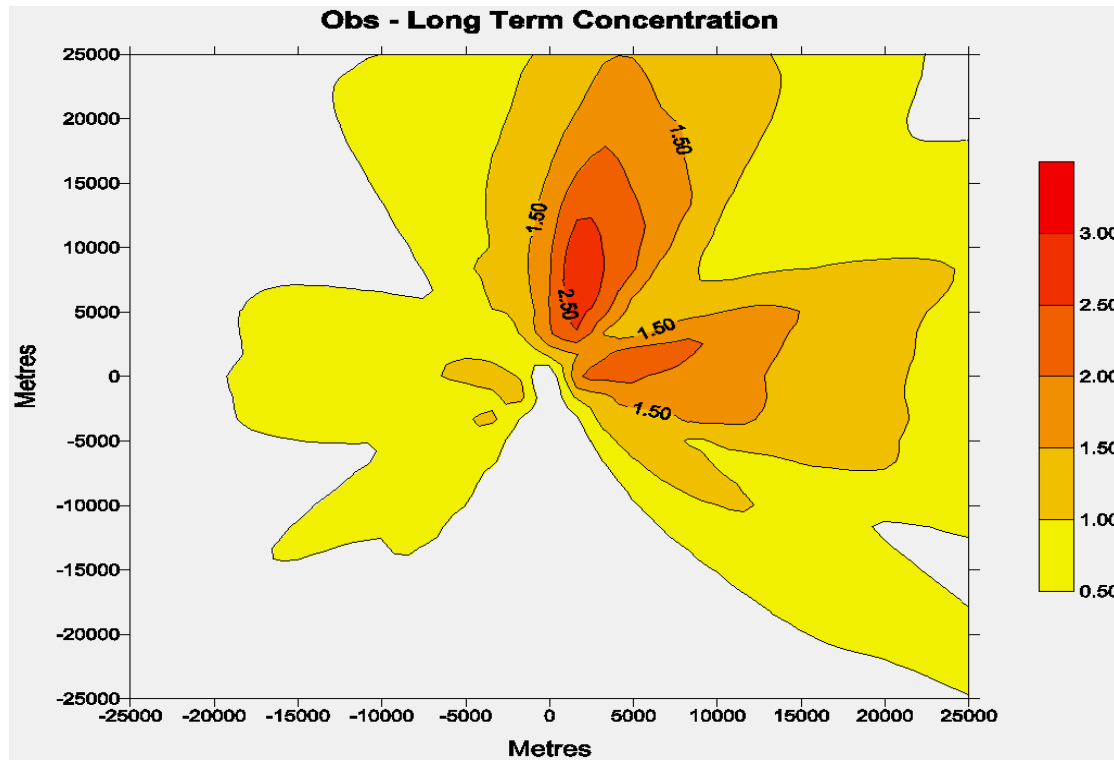


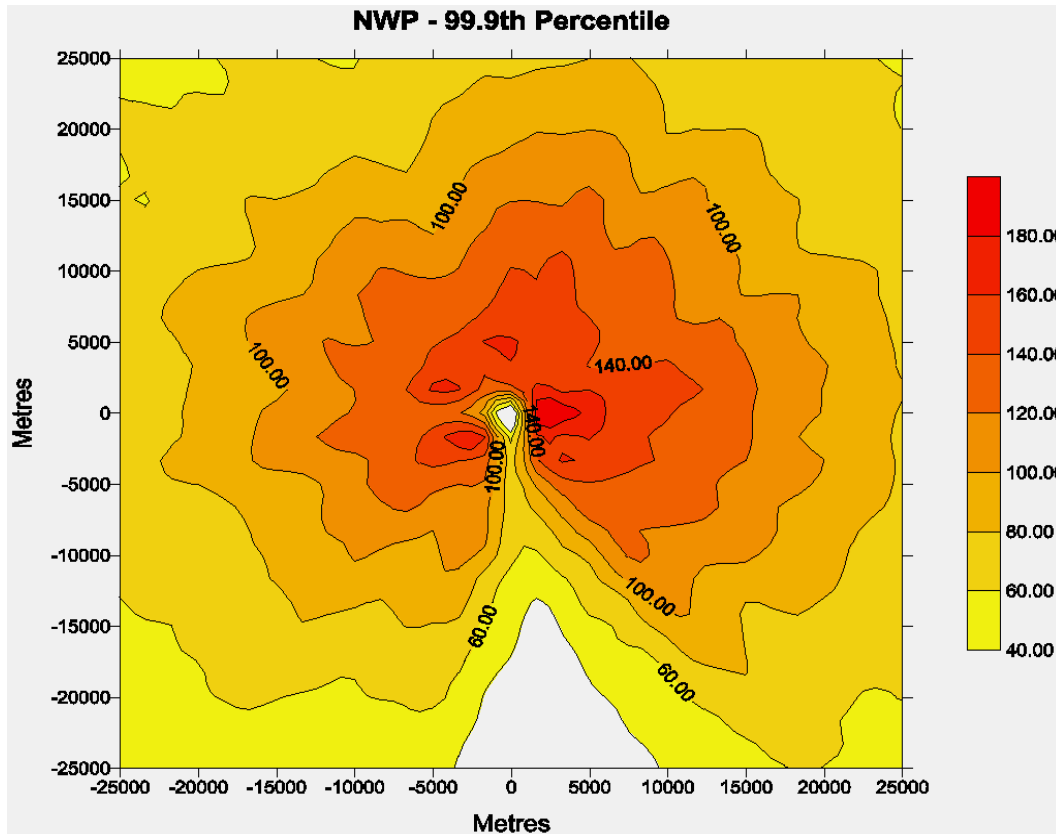
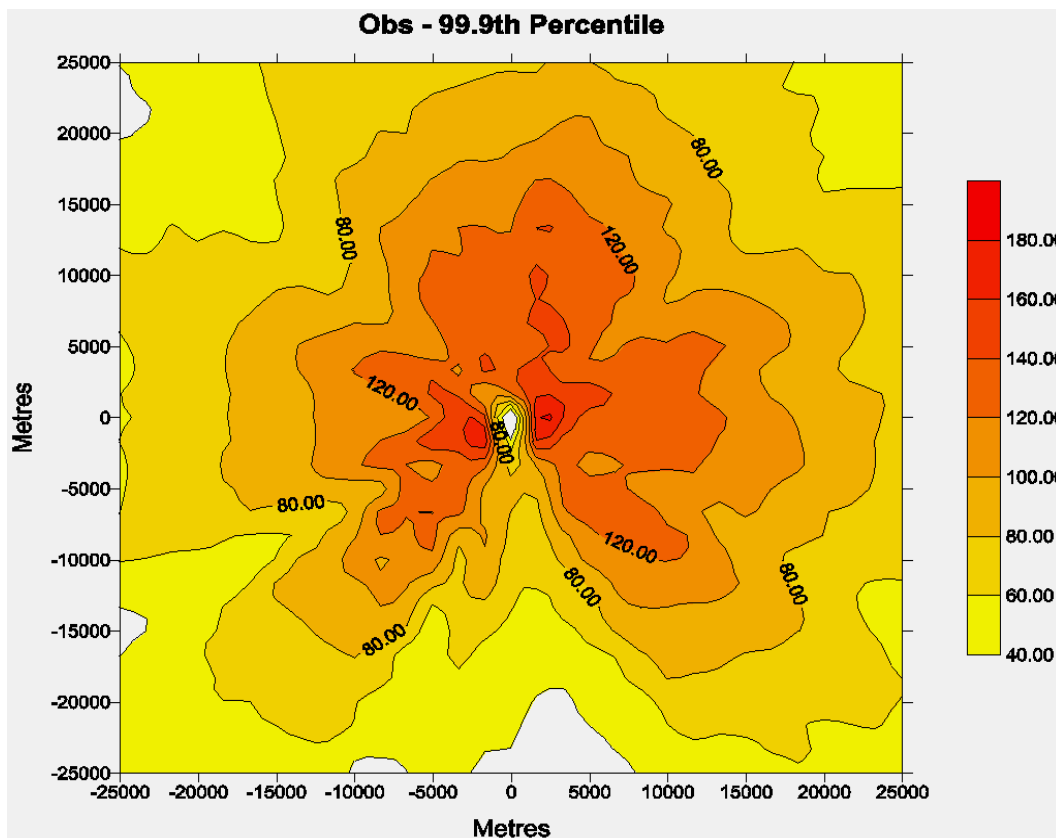


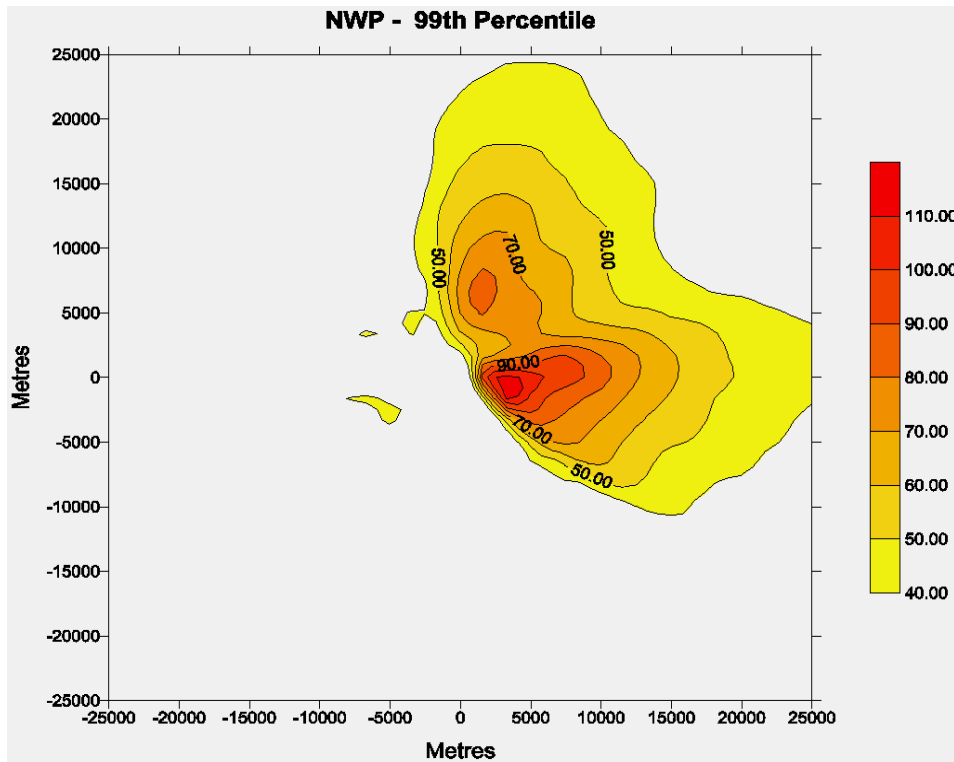
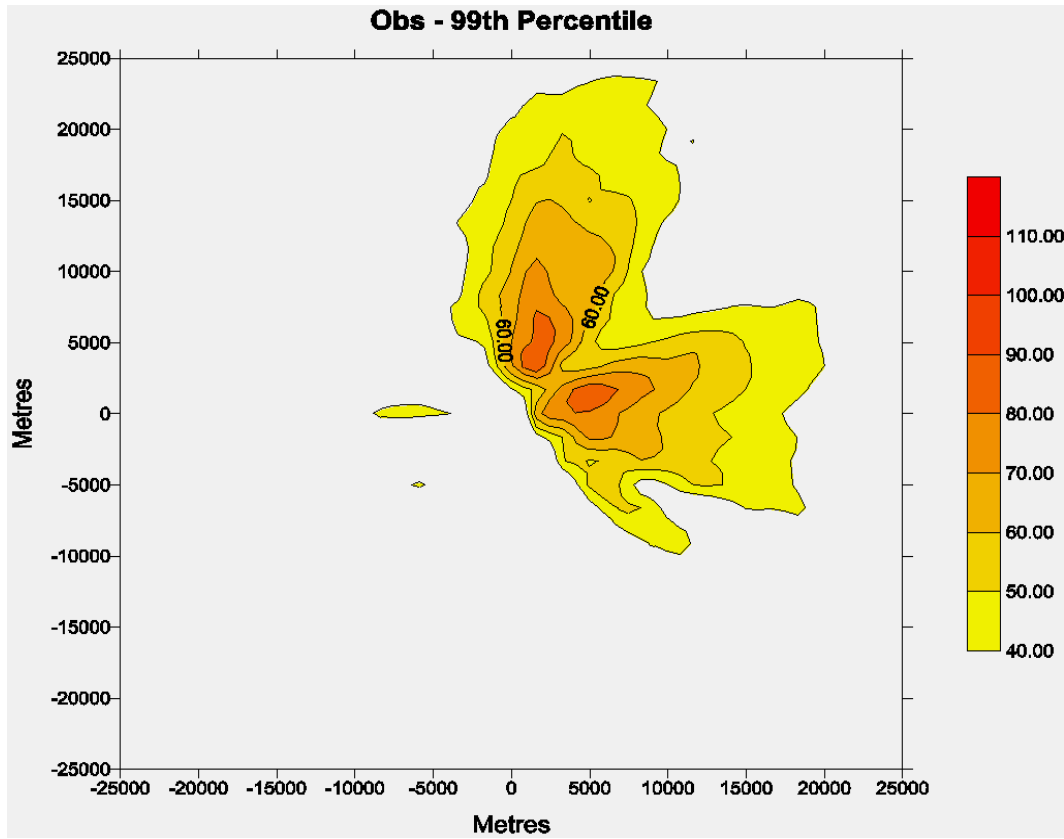


Dispersion plots produced by Observed and NWP data for Ringway at Source 1

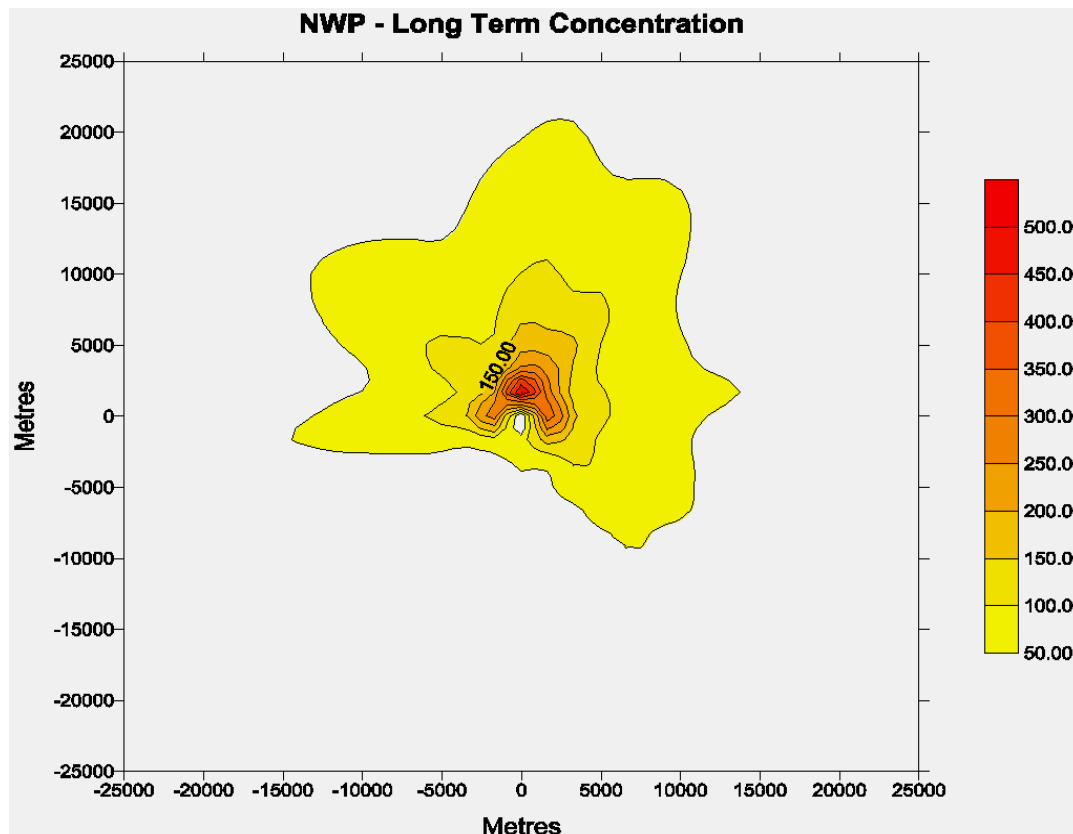
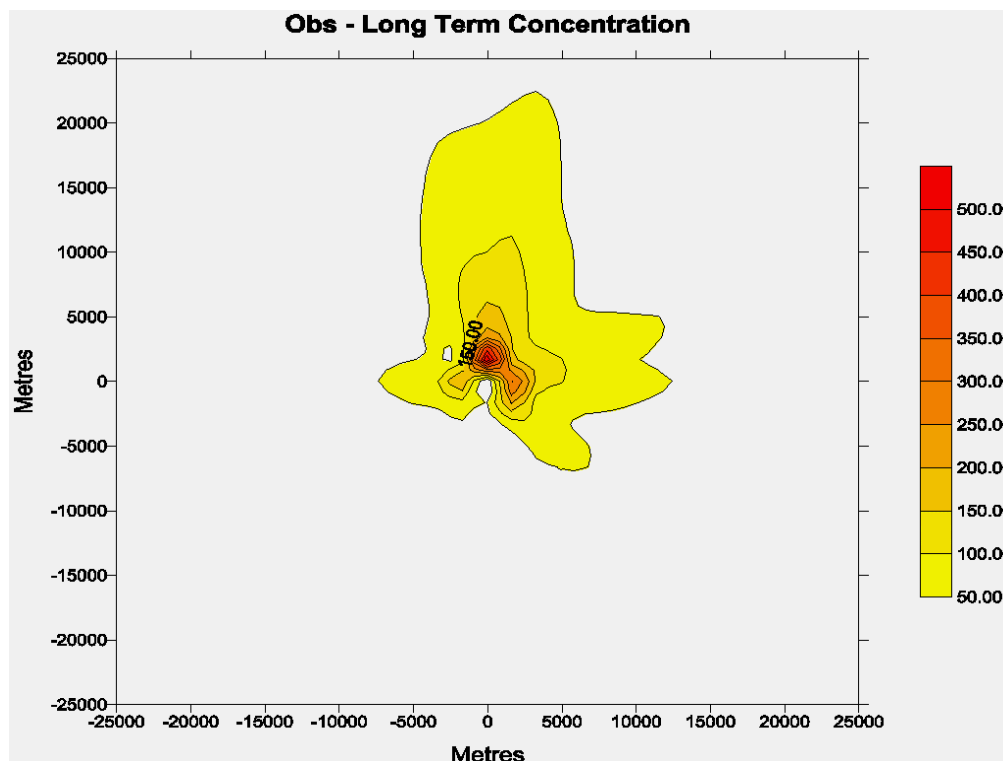
These plots are of the long term concentration, the 99.9th percentile, or the 99th percentile at each grid point. Where there is no shading on the plots the values are less than the lowest contour value plotted.

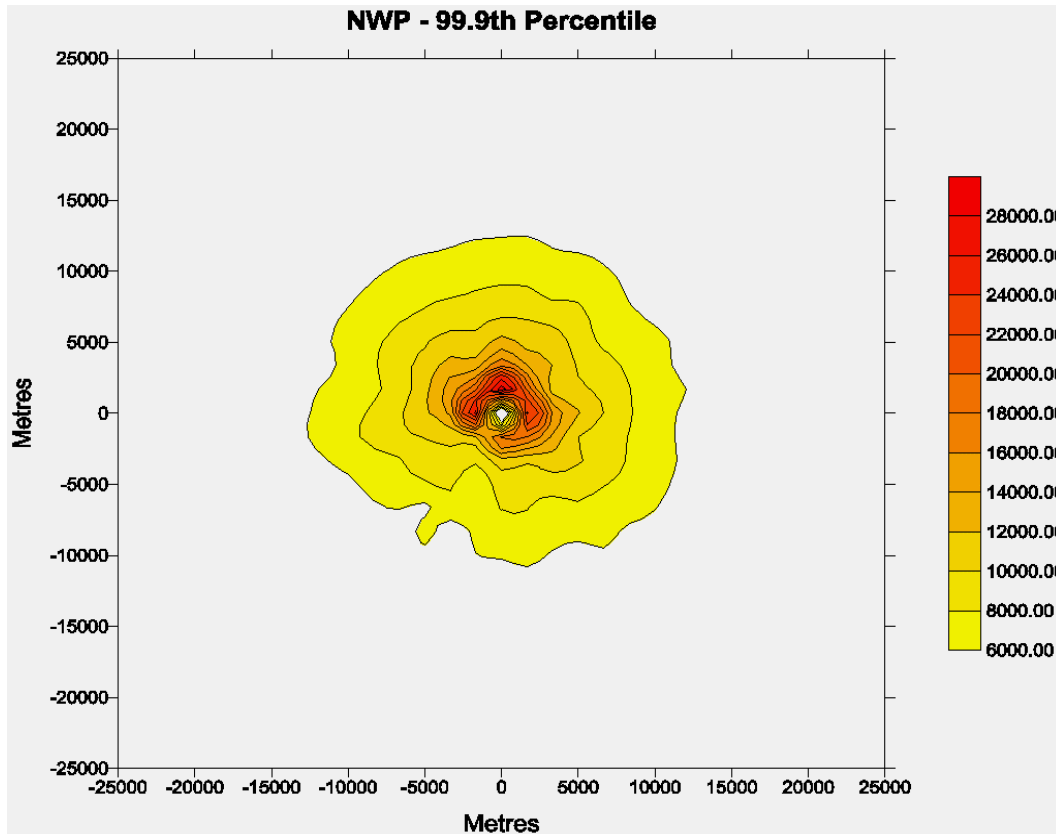
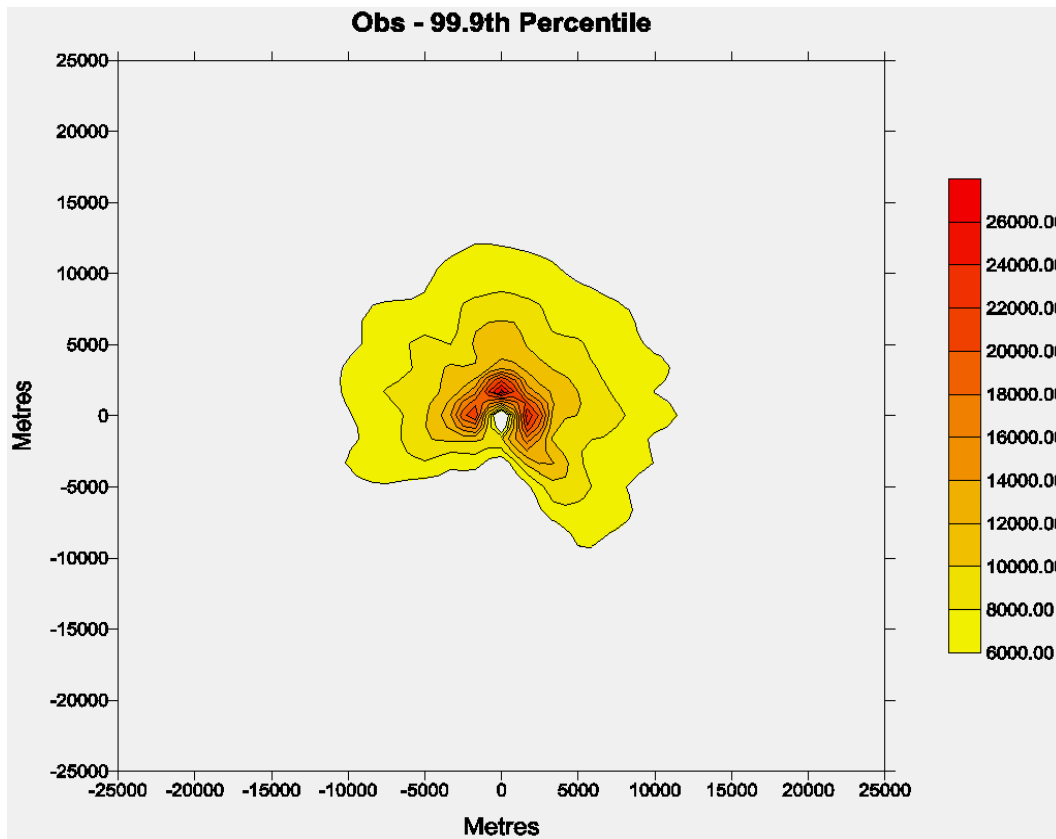


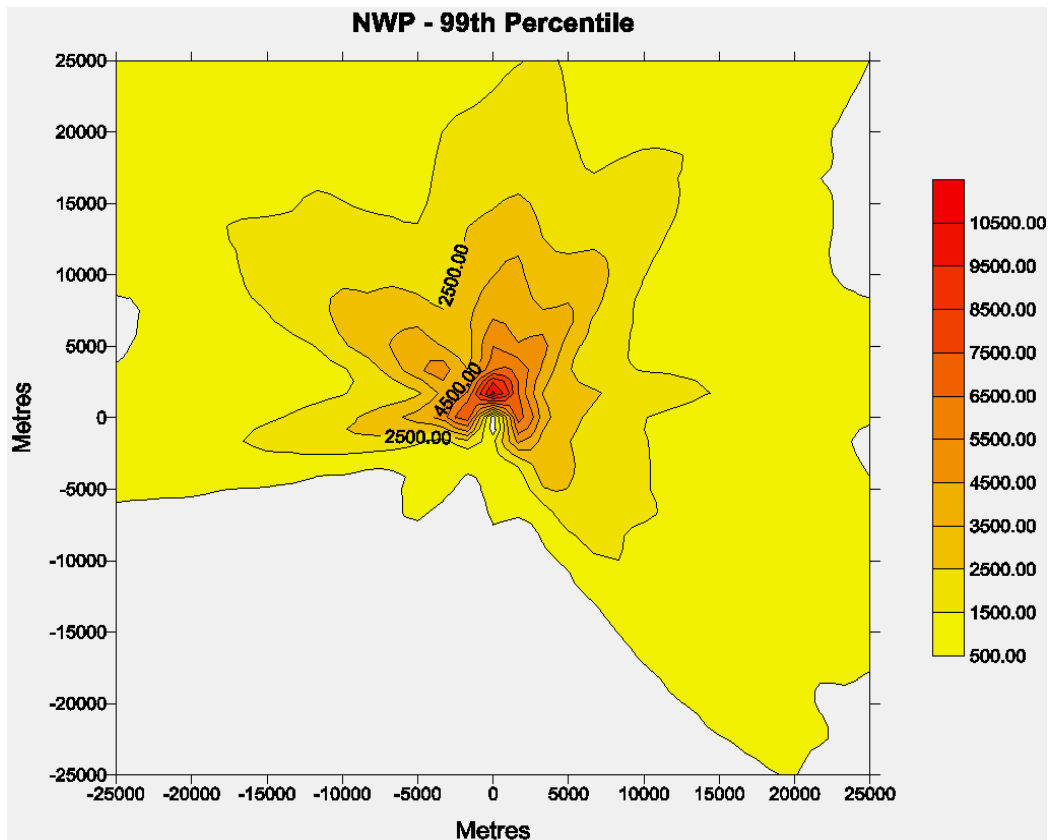
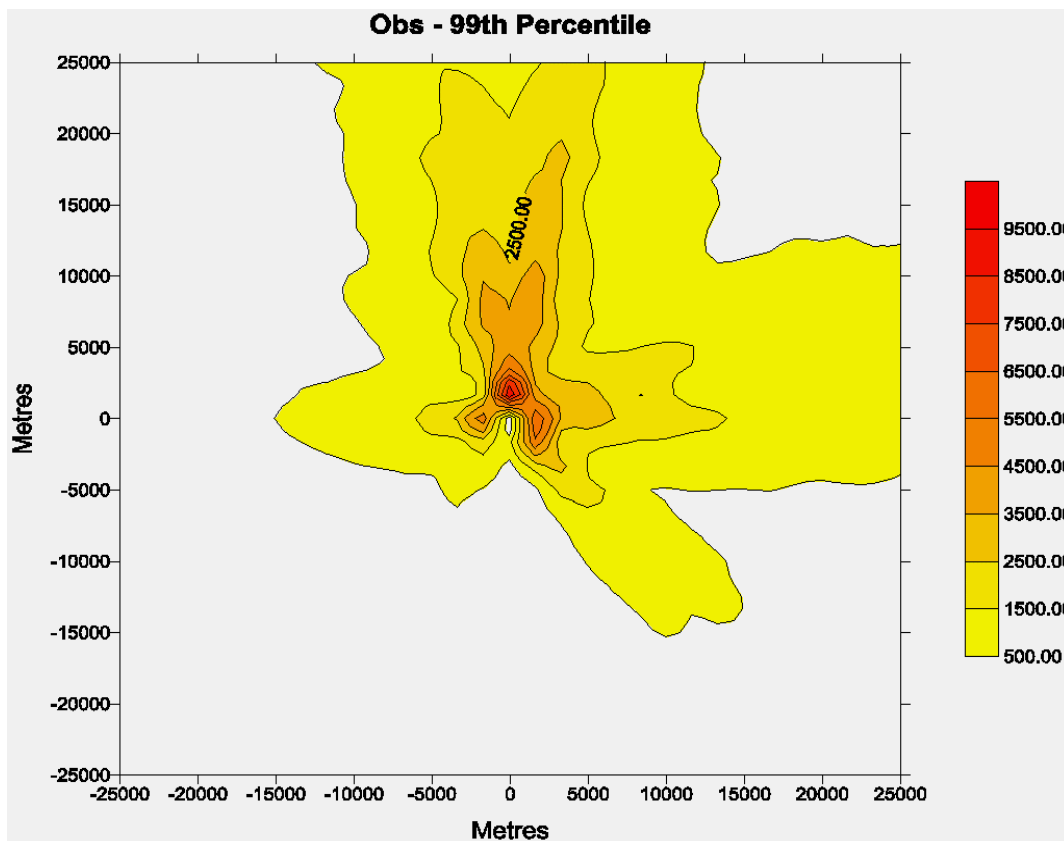




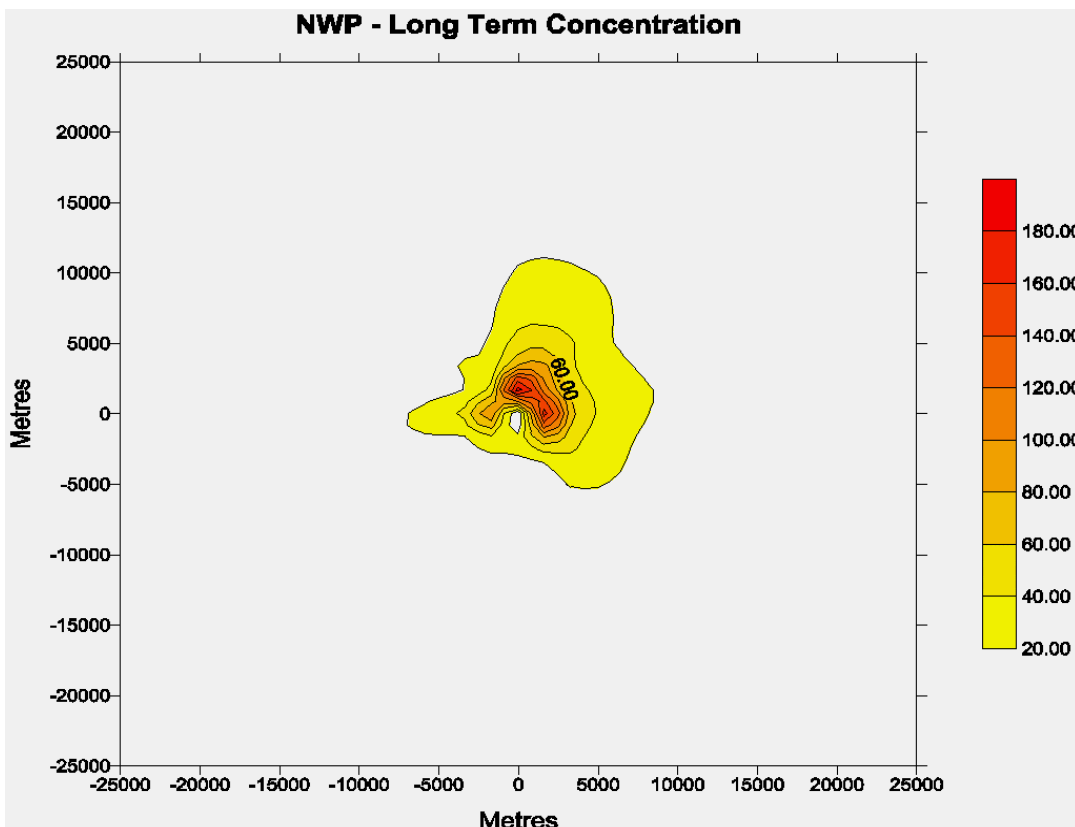
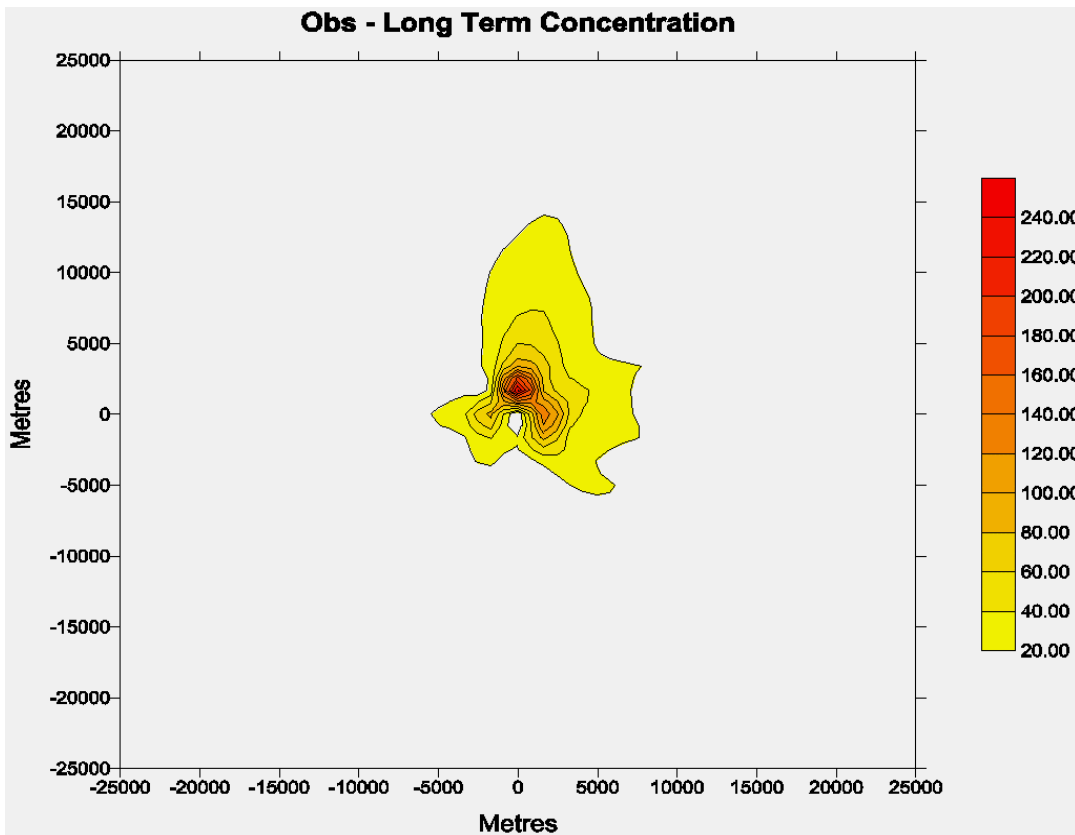
Dispersion plots produced by Observed and NWP data for Ringway at Source 2

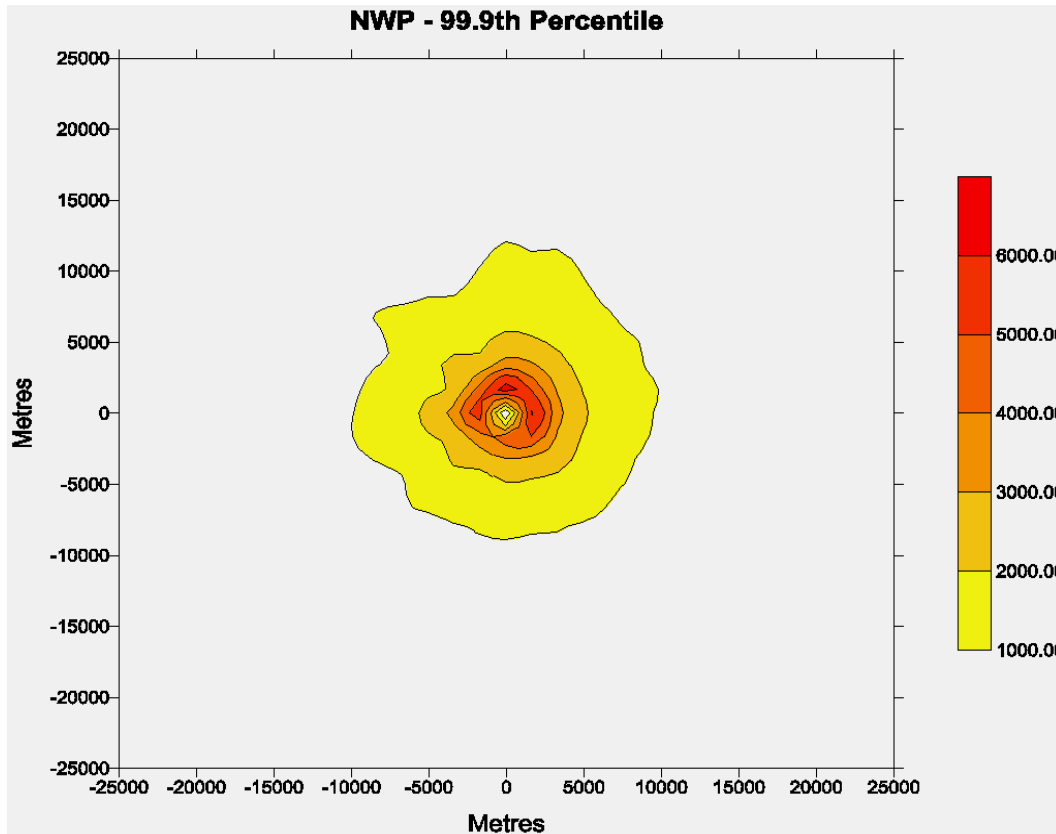
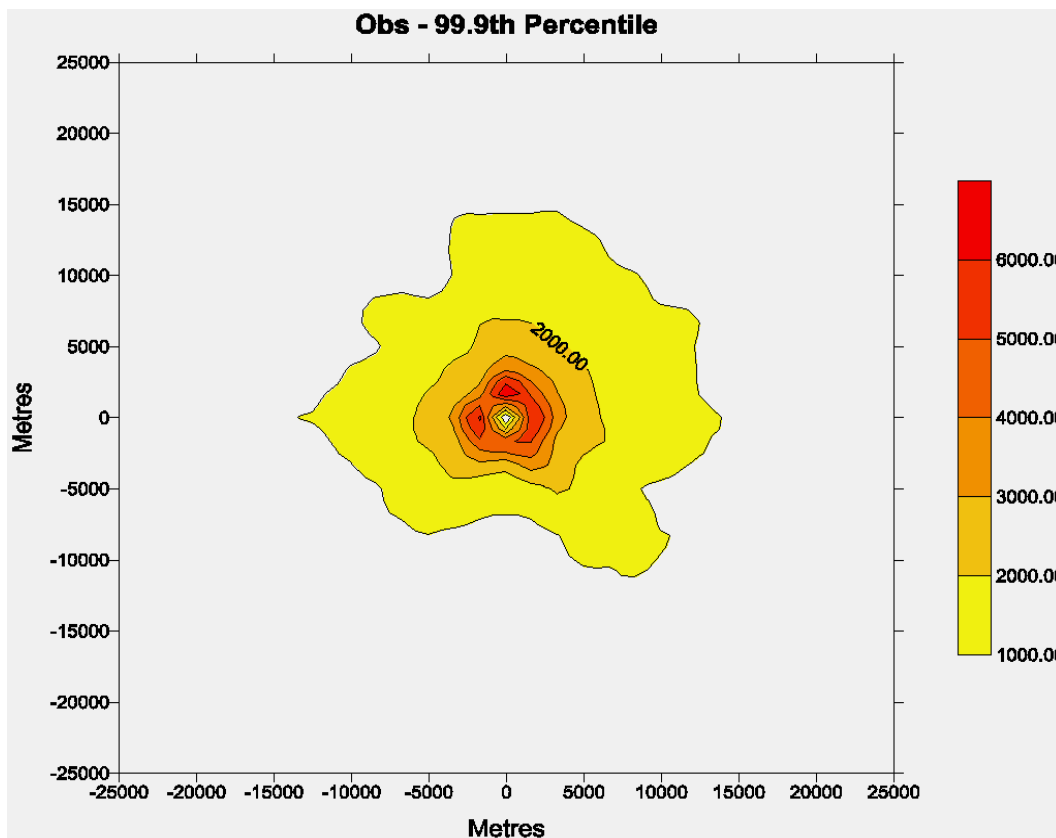


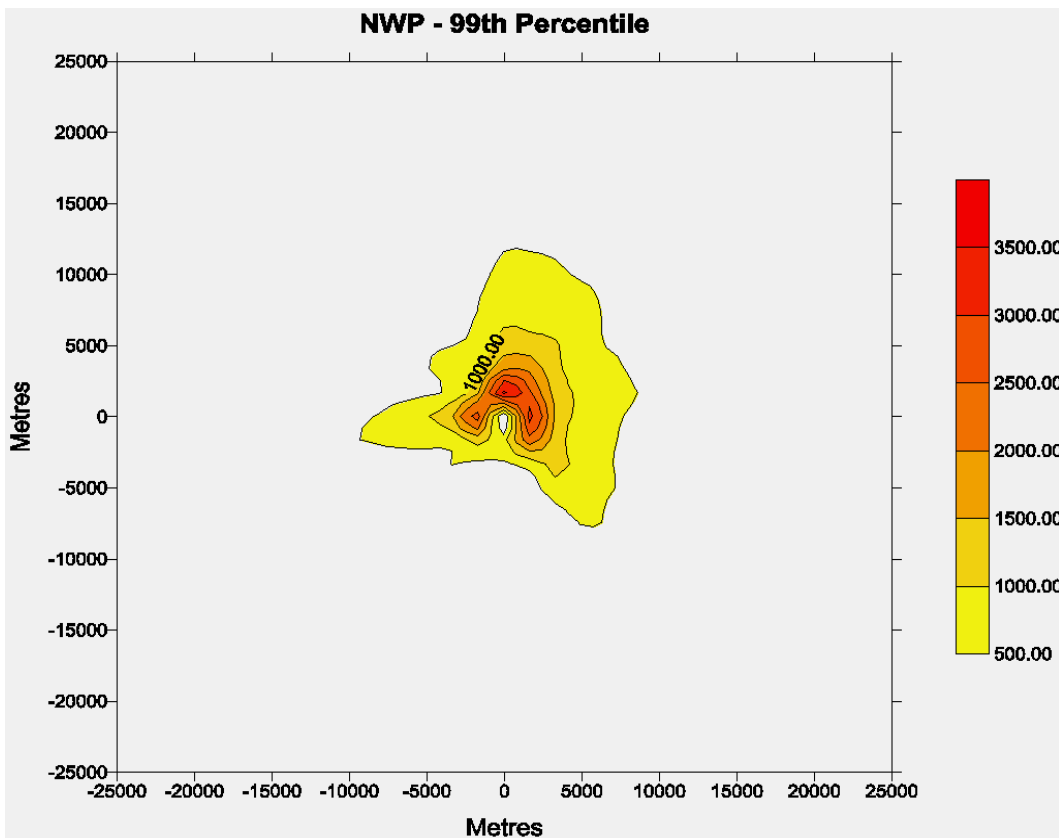
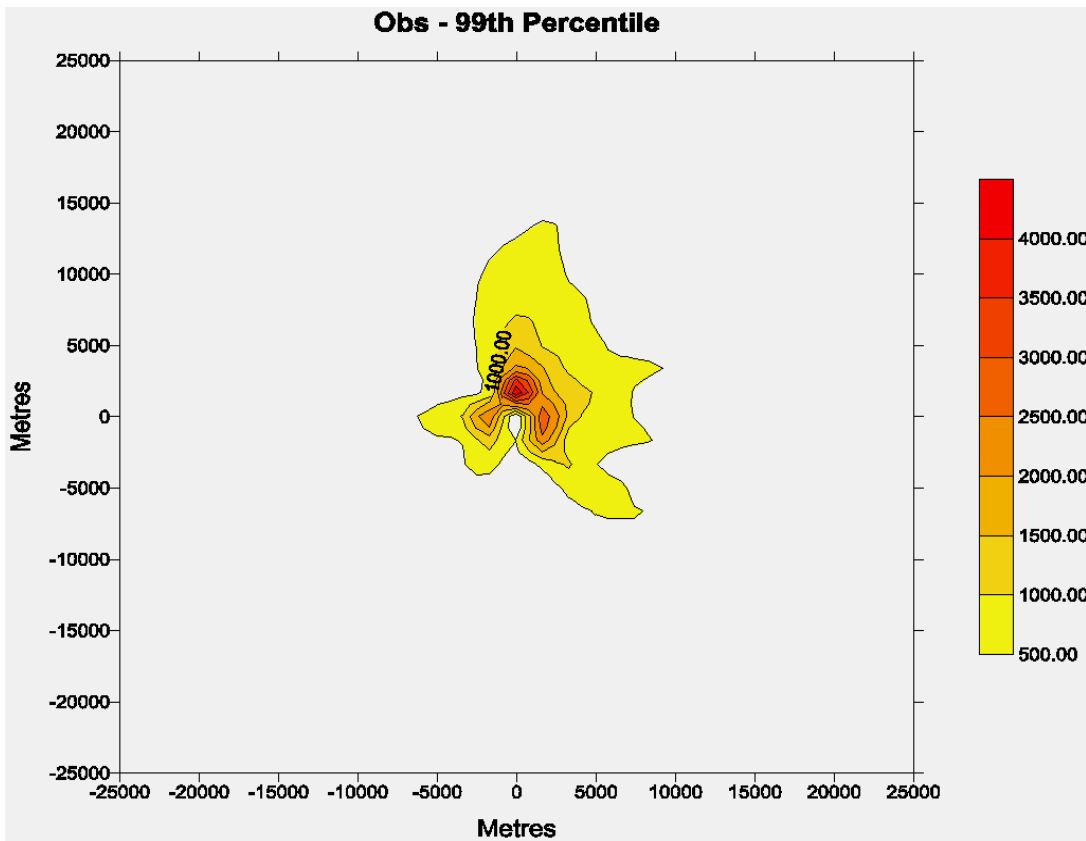




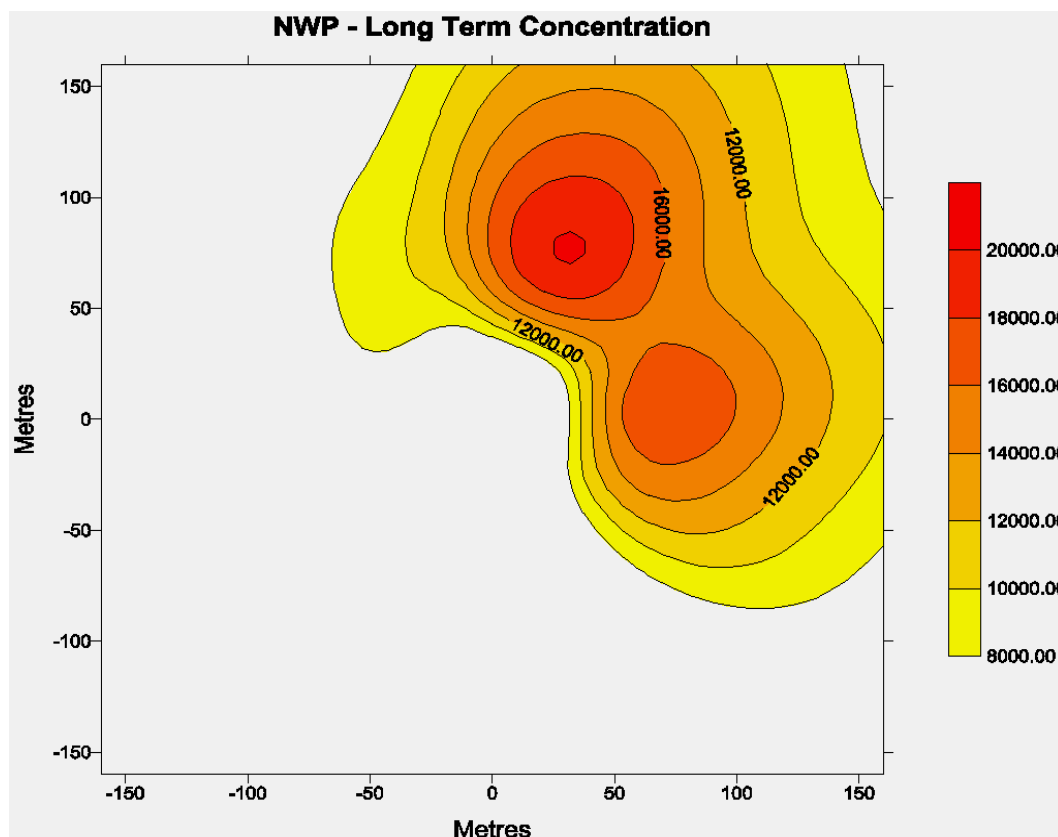
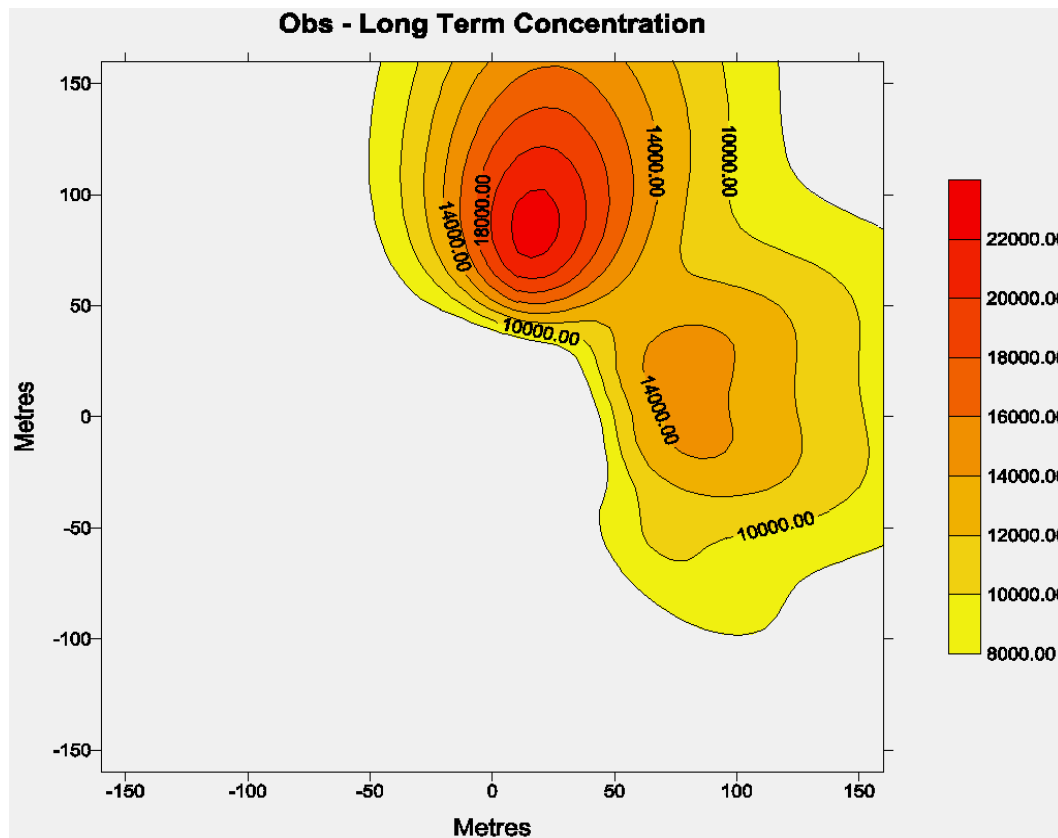
Dispersion plots produced by Observed and NWP data for Ringway at Source 3

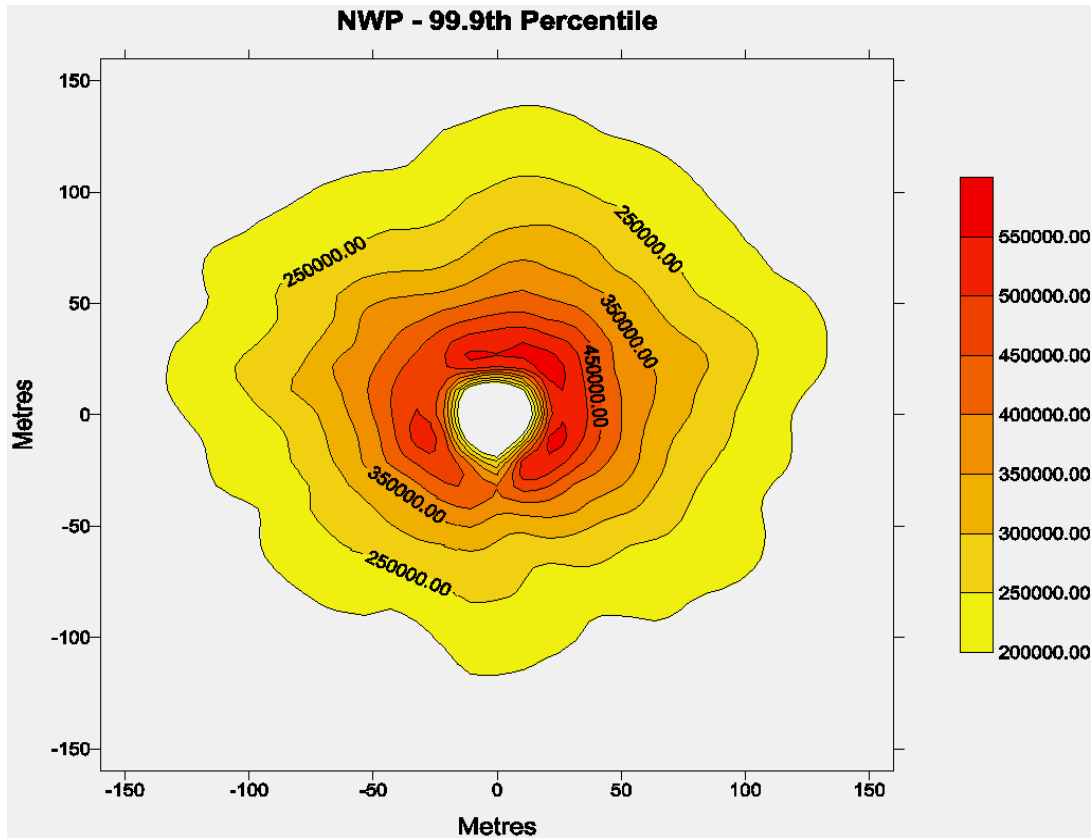
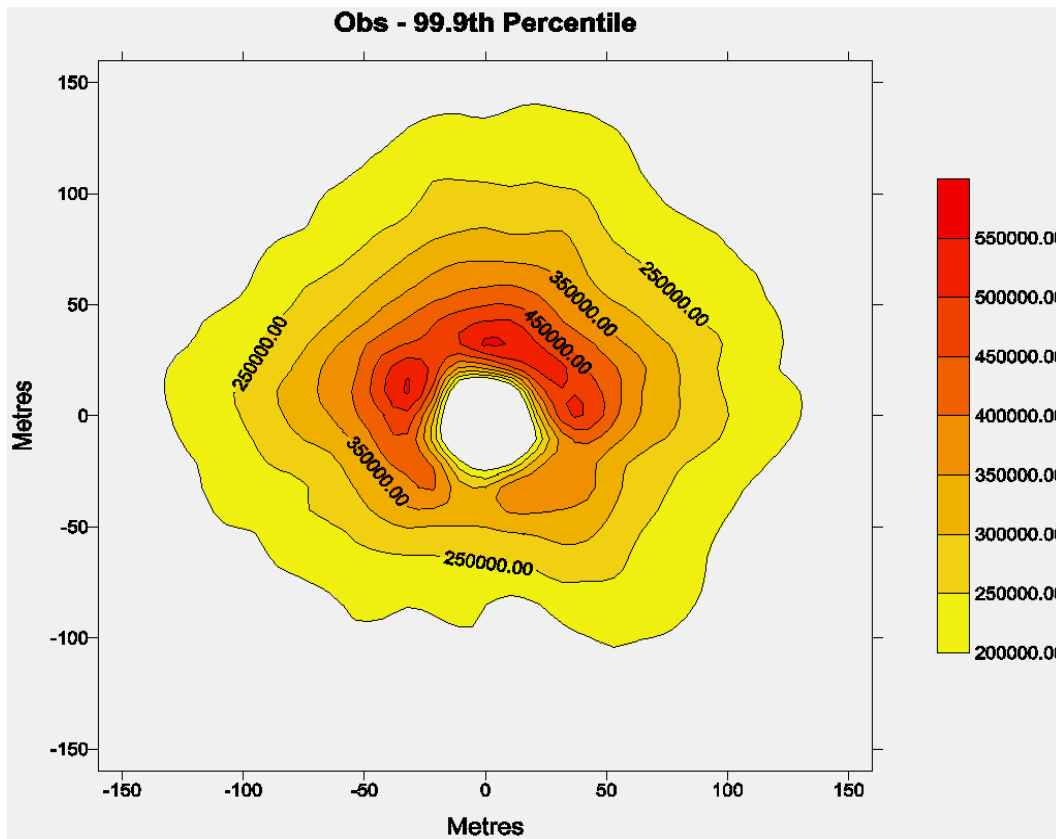


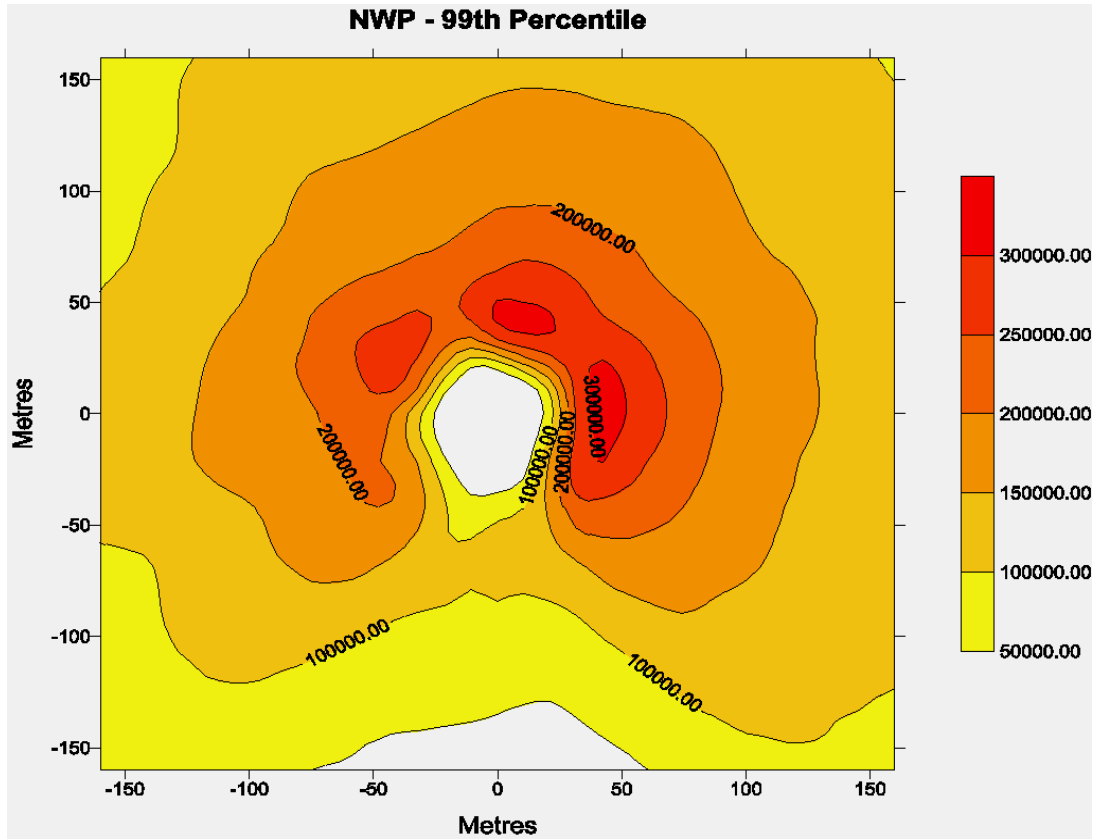
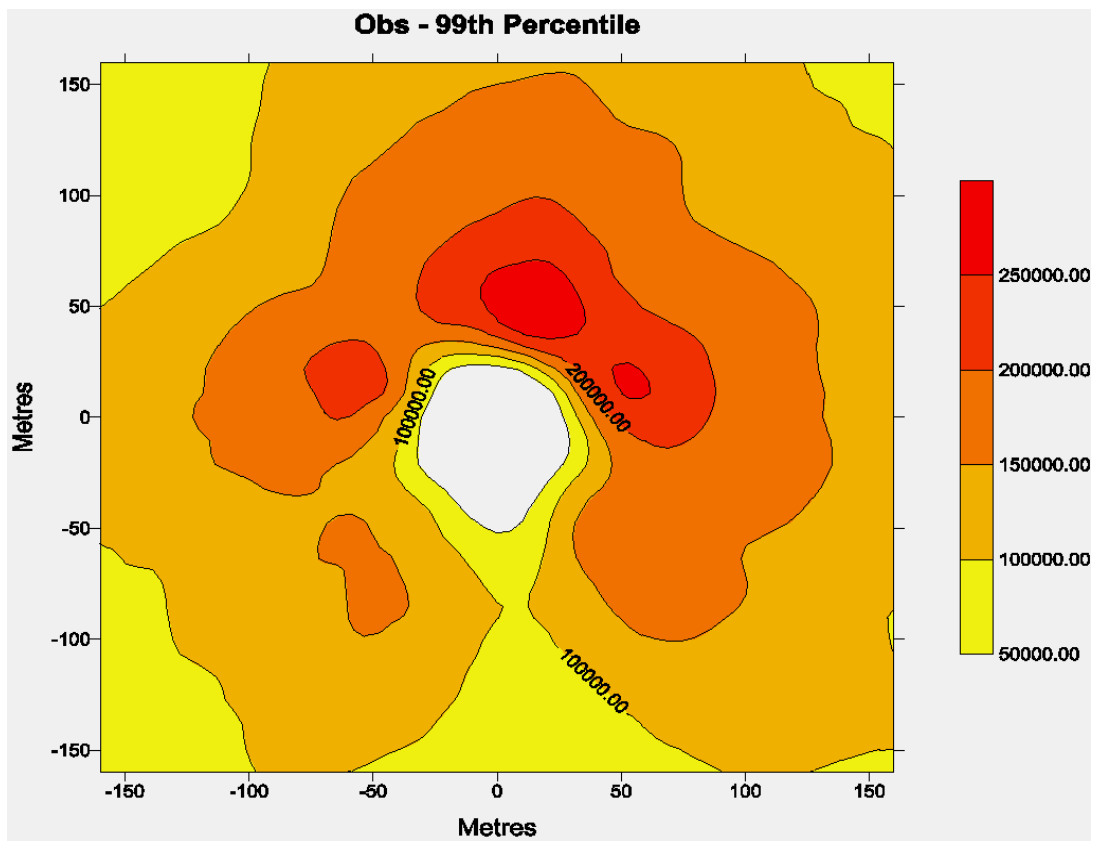




Dispersion plots for Source 2 at Ringway (smaller modelling area)

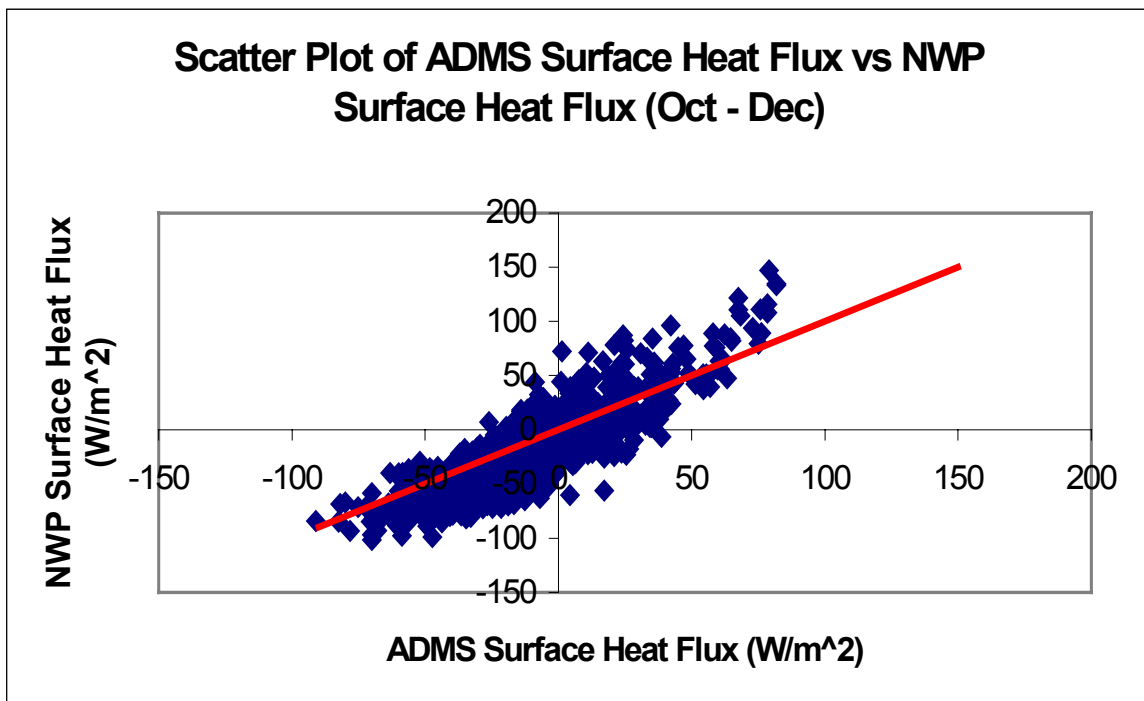
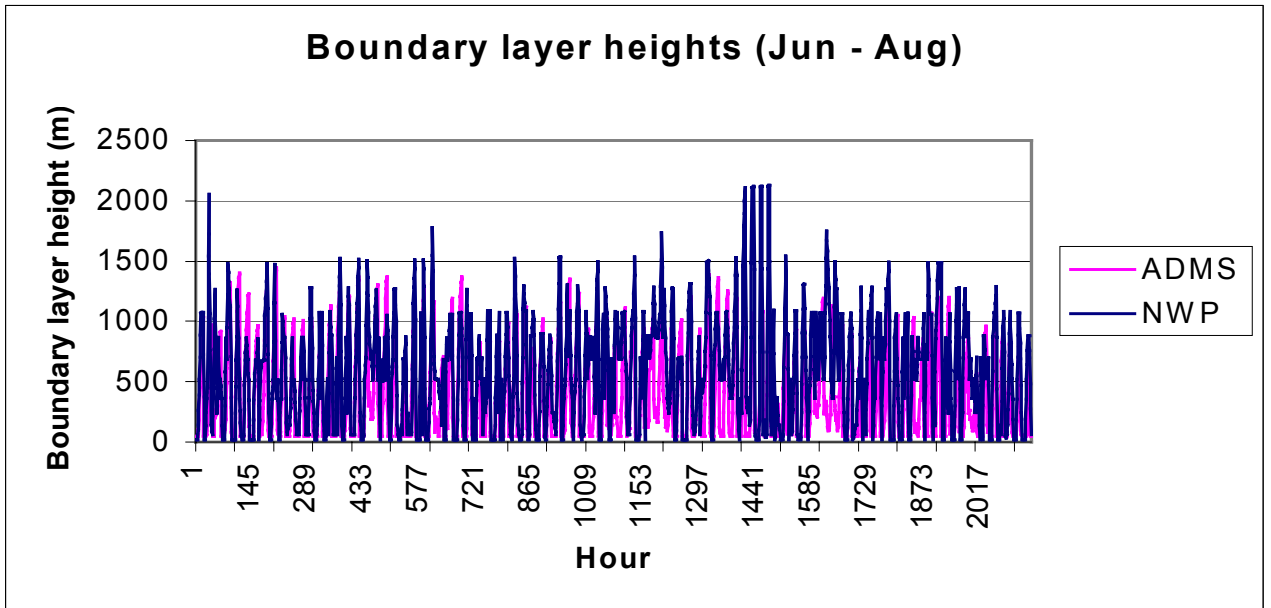


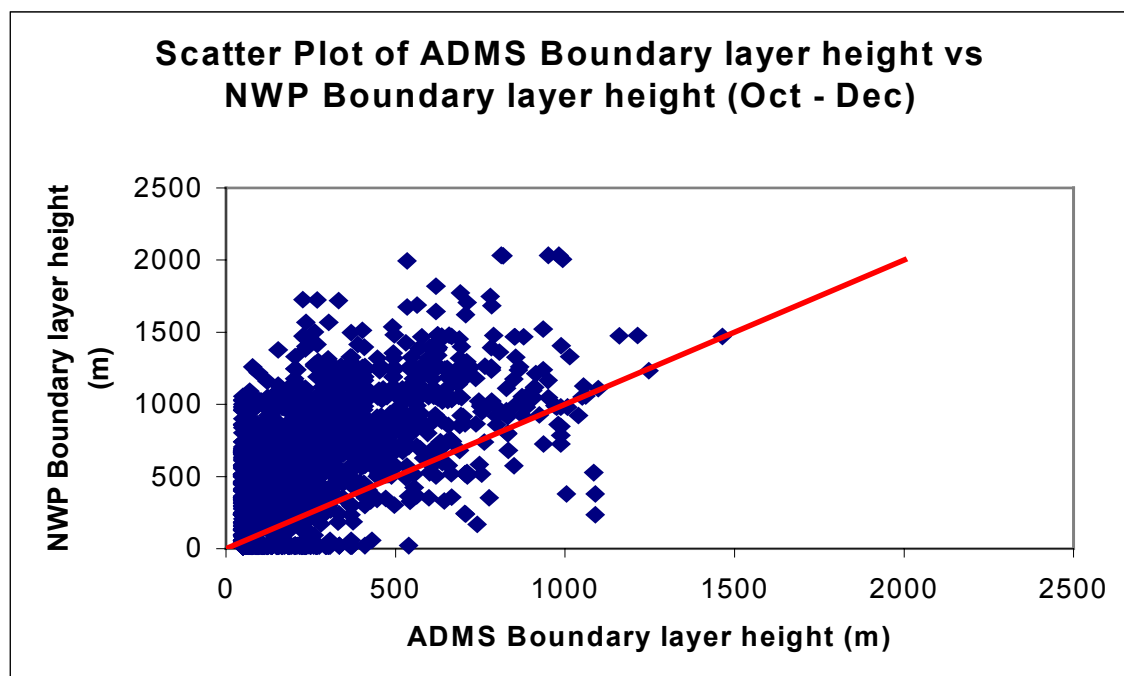
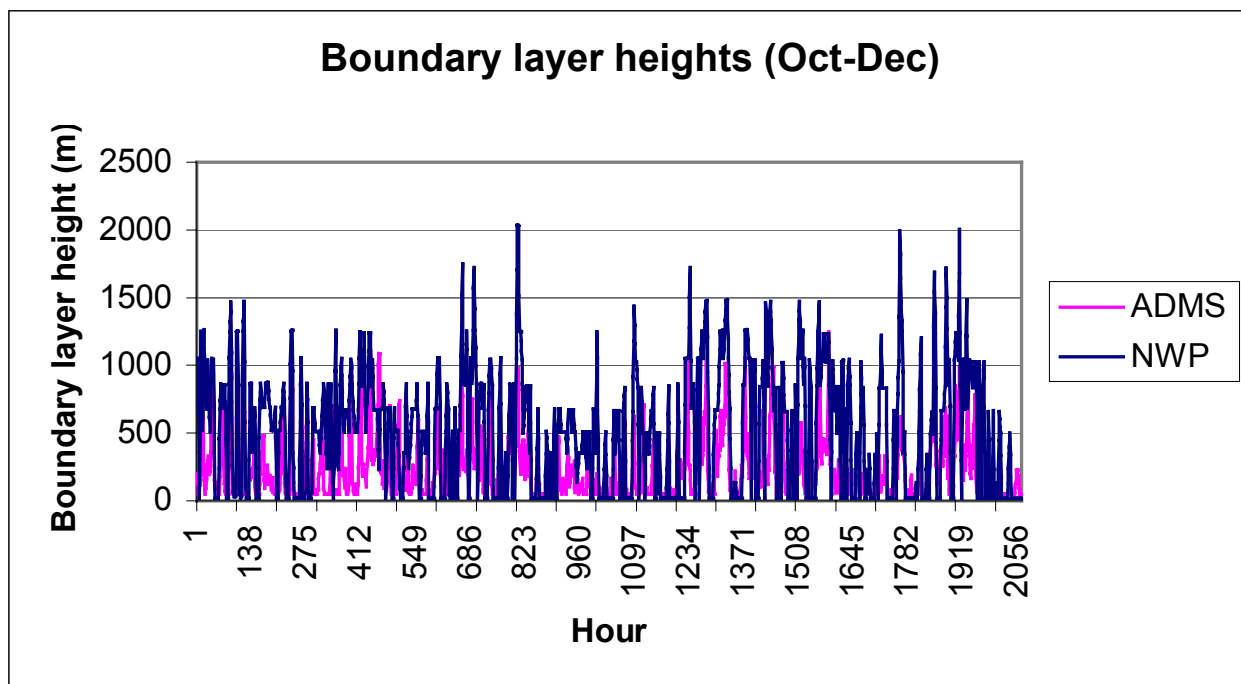


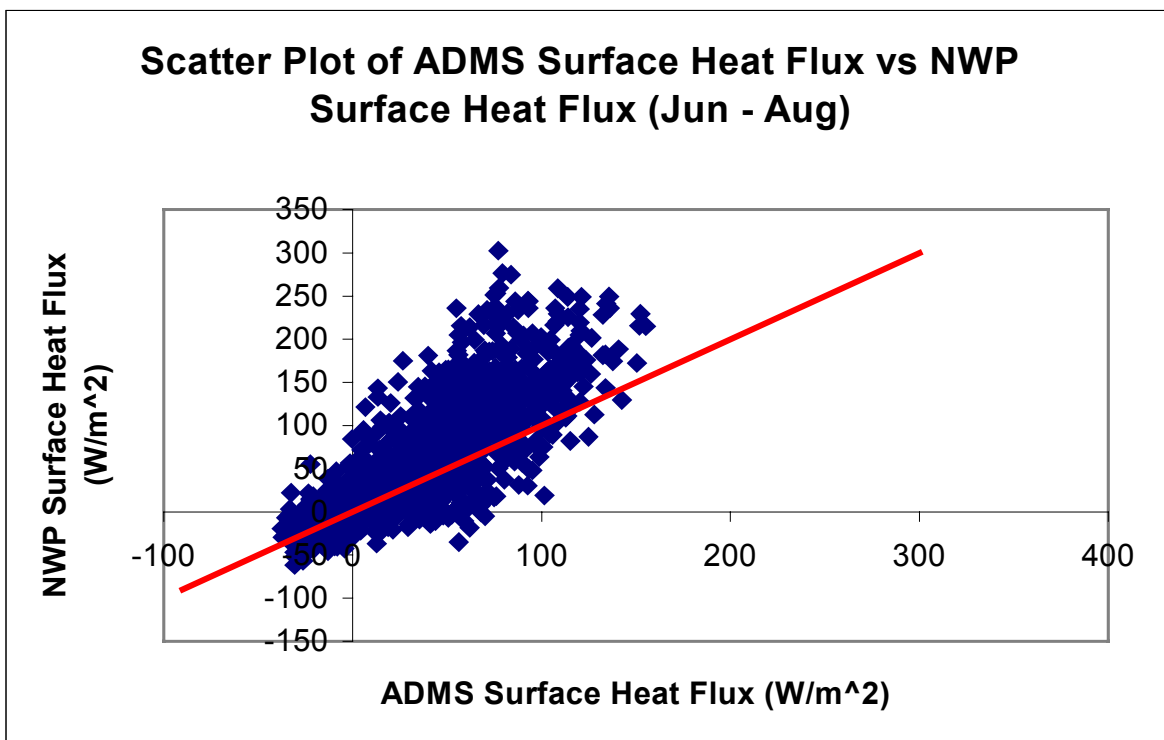
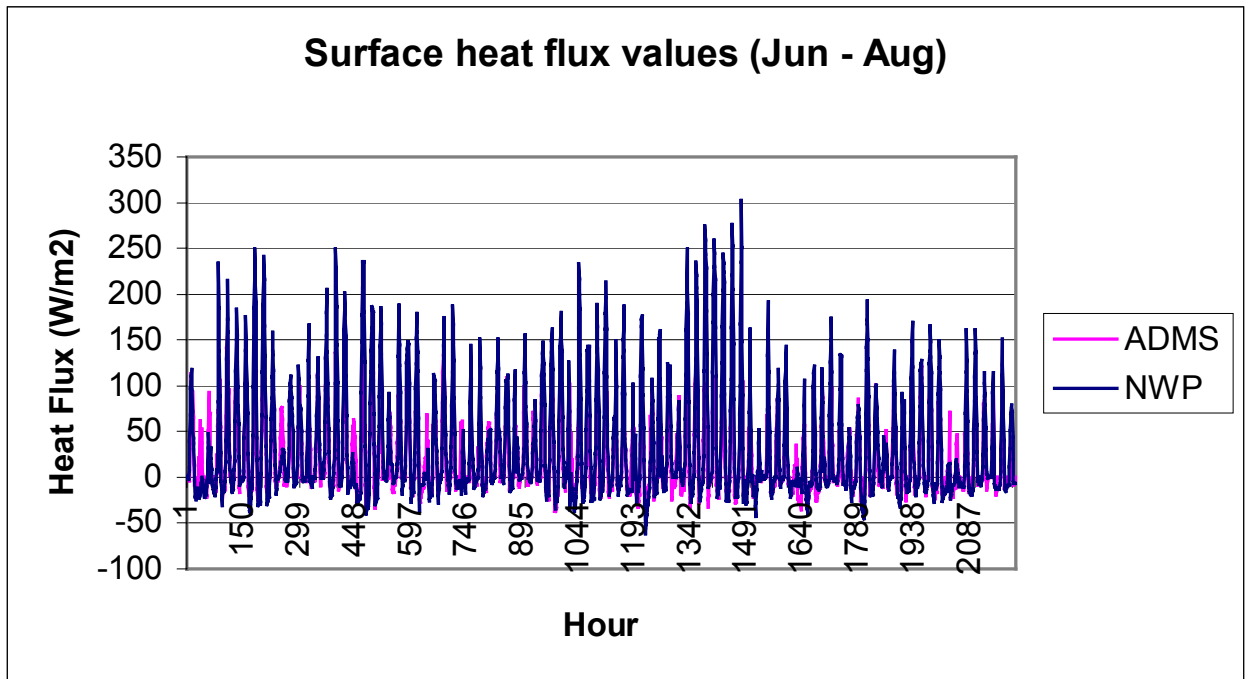


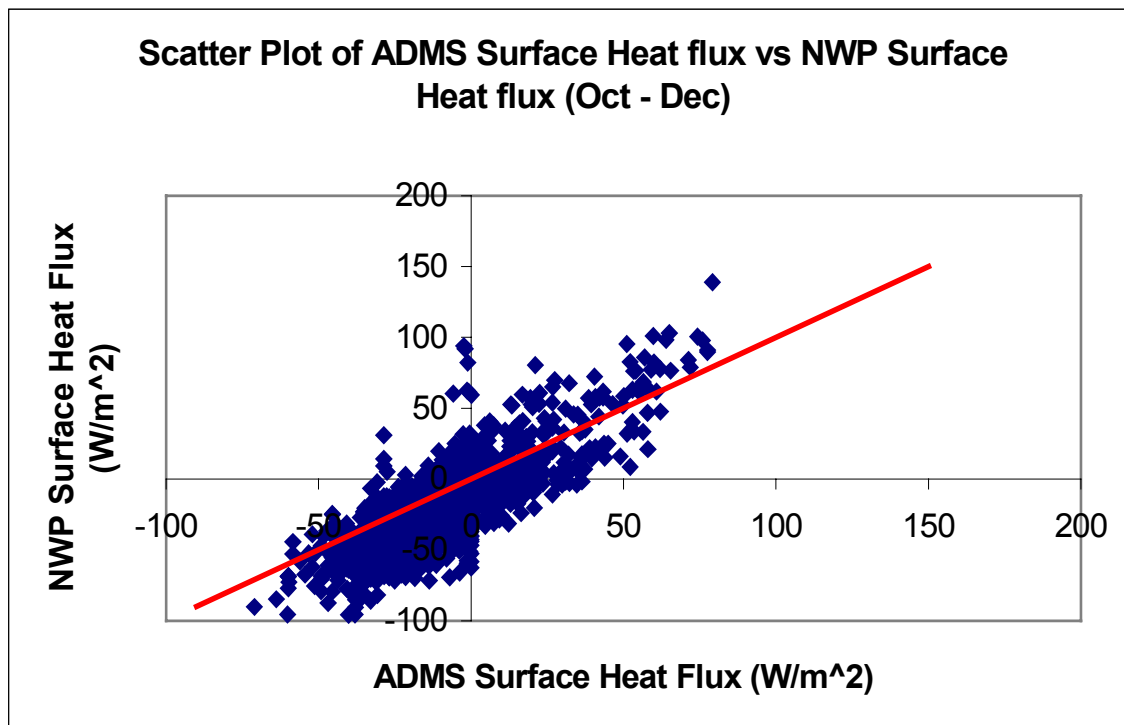
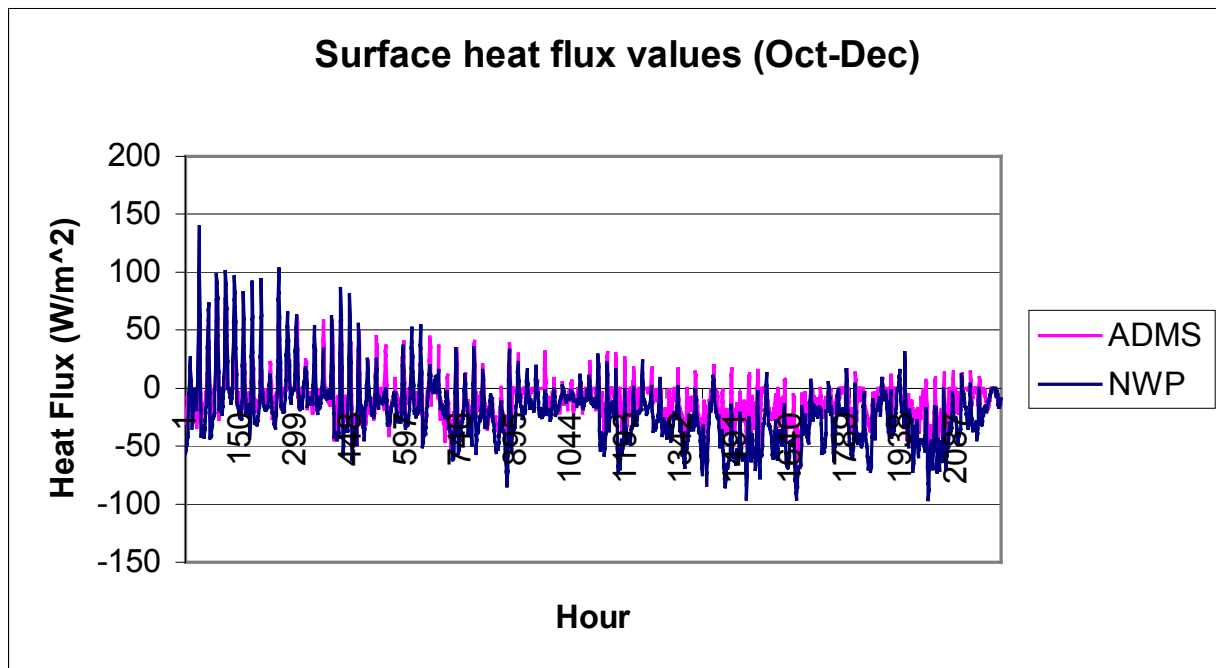
APPENDIX C
Comparison of boundary layer heights and surface heat fluxes

Ringway

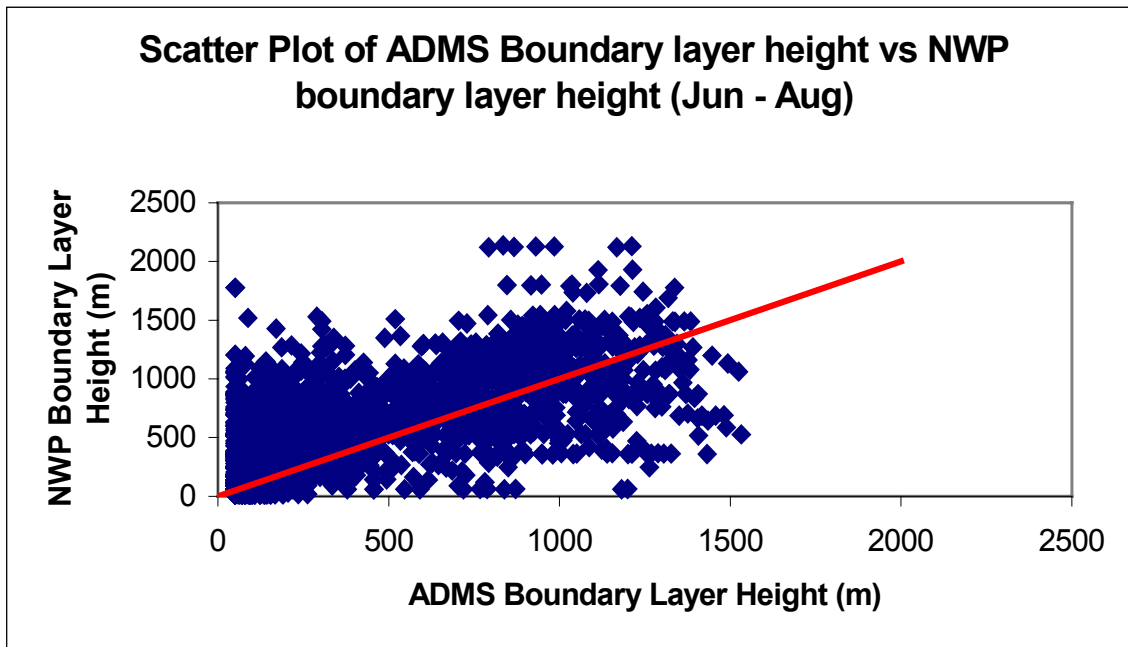
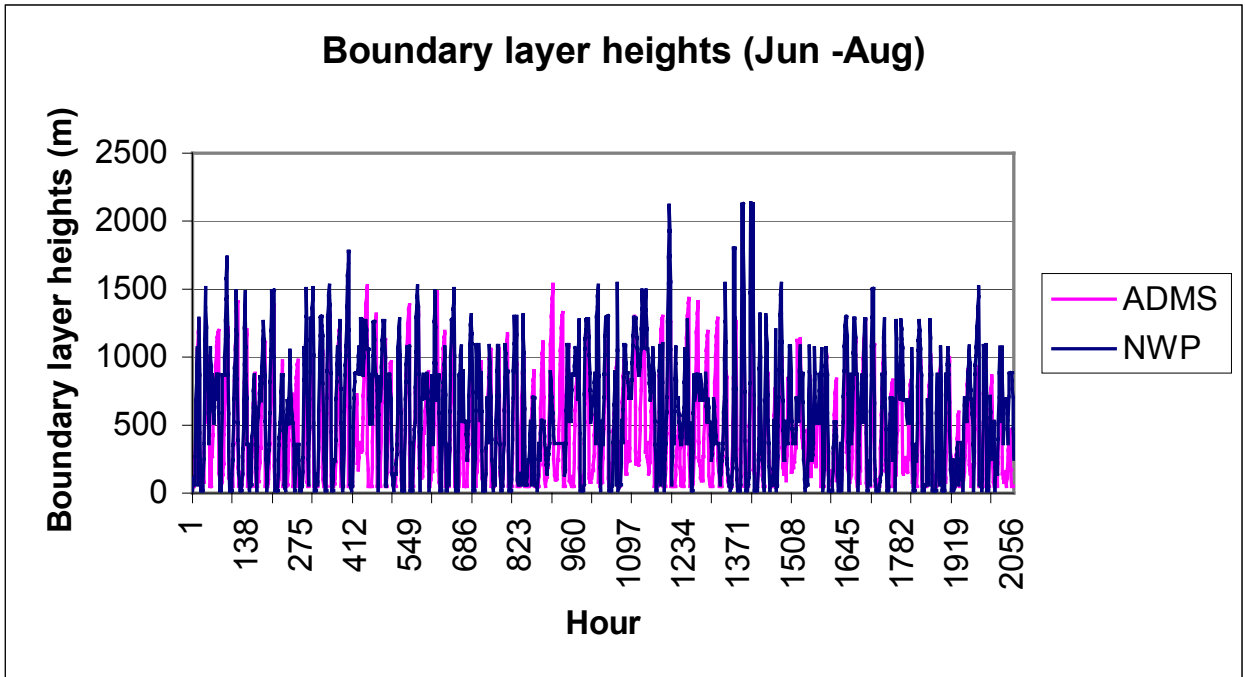


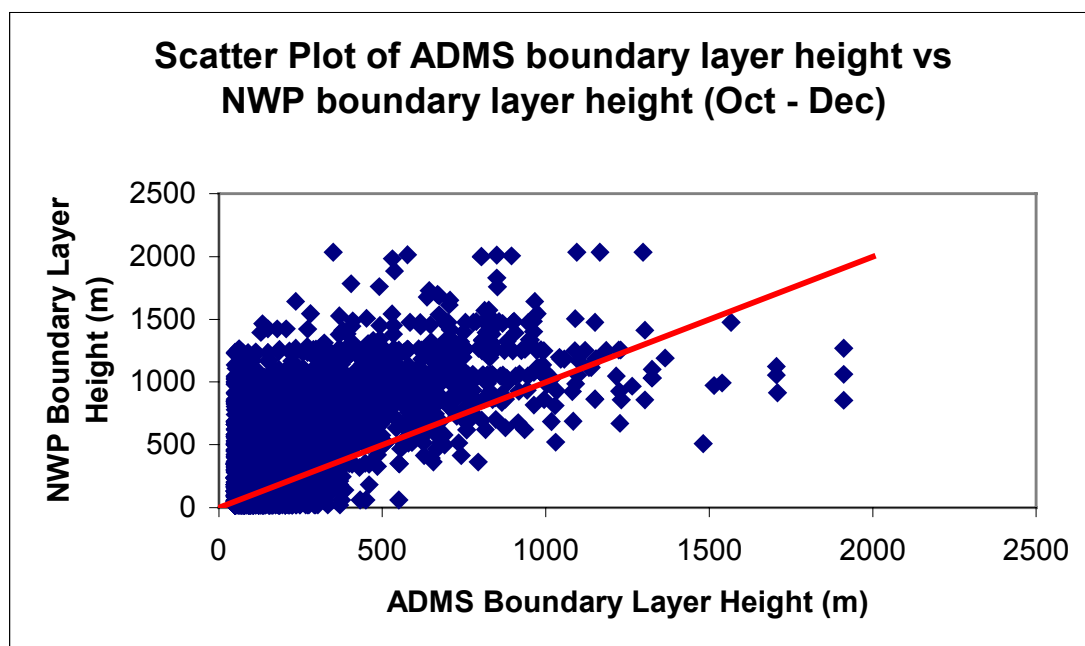
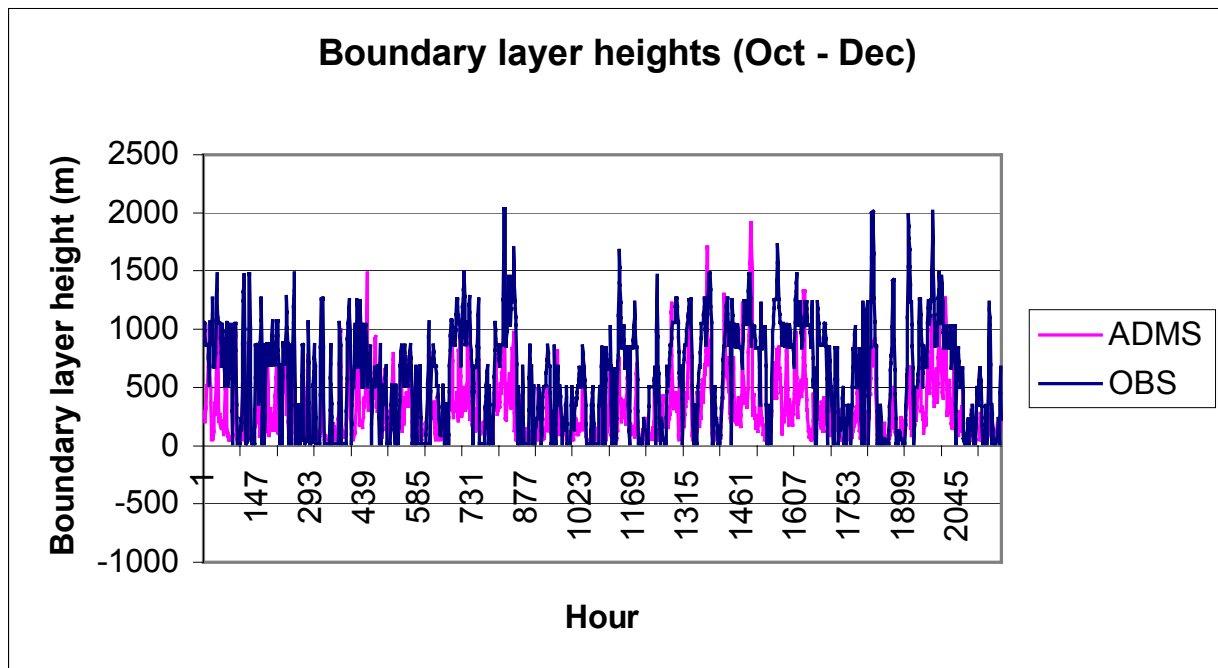


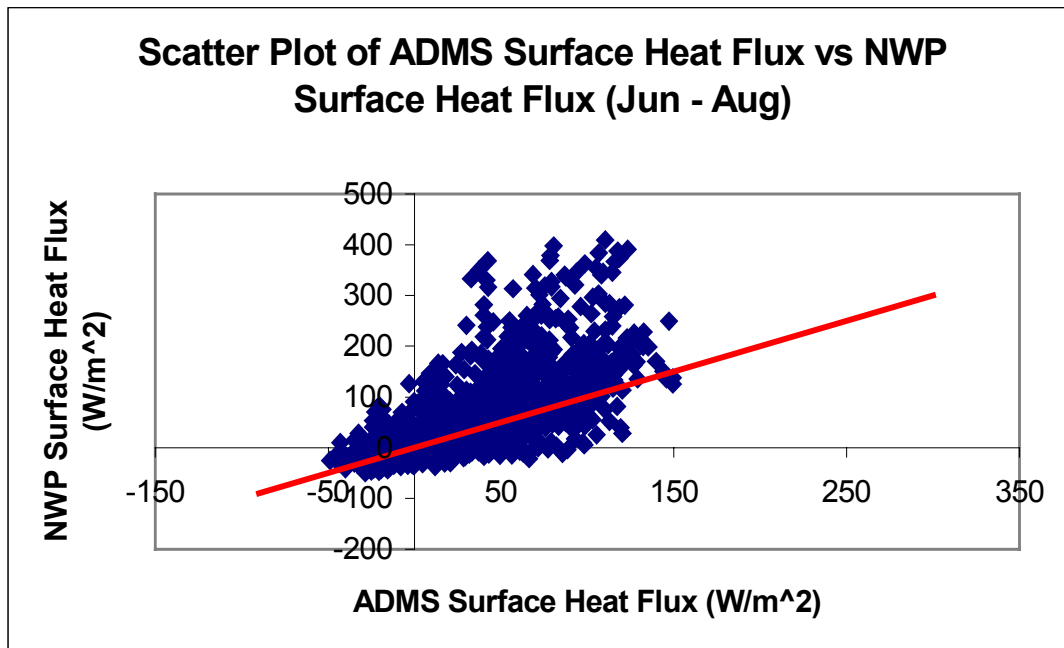
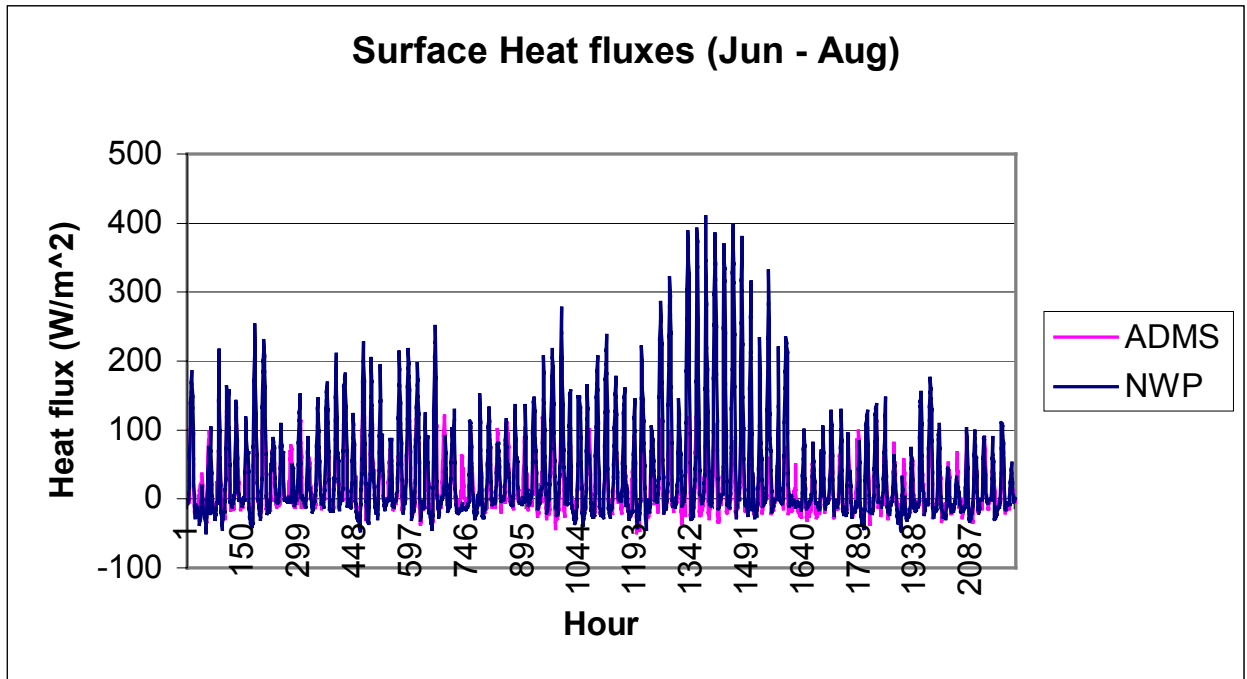


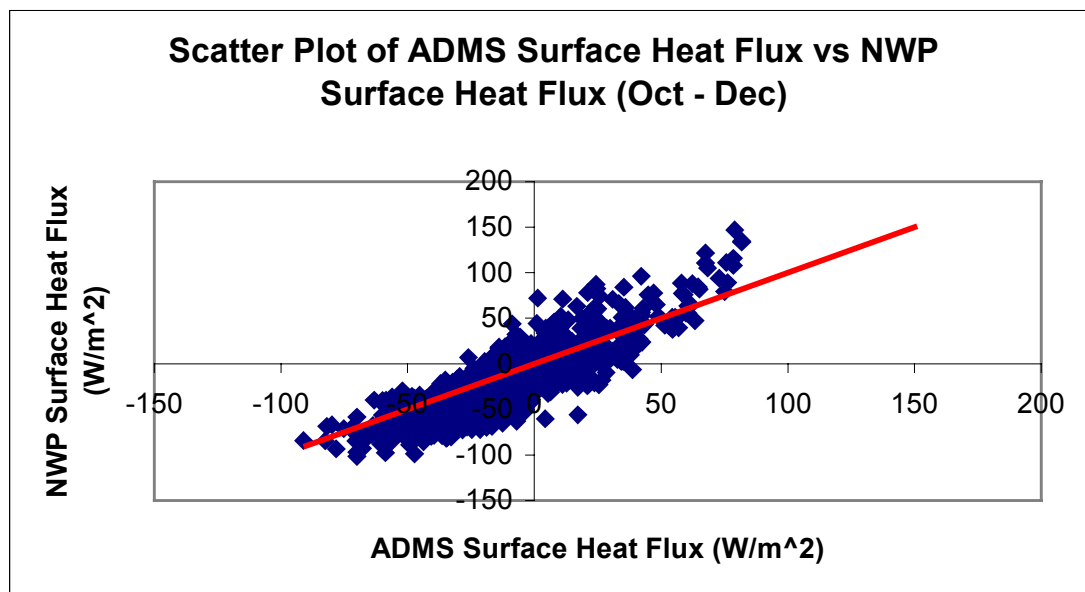
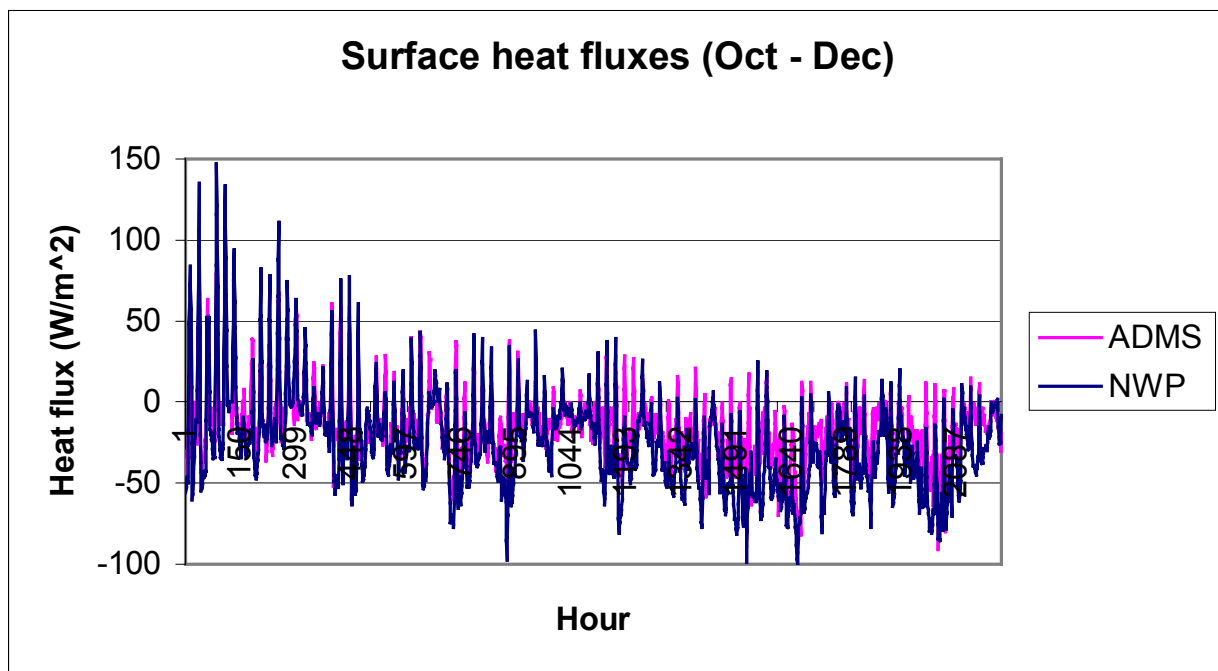


Waddington





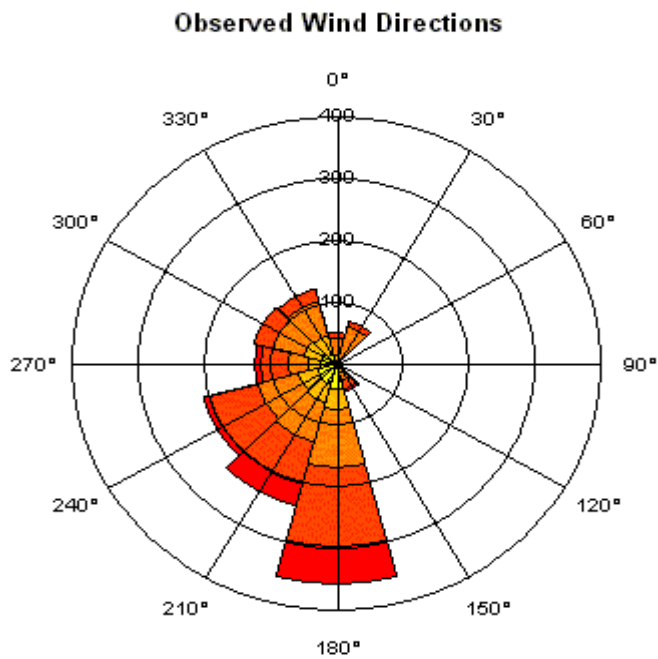




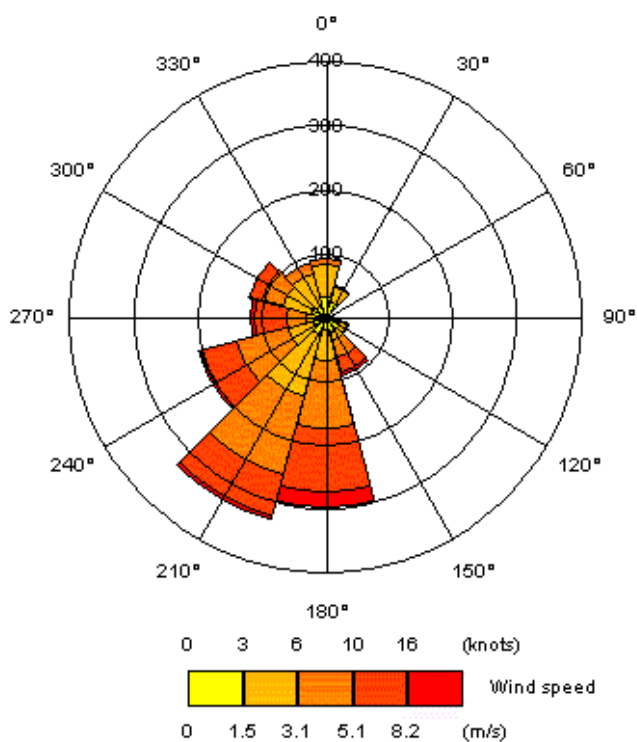
APPENDIX D

Comparison of SSM data with that observed at Ringway

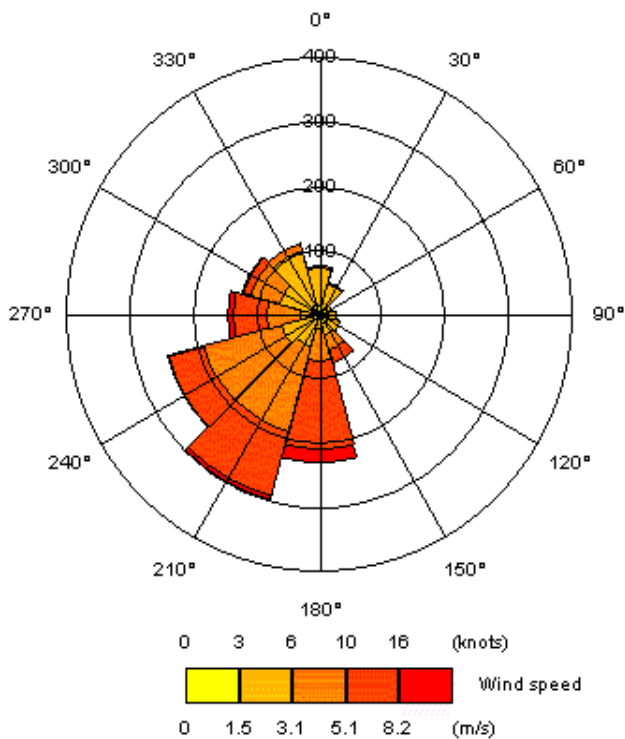
Wind roses for November and December at Ringway



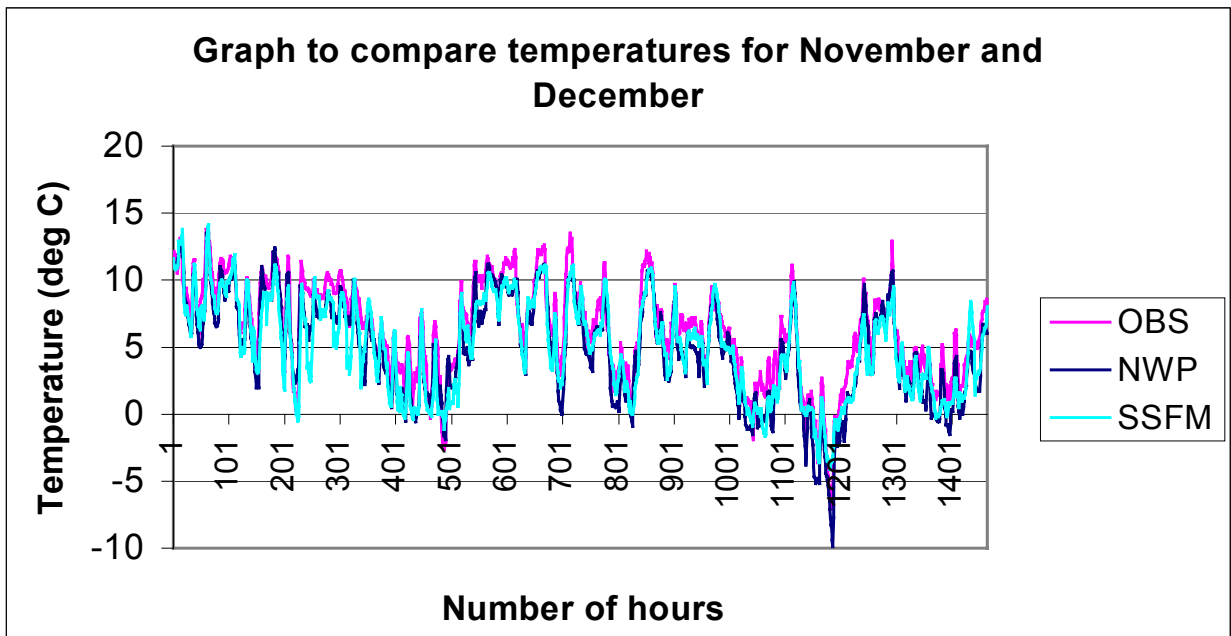
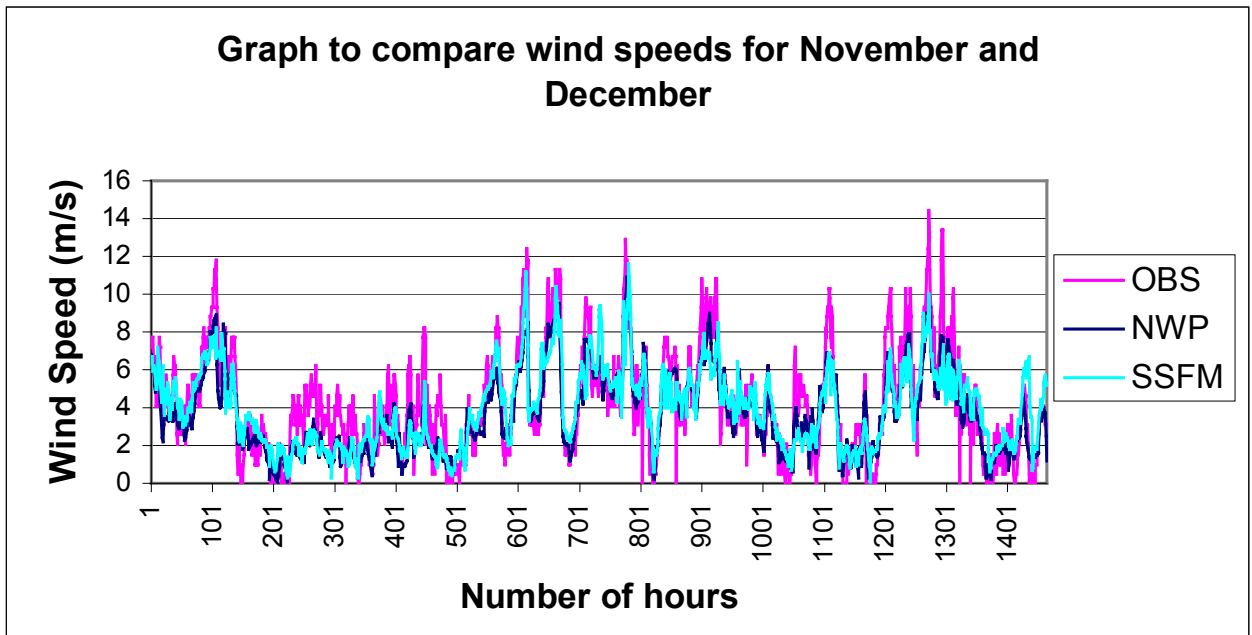
NWP Wind Directions



SSFM Wind Directions



Graphs to compare various weather parameters



ANNEX B

METHODS FOR UNDERTAKING UNCERTAINTY ANALYSES

Members of ADMLC expressed an interest in methods of undertaking uncertainty analyses. The Committee did not place a contract for work in this area, but arranged for a series of presentations at one of its meetings. The speakers did not provide written papers, but the following are the overheads used in their presentations.

Uncertainty analysis

- Introduction (Arthur Jones)
- Expert judgement (Simon French)
- Monte Carlo techniques (Arthur Jones)
- Other methods (Tony O'Hagan)
- (Data assimilation)



Aims of uncertainty analysis

- Models predict values of one set of quantities (results) from values of other quantities (inputs)
- Input values (and models) are uncertain
- Therefore results are uncertain
- Uncertainty analyses aim to quantify the uncertainty on the results from that on the parameter values



Stages in an uncertainty analysis

- Identify the uncertain input parameters and provide distributions on the values
- Propagate the distributions through the model to give uncertainty distributions on model prediction
- Identify the inputs whose uncertainty make large contributions to overall uncertainty



Distribution on input parameters

- All model applications involve some averaging
- "Default values" are the best estimate of the average over the range of conditions involved
- Uncertainty distribution must reflect lack of knowledge of this best estimate on the average, not the possible range



Uncertainty or variability

- Uncertainty on the best estimate prediction represents lack of knowledge of the most appropriate values for the model parameters
- Variability reflects effects of differences between members of a population, transfer processes, etc.



Propagate uncertainty through models

- Analytical techniques
- Monte Carlo
- Bayesian techniques



Identify important uncertainties

- Derive statistical parameters to describe relationships between input and output values

NRPB

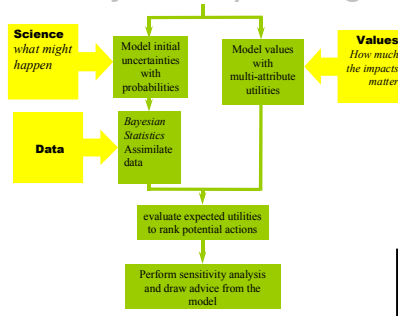

Expert Judgement, Data Assimilation, Uncertainty and Risk Analysis

Simon French
simon.french@mbs.ac.uk



MANCHESTER BUSINESS SCHOOL

The Bayesian paradigm

MANCHESTER BUSINESS SCHOOL

Players

Science

Experts

Forecasts of what might happen

Values


Stakeholders

Accountabilities and responsibilities

Decision Makers

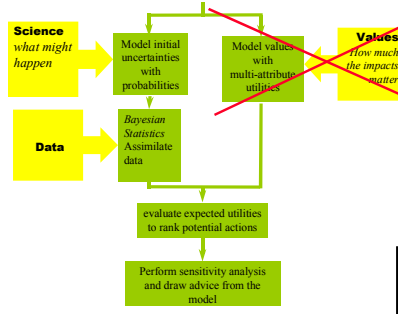

Process expertise

Analysts



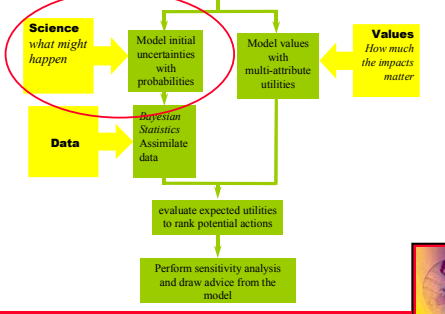

MANCHESTER BUSINESS SCHOOL

Ignore value issues

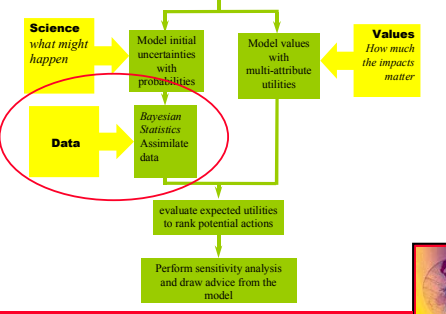

MANCHESTER BUSINESS SCHOOL

Expert judgement

MANCHESTER BUSINESS SCHOOL


Data assimilation

MANCHESTER BUSINESS SCHOOL

Bayes theorem

Posterior probability
 \propto likelihood \times prior probability




MANCHESTER BUSINESS SCHOOL

Bayes theorem

Posterior probability
 \propto likelihood \times prior probability

↑
 Probability distribution of parameters




MANCHESTER BUSINESS SCHOOL

Bayes theorem

Posterior probability
 \propto likelihood \times prior probability

↑
 likelihood of data given parameters




MANCHESTER BUSINESS SCHOOL

Bayes theorem

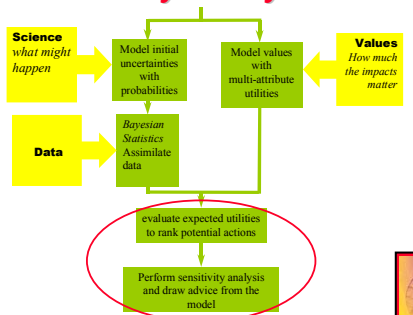
Posterior probability
 \propto likelihood \times prior probability

↑
 Probability distribution of parameters given data




MANCHESTER BUSINESS SCHOOL

Uncertainty analysis



```

    graph TD
        Science[Science: what might happen] --> ModelUnc[Model initial uncertainties with probabilities]
        Data[Data] --> Bayesian[Bayesian Statistics: Assimilate data]
        ModelUnc --> Bayesian
        Bayesian --> ModelVal[Model values with multi-attribute utilities]
        Values[Values: How much the impacts matter] --> ModelVal
        ModelVal --> Eval[evaluate expected utilities to rank potential actions]
        Eval --> Sens[Perform sensitivity analysis and draw advice from the model]
    
```




MANCHESTER BUSINESS SCHOOL

Decision support and risk analysis

Supporting decisions
 \neq
 Building scientific understanding

Related, similar but **NOT** identical

Requisite modelling and analysis



MANCHESTER BUSINESS SCHOOL

Expert judgement

MANCHESTER BUSINESS SCHOOL



Why use probabilities

- combines expert judgement and other data
- communication via *operational* definitions
- unified approach
- computationally feasible
- fit with decision making
- axiomatic foundations

MANCHESTER BUSINESS SCHOOL



The expert problem

- Focus of a decision
- Expert provides advice to decision maker.
- Expert not responsible for decision



MANCHESTER BUSINESS SCHOOL



Group decision problem

- Focus of a decision
- Whole group jointly responsible for decision
- No external experts



MANCHESTER BUSINESS SCHOOL



Eliciting from a group of experts

Group process → consensus
(behavioural aggregation)

versus

Individual elicitation and combination
(mathematical aggregation)

MANCHESTER BUSINESS SCHOOL



Substantive expertise


- in identifying the events and hazards to be considered;
- in choosing the assumptions to make;
- in selecting the form of the model(s) to use;
- in identifying datasets and experts to draw upon in determining parameters to use in the model;
- in judging the parameters in the model;

MANCHESTER BUSINESS SCHOOL



Process expertise

- in analysing the models;
- in performing stochastic uncertainty analyses, including selecting the form of distributions, moments and correlations to model the uncertainties arising from expert judgement and the analysis of empirical data;
- in performing deterministic sensitivity analyses including selecting variables and the ranges over which to vary them;
- in knowing when to stop.



MANCHESTER BUSINESS SCHOOL

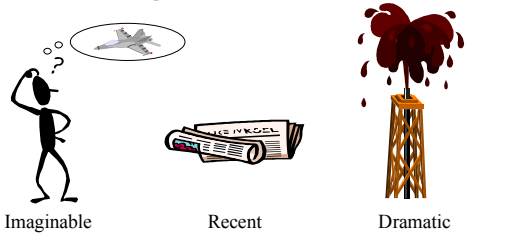
Psychological 'biases'

- Availability
- Ignoring base rates
- Misconceptions of randomness
- Anchoring
- Framing
- ...




MANCHESTER BUSINESS SCHOOL

Availability

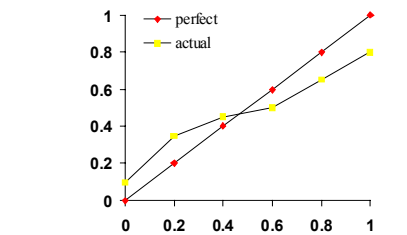



Imaginable Recent Dramatic



MANCHESTER BUSINESS SCHOOL


Calibration

MANCHESTER BUSINESS SCHOOL

Elicitation

- Need to be mindful of experts' cognitive characteristics when eliciting probabilities
- training essential
 - for elicitors
 - for experts
- Use of calibration data set is valuable




MANCHESTER BUSINESS SCHOOL

Mathematical aggregation

- Expert judgements are really *data* to decision maker
 - ⇒ could use Bayes to learn from experts

But methodological and cognitively difficult

- Cooke has developed a very successful alternative



MANCHESTER BUSINESS SCHOOL

Cooke's method

$$P_X^{DM}(x) = \sum_{e=1}^{n_{\text{experts}}} w_e P_X^e(x)$$

where

$$w_e = \begin{cases} C_e/H_e & \text{if expert is 'good'} \\ 0 & \text{if expert is 'poor'} \end{cases}$$



MANCHESTER BUSINESS SCHOOL

Process expertise

Need to ensure that:

- disagreements among the experts are continually revisited and explored
- the DMs gain understanding of
 - the model,
 - the importance of selected inputs,
 - the distributional assumptions on inputs,
 - effects of changes in distributional assumptions
- The DMs understand the importance of inputs identified as key during sensitivity and uncertainty analyses in determining the outcome of the analysis.



MANCHESTER BUSINESS SCHOOL

Bayes_RIMPUFF and Bayes_MATCH



MANCHESTER BUSINESS SCHOOL

Data assimilation in ADMs

- DM's questions concern *downwind* predictions
 - Where is it going (*roughly*)?
 - When is it going to get there (*roughly*)?
 - How much is going to get there (*roughly*)?
- ⇒ need to estimate/predict contamination along plume at current and future times *not* source term at release.
- Variety of monitoring observations
- Need to balance detail of science with quick and dirty robust forecasting.



MANCHESTER BUSINESS SCHOOL

Data Assimilation in Early Phase ADM

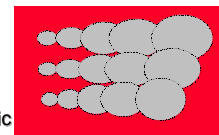
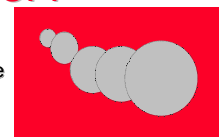
- **Real Time:** At least for early phase
 - ⇒ tough computational constraints
- **What-if?** Need to explore possibilities
 - ⇒ tougher computational constraints
- **Disordered data:** arrival out of chronological order
 - ⇒ constraints on computational logic
- Decision making will be conservative
 - ⇒ detail less important than broad brush picture.



MANCHESTER BUSINESS SCHOOL

Bayes_RIMPUFF

- Based on RIMPUFF
- Runs nine models: 3 wind dir'ns × 3 release hts.
- But**
- Single nuclide
- Single windfield
- Only instantaneous air concs.
- Computational size and logic of new deterministic RIMPUFF an issue




MANCHESTER BUSINESS SCHOOL

Bayes_RIMPUFF validation

Sensitivity, efficiency and reliability

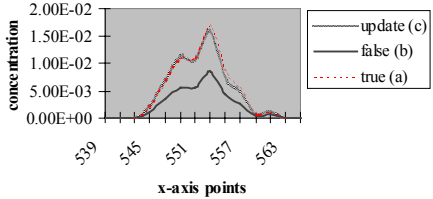
- of forecasts to deterministic parameter settings
- to correct misspecified parameter settings
- to statistical assumptions:
 - (normality vs Poisson) & multi-process modelling
- to computational algorithms: KF v Bayes Net
- to model misspecification; ATSTEP v RIMPUFF

Data Lundtoffe, Nord Siesta (adapted), Karlsruhe,




MANCHESTER BUSINESS SCHOOL

Updated and non-updated concentrations

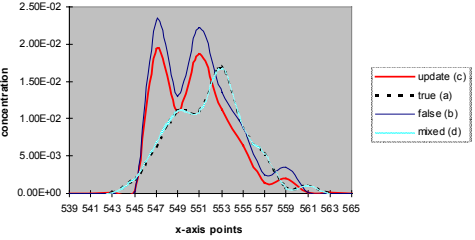



Downwind cross-section of the plume (y-co-ordinate fixed) based on 25 observed air concentrations



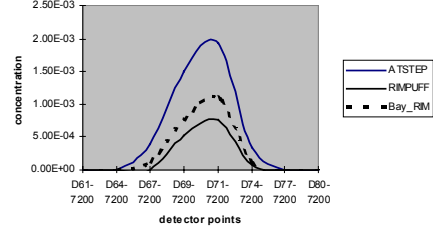

MANCHESTER BUSINESS SCHOOL

Effect of mixing models

MANCHESTER BUSINESS SCHOOL


Not all good news

MANCHESTER BUSINESS SCHOOL

Some results

- Sensitive Parameter Values:
 - Wind field, release height, wind velocity
- Multi-process models converge to true (true-ish) model remarkably quickly provided that different models predict different profiles at observation sites
- Gaussian (normal) assumptions usually outperform Poisson




MANCHESTER BUSINESS SCHOOL

Data assimilation in long range ADMs

- Computational demands of model more stringent

But

- Real-time demands less strong
- Disordered data issue less important
- ETEX and RTMOD provide a backdrop on the accuracy that is obtainable



MANCHESTER BUSINESS SCHOOL

Bayes_MATCH

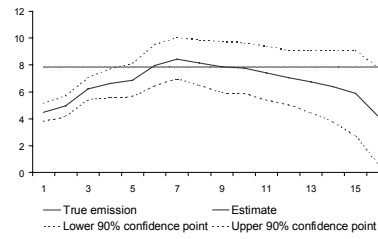
- Have adapted the adjoint version of MATCH to use Kalman Filtering
- Compatible approach with Bayes_RIMPUFF
- Hence will be able to have a model chain:

RODOS_STM → BayesRIMPUFF → BayesMATCH



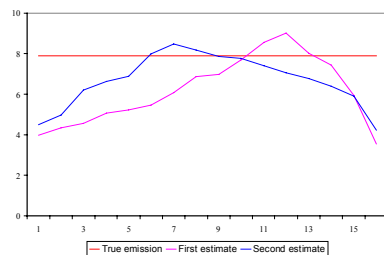
MANCHESTER BUSINESS SCHOOL

Example: Bayes MATCH (1)



MANCHESTER BUSINESS SCHOOL

Example: Bayes MATCH (2)



MANCHESTER BUSINESS SCHOOL

Monte Carlo Methods

Arthur Jones

MRPB

Steps in Monte Carlo analyses

- Specify distributions on inputs
- Select sets of values
- Run model (many times)
- Generate uncertainty distributions on outputs
- Identify important uncertainties

MRPB

Select sets of input values (sampling)

- Simple random sampling
can easily add more samples if needed
- Stratified sampling
sub-divide range, sample from sub-ranges
harder to add more samples
- Latin hypercube sampling
divides range into equal probability segments
cannot add more samples

MRPB

Correlated input parameters

- Sampling techniques exist for use with Monte Carlo or LHS samples
- Problems and limitations, but these should be considered alongside the difficulties of specifying correlations

MRPB

Steps in Monte Carlo analyses

- Specify distributions on inputs
- Select sets of values
- Run model (many times)
- Generate uncertainty distributions on outputs
- Identify important uncertainties

MRPB

Advantages and Disadvantages of Monte Carlo methods

- ✓ Easy to understand - simply involves running the model (very) many times
- ✓ Works for any model, however complex or non-linear
- × May require very large amounts of computer time
- × Hard to know if enough runs have been done

MRPB

Steps in Monte Carlo analyses

- Specify distributions on inputs
- Select sets of values
- Run model (many times)
- Generate uncertainty distributions on outputs
- Identify important uncertainties



Quantifying the uncertainty

The analysis provides n values for each of the model outputs (results) considered

These values are the uncertainty distribution on the model output, and may be described in many ways (mean, variance, percentiles ...)



Steps in Monte Carlo analyses

- Specify distributions on inputs
- Select sets of values
- Run model (many times)
- Generate uncertainty distributions on outputs
- Identify important uncertainties



Identifying important uncertainties

- Correlation coefficients measure strength of linear relationships
- Regression coefficients measure slope of linear relationships
- Other methods do not assume linearity, but may require more runs
- Method based on analysis of scatter plots on ranks of variables



Bayesian Uncertainty Analysis

Tony O'Hagan

University of Sheffield

Basics

- A computer model converts inputs x to outputs y
- We can represent this as a function

$$y = f(x)$$
- We have uncertainty about x so it has a probability distribution
- This induces a distribution for y

The problem

- We wish to characterise, in some suitable way, the uncertainty about y
 - The expectation $E(y)$
 - The variance $Var(y)$
 - The uncertainty distribution $P(y < c)$
 - The density function
- Also to partition the uncertainty to show contributions due to each uncertain input

Methods

- Exact solution
 - might be feasible if $f(x)$ is a simple function
- Monte Carlo
 - requires many code runs
 - impractical if each run takes substantial time
 - some progress possible with more efficient MC
- Bayesian methods

Bayesian approach

- Uses code runs more efficiently
- Key idea - f is a smooth function
 - $f(x)$ changes smoothly as x changes
 - Observing $f(x)$ provides information about $f(x^*)$ for x^* close to x
- Instead of running the code at random inputs, we choose x values well spread out

Some details

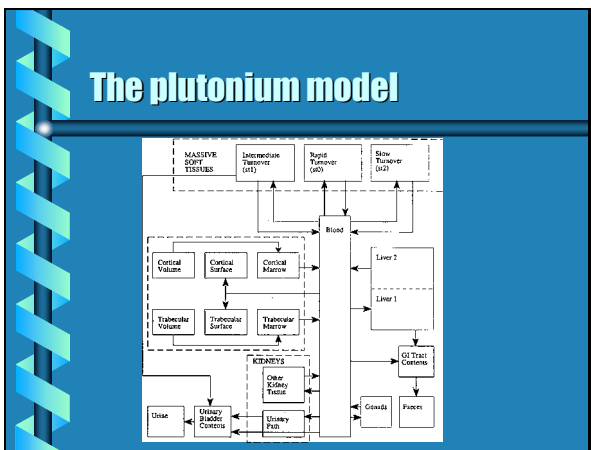
- Treats code as a black box
- Prior information specifies smoothness
 - Prior distribution for the whole function $f(\cdot)$
- Data $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$
 - Used to learn about $f(\cdot)$
 - And about unknown parameters in the prior distribution

Benefits

- Number of runs, n , can be much smaller
- Full posterior probability distribution of uncertainty characteristics
- Flexible framework encompasses many related problems
 - Interpolation / emulation
 - Calibration
 - Multi-level codes ...

Example

- Plutonium internal dosimetry model
- 14 uncertain inputs
 - e.g. rate of clearance from blood
- Conventional UA performed by NRPB
 - 500 point Latin Hypercube sample
 - Estimated mean K and variance L of uncertainty distribution



Results

Method	n	Est. K	+/-	Est. L
LHC	500	5.35 E-4	7.4 E-6	7.20 E-9
Bayes	50	5.27 E-4	6.6 E-6	5.24 E-9
Bayes	100	5.24 E-4	4.1 E-6	5.16 E-9
Bayes	500	5.29 E-4	1.7 E-6	6.48 E-9

Conclusions

- Bayes on 50 observations is as precise as LHC on 500
- LHC seems to be biased in this example
- Bayes estimates validated thoroughly
 - Sensitivity analysis to hyperparameters
 - When $n = 50, 100$ predict rest of original LHC sample of $n = 500$

Emulation

- A simpler idea than uncertainty analysis
- Estimate $f(x)$ for an x that has not been tried yet
 - Provides a quick approximation to the code itself
 - Can be adequate if there is enough data from runs with 'nearby' x values
 - Can replace the code for multiple 'what ifs' and searches

Calibration

- **Code represents some real phenomenon**
- **Observations of the real thing allow us to learn about uncertain inputs**
 - Conventional practice - trial and error fit of model to data, then act as if fitted parameters known
 - Bayesian formulation fully recognises all the uncertainties

Model tuning

- **Calibration itself is a way of tuning the model to data**
- **But in practice no values for calibration parameters will correctly reproduce data**
 - Model inadequacy - all models are wrong!
 - Learn about model inadequacy
 - Tunes model even further by statistical correction

The reality

- **Bayesian approach looks very promising**
 - RSS ordinary meeting on December 13
- **But still relatively untried**
 - Need much more practical experience on a variety of codes
- **And limited in size of problem that can be tackled**
 - Need more technical development