



**REPORT ON FORECASTING
THE PROBABILITY OF NO₂ AIR QUALITY EVENTS
USING MODEL OUTPUT STATISTICS.**

BY CATHERINE STRUDWICKE

CENTRAL FORECASTING DIVISION

MARCH 1994

TECH NOTE CF 17

Headquarters, Bracknell

Table of Contents

1. Introduction	1
2. Preliminary Analysis	2
2.1 Scatter plots	2
2.2 Tabulation	5
3. Further Analysis	7
3.1 Linear regression	7
3.2 Discriminant analysis	11
3.3 Restricted and unrestricted models	13
4. Verification of the model	15
4.1 Tabulated verification	15
4.2 Verification using histograms	17
5. Conclusions	19
Appendix One: Description of discriminant analysis	23
Appendix Two: Cost analysis	27
References	29

Forecasting the probability of NO₂ air quality events
using Model Output Statistics.

1. Introduction

This investigation of NO₂ air quality data was commissioned by the Environmental Consultancy Services section at the Meteorological Office for Bristol County Council. The investigation was carried out within Central Forecasting Division.

The aim of this analysis was to investigate the amount of NO₂ present in the atmosphere and its possible relationships with meteorological parameters. The objective was to develop forecast equations that predicted the probability of NO₂ classified into three groups: 'OK', 'poor' and 'very poor' events. The data used in this investigation came from two sources. The first source consisted of hourly totals of NO₂, (measured in parts per million) from Old Market in Bristol. The second source of data was the Central Forecasting MOS archive, which contained parameters from the NWP forecast. The NWP data, interpolated to the Bristol Weather Centre site (03726), were taken from the Limited Area model for T+24 and T+30. The parameters considered are shown in Table 1a.

Parameters	Level in the model
Mean sea level pressure (MSLP)	Sea level
Surface pressure	Surface
Surface temperature	Surface
Wind direction	4
Verification hour	-
Verification day	-
Verification month	-
Verification year	-
Temperature	1,2,3,4,5,6,7
Temperature advection	1,2,3,4,5,6,7
Dew point	1,2,3,4,5,6,7
Dew point advection	1,2,3,4,5,6,7
Wind speed	1,2,3,4,5,6,7
Wind shear	1,2,3,4,5,6,7
Divergence	1,2,3,4,5,6,7
Vorticity	1,2,3,4,5,6,7

Table 1a: Table of parameters considered in the analysis and the level at which they were taken from the Limited Area model (level one ~ 25m, level seven ~5km).

Since most of these parameters maybe correlated only indirectly with NO₂ levels, other parameters were derived from meteorological experience and an expectation of how NO₂ and meteorology may interact. These created parameters are indicators of temperature and dew point inversion, and parameters that are time oriented, such as season and day of the week for which the observation was valid, etc.

The observations were divided into four six hour periods centred around 00, 06, 12 and 18Z. The mean amounts of NO₂ for these periods were calculated from the hourly observations and then matched to the appropriate verification time for the forecast. The criteria used to classify the events are the same as those used by the City of Bristol Health and Environment (1993) for eight hourly means, Table 1b.

Category	µg/m ³	Parts per billion	parts per million
OK	<150	<78.75	<0.07875
Poor	150-200	78.75-105	0.07875-0.105
Very Poor	>200	>105	>0.105

Table 1b: Criteria used for classifying NO₂ amount into categories.

The observational NO₂ data supplied were from November 90 to October 93. However the change to the Unified Model occurred in June 91 and greatly affected some of the forecast parameters. Hence only data from July 91 were considered, giving approximately two and a half years of data.

2. Preliminary Analysis.

The analysis was divided into two parts, the first consisted of an initial exploration of the data. This gives the analyst an insight into the behaviour of the data and a guide to the appropriateness of some of the more sophisticated statistical techniques available at the second stage. The analysis was performed using BMDP, a statistical software programming package.

2.1 Scatter plots.

An initial investigation of the data involved scatter plots of the NWP parameters against the mean NO₂ values. These showed unsurprisingly that there was no one parameter that adequately described the relationship with mean NO₂. Some of the NWP parameters have scatter plots from which a pattern emerges. This consisted of a restricted range of the NWP parameter within which the NO₂ events occurred. (A

NO₂ event as defined here is an event that falls into the 'very poor' or 'poor' category.) A typical example of this is mean sea level pressure (MSLP), Figure 1.

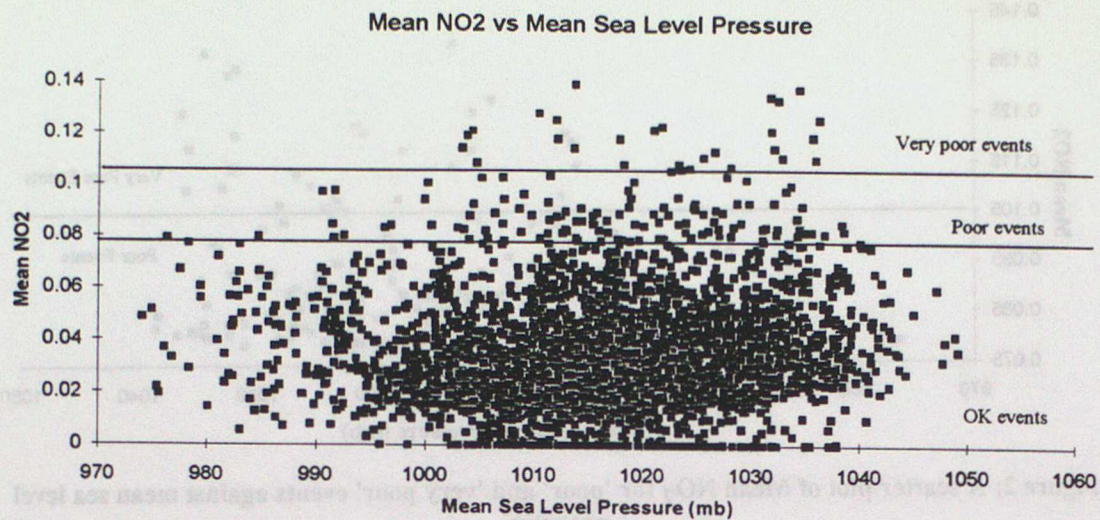


Figure 1: A scatter plot of mean NO₂ events against mean sea level pressure.

Figure 1 shows that comparatively few observations fell into the 'very poor' and 'poor' categories. There were approximately 3500 observations of which 3% were 'poor' and 1% 'very poor'. The majority of the data fell below the line separating the 'poor' and 'OK' NO₂ events, within a range of 975-1050 mb. As the pressure dropped below 1000 mb, the 'very poor' NO₂ events cease with a corresponding decline in the 'poor' events. From 1000-1035 mb the distribution of NO₂ events was relatively even, with as expected more 'poor' events than 'very poor'. Since it is known that in a stable high pressure system with little or no wind, foggy or misty weather is quite likely, these are also likely to have associated NO₂ 'poor' air quality events. However the scatter plot shows that the NO₂ events are not simply confined to high pressures. Enlarging the scatter plot for NO₂ events clarifies the distribution and range in which these events occurred, Figure 2.

Very Poor and Poor Events vs Mean Sea Level Pressure

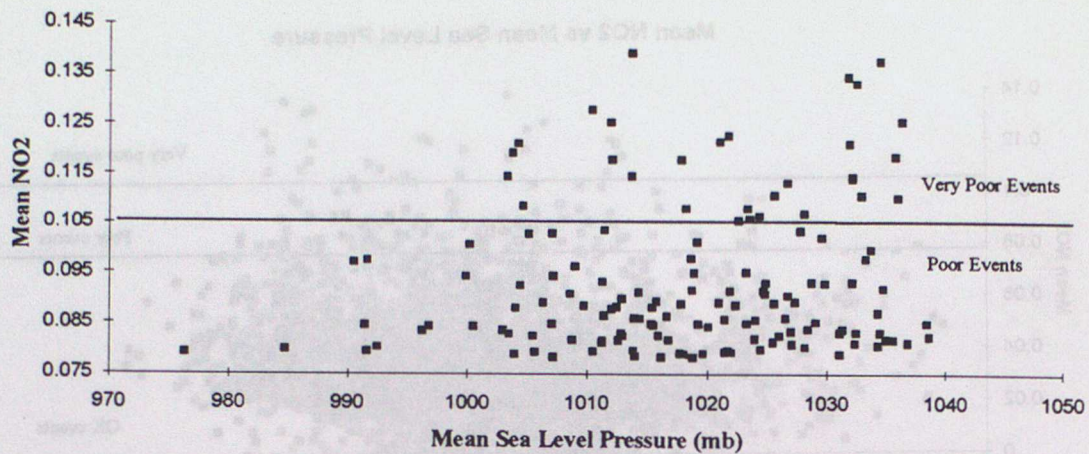


Figure 2: A scatter plot of Mean NO_2 for 'poor' and 'very poor' events against mean sea level pressure.

A further example of the typical pattern observed was the scatter plots for wind speed, at level two ($\sim 975\text{hPa}$) in the model, against mean NO_2 , Figure 3. Here, 'very poor' NO_2 events occurred only when the wind speed was below 10m/s , and for a 'poor' NO_2 event the wind speed should be below 16m/s . Days on which the air quality was 'OK', had a range of wind speeds from 0 to 30m/s , the majority of which occurred within the range 0 to 20m/s .

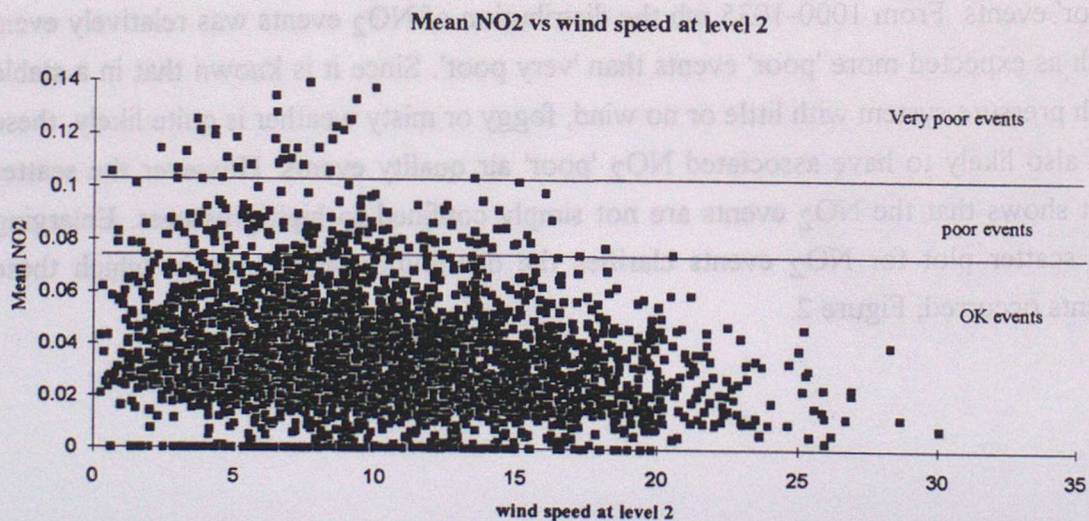


Figure 3: A scatter plot of mean NO_2 against wind speed at level 2.

It has been suggested that NO_2 events at relatively high wind speeds of $>10\text{m/s}$ may not be due to local sources of NO_2 . Such a complication is likely to weaken the analysis discussed in this paper, since the combination of local and distant sources may need a mixed distribution of two types of weather system to describe it.

2.2 Tabulation

Tabulating the data, cross-classified by a selection of NWP parameters is another fruitful approach to sifting the data. This involves categorising the data into mutually exhaustive and exclusive groups. Table 2 shows the results for tabulating the pressure by month and NO₂ event.

	Poor Events				Very Poor events			
	< 1000 mb	1000 - 1030 mb	> 1030 mb	Total	< 1000 mb	1000 - 1030 mb	> 1030 mb	Total
Jan	0	7	0	7	0	1	0	1
Feb	0	0	6	6	0	0	0	0
Mar	0	0	1	1	0	0	1	1
Apr	0	4	0	4	0	0	0	0
May	0	0	0	0	0	0	0	0
Jun	0	3	0	3	0	0	0	0
Jul	0	3	0	3	0	0	0	0
Aug	0	1	0	1	0	0	0	0
Sep	1	11	1	13	0	0	0	0
Oct	7	25	1	33	0	10	0	10
Nov	0	20	0	20	0	5	2	7
Dec	2	8	12	22	0	8	8	16
Total	10	82	21	113	0	24	10	34

Table 2: Table showing the months and range of pressures at which NO₂ events occur for 'poor' and 'very poor' classification.

The margins for MSLP (bottom row) highlight the results obtained from the scatter plot, i.e. the majority of NO₂ events occurred when the pressure was greater than 1000 mb, within the preferred range of 1000-1030 mb. The margins for the monthly variation (total columns), show that season has a definite effect on NO₂ concentration. The season for 'poor' events is from September to February, whereas the season for 'very poor' events is from October through to December. November shows some discrepancies when compared to October and December, with fewer occurrences in both NO₂ categories. This may be a result of a shift in the pattern between October and December. In October, the majority of the NO₂ events occur when the pressure is between 1000- 1030 mb. However by December, this pattern has shifted and now the majority of the events occur when the pressure exceeds 1030 mb.

It must be borne in mind that these results may be a quirk of the data, since there are only two and a half years of observations, and hence only a limited range of conditions are present within the individual months. After December, there is a marked

decline in the number of NO₂ events occurring. Virtually no 'very poor' NO₂ events occur outside the season and few 'poor' NO₂ events.

Considering this in terms of meteorology poses a problem. Climatologically there is little difference in the weather between these adjacent months, i.e. December and January, and the difference may be due to the data sample. However it also raises the question of whether there is an external factor influencing the build up of NO₂.

Table 3 shows how the NO₂ events relate to the wind speed at level 2. Most of the NO₂ events occurred when the wind speed was between 5-10 m/s. Again, 'very poor' events have a more restricted range than 'poor' events, with no 'very poor' events observed when the wind speed was greater than 15 m/s and only one event occurs in the 10 - 15 m/s category.

	< 5 m/s	5 - 10 m/s	10 - 15 m/s	> 15 m/s	Total
Poor events	28	45	30	6	109
Very poor events	9	20	1	0	30
Total	37	65	31	6	139

Table 3: A table showing the distribution of 'very poor' and 'poor' events when compared to wind speed at level 2 in the model.

3. Further Analysis.

Scatter plots and tables only reveal a limited amount of information about possible relationships between the observed events and the individual parameters. They will become complicated if more information is added, or the effects of several parameters are studied. For a deeper study, a different approach to the problem is needed.

3.1 Linear regression.

There are a number of different approaches which could be considered with this data set. One approach would involve the transformation of the observed NO₂ data to compensate for possible multiplicative errors. However, this would involve an in-depth analysis of the data, which is probably far more intricate than necessary at this stage in the analysis. The primary aim at this stage of the analysis was the selection of a subset of the original sixty-four plus variables. One such robust approach to this problem was linear regression. It is assumed that any parameter which is highly influential in the model using linear analysis would also be of importance in any model developed upon transformed values of NO₂ amount. The basic idea behind regression is to model the data in such a manner that minimises the differences between observed and predicted values of NO₂ amount. The model should have the minimum number of parameters which are necessary. There are two main reasons why the model should be conservative in terms of the number of parameters included; the first is in the ease of interpretation of the model, the second is in the stability of the model under different samples.

Stepwise regression is a method by which parameters are added into the model based on their ability to explain the variation in amount of NO₂. The parameters are ranked in terms of the amount of variation that each explains, the parameter with the largest amount of variation explained is then added to the model. The parameters left are reassessed in terms of the remaining variation, and again the parameter with the highest explanation of the variation (in terms of the F statistic) is added. This is repeated until the variation explained by the parameter falls below a specified threshold, in this case the 5% significant level. In addition, at every step, parameters in the model are tested to ensure that they still explain a significant amount of the variation, parameters that fail the test are discarded from the model.

Fitting the data using stepwise regression on purely meteorological parameters yields a model with nineteen parameters, and explains 47% of the variation between the predicted value for NO₂ and the observed value. The parameters selected for inclusion and exclusion from the model are shown in Table 4.

Step No.	Variable	Multiple R	RSQ	Change in RSQ	F	No. of variables included	Final coefficients
1	Dew point 2	0.3038	0.0923	0.0923	329.33	1	-0.00086
2	Wind speed 2	0.3834	0.1470	0.0547	207.39	2	-0.00117
3	Wind shear 1	0.4013	0.1611	0.0141	54.41	3	
4	Mean sea level pressure	0.4095	0.1677	0.0066	25.57	4	-0.00014
5	Direction 4	0.4151	0.1723	0.0047	18.21	5	0.000655
6	Dew point advection 2	0.4208	0.1771	0.0048	18.70	6	-0.00002
7	Surface temperature	0.4270	0.1823	0.0052	20.62	7	
8	Wind shear 2	0.4323	0.1869	0.0046	18.39	8	-0.00414
9	Temperature 4	0.4362	0.1903	0.0034	13.42	9	-0.00068
10	Curl 7	0.4399	0.1935	0.0032	13.00	10	-0.00041
11	Curl 4	0.4425	0.1958	0.0023	9.21	11	0.00062
12	Curl 1	0.4506	0.2031	0.0072	29.33	12	-0.00151
13	Temperature advection 3	0.4530	0.2052	0.0021	8.54	13	-0.00056
14	Curl 2	0.4558	0.2077	0.0026	10.40	14	0.00077
15	Divergence 6	0.4579	0.2097	0.0019	4.95	15	-0.00087
16	Wind speed 6	0.4599	0.2115	0.0018	7.48	16	-0.00054
17	Temperature 1	0.4615	0.2130	0.0015	6.03	17	0.00250
18	- (Surface temperature)	0.4610	0.2125	-0.0005	-2.01	16	
19	Temperature 2	0.4645	0.2158	0.0033	13.47	17	-0.00169
20	Dew point 3	0.4670	0.2181	0.0023	9.40	18	-0.00043
21	Dew point 1	0.4685	0.2195	0.0014	5.75	19	-0.00056
22	- (Wind shear 1)	0.4675	0.2186	-0.0009	-3.66	18	
23	Wind speed 5	0.4690	0.2200	0.0014	5.85	19	0.00050
24	Constant term					20	0.19465

Table 4: A summary table of parameters added and dropped from the model, along with the coefficients of the final model selected.

This is a cumbersome model and unlikely to be stable if a different sample of data is taken. It is a well-known result that adding parameters into the model will increase the R^2 term and correspondingly increase R , without necessarily improving the model. The stepwise method also assumes independence across cases. Since the data are clearly correlated in time, these tests will not be conservative.

Hence only the first three parameters are studied in detail, these contain most of the variation explained at 40%. Of more interest, are the type of parameters and the order in which they are selected for inclusion in the model. The first parameter chosen is dew point temperature at level 2. This is the temperature at which water is condensed out of the air; it is always lower or equal to the air temperature. (When it rains, temperature and dew point temperature are equal.) Dew point temperature is a highly seasonal parameter; it is generally lower in winter than summer and thus may be accounting for seasonal variation within the model. The wind speed might reflect the stability of the air mass; a low wind speed indicates a more stable atmosphere. As the wind speed rises, so the predicted value of NO_2 falls due to the wind dispersing the build up of NO_2 . This is indicated by the negative coefficient for the wind speed terms. The parameters selected so far have all been at level two in the Limited Area model.

This level may have been chosen in preference to level one as it is sufficiently above the ground to be unhindered by friction and channelling effects. The third parameter fitted (but later dropped from the model) is wind shear at level 1. This is a measure of the amount of twisting in terms of wind direction in the vertical direction and again can be thought of as a stability indicator.

Nineteen variables fitted in the model are too many and would make for a complicated and unwieldy model. The next step is to reduce the number of parameters. This may be overcome if some of the variables added into the model are compensating for a missing factor such as season. If the data are restricted to events where the wind speed at level 2 was less than 15 m/s and the season was autumn or winter, there may be more of a chance of observing the underlying mechanism. In addition there are several other parameters which prior experience suggests might be important, these are indicators of temperature and dew point inversion as well as dew point depression. Dew point depression is the difference between the actual temperature and the dew point temperature; low values indicate that a large amount of water is present within the air mass and hence may lead to low values of NO₂ being recorded, due to the NO₂ being washed out of the air. The model obtained, when these parameters were considered using stepwise regression on the restricted data set, is shown in Table 5.

Step No.	Variable	Multiple R	Multiple RSQ	Change in RSQ	F statistic	No. of variables included	Final Coefficients
1	Constant						0.34800
2	Time of day	0.4118	0.1696	0.1696	265.29	1	0.00141
3	Temperature 4	0.4594	0.2110	0.0414	68.16	3	-0.00057
4	Wind speed 2	0.5036	0.2537	0.0426	74.10	4	
5	Surface pressure	0.5088	0.2589	0.0053	9.20	5	-0.00029
6	Dew point 2	0.5197	0.2700	0.0111	19.73	6	-0.00217
7	Divergence 1	0.5244	0.2750	0.0049	8.83	7	0.00082
8	Surface Temperature	0.5293	0.2801	0.0052	9.27	8	0.00120
9	Dew point depression 1	0.5357	0.2869	0.0068	12.32	9	-0.00126
10	Wind speed 1	0.5385	0.2900	0.0030	5.50	10	-0.00239
11	-(Wind speed 2)	0.5382	0.2896	-0.0004	-0.65	9	

Table 5: Summary table of parameters included into the model when using a restricted data set and time dependent variables.

The first three parameters added explain over 50% of the variation between the observed and the predicted values of NO₂. Again a study of the order of fit shows that the time of day from which an observation came is highly relevant; this is verified by Table 6. (The margin totals are different from those of tables 2 and 3 as only observations that have a wind speed less than 15 m/s and occur in autumn or winter are used.) It is noticeable that NO₂ events mainly occur during the day between 9 am

and 9 pm. Events that occur outside this period may be due to carry over effects from the preceding six hour period. Restricting the observations to those available at 12 and 18Z, may help discover the underlying factors affecting NO₂ concentration, without masking the relationship by looking at similar meteorological conditions which have no NO₂ events.

	Poor	Very Poor	Total
06Z	4	0	4
12Z	36	15	51
18Z	37	14	51
00Z	6	1	7
Total	83	30	113

Table 6: A frequency table of NO₂ events that occur throughout the day.

The second parameter fitted into the model is temperature at level 4. Again this could be indicative of seasonal effects, perhaps the individual months should be added to the model? Wind speed at level 2 is the third parameter of importance fitted. It is observed that no parameter selected is above level 4 in the Limited Area model. The equation of the model now has eight parameters and takes the form:

$$\hat{y} = 0.3480 + 0.0014X_{\text{timeofday}} - 0.0006X_{\text{Temp4}} - 0.0003X_{\text{Surfacepressure}} - 0.0022X_{\text{Dewpoint2}} + 0.0008X_{\text{Divergence1}} + 0.0012X_{\text{Surfacetemperature}} - 0.0013X_{\text{Dewpointdepression1}} - 0.0024X_{\text{Windspeed1}}$$

This type of analysis was repeated several times, until a subset of parameters was decided upon which fitted well with prior expectation. The data set was restricted to observations that were valid only when it was daytime, the wind speed was less than 15 m/s and the season was either autumn or winter. This subset consisted of month, season, time of day, pressure (mean sea level and surface), dew point inversion indicator and temperature inversion indicator at or below level three in the model. In addition, the following parameters at or below level three were considered: wind speeds, wind shears, dew point temperatures, dew point depressions and temperatures. Wind direction rarely came into the stepwise analysis suggesting that there is at least no dominate distant source.

3.2 Discriminant analysis.

The next stage in this analysis was to develop an equation that would classify the observations into categories. This was achieved by first transforming the NO₂ amount into three groups as defined by Table 1b. Discriminant analysis is a technique which aims to maximise the separation between these groups of individuals. A more detailed explanation of discriminant analysis is given in Appendix One. Again this method was approached in a stepwise manner, using the selected subset of data. Table 7 shows the results obtained when the 'very poor' and 'poor' groups are combined together to form a super-group called 'event'.

Variable	F statistic to remove from equation	Classification Function	
		OK	Event
MSLP	4.19	8.65	8.63
Temperature 3	9.51	0.02	-0.09
Wind speed 1	29.16	13.75	13.47
Wind Shear 1	5.89	-34.15	-32.14
Wind Shear 2	5.98	55.30	56.62
September	8.55	2.05	0.99
December	33.12	23.55	25.69
Autumn	23.55	125.46	127.16

Table 7: A summary table for the model fitted under discriminant analysis.

The 'F statistic to remove' examines whether or not a parameter contributes significantly to the explanation of the model. Values greater than 4 are taken to be significant at the 5% level. The parameters selected fell into two groups those that were meteorologically dependent and those that are time dependent. At first sight it seems peculiar to have so many time variables included. However an examination of the summary data for months, for wind speeds at level 2 less than 15m/s, suggest that this result is reasonable. Autumn as defined consisted of the months from September to November, and winter as December to February. If the margin totals are examined, it is observed that September has comparatively few events compared to October and November. In this particular sample of data December has many more events compared to January and February. It may be better to define time in terms of occurrences of NO₂ giving a high, middle and low season. High season consists of the months October to December, middle is September, January and February and low is April through to August.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Poor	4	4	1	3	0	2	3	1	13	32	20	20	103
Very poor	1	0	0	0	0	0	0	0	0	9	5	15	30
Total	5	4	1	3	0	2	3	1	13	41	25	35	133

Table 8: A summary table of the occurrences of NO₂ events by months and categories.

A classification function was obtained for the two groups, which is the linear combination of parameters that best discriminates the groups. A new observation is defined to belong to the group that yields the greatest value.

There are several different methods of verifying the classification functions; one of the better methods is the Jack-Knife classification matrix. Each observation was grouped according to the classification function computed from the permissible data excluding the case under consideration, Table 9.

	Classified OK	Classified Event	Total	% Correct
OK	376	185	561	67
Event	29	83	113	74
Total	405	268	673	68
% Misclassification	7	69	32	

Table 9: Verification of the model developed using discriminant analysis.

The results show that there is a 74% chance of accurately classifying an event (known as 'probability of detection'). However, on 69% of occasions when an NO₂ event was forecast, the actual value of NO₂ fell into the 'OK' category; this type of misclassification is known as the 'false alarm rate'. Seven percentage of events classified as 'OK' were actually NO₂ events, known as the 'miss rate'.

This type of analysis was repeated a number of times redefining the seasons, using a different definition of an event, i.e. combining the 'OK' and 'poor' events together and comparing with the 'very poor'. A new variable was introduced indicating whether or not the day was a week day. Finally a model was selected that was valid for a data set which contained observations that had the following properties: wind speeds less than 15m/s, peak season as October to December and week days during the hours of 9 am and 9 pm. This model was a hybrid model defined from the past results and in conjunction with meteorological experience. The results from this model are shown in Table 10. These restrictions are chosen to improve sensitivity to NO₂ events and to

reduce the bulk of cases where no events were observed in the sample data set. This model, applied to any occasion, is known as the restricted model, and will give probabilities that are consider to be too high outside of the conditions for which the restrictions are valid. Applying extra weightings which are related to the prior likelihood of an event occurring will reduce the probabilities to more realistic levels. The weights change the restricted model to an unrestricted model.

Variable	F statistic to remove	Classification Function		
		Very poor	Poor	OK
Mean sea level pressure	5.72	7.69	7.71	7.74
Wind speed 1	11.01	11.95	12.17	12.35
Wind shear 2	6.32	-5.51	-9.11	-9.91
Dew point 2	22.23	6.20	6.39	6.57
Temperature inversion indicator	10.08	-87.02	-90.49	-90.48
Constant		-3935.75	-3953.76	-3990.56

Table 10: Summary table for final model selected.

It is observed that the removal of any one of the variables would result in a significant reduction in the variation explained by the model. The discriminant equations for the model are as follows;

$$\begin{aligned}
 B_{OK} &= -3990.56 + 7.74X_{MSLP} + 12.35X_{Windspeed\ 1} - 9.91X_{Windshear\ 2} + 6.57X_{Dewpo\ int\ 2} - 90.48X_{tempinversion} \\
 B_{Poor} &= -3953.76 + 7.71X_{MSLP} + 12.17X_{Windspeed\ 1} - 9.11X_{Windshear\ 2} + 6.39X_{Dewpo\ int\ 2} - 90.49X_{tempinversion} \\
 B_{Verypoor} &= -3935.75 + 7.69X_{MSLP} + 11.95X_{Windspeed\ 1} - 5.51X_{Windshear\ 2} + 6.20X_{Dewpo\ int\ 2} - 87.02X_{tempinversion}
 \end{aligned}$$

The B's are the discriminant scores for the respective groups. Values obtained from the scores are proportional to the log of the probability of an observation belonging to that group. With a little algebraic manipulation it is possible to derive an equation that will give the probability of an event for each category (see Appendix One).

The model is valid for the peak season under a very restrictive set of conditions. What is needed is a generalisation of this model to fit all seasons and conditions. In this, a heuristic approach was taken using the summary statistics described earlier, i.e. the distribution of occurrences of NO₂ events external to the restrictive conditions were studied and a subjective weighting applied.

3.3 Restricted and unrestricted models.

The restrictions on the discriminant model above (winds less than 15 m/s; daytime; October to December and a weekday) were chosen to pre-select those cases

in the sample with the greatest number of NO₂ events. It restored, a little, the imbalance between 'OK' and 'Poor' or 'Very Poor' events. However, formally, the discriminant model chosen applies only to the restricted set (i.e. cases included in the analysis), and does not apply to the excluded cases.

If the discriminant model is used on cases, either in the sample or on independent cases where the restrictions excluded that case from the analysis, then a model, *conditional on the restrictions*, to a case which it does not cover is applied. This model is called the Restricted model, for convenience.

It is assumed that the model is approximately true; that the NWP terms chosen in the restricted set are still important in describing the occurrence of NO₂ in the excluded set. However, since there are very few cases in the excluded set, estimates cannot be derived for the excluded model in the manner used for the restricted set. It assumes that the analysis can be extrapolate from the restricted set of data to all cases. If it is assumed that the NWP terms are still likely to be important, a mechanism needs to be found to convert the probabilities estimated using the restricted model to values which more accurately describe the excluded cases.

The restricted probabilities are found from the equations of Appendix One (19 to 21). These probabilities will be too high for the excluded set, for the prior knowledge of the proportions of the three events are different.

By using this extra knowledge, weightings can be derived which adapt the "Restricted" probabilities to "Unrestricted" probabilities. The unrestricted probabilities should be the same as the restricted probabilities for the restricted data set, but should reflect the reduced likelihood of finding an event in the excluded data set.

The weightings can be allocated separately for each of the restrictions, and the modified equations of Appendix One, of the style of equation 26 applied. These weightings should be close to the ratio of the proportions of the event in the included (restricted) and excluded cases.

Extra knowledge can also be used to modify the weights. For example, NO₂ events almost solely apply to data from 12Z and 18Z. However the model is applied case by case, and the true probabilities for the period following a known event are likely to be higher than estimated by the model, because NO₂ conditions are likely to linger into the evening.

Similarly, there were relatively few cases in February, in the two and a half years of the sample. Meteorological experience, however, can suggest no good reason a priori why a February NO₂ event would be much less likely to occur than a January one. So experience suggests that the February weighting probably ought to be the same as for January, despite the numbers in the sample.

4. Verification of the model.

The model verification is split into two sorts, that for the unrestricted model with no restrictions and that for the restricted model. The verification used for the unrestricted model is likely to be overly optimistic, because the data used to verify the model is the same data as that used in the development stage. However since the entire analysis has used NWP model forecast fields at T+24 and T+30, any model forecast errors have already been introduced.

4.1 Tabulated verification.

For the restricted data the Jack-Knife classification matrix was used; the results of which are shown in Table 11.

Actual group	Classified groups			Total	% correct
	OK	Poor	Very poor		
OK	160	75	27	262	61
Poor	23	29	10	62	47
Very poor	1	11	15	27	56
Total	184	115	52	351	60
% over classified	-	65	71		
% under classified	13	10	-		
% misclassified	13	75	71		

Table 11: Verification results of the final model selected.

These results show that 60% of the observations are accurately classified. The model tends to over-predict events, i.e. predict an event that is more severe than the one that actually occurs.

Three different occasions are studied for conditions that fall outside the restrictions set in the development of the model. Results for both restricted and unrestricted forecasts are studied. The first case studied is for peak season on a weekend at night, with a wind speed, at level two in the model, forecast as less than 9 m/s. The restricted model forecast 4 'OK' events, 11 'poor' events and 36 'very poor' events. The unrestricted model down-weights the 'very poor' and 'poor' events to 'OK', a total of 51; no NO₂ events were actually observed.

The second occasion examined is for wind speed at level 2 greater than 9m/s, during the winter on a weekday between 9 am and 9 pm, Table 12.

Actual observation	Unrestricted Forecast				Total	Restricted Forecast			
	OK	poor	Very poor	% correct		OK	Poor	Very poor	% correct
OK	134	11	3	91	148	103	39	6	70
Poor	20	11	2	33	33	15	15	3	45
Very Poor	1	0	1	50	2	0	0	2	100
% under forecast	16	0	-			15	0	-	
% over forecast	-	50	83			-	72	81	
% misclassified	16	50	83			15	72	81	

Table 12: Verification results for restricted and unrestricted forecasts for when the wind at level 2 was greater than 9 m/s, winter, daytime and a weekday.

Comparing the unrestricted forecast to the restricted forecasts gives an increase in the percentage forecast as 'OK', but this is achieved at the expense of the percentage correct for the 'very poor' and 'poor' categories. The other difference between the two forecast types is that the percentage over-classified decreases from the restricted to the unrestricted, i.e. the number of false alarms fell.

The final case is for low and middle season with a wind speed greater than 9 m/s on a weekday during 9 am and 9 pm, Table 13.

Actual observation	Unrestricted Forecast				Total	Restricted Forecast			
	OK	Poor	Very poor	% correct		OK	Poor	Very poor	% correct
OK	445	0	0	100	445	379	57	9	85
Poor	7	0	0	0	7	5	2	0	13
Very Poor	0	0	0	100	0	0	0	0	100
% under forecast	2	0	-			1	0	-	
% over forecast	-	0	0			-	97	100	
% misclassified	2	0	0			1	97	100	

Table 13: Verification results for low and middle season, wind speed less than 9 m/s on a weekday in daytime.

Here the unrestricted forecasts dramatically decreases the number of over-forecast events. This is achieved at the expense of the model's ability to forecast 'very poor' and 'poor' events.

4.2 Verification using histograms.

Figure 14 shows a series of six histograms, each of these histograms represent the frequency of the quantity of NO_2 observed separated into 0.01 intervals for both the restricted and unrestricted forecast groups. Histograms for the separate event groups (in different rows) have different ordinate scales but each pair of histograms in a row have the same scale. This has meant cutting short some of the group values for the restricted model for 'poor' and 'very poor' events to highlight the effect at high observed NO_2 levels. Where this has occurred an arrow is printed above the appropriate bar pointing to a number which is the actual frequency.

The histograms highlight the distinct differences between the restricted and unrestricted models within the forecast groups. It is also noticeable that the effect of weighting the restricted model to obtain the unrestricted model has resulted in a higher proportion of events categorised as 'OK' for the unrestricted model. The consequence of this is a rise in the number of forecasts under-classified. (These forecasts are shown from approximately 0.08 to 0.15 in the histograms.)

The middle pair of histograms in Figure 14 are for observations forecast as 'poor'. (A poor event corresponds to the approximate region of 0.07 to 0.10.) It is observed that the restricted forecast is bad at predicting 'poor' events. The unrestricted forecasts performs much better. A further point to note is the steady decline in the number of forecast events, as the amount of observed NO_2 falls.

The final pair of histograms are for observations forecast as 'very poor'. Again it is noticeable that the restricted model tends to over-classify the events. This has decreased in the unrestricted model, but there is still a large number of forecasts misclassified. The bimodal appearance of the unrestricted histogram is unexpected and is probably a result of this particular sample of data.

Care should be taken when examining the histograms from different event classes due to the differences in scale. A careful study reveals that the gain in accuracy for forecasting 'very poor' and 'poor' events is at the expense of the accuracy in forecasting 'OK' events.



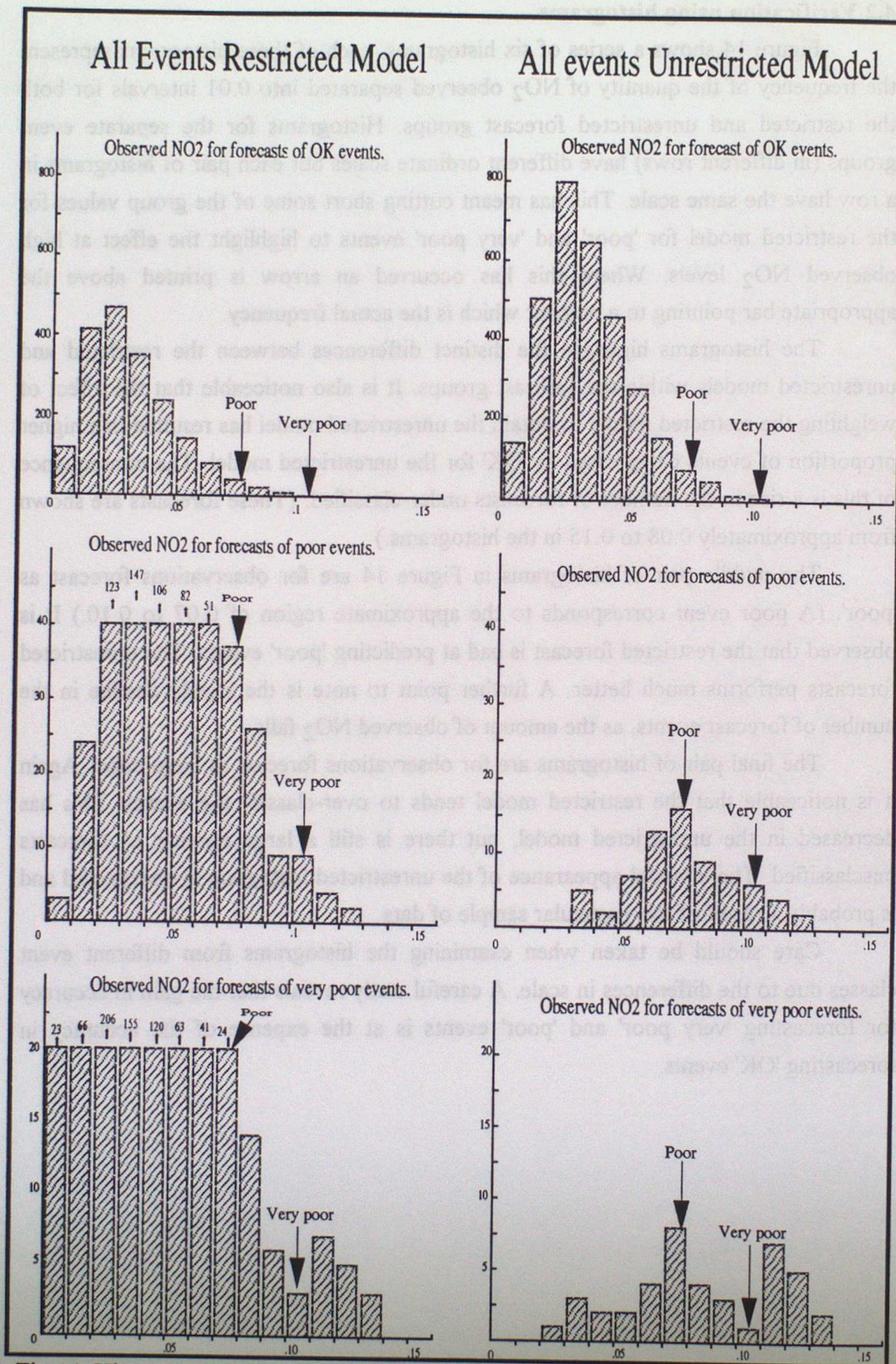


Fig 14: Histograms of NO₂ for restricted and unrestricted forecasts

5. Conclusions.

The results of this analysis have shown that there is skill in predicting a NO₂ event under certain conditions. These conditions are a weekday from 9 am to 9 pm during the months October through to December, with little or no wind forecast at level two in the Limited Area model. The model was generalised for any set of conditions, by means of heuristically derived weights to form an unrestricted forecast. These weights can easily be altered should further data or information be obtained. An investigation into NO₂ levels in the atmosphere by the United Kingdom Photochemical Review group (1990) has shown that NO₂ events can occur all year round, with little seasonal change. This is very much at variance with the sample data used in this investigation.

In the verification, this unrestricted model was shown to have some skill in predicting the type of NO₂ day, but was not as good at predicting NO₂ events. However this loss of accuracy was compensated by the decrease in the number of events misclassified, especially events that were over-classified.

The large number of events which were over-forecast could indicate that an external parameter, (external to the ones already present in the model, meteorological or otherwise) is influencing the accumulation of NO₂. The present model might accurately predict days on which the conditions are right for an event to occur, provided the external factor is present. It has been suggested that one such possible factor could be the presence of up wind NO₂ sources; these may have a confounding effect with local NO₂ sources. The effects of the two sources (local and distant) may be combining together so that the cut point for a NO₂ event is exceeded. If this is the case then an obvious parameter to include into any further study would be that of wind direction. This term was considered early on in the analysis, but rejected at the subset selection stage. Wind direction as stored in the MOS archive is at level four (~870hPa) only. While this may not be the most relevant level in the model atmosphere for forecasting NO₂ events, at high wind speeds it should not be badly wrong. An alternative explanation could be that the NO₂ events are extremely local, and when an event is not observed at the Old Market site, it is possible that one does occur in the local vicinity.

Both the restricted and unrestricted model consider the six hour periods for which the mean values of NO₂ are calculated as independent of one another. Hence categories may need to be adjusted, if it is considered that there is a carry-over effect from one period to the next, i.e. the underlying meteorological conditions remain constant and a NO₂ event occurs in the first period, but is not forecast for the second. The weights applied to the restricted probabilities are based upon a limited period of

data and should be adjusted if the forecaster considers it necessary. One such adjustment needed is for Bank holidays, where a weighting corresponding to the weekend rather than the weekday might be more appropriate.

An example of a forecast for 15th February 1994 from the midday run of the Limited Area model with a lead time of 0Z is shown in Figure 4.

AIR QUALITY FORECASTS - NO2 FOR BRISTOL		
~~~~~		
VALID FOR 15/ 2/1994 AT 12Z: FORECAST LEAD TIME 0Z		
~~~~~		
	CATEGORIES	
	POOR	VERY POOR
RESTRICTED PERCENT	30%	68%
WEIGHTS FOR WINDSPEED AT LEVEL 2	1.00	1.00
WEIGHTS FOR WEEKDAY/END	1.00	1.00
WEIGHTS FOR TIME OF DAY	1.00	1.00
WEIGHTS FOR MONTH	0.20	0.01
UNRESTRICTED PERCENT	6%	1%

Figure 4: An example of an NO₂ forecast for 12Z 15/2/1994, T+00Z.

The restricted forecasts for the 'poor' and 'very poor' categories are 30 and 68 percent respectively. The unrestricted forecast has down-weighted these results to 6% for a 'poor' event and 1% for a 'very poor' event. The weights work by assessing which conditions are valid for the restricted forecast. The wind speed at level 2 forecast for 12Z 15/2/94 was 5.3m/s. This is within the permissible range for restricted forecasts and hence the weighting applied was one for both NO₂ events. The 15th was a Tuesday, which again falls within the restrictions set for the restricted model, as does the time of day at 12Z. However, February falls outside the restricted set used to develop the restricted model and consequently has a weight assigned for the month of 0.2 and 0.01 respectively for 'poor' and 'very poor' events, as mentioned earlier.

Another example is shown in Figure 5 for 0Z on 24th October 1993 with a lead time of T+24Z.

AIR QUALITY FORECASTS - NO ₂ FOR BRISTOL		
~~~~~		
VALID FOR 24/10/1993 AT 0Z: FORECAST LEAD TIME 24Z		
~~~~~		
	CATEGORIES	
RESTRICTED PERCENT	POOR	VERY POOR
	15%	63%
WEIGHTS FOR WINDSPEED AT LEVEL 2	0.80	0.10
WEIGHTS FOR WEEKDAY/END	0.20	0.05
WEIGHTS FOR TIME OF DAY	0.10	0.05
WEIGHTS FOR MONTH	1.00	1.00
UNRESTRICTED PERCENT	0%	0%

Figure 5: An example of an NO₂ forecast for 0Z 24/10/1993, T+24Z.

The restricted forecasts gave probabilities of 15 and 63 percent for 'poor' and 'very poor' events respectively. The forecast wind speed at level 2 was between 12 and 13 m/s and is assigned weights of 0.8 and 0.01 for 'poor' and 'very poor' categories. The 24th October 1993 was a Saturday. Both Saturday and a validity time of 0Z are outside the set of restrictions used to develop the restricted model, so appropriate weights are applied for each parameter and event. October falls within the peak NO₂ season, and therefore is assigned weights of one. The weights for the appropriate NO₂ category are multiplied together and applied to the restricted probability to obtain the unrestricted probability of an event occurring, in this case 0% for both categories. This should not be taken as a definite non-occurrence rather that the chances of an event occurring are extremely small. A similar point is valid if a forecast of 100% was obtained.

There are limitations to the model developed; these depend ultimately on the end user's requirements and cost analysis. Appendix Two gives further details about cost analysis. The user must consider the worse case in the cost analysis: that of acting on a forecast which is a false alarm compared to not acting when a NO₂ event occurs. The outcome of this decision may well prefer either the restricted or the unrestricted model. However it is more likely that in certain situations one model is more appropriate than the other; this knowledge can only be gained through experience. As the forecasters gain more experience of the scheme, it is hoped that they might adjust the weights applied to the restricted forecast, i.e. if the meteorological conditions in

January were similar to those in December the forecaster may consider adjusting the weight applied to the month towards that of December.

To summarise, the unrestricted model generally forecasts 'poor' and 'very poor' events well. However, this is at the expense of forecasting 'OK' events. The effects of this are twofold. Firstly there will be comparatively few false alarms, that is forecasting a NO_2 event when none occurs. Secondly, there will be occasions when an event forecast as 'OK' will be classified as 'poor' or 'very poor'. The consequences of this could be the loss of any action usually taken on receipt of a warning, such as warning the general public and the possible loss in confidence of the model.

The next stage in this investigation should be a trial period when the models can be tested on actual data. This will give both the forecaster and the end user time to gain experience of the model and assess its usefulness. Ideally the model should be assessed in a variety of seasons.

APPENDIX 1.

DESCRIPTION OF DISCRIMINANT ANALYSIS.

Every NO₂ observation was classified into three mutually exclusive and exhaustive groups: 'OK', 'Poor' and 'Very poor'. The data set of observations, NWP and related parameters was used as a training set to find rules that could allocate future observations to one of the three groups. The procedure followed is described below:

Let \underline{X} be a vector of NWP and related parameters

$f_i(\underline{X})$ is the density function of \underline{X} in group i . $i = \text{'OK', 'Poor', 'Very poor'}$

π_i is the proportion of population in group i

$\pi_i f_i(\underline{X})$ is the proportion of the total population with $\underline{X} = \underline{x}$ that are in group i .

The expected number of misclassifications are minimised if observations with $\underline{X} = \underline{x}$ are allocated to the group that has the greatest $\pi_i f_i(\underline{x})$. However, $f_i(\underline{x})$ and π_i are unknown and are estimated from the data set. It is assumed that the distribution of \underline{X} within the NO₂ groups is a conditional gaussian distribution (that is a distribution where the marginal distribution of the discrete variables is multinomial, and the conditional distribution of the continuous variables given the discrete variable is multivariate normal with parameters that depend on the discrete variables) with mean vector $\underline{\mu}_i$ and covariance matrix Σ_i . (This is approximately true for all the continuous NWP model parameters which were used.) There are three distinct pairings of the groups these are 'OK' and 'Poor', 'OK' and 'Very poor' and 'Poor' and 'Very poor'. In each NO₂ group the multivariate normal function has the form:

$$f_i(\underline{x}) = (2\pi)^{-p/2} |\Sigma_i|^{-1/2} \exp \left\{ -1/2 (\underline{x} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{x} - \underline{\mu}_i) \right\} \quad -(1)$$

Where p is the number of continuous variables. The optimal allocation rule is to allocate to group i if;

$$\pi_i f_i(\underline{x}) > \pi_j f_j(\underline{x}) \quad i, j = \text{'OK', 'Poor', 'Very Poor'} \quad i \neq j \quad -(2)$$

This is equivalent to;

$$A = \log \left\{ \frac{\pi_i f_i(\underline{x})}{\pi_j f_j(\underline{x})} \right\} > 0 \quad -(3)$$

$$= \log \left\{ \frac{2\pi_i \Pi^{-\frac{p}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\underline{x} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{x} - \underline{\mu}_i) \right\}}{2\pi_j \Pi^{-\frac{p}{2}} |\Sigma_j|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\underline{x} - \underline{\mu}_j)' \Sigma_j^{-1} (\underline{x} - \underline{\mu}_j) \right\}} \right\} > 0 \quad -(4)$$

$$= \frac{1}{2} (\underline{x} - \underline{\mu}_j)' \Sigma_j^{-1} (\underline{x} - \underline{\mu}_j) - \frac{1}{2} (\underline{x} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{x} - \underline{\mu}_i) + \log \left(\frac{\pi_i}{\pi_j} \right) - \frac{1}{2} \log \left\{ \frac{|\Sigma_j|}{|\Sigma_i|} \right\} > 0 \quad -(5)$$

The method described is not very robust. This is overcome by assuming a common covariance function, Σ .

$$\Rightarrow A = \frac{1}{2} (\underline{x} - \underline{\mu}_j)' \Sigma^{-1} (\underline{x} - \underline{\mu}_j) - \frac{1}{2} (\underline{x} - \underline{\mu}_i)' \Sigma^{-1} (\underline{x} - \underline{\mu}_i) + \log \left\{ \frac{\pi_i}{\pi_j} \right\} > 0 \quad -(6)$$

Equation 6 is called the linear discriminant function. When parameters Σ and $\underline{\mu}$ are unknown they are estimated by S and $\hat{\underline{\mu}}$ respectively from the sample data set.

$$\Rightarrow \hat{A} = \underline{x}' S^{-1} (\hat{\underline{\mu}}_i - \hat{\underline{\mu}}_j) - \frac{1}{2} \left(\hat{\underline{\mu}}_i' S^{-1} \hat{\underline{\mu}}_i - \hat{\underline{\mu}}_j' S^{-1} \hat{\underline{\mu}}_j \right) + \log \left(\frac{\hat{\pi}_i}{\hat{\pi}_j} \right) > 0 \quad -(7)$$

Let B_i be defined as;

$$B_i = \underline{x}' \Sigma^{-1} \underline{\mu}_i - \frac{1}{2} \underline{\mu}_i' \Sigma^{-1} \underline{\mu}_i + \log \pi_i \quad -(8)$$

$$\Rightarrow B_{OK} = \underline{x}' \Sigma^{-1} \underline{\mu}_{OK} - \frac{1}{2} \underline{\mu}_{OK}' \Sigma^{-1} \underline{\mu}_{OK} + \log \pi_{OK} \quad -(9)$$

$$\Rightarrow B_{Poor} = \underline{x}' \Sigma^{-1} \underline{\mu}_{Poor} - \frac{1}{2} \underline{\mu}_{Poor}' \Sigma^{-1} \underline{\mu}_{Poor} + \log \pi_{Poor} \quad -(10)$$

$$\Rightarrow B_{Very\ poor} = \underline{x}' \Sigma^{-1} \underline{\mu}_{Very\ poor} - \frac{1}{2} \underline{\mu}_{Very\ poor}' \Sigma^{-1} \underline{\mu}_{Very\ poor} + \log \pi_{Very\ poor} \quad -(11)$$

The group with the largest value of B (Equations 9-11) is allocated the observation. The B_i 's are proportional to $\log \{ \text{probability} \}$ of group membership. Hence;

$$B_{OK} \propto \log \{ p_{OK} \} \quad -(12)$$

$$B_{Poor} \propto \log \{ p_{Poor} \} \quad -(13)$$

$$B_{Very poor} \propto \log \{ p_{Very poor} \} \quad -(14)$$

Re-organising these equations:

$$p_{OK} \propto \exp \{ B_{OK} \} \quad -(15)$$

$$p_{Poor} \propto \exp \{ B_{Poor} \} \quad -(16)$$

$$p_{Very poor} \propto \exp \{ B_{Very poor} \} \quad -(17)$$

As the NO_2 categories are exhaustive and exclusive the sum of the group probabilities will sum to one.

$$p_{OK} + p_{Poor} + p_{Very poor} = 1 \quad -(18)$$

Thus if Equations 15-17 are normalised the following probability equations are obtained;

$$p_{OK} = \frac{e^{B_{OK}}}{e^{B_{OK}} + e^{B_{Poor}} + e^{B_{Very poor}}} = \frac{1}{1 + e^{B_{Poor} - B_{OK}} + e^{B_{Very poor} - B_{OK}}} \quad -(19)$$

$$p_{Poor} = \frac{e^{B_{Poor}}}{e^{B_{OK}} + e^{B_{Poor}} + e^{B_{Very poor}}} = \frac{e^{B_{Poor} - B_{OK}}}{1 + e^{B_{Poor} - B_{OK}} + e^{B_{Very poor} - B_{OK}}} \quad -(20)$$

$$p_{Very poor} = \frac{e^{B_{Very poor}}}{e^{B_{OK}} + e^{B_{Poor}} + e^{B_{Very poor}}} = \frac{e^{B_{Very poor} - B_{OK}}}{1 + e^{B_{Poor} - B_{OK}} + e^{B_{Very poor} - B_{OK}}} \quad -(21)$$

So the effect of normalising equations 15-17 means that the three discriminant functions, equations 9-10 can be reduced to two terms:

$$B_{Poor} - B_{OK} = \bar{x}' \Sigma^{-1} (\bar{\mu}_{Poor} - \bar{\mu}_{OK}) - \frac{1}{2} (\bar{\mu}_{Poor}' \Sigma^{-1} \bar{\mu}_{Poor} - \bar{\mu}_{OK}' \Sigma^{-1} \bar{\mu}_{OK}) + \log \left(\frac{\pi_{Poor}}{\pi_{OK}} \right) \quad -(22)$$

$$B_{Very\ poor} - B_{OK} = \bar{x}' \Sigma^{-1} (\bar{\mu}_{Very\ poor} - \bar{\mu}_{OK}) - \frac{1}{2} (\bar{\mu}_{Very\ poor}' \Sigma^{-1} \bar{\mu}_{Very\ poor} - \bar{\mu}_{OK}' \Sigma^{-1} \bar{\mu}_{OK}) + \log \left(\frac{\pi_{Very\ poor}}{\pi_{OK}} \right) \quad -(23)$$

The equations 19-21 define the probability of an event occurring for a value of X which is of comes from the distribution described by the sample - the restricted case as described in the text.

If however the event associated with X occurs in a sample from a distinct population to that described above which contains a different proportion of 'OK', 'Poor' and 'Very poor' events π'_{OK} , π'_{Poor} and $\pi'_{Very\ poor}$ say, then the probabilities can be adapted with:

$$\rho_{OK} = \frac{\pi'_{OK}}{\pi_{Poor}} \text{ etc.} \quad -(24)$$

Adapting equations 15-21:

$$P_{OK} \propto \rho_{OK} \exp\{B_{OK}\} \text{ etc.} \quad -(25)$$

$$P_{OK} = \frac{\rho_{OK} e^{B_{OK}}}{\rho_{OK} e^{B_{OK}} + \rho_{Poor} e^{B_{Poor}} + \rho_{Very\ poor} e^{B_{Very\ poor}}} = \frac{1}{1 + \frac{\rho_{Poor}}{\rho_{OK}} + \frac{\rho_{Very\ poor}}{\rho_{OK}}} \text{ etc.} \quad -(26)$$

The ratio $\frac{\rho_{Poor}}{\rho_{OK}} = L_{Poor}$ and $\frac{\rho_{Very\ poor}}{\rho_{OK}} = L_{Very\ poor}$ are the weights used to convert the restricted probabilities to the unrestricted probabilities.

APPENDIX TWO.

COST ANALYSIS

Cost analysis (Decision Analysis, Risk Analysis etc.) is a method to analyse the decision process based on forecast probabilities, by assigning costs to each action and outcome. The probabilities derived from discriminant analysis can be used with the costs of each action and outcome to estimate the expected cost of a decision. In the analysis of this paper, a default cost was assumed, i.e. each cross classification and outcome had identical costs. So the way to classify a forecast event was to choose that event which had the greatest forecast probability.

However, in real life costs are rarely equal. For simplicity, consider the binary case, where an NO₂ event does or does not occur. Consider forecasts of NO₂ events in terms of one occurring, ✓, or not occurring, ✗. Table 14 shows the four possible outcomes.

Actual Event	Warning issued	Cost of Action
✗	✗	$C_{NW, No}$
✗	✓	$C_{W, No}$
✓	✗	$C_{NW, Yes}$
✓	✓	$C_{W, Yes}$

Table 14: Four possible outcomes available when forecasting a NO₂ event.

There are different costs involved in taking action on the basis of the above forecasts and outcomes. (The costs described here are not necessarily monetary costs. In practice they should be assessed to meet the end user's requirements and expectations. The cost of making the forecast, although not zero, is constant across all four classes, and therefore can be ignored.)

The cost of accurately predicting an 'OK' day consists of only the basic cost, and if this is subtracted from all classes, can be set zero. $C_{NW, No} = 0$ (No Warning, No event)

If a NO₂ event is forecast and such an event does not occur, the extra cost is that of the action taken (e.g. more communication costs, warning people to stay indoors, laying on more public transport) plus the cost of the loss of confidence in the forecast by the general public when the event did not, in fact, occur. Let this cost be $C_{W, No}$ (Warning, No event)

The other type of misclassification is that of a NO₂ event occurring where none was forecast. This might include the loss of public confidence in the forecast, plus the

health costs to the general public which occurs as a result of an NO₂ event. Cost $C_{NW, Yes}(\text{No Warning, Event})$

The final case is that of accurately forecasting a NO₂ event. The cost incurred would be the action taken as a result of the warning, plus the health costs of an event, minus any mitigating health benefit of public actions taken on the basis of a warning. There might also be a reduction in cost attributable to the approval of the public for getting a forecast right! The cost is $C_{W, Yes}(\text{Warning, Event})$

The aim of cost analysis is to minimise the expected overall cost in terms of a probability weighted classification of events, assessing the costs on a common scale. Effectively we are reassessing the probability level at which we classify an event.

The expected cost of making the decision to issue a warning is:

$$\text{Prob}(\text{event}) * (C_{W, No} + C_{W, Yes})$$

The expected cost of not issuing a warning is:

$$(1 - \text{Prob}(\text{event})) * (C_{NW, Yes} + 0)$$

The classification probability is the value at which the two expected costs are equal. If the forecast probability of an event is greater than this value, then the cost analysis says that a warning should be issued.

However, these costs are not always easy to quantify. Nor are they the same for all observers. The costs will be different for the advisory services, compared to different users.

An asthmatic who listens to the forecast probability, and who knows he (or she) suffers very badly in an occasion of high NO₂, might decide to stay in even if the probability was only 0.1. His cost of being caught off guard by an NO₂ event ($C_{NW, Yes}$) might be ten times higher than his cost of taking precautions and staying in ($C_{W, No} + C_{W, Yes}$).

The opposite case might be a manufacturer whose action on a warning is to shut down his plant in advance. He might perceive his only cost of being caught off guard (perhaps not having the time to shut down before the event passes by), is any adverse publicity. This might even be zero if he could pass the blame on to the advisory service for not predicting an event! However his cost of shutting down either correctly or unnecessarily ($C_{W, No} + C_{W, Yes}$) grossly outweighs not taking action at all ($C_{NW, Yes}$). His decision probability would be so near 1.0, (a difference much smaller perhaps, than the error in the forecast probability), that he would never take action.

REFERENCES

BMDP Statistical Software, Inc. (1990) *BMDP*.

City of Bristol Health and Environmental Services Technical Services Section, (1993).
Air Quality in Bristol.

United Kingdom Photochemical Oxidants Review Group, (1990). *Oxides of Nitrogen in the United Kingdom*. Department of the Environment & Department of Transport.