THE VERIFICATION OF ROAD WEATHER FORECASTS AND PERFORMANCE RELATED ROAD WEATHER CONTRACTS

Dr John E. Thornes
Director of the Birmingham Climate and Atmospheric Research Centre (BCAR)
School of Geography
University of Birmingham
Birmingham UK
B15 2TT
J.E.Thornes@bham.ac.uk
Phone/Fax: 121-414-5556

Introduction

Highway Engineers (Road Masters) are operating within an increasingly commercial environment and have to attempt to prove that they are getting value for money from all expenditure. It is difficult to independently assess the quality and accuracy of road weather forecasts and most highway authorities rely on statistics provided by the weather service providers. This paper sets out the commonly used indicators of weather forecast quality and accuracy so that highway engineers can assess for themselves the quality of the service they are getting. This also allows for the setting up of performance related contracts. For example if a winter road weather contract is priced at £25,000 for a percent correct frost prediction record of 86%, it could be stipulated that for every percent that the road weather forecast is below 86% : £1,000 is deducted from the value of the contract and for every percent above 86% : £1000 is added to the value of the contract. This would provide an incentive to the weather forecast provider to give the best possible service and would also ensure that the highway authority is getting value for money. The other important use of verification statistics is to help decide on which weather forecast supplier to use. It is now common in many countries that the National Met. Service is having to face competition from private weather companies. Verification can also be used as a safeguard if the service provider is new and has no track record.

A detailed discussion of the assessment of weather forecast quality is given in Stanski et al (1989) and Thornes (1995,1996). This paper is just concerned with the verification of road weather forecasts and how that information can be used by Highway Engineers to keep a 'sharp eye' on their road weather forecast suppliers.

The simplest way of assessing the quality of road weather forecast is to construct a simple 2 X 2 contingency table. This is a very flexible technique that can be used for example for verifying road surface temperature data, road wetness data, snowfall data, or timing data. Currently in the United Kingdom road surface temperature data is normally used, as measured by road sensors, but obviously a variety of contingency tables can be constructed from road weather information systems (RWISs). Ten statistical measures that can be calculated from the Contingency Table are given below. It is up to the Highway Engineer to decide which are the most appropriate for performance related contracts.

The Contingency Table

There are many different statistical measures of forecast quality that can be derived from the contingency table. Some are more relevant to road weather forecast verification than others as will be discussed below. Let us take the normal example of assessing whether or not the weather forecast predicts correctly if there will be a road frost, i.e. that the road surface temperature will fall to zero degrees Celsius or below. For simplicity the example is for a sample of 100 winter nights.

		FORECAST		
		FROST	*NO FROST*	*TOTAL*
	FROST	F/F (20)	NF/F (10)	O(F) (30)
OBSERVED	*NO FROST*	F/NF (12)	NF/NF (58)	O(NF) (70)
	TOTAL	P(F) (32)	P(NF) (68)	Total (100)

Key: F/F : Frost Forecast and Frost Observed (20 nights) Forecast correct

NF/F : No Frost Forecast but Frost (10 nights) also called a Type 1 error

F/NF : Frost forecast but No Frost Observed (12 nights) also called a Type 2 error

NF/NF : No Frost Forecast and No Frost Observed (58 nights) Forecast correct

O(F) : Total number of nights Frost Observed (30 nights)

O(N/F) : Total Number of Nights Frost not Observed (70 nights)

P(F) : Total number of nights Frost Forecast (32 nights)

P(N/F) : Total number of nights Frost Not Forecast (68 nights)

From all this information how can we assess the forecast quality? Here are a number of statistical measures that are easily calculated from these numbers:

1) **Percent Correct:** This is the most commonly used statistic and is calculated as follows:

$$\text{Percent Correct} = \frac{(F/F + NF/NF) \times 100}{\text{Total}} = \frac{(20 + 58) \times 100}{100} = \mathbf{78\%}$$

In other words it is the percentage of correct forecasts. In the United Kingdom the target for Percent Correct is 86%.

2) **Hit Rate (HR):** This calculates how many of the observed frosts were actually Forecast. This statistic can be misleading. For instance if a frost was forecast every night then the Hit Rate would be 100%!

$$\text{Hit Rate} = \frac{(F/F) * 100}{O(F)} = \frac{(20) * 100}{30} = 67\%$$

3) **Miss Rate (MR):** This calculates how many of the observed frosts were missed by the forecast. This is an important statistic as it includes all the Type 1 errors, in other words those nights when no frost was forecast but a frost actually occurred which could lead to accidents on icy roads.

Miss Rate = (NF/F) * 100 = (10) * 100 = 33%

-------- ------

O(F)	30

4) **Probability of Detection (POD):** The hit and miss rates are calculated in relation to the observed number of frosts. The probability of detection and the false alarm rate (see 5) below) are calculated from the total number of forecast frosts (32).

$$\text{Probability of Detection} = \frac{(F/F)}{P(F)} * 100 = \frac{20}{32} * 100 = 63\%$$

5) **False Alarm Rate (FAR):** This is an important statistic as it looks at the number of Type 2 errors, ie the number of nights that a frost was forecast but did not occur. These nights are when roads may be salted unnecessarily. The UK Met.Office has set a target of achieving a False Alarm Rate of less than 33% for its Open Road Forecasts.

$$\text{False Alarm Rate} = \frac{(F/NF)}{P(F)} * 100 = \frac{12}{32} * 100 = 38\%$$

It is important to consider which of these last four statistics to use as they appear to give similar results. Consider the extreme case of where a frost was forecast every night then the **Hit Rate** would be 100% and the **Miss Rate** would be 0%! However the **Probability of Detection** would only be 30% and the **False Alarm Rate** would be 70%. Therefore it is always preferable to consider **POD** and **FAR** as they give a much better picture of the performance of the forecast provider.

6) **Bias in the Frost Forecasts:** If the forecast provider is forecasting too many frosts in order to cover themselves then the Bias will also pick this up. It calculates the total number of nights when a frost was forecast (32 nights) compared to the total number of frosts observed (30 nights).

$$\text{Bias} = \frac{P(F)}{O(F)} * 100 = 32 * 100 = 107\%$$

The Bias should be as close to 100% as possible. If it is greater than 100% then more frosts are been forecast than observed, and similarly if the Bias is less than 100% then too few frosts are being forecast.

7) **Total Alarm Rate (TAR):** This computes the total percentage of nights when a frost was forecast. This gives an indication of the severity of the climate and can be used to check if too many frosts are been forecast compared to your perception of the climate:

$$\text{Total Alarm Rate} = \frac{P(F)}{\text{Total}} * 100 = \frac{32}{100} * 100 = 32\%$$

8) **Critical Success Index (CSI):** The Critical Success Index compares the number of correct frost forecasts with the number of nights when some threat of a frost was present. This index therefore takes into account both the number of Type 1 errors (NF/F) and Type 2 errors (F/NF).

$$\text{Critical Success Index} = \frac{F/F}{(F/F + F/NF + NF/F)} * 100 = \frac{20}{(20+12+10)} * 100 = 48\%$$

Thus if all the frost forecasts were correct then the Critical Success Index would be 100%.

9) **Skill Score (SS):** It is always useful to check the skill exhibited in the road weather forecasts. There are several different Skill Scores but the simplest is given by:

$$\text{Skill Score} = \frac{\text{Percent Correct - Climatology/Persistence/Chance}}{\text{Total - Climatology/Persistence/Chance}}$$

In other words if the weather is the same as you had on average on that day over the last 30 years (climatology); or the same as the day before (persistence); or the same as one could predict by, for instance, throwing a dice (chance) then the forecast is easy. For example the minimum road surface temperature may not change that much in the middle of winter from day to day. In the United Kingdom persistence can predict the minimum road surface temperature within 2 degrees Celsius on about 70% of the nights. If the forecast supplier is achieving an 84% accuracy then the skill for 100 nights would be:

$$\text{Skill Score} = \frac{(84 - 70) * 100}{(100 - 70)} = \frac{1400}{30} = 47\%$$

Note that if the forecast was correct on 99% of the days then SS = 97% whereas if the forecast was correct on only 71% of the days then SS = 3%. The basic skill score is fine for comparisons of different forecast suppliers, or comparisons from year to year.

10) **Heidke Skill Score (HSS):** Sometimes it is not possible to get hold of climatological data or even to have data to calculate persistence. However it is possible to use the contingency itself to calculate how much better than chance is the forecast. The formula to calculate this statistic looks complicated but it is logical:

$$\text{Heidke Skill Score} = \frac{\text{Number of correct forecasts - Number expected by chance}}{\text{Total Number of forecasts - Number expected by chance}}$$

$$HSS = \frac{((F/F + NF/NF) - ((O(F)*P(F) + O(NF)*P(NF))/TOTAL)) * 100}{TOTAL - ((O(F)*P(F) + O(NF)*P(NF))/TOTAL)}$$

$$= \frac{(78 - 57.2) * 100}{100 - 57.2} = 49\%$$

The Heidke Skill Score varies between 0% and 100% - the higher the score the better. It is the most popular of the Skill Scores because it is so easy to calculate.

Out of the 10 statistical measures given above the best ones to use for the highway engineer are probably as follows, together with suggested targets:

For Frost/No Frost Forecasts:

	Target
Percent Correct	>= 86%
Miss Rate	<= 20%
False Alarm Rate	<= 20%
Critical Success Index	>= 75%
Heidke Skill Score	>= 75%

Targets should be set based on experience and these could then be incorporated into a performance related contract.

The Base Rate Effect

Some weather events occur much more frequently than others and this can effect how well they are forecast. This is called the Base Rate Effect. Mathews (1997) uses this effect to show that when 'rain' is forecast it is less likely to be accurate than when the forecast is 'no rain'. We can use the 2x2 contingency table to show why this is so:

		FORECAST		
		SNOW	NO SNOW	TOTAL
	SNOW	8	2	10
OBSERVED	NO SNOW	18	72	90
	TOTAL	20	74	100

Percent Correct	= 80%
Hit Rate	= 80%
Miss Rate	= 20%
POD	= 40%
FAR	= 90%
Bias	= 200%
TAR	= 20%
CSI	= 29%
HSS	= 36%

Thus although the first three calculated statistics look good with a percent correct of 80% and a Hit Rate of 80% with a small Miss Rate of 20%, the rest of the statistics tell a very different story. There is a very large Bias of 200% and the False Alarm Rate is 90%! Thus the Snow forecasts are far from reliable. This is due to small Base Rate of only 10 days when snow fell out of 100 days. Having said that the 'No Snow' forecasts are much better with 72 of the 74 'No Snow' forecasts being correct i.e. 97%!

Conclusion

With the use of a simple contingency table a number of very useful statistics can be calculated by the Highway Engineer. These results can be written into performance related contracts, or at the very least be demanded from the weather forecast service providers at the end of the season.

References

Matthews, R. 1997 How Right Can You Be?, ***New Scientist***, No 2072, 28-31

Stanski, et al 1989 Survey of common verification methods in meteorology, WMO, ***World Weather Watch Technical Report***, No. 8

Thornes, J.E. 1995 A comparative real-time trial between the Met. Office and Oceanroutes to predict road surface temperatures, ***Meteorological Applications***, 2, 113-119

Thornes, J.E. 1996 The Quality and Accuracy of a Sample of Public and Commercial Weather Forecasts in the UK, ***Meteorological Applications***, 3, 63-74