Can you supply? no cond future date

Title: Meteorological applications.
Imprint: Cambridge, England; New York, NY: Cambridge University Press, 1994-
Article: Thornes, J.E. "The quality and accuracy of a sample of public and commercial weather forecasts in the UK"
Vol: 3
No:
Date: 1996
Pages: 63-74
Verified: OCLC ISSN: 1350-4827 [Format: Serial]
Patron: He, Wei
Ship to: Kansas State University/Interlibrary Services/137 Hale Library/Manhattan KS 66506-1200
Bill to: Same
Ship via: Ariel/Fax/First class
Max cost: $25 ifm
Copyrt compliance: CCL
Fax: (785) 532-6144
Phone: (785) 532-7440
E-mail: Ariel: ariel.lib.ksu.edu or 129.130.36.208
E-mail: ill@ksu.edu
Billing notes: We require patron name, ILL# and/or copy of request with invoice for payment. Pls. include FFEIN # and "ATTN: Sharon Van Riselburghe" on invoice.

Borrowing notes: ClioID: C0023447//
Affiliation: GWLA (Big-Twelve Plus)
Lending charges:
Shipped:
Ship insurance:
Lending restrictions:
Lending notes:
At a time when the commercialisation of weather forecasting is proceeding at an alarming rate around the world, it is essential to review the quality and accuracy of weather forecasts to ensure that current standards are at least maintained. The privatisation of meteorological services in some countries (e.g. New Zealand) and the vigorous commercialisation of established Met. Offices such as in the United Kingdom, France and Sweden suggests that customer satisfaction, increased income and reduced expenditure will dominate the weather forecasting scene over the next decade. It is important that the quality of weather forecasts is closely monitored before and after commercialisation to ensure that the accuracy of weather forecasts continues to improve and that public weather forecasting does not suffer. The improved technologies for data collection such as weather satellites, weather radar, wind profilers, road weather information systems, sferics, etc. should make it easier to assess forecast accuracy – but with the exception of road weather forecasts there is little sign that the new technology has been used for this purpose. The current accuracy of a sample of weather forecasts in the UK is discussed and a framework for publishing future weather forecast verification statistics is presented.

1. Introduction

The accuracy of weather forecasts is not widely reported, indeed it is very difficult to obtain information on the subject despite the fact that thousands of weather forecasts are issued around the world every day. There has been some discussion of weather forecast accuracy in the literature (e.g. Lynagh, 1990; Murphy, 1991; Murphy & Sabin 1986; Murphy & Winkler 1987) but there have been few published statistics in the UK. There is a generally broadcast belief by the UK Met. Office that 24-hour weather forecasts in the UK are about 85% accurate but there are no independently published figures to confirm this. Also, it is not clear what is meant by the word accuracy in the context of weather forecasts. Stanski et al. (1989), in attempting to define the quality of a weather forecast, list accuracy as just one of six attributes to be considered, together with skill, reliability, resolution, sharpness and uncertainty. They define accuracy as 'the general agreement between the forecast and actual weather normally using some statistical measure of error in the forecast' and skill as 'the accuracy of the forecast in relation to the weather situation taking into account persistence and climatology'. These terms will be considered in section 4.2 below but the main aim of this paper is to assess independently current UK weather forecast accuracy in the 1990s and to suggest a framework for assessing and monitoring future weather forecast quality. There are insufficient data at present to enable an objective review of weather forecast quality.

The increased commercialisation of weather forecasts around the world has led to fears of reduced accuracy and therefore it is essential that forecast accuracy is monitored. The privatisation of meteorological services in some countries (e.g. New Zealand) and the vigorous commercialisation of established Met. Offices such as in the United Kingdom, France and Sweden suggests that customer satisfaction, increased income and reduced expenditure will dominate the weather forecasting scene over the next decade. Tennekes (1988) has already warned that weather forecasts are likely to become more mercenary and the label 'cheap but cheerful' may become the catch phrase of the weather forecasts of the 1990s. As the Head of one National Weather Service recently stated (pers.comm.): 'it is better to be roughly right than precisely wrong'. On the other hand, with the advent of new technology it may be possible to be cheap, cheerful and accurate.

Weather forecasting accuracy has undoubtedly improved steadily over the last 40 years due to improved computer processing, improved numerical models, better data assimilation, introduction of weather satellites and other remote sensing devices such as weather radar and road weather outstations (ESA, 1993). Is this improvement in accuracy about to level off or indeed fall?

Weather forecasts are a fundamental provision of most national and private meteorological services. The UK Meteorological Office, in its annual review for 1990/
J E Thornes

1991 (UK Met. Office, 1991) presented a graph entitled 'How the forecasts have improved', shown in Figure 1. From the graph there is no doubt that 24-hour, 48-hour and 72-hour forecasts of height-change correlation by the coarse-mesh (scale) model for north-west Europe and the eastern North Atlantic have improved over the last 20 years. This suggests that numerical weather prediction (NWP) has improved, but this information still has to be translated into weather forecasts for site-specific locations that the general public and commercial customers understand. As weather services become more commercial it is likely that less emphasis will be placed on public weather forecasts and more emphasis will be given to commercial customers. However, it is important for weather forecast providers, in the light of increasing commercial competition, that weather forecasts are seen to be getting more accurate for both public and commercial consumption. There are several reasons for this:

- Commercial customers are now paying the real cost of weather forecast production and the forecast providers are therefore more vulnerable to possible legal action if a forecast is incorrect (Millington, 1987).
- In a competitive market it is important to be able to show that your forecasts are more accurate and more likely to improve than your competitors.
- Perceived incorrect forecasts — such as those associated with the UK storm of 1987 — can attract national adverse publicity which may affect future sales of forecasts.
- In order to encourage further investment in new technology, such as weather satellites, it is necessary to demonstrate that similar previous investments have produced improvements in forecast accuracy.
- The public are paying through direct taxation for the provision of public weather forecasts, and governments have the power to decide that public weather forecasts are not cost effective. This may reduce the funding of national Met. Offices or encourage privatisation.

Not only is it important that weather forecasting accuracy is seen to improve, for instance in percentage terms, it is also equally important that the time period for which a weather forecast remains accurate is extended.

2. A brief history of weather forecast provision

The development of numerical prediction models for weather forecasting this century has been remarkable (Shuman, 1989; Mesinger, 1990). Starting with the ideas of Bjerknes (1904) and continued by the perseverance and vision of Richardson (1922), who envisaged that a team of 64,000 mathematicians would be needed to keep pace with the weather, there has been a steady increase in the effort and money put into numerical weather prediction. It was not until the advent of high-speed digital computers in the 1950s that Richardson's dream of using a set of physical equations to predict the behaviour of the atmosphere became a reality.

The first operational numerical forecasts using a computer were issued in 1955 in the United States, but the numerical predictions were not available in time to affect the subjective ‘man-made’ forecasts. The introduction of new models and more powerful computers,
together with more input data from satellites and other sources, has led to a steady reduction in errors.

The UK Met. Office bought its first computer in 1959 – a Ferranti Mercury which could do 3000 multiplications per second (Burroughs, 1991). The subsequent development of new models and computers is shown Table 1. Progressively, the number of vertical levels has been increased and the grid size has been reduced, but as Hunt (1994a) points out the reduction in grid-box size is no longer the main priority of increased computer power and that:

present practice is to use increased computer capacity to improve the physical modelling and data analysis, rather more than reducing the grid-box size. (p. 313)

The models are now global and probably the most sophisticated one is the spectral model that is run by the European Centre for Medium-range Weather Forecasting, which forecasts the wind, air temperature and humidity at 4154868 points throughout the atmosphere (ECMWF, 1993).

Smaller-scale models with a much higher resolution, such as the mesoscale model currently being used by the UK Met. Office, which has 31 layers with a 17 km grid, are being developed in an attempt to bridge the gap between global models and the commercial requirements for local site-specific weather forecasts, review of current European global and regional models is given by Wergen & Majewski (1993).

In order to produce a forecast of the weather at a specific point and at a specific time a weather forecaster has to interpret the inputs and outputs (e.g. data and pressure patterns respectively) produced by the NWP models. Figure 2 shows, for the UK Met. Office, how an intervention bench can check data used to initialise the model and then fine tune the outputs before the weather forecasts are sent to the general public or commercial customers. The NWP models are not yet capable of predicting all the weather elements by themselves but some outputs, e.g. wind velocity for aircraft, can be inferred directly.

The weather that is observed at a point in space and time is an aggregate of temperature, cloud, wind, precipitation, sunshine, visibility and humidity. The person receiving a weather forecast is only interested in the weather parameters that affect his or her personal comfort, leisure interest or commercial activity. Eventually, it may be possible to produce numerical forecasts on a small enough scale to provide all these elements to a client at a given location – on perhaps a

![Quality and accuracy of weather forecasts in the UK](image)

Table 1. New computers and development of new models at the UK Met. Office

<table>
<thead>
<tr>
<th>Computer</th>
<th>Number of height levels and spatial grids</th>
</tr>
</thead>
<tbody>
<tr>
<td>1959 Ferranti Mercury KDF9</td>
<td>3-level model</td>
</tr>
<tr>
<td>1972 IBM 360/195</td>
<td>10-level model</td>
</tr>
<tr>
<td></td>
<td>Octagon (300 km)</td>
</tr>
<tr>
<td></td>
<td>Rectangle (100 km grid)</td>
</tr>
<tr>
<td>1982 Cyber 205</td>
<td>15-level model</td>
</tr>
<tr>
<td></td>
<td>Coarse-mesh (1.5° latitude, 1.875° longitude)</td>
</tr>
<tr>
<td></td>
<td>Fine-mesh (0.75° latitude, 0.9375° longitude)</td>
</tr>
<tr>
<td>1991 Cray</td>
<td>Unified model</td>
</tr>
<tr>
<td></td>
<td>Global (19 levels, 0.83° latitude, 1.28° longitude) (approx. 95 km grid over UK)</td>
</tr>
<tr>
<td></td>
<td>Limited area model (19 levels, 49 km mesh)</td>
</tr>
<tr>
<td></td>
<td>Mesoscale model (31 levels, 17 km grid)</td>
</tr>
</tbody>
</table>
1 km or 2 km grid. Golding (1989) predicted that by the mid 1990s:

... it should be possible to resolve at least the lowest layers of the atmosphere with a resolution of one or two kilometres, sufficient to represent most hills and valleys. Provided that the inevitable difficulties in representing the physical processes operating at this scale are overcome, it should provide direct prediction of most of the variables required by users.

(p. 51)

Site-specific forecasts are increasingly required for commercial customers. This may involve the use of other more detailed models for specific applications such as the forecasting of a road or runway temperature (e.g. Thornes & Shao, 1991). Outputs from the UK Met. Office mesoscale model have been fed directly into an energy balance model to predict road-surface temperatures. The results are almost as good as those produced by the forecaster interpreting the inputs to the energy balance model (Thornes & Shao, 1992; Astbury, 1995) and it is obviously less time consuming to get one model to feed another. In the near future it will certainly become more cost effective to interpret mesoscale models for site-specific forecasts.

3. The use of meteorological data systems for forecast verification

Observations of weather elements, such as temperature, rainfall, pressure, etc., have been made systematically since the middle of the last century. Most standard observations are land based but increasingly reliance is made on satellites, aircraft, ships, radiosondes, radar etc., in order to give a reasonable coverage over the oceans and in the upper atmosphere (Browning & Szejwach, 1994).

The acquisition of observations of the weather around the world in order to initialise the forecast models is of great importance. Altogether, there are about 100,000 sets of weather observations made every day around the globe, but perhaps only about one third of these are fed into a typical NWP forecast. Improved 'data assimilation' techniques are constantly being developed to check the data and to allow more data to be used.

It is to be expected that the phenomenal growth in new technology for data observation and collection should result in better assessments of weather forecast quality. Data which previously had to be estimated, such as road-surface temperatures, are now being measured. The exact locations of thunderstorms and bands of rain are now being obtained. It should therefore be easier to determine if a weather forecast for rain or thunderstorms was correct. Unfortunately, there has been little sign of this happening. The National Ice Prediction Network in the UK is still the only system that displays in real time observations versus forecast. The user can see at a glance the accuracy of the forecast road surface condition in real time. Also, the Department of Transport has laid down a minimum required accuracy for road weather forecasts, which requires that the forecast providers produce statistical summaries of forecast accuracy (Thornes, 1995).

Figure 3 shows the improvements in forecast accuracy of Open Road forecasts issued by the UK Met. Office over the last five winters (Astbury, 1995). This is a good example of user feedback and it is to be hoped that similar information will be made available for other forecast services in the future.

4. Improvements in weather forecast accuracy

For the purpose of this article the accuracy of weather forecasts as published by the UK Met. Office (1995) has been examined, together with data extracted from media weather forecasts (Fleming, 1994; Redfern, 1995).

The principal outputs from NWP models include projected pressure levels, clouds and winds from which the forecast weather charts are produced. The fore-
caster has to interpret the likely weather for a particular place or region in terms of expected winds, air temperature, likely precipitation and cloud/sunshine. According to Morris (1991):

Weather forecasts are largely based upon NWP model output, with the (human) forecaster having the role of fine-tuning the product and adding value to it. There are two fundamental reasons why NWP models do not yield perfect forecasts: firstly there is the limitation of the modelling itself, e.g. simulation of the complex processes of transferring heat, momentum and moisture through the atmospheric boundary layer, and secondly the starting conditions, from which the complex equations are integrated forwards in time, are imperfectly specified.

Ideally, NWP models need about one million data points for initialisation, and even if all observations from around the world could be used, the number of points would be at least an order of magnitude less.

If the model output is incorrect then obviously the forecaster is likely to pass on these errors in the weather forecast. However, there is a skill in interpreting the model output which has led to the development of, for instance, skill scores and in some situations the forecaster may use experience to overcome the model deficiencies.

4.1. The assessment of the quality and accuracy of weather forecasts

The quality and accuracy of a weather forecast is not an easy thing to judge. It depends upon which weather element is considered, and what time period and what area are chosen for analysis. For instance, the UK Met. Office (1993) claim:

Basic forecasting accuracy is stable at around 84% as measured by the 24-hour forecast at 1755 on BBC Radio 4. This met the target set. (p. 13)

It is not stated how this accuracy was determined, but apparently regional weather centres routinely assess the elements of the forecast as being right or wrong on a daily basis. Is an accuracy of 84% to be considered good? It means that on 16% of occasions, that is on 58 days in the year, i.e. for the equivalent of nearly two months of the year, the forecast was considered not to be accurate. Also, it could be implied that the accuracy is probably much worse than that because there are many occasions in a year when the weather is static and remains the same for several days (persistence). Also, the weather may be average for the time of year which would also be easy to forecast (climatology). The importance of ‘persistence’ and ‘climatology’ are discussed below. Also, if the forecast misses a number of critical weather events (e.g. the

Quality and accuracy of weather forecasts in the UK

wrong side of 0°C) the quality is low even though the accuracy may be relatively high.

Tennukes (1988) suggests that weather forecasters should now concentrate on improving the presentation and communication of weather forecasts as the numerical weather prediction models near their theoretical limits. There are dangers in the over-simplification of weather forecasts, such as the use of phrases like ‘scattered showers’ which bewilder the general public. However it should be noted that presentation could be used to mask uncertainty and that should also be avoided. For instance, the use of probability forecasts (scattered showers can be presented as a 10% chance of a shower) may sound scientific but mean little to the general public. Also, it is not possible to assess the accuracy of a single probability forecast, as in crude terms a probability forecast cannot be incorrect!

In order to assess whether or not the accuracy of weather forecasts has improved and if so why, it is first necessary to establish a framework that contains spatial, temporal and forecast element information. Before we can class a forecast as accurate we have to define:

- The place or region for which the forecast was issued.
- The time period covered by the forecast.
- The meteorological elements considered, such as pressure, temperature, wind, precipitation, visibility, etc.
- The existence of thresholds such as 0 °C (possible ice on roads, frost damage to crops), 1000 metres visibility (definition of fog), wind speed above 20 ms⁻¹ (damage to structures) etc.

If a weather forecast is on the wrong side of a threshold and damage occurs that could have been prevented, the forecast should be considered incorrect even if the margin of error is small. For instance, if the forecast minimum road surface temperature is +0.5 °C and the actual minimum is −0.2 °C (and a road becomes icy) then the forecast is obviously to be considered incorrect even though the error was less than 1 °C. One must be careful that percentage correct figures are calculated consistently.

4.2. The verification of weather forecasts

In order to assess objectively whether or not weather forecasts have improved over a given time period we have to define a verification procedure. There are three main purposes for verification: firstly administrative, secondly, scientific and thirdly, for the general public (e.g. for a Citizens’ Charter). Administrative verification is used, for example, to justify the cost of a weather service to politicians; to provide support for the purchase of new equipment, new computers, etc. and to monitor the overall quality of the forecasts.
Scientific verification is concerned with research and development, for instance the tuning of forecast models.

Stanski et al. (1989) have produced the most comprehensive review of verification methods. They suggest six attributes of a weather forecast that have to be examined in order to be able to define the quality of a weather forecast. These are:

- **Accuracy.** The general agreement between the forecast and actual weather normally using some statistical measure of the error in the forecast (e.g. standard deviation or root mean square error).
- **Skill.** The accuracy of the forecast in relation to the weather situation taking into account persistence and climatology.
- **Reliability.** Equivalent to 'bias' of the forecasts - is there consistent over or under forecasting of a parameter over time or space?
- **Resolution.** The ability to distinguish different weather events consistently, both spatially and temporarily. Low resolution would be when observations of a particular weather event (e.g. frost) were associated with a wide range of forecasts.
- **Sharpness.** The ability to forecast within a narrow band, for instance, to be on the correct side of a threshold.
- **Uncertainty.** The variance of the observations, which obviously does not depend upon the forecasts, e.g. the diurnal temperature range is much smaller in winter than in summer.

Note that the skill of the forecaster needs to be compared with that of an unskilled forecaster. In practice this is usually achieved by comparing the forecast with chance, persistence or climatology.

Chance represents a pure guess and requires no prior knowledge; persistence is a 'no change' forecast and requires knowledge of only the initial weather conditions; climatology is a forecast of the long-term average weather, and requires a knowledge of the history of the weather. Persistence and climatology can be used in combination to achieve apparently reasonable results for some weather parameters.

Some estimates of the accuracy of 24-hour forecasts that can be achieved using climatology and persistence, for maximum and minimum air temperature forecasts, taken from data for The University of Birmingham, are shown in Table 2. For the period 1961–1990 the daily maximum and minimum air temperatures were estimated for each day from the climatological average (climatology) and from the previous day's value. These values were then compared with the actual minimum and maximum air temperature for each of the 11 680 days.

<table>
<thead>
<tr>
<th>Temperature range (°C)</th>
<th>Climatology (%)</th>
<th>Persistence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_{\text{max}}$</td>
<td>$T_{\text{min}}$</td>
</tr>
<tr>
<td>±0.5</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>±1.0</td>
<td>24</td>
<td>29</td>
</tr>
<tr>
<td>±2.0</td>
<td>46</td>
<td>53</td>
</tr>
<tr>
<td>±3.0</td>
<td>64</td>
<td>71</td>
</tr>
</tbody>
</table>

Table 2. Percentage accuracy of 24-hour maximum and minimum air temperature forecasts for Birmingham (University) using climatology and persistence for 1961–1990 (11 680 days)

temperature forecasts for Birmingham and that for climatology alone minimum air temperature forecasts are better whereas for persistence alone maximum temperature forecasts are better. An accuracy of ±2 °C is normally considered adequate for air temperature forecasts and an accuracy of about 60% is possible in Birmingham using persistence. This is equivalent to getting the forecast correct 4 days a week, whereas an 85% accuracy is equivalent to 6 days per week. The UK Met. Office (1993: p. 13) claim a 97% accuracy in forecasting air temperatures for the Gas Board to within ±3.0 °C and it can be seen that this is about 20% better than might be achieved by persistence alone.

4.3. The statistical measures of forecast accuracy

The main methods and statistical measures of forecast accuracy and reliability for normal weather forecasts include: scatter plots, bias or mean error, mean absolute error, root mean square (RMS) error, reduction of variance, contingency tables and percentage correct.

Stanski et al. (1989) give a number of examples of the above and also discuss the assessment of the quality of NWP models using similar statistics, but there are added problems. Normally, NWP output is in a gridded format and therefore it can be difficult to compare with actual observations (Flood, 1985). In order to compare the performance of four national NWP models Geleyn & Dupont (1992) use the RMS error of the 72-hour forecast of the height of the 500 hPa level, and the bias in the height for the same time frame. The 72-hour forecast has to be compared with the inferred data at the grid points calculated from observations 72 hours later.

The main methods and statistical measures of forecast accuracy and reliability for NWP are: bias or mean error, mean absolute error, standard deviation error and anomaly correlations.

Thus percentage correct, which is probably the easiest statistic to understand, is only one of many methods of accuracy measurement. Its use can give misleading information unless the appropriate climatology and
5. Trends in weather forecast accuracy

The accuracy of weather forecasts and whether they have improved over time will now be assessed. First, let us examine the verification data published by the UK Met. Office (1995).

5.1. Numerical weather prediction accuracy

Over the last 25 years or so, the accuracy of NWP model output does appear to have improved considerably, as shown in Figure 1. For the UK Met. Office coarse-mesh model, forecasting 24 hours ahead, the height-change correlation for 1000 hPa has increased from 0.79 to 0.94, and the RMS wind vector error has reduced by more than 30%. Table 3 shows that over the period 1969–1994, for the same coarse-mesh model there has been an increase in height-change correlation for the 1000 hPa level at all time scales considered. By 1994 the 48-hour and 72-hour forecasts have become almost as good as the 24-hour and 48-hour forecasts respectively were in 1989. There is no sign yet of a levelling off of forecast accuracy but of course height-change correlations mean very little to the general public.

5.2. Central Forecast Office (CFO), media and commercial weather forecast accuracy

(a) Precipitation

The location and timing of precipitation are probably the most difficult and yet important weather elements to forecast. The combination of satellite and weather radar images should be improving the accuracy of precipitation forecasts (Collier et al., 1989) but it appears that there is still much room for improvement.

The UK Met Office (1995) verification statistics (p. 23) gives the accuracy of precipitation forecasts for London made by both the Central Forecasting Office (CFO) and the limited area model (LAM). These forecasts are simply of a yes/no nature for up to 24 hours ahead. Figure 4 compares the accuracy of rain forecasts with no rain forecasts since 1962. In other words the rain forecast is the percentage correct of those days when rain was forecast for London as measured by London Weather Centre and Heathrow. The mean accuracy since 1962 is 71% correct and the upward trend in accuracy is very shallow. Indeed, there appears to have been a decline in accuracy since 1986. The mean accuracy for no rain since 1962 is 81% and there does appear to have been a significant improvement in accuracy up to about 1980, since when the accuracy appears to have levelled off. It is interesting to note that since 1980 forecasts of no rain have been more accurate than forecasts of rain by about 12%, i.e. by nearly one day in seven.

The limited area model has also been used to predict precipitation over London. Figure 5 shows that since 1984 when statistics were first kept the average per-

![Figure 4. Accuracy of London rain forecasts - Central Forecasting Office (UK Met. Office, 1995).](image4.png)

![Figure 5. Accuracy of London rain forecasts - Limited Area Model (UK Met. Office, 1995).](image5.png)
percentage correct for rain forecasts is 71% and for no rain forecasts is 79%. Surprisingly, the LAM forecasts of rain have deteriorated since the mid 1980s. It is clear that the LAM forecast is not as accurate as that from CFO. The decline in rain forecast accuracy since the mid 1980s is alarming, although no rain forecasts have improved over that period.

The decline in the accuracy of rain forecasts since the mid 1980s by both the CFO and the LAM is surprising but may be related generally to the dry weather conditions that have prevailed in south-east England in the last decade. Accuracy figures for the UK mesoscale model for precipitation forecasts are available for 1993 and 1994. The percentage of correct forecasts, based on rain/no rain criteria for a selection of UK stations, in 1993 was 72.9% for the 24- to 30-hour forecast period and was 71.3% in 1994.

Forecasts of rain, which are often the most notable part of a weather forecast, are accurate approximately five days out of seven and then only on a yes/no basis. In terms of site-specific rain forecasts including the timing and amount of rain the accuracy is probably less than 50%, but there are no statistics to check this.

(b) Surface air temperature

Table 4 shows that for BBC Radio 4 forecasts of maximum air temperature within ±2 °C the UK Met. Office claim an accuracy for same day maximum temperature forecasts of 88.4% on average and 84.6% on average for the next day (UK Met. Office, 1995). There is some indication that the next day forecasts are getting better but the same day forecasts have only improved slightly.

(c) Summary data

The UK Met. Office also monitors the total accuracy, in percentage form, of the BBC Radio 4 weather forecasts issued at 0755 and 1755. Four criteria are checked: wind, weather, sky and maximum temperature. The average accuracies for the period 1980 to 1994 are also given in Table 4. These are the figures quoted in the Met. Office Annual Reports and Accounts (e.g. UK Met. Office 1993). Unfortunately, it is not clear how good a representation of overall forecast accuracy this measure is because the accuracy criteria are not stated. Although the maximum temperature within ±2 °C is easy enough to measure accurately, the state of the sky is a rather poor substitute for precipitation forecasts, which we have already seen are much less accurate. A better standard verification statistic is required.

The morning forecast for the day has improved slightly over the time considered (from 86% in 1980/81 to 89% in 1993), averaging 87% whereas the evening forecast for the next day has improved by 6% (from 80% in 1980/81 to 86% in 1993), averaging 83%. It is not clear why the 'today' forecast has not improved very much other than perhaps it is reaching an accuracy limit beyond which it is very difficult to progress.

(d) Commercial forecasts

Other more specific types of forecast have shown marked improvements in accuracy in recent years, for example the 'Open Road' forecasts issued by the Met. Office to Highway Authorities. The percentage of correct forecasts as to whether or not road surface temperatures would fall below 0 °C was 88% on average for 260 forecast sites during the winter of 1994/95 (Astbury, 1995). Figure 3 shows how the RMS error of predicted road surface temperatures has fallen over the last five winters.

6. An independent check on forecast accuracy

Fleming (1994) carried out a comparison of various media weather forecasts with actual observed weather in the Midlands. Obviously, it is often a subjective judgement as to whether or not a weather forecast is judged to be accurate, as already discussed. The forecast maximum and minimum air temperature is the simplest one to test in that the range ±2 °C is generally considered acceptable. Also, precipitation forecasts are simple to verify in terms of precipitation/no precipitation.
6.1. Nottingham Evening Post weather forecasts

The forecast printed by the Nottingham Evening Post was examined for the three years 1990 to 1992 and the forecast minimum and maximum temperature for the next day was compared with observations at Watnall (Nottingham Weather Centre). The forecast was prepared at Nottingham Weather Centre and transmitted at about noon and hence the minimum temperature forecast is about an 18-hour forecast and the maximum temperature forecast is about a 26-hour forecast on average. The accuracy within ±2 °C was 76% for the minimum temperature and 68% for the maximum temperature. Figures 6 and 7 show the monthly distribution of the percentage correct together with the percentage correct that would have been achieved using persistence (i.e. using today's minimum and maximum temperature to predict tomorrow's). On average, persistence was correct for 56% of the minimum temperatures and 59% of the maximum temperatures (just 3% lower than the Birmingham figures quoted earlier). Thus, the newspaper forecast of minimum temperature was 20% better than persistence, whereas the maximum temperature forecast was only 9% better than persistence.

![Minimum Temperature Forecasts Graph](image)

*Figure 6. Accuracy of Nottingham Evening Post weather forecasts – Tmin (Fleming, 1994).*

![Maximum Temperature Forecasts Graph](image)

*Figure 7. Accuracy of Nottingham Evening Post weather forecasts – Tmax (Fleming, 1994).*

Quality and accuracy of weather forecasts in the UK

The accuracy of precipitation forecasts was checked in a similar way and Figure 8 shows the comparison with persistence for Watnall. The accuracy was calculated simply by summing the number of correct forecasts in terms of a yes/no forecast for any precipitation on the day in question in the Watnall region and a yes/no observation. The mean accuracy for the three-year period was 73% compared with a persistence accuracy of 68% – an improvement of just 5%.

These results confirm the generally held view that newspaper forecasts are among least accurate of media forecasts (Thornes, 1992), but it is rather surprising that the forecasts of maximum temperature and precipitation were not much better than persistence.

6.2. Radio 4 weather forecasts

Whereas the forecasts for newspapers are normally issued before the noon run of the fine-mesh model the Radio 4 forecast at 1755 has the benefit of the noon run, and is prepared shortly before broadcast. Hence, the minimum temperature forecast for the coming night is only about a 12-hour forecast and the maximum temperature forecast is about a 20-hour forecast.

The study period was 23 March to 10 April plus 13 July to 20 August 1993, a total of 58 days. The accuracy of the temperature forecasts ±2 °C was 84% for both minimum and maximum temperature forecasts using Watnall temperatures for verification. The claimed accuracy by the UK Met. Office (1995) for the Midlands for the whole of 1993 was 87% for maximum temperature. These figures are encouraging close. The corresponding persistence accuracy was 65%.

![Rainfall Forecasts Graph](image)

*Figure 8. Accuracy of Nottingham Evening Post weather forecasts – Rainfall (Fleming, 1994).*
6.3. Other media weather forecasts

Redfern (1995) carried out a more detailed comparison of media weather forecast sources for a period of 57 days between 16 October 1994 and 11 December 1994. Again using maximum and minimum air temperature, forecasts were verified as accurate if within ±2°C of the air temperature recorded at the University of Birmingham weather station (Winterbourne). November 1994 was the warmest November in the Midlands since records began and a blocking anticyclone meant that air temperature forecasting was a little easier than one might expect on average. Table 5 shows that whilst climatology was only correct for 38% of the days, persistence was accurate for maximum temperature forecasts for 79% of the days. Indeed, persistence was better than Radio 4 and The Guardian for maximum temperature forecasts. Persistence only achieved an accuracy of 41% for minimum temperature forecasts, however. Overall, only the local BBC forecast and Central TV achieved the target accuracy of 85%.

7. Conclusion

The belief that 24-hour weather forecasts are about 85% accurate in the UK has been shown to be misleading, especially for precipitation forecasts, which even on a yes/no basis are only about 70% correct. Site-specific rain forecasts including the timing and amount are probably less than 50% accurate. The figure of 85% represents an upper limit of current accuracy rather than the average. Also, it has been shown that the accuracy depends upon which weather elements are being considered and how the percentage accuracy is calculated. This paper presents only a limited set of results but it is clear that much more comprehensive monitoring of weather forecasts is required before objective statements of accuracy and quality can be made. Much further research is required to be able to give the public and commercial customers reliable information on current weather forecast accuracy.

Very few verification statistics are published and yet the accuracy of weather forecasts is a vital component of their value. There are no published verification statistics on how the accuracy of weather forecasts diminishes with time and yet this information is also critical if the weather forecasts are to be used for operational decision making. It is also clear that the accuracy of weather forecasts has improved significantly since the introduction of computers in the 1950s, but as an incentive to ensure that accuracy continues to increase verification statistics should be widely published. Also, to ensure that weather forecasts provided by the private sector are of an equivalent standard to National Meteorological Services their verification statistics should also be published.

It does seem sensible to set a target accuracy for different weather forecasts although it should be recognised that the quality of a weather forecast is dependent upon what the client, be it a commercial organisation or member of the general public, wants.

Although the UK Met Office do now publish a few statistics as part of their commitment to the Citizens' Charter it is not a healthy situation for forecast providers to publish their own statistics. The private sector forecast providers do publish a few accuracy figures as part of their advertising material but there is no consistent appraisal.

With the growth in commercialisation and the increasing role of private sector weather forecast providers, it is timely to suggest that an independent organisation should be encouraged to provide a representative set of verification statistics for the UK that are published on a regular basis. The appropriate verification statistics could be agreed with the Met Office and the Independent Group recently set up by the private sector.

Table 5. Percentage accuracy of air temperature forecasts (±2°C) for 57 days in 1994 for different forecast suppliers in Birmingham

<table>
<thead>
<tr>
<th>Forecast source</th>
<th>Time</th>
<th>Forecast provider</th>
<th>Percentage correct</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local BBC</td>
<td>1855</td>
<td>UK Met. Office</td>
<td>83</td>
<td>(weekdays only)</td>
</tr>
<tr>
<td>Central TV</td>
<td>1850</td>
<td>The Weather Dept.</td>
<td>93</td>
<td>(7 days/week)</td>
</tr>
<tr>
<td>Radio 4</td>
<td>1755</td>
<td>UK Met. Office</td>
<td>77</td>
<td>(7 days/week)</td>
</tr>
<tr>
<td>National ITV</td>
<td>1755</td>
<td>UK Met. Office</td>
<td>88</td>
<td>(7 days/week)</td>
</tr>
<tr>
<td>National BBC</td>
<td>1830</td>
<td>UK Met. Office</td>
<td>79</td>
<td>(7 days/week)</td>
</tr>
<tr>
<td>Cefax</td>
<td>1800</td>
<td>Oceanroutes</td>
<td>80</td>
<td>(7 days/week)</td>
</tr>
<tr>
<td>Local Teletext</td>
<td>1800</td>
<td>Oceanroutes</td>
<td>81</td>
<td>(7 days/week)</td>
</tr>
<tr>
<td>Birmingham Evening Mail</td>
<td>1700</td>
<td>Birmingham Env. Svcs</td>
<td>80</td>
<td>(6 days/week)</td>
</tr>
<tr>
<td>The Guardian</td>
<td>0900</td>
<td>The Met. Office</td>
<td>75</td>
<td>(6 days + observer)</td>
</tr>
<tr>
<td>Persistence</td>
<td></td>
<td></td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Climatology</td>
<td></td>
<td></td>
<td>37</td>
<td></td>
</tr>
</tbody>
</table>

Source: Redfern (1995),
Such an initiative would be welcomed and hopefully copied in other parts of the world.

One organisation in the UK that would be suited to this independent role might be the Royal Meteorological Society, who could publish the results in Weather and/or Meteorological Applications. The World Meteorological Organization should be encouraged to instigate the setting up of similar schemes in other countries.

Alternatively, in the UK, an independent organisation, which could be the equivalent of an ‘OFMET’, could carry out forecast verification as well as regulating the pricing of meteorological data and products (Hunt, 1994b). The main advantage of an independent organisation carrying out such monitoring would be to avoid public and private organisations concentrating on particular weather forecasts that are being monitored for citizens’ charts or trials. An independent organisation could choose services to monitor at random, but using agreed verification techniques.

If there is not the means or the will to set up such an independent organisation, then weather forecast providers should be required to provide more verification statistics. Commercial customers can, and should, insist on the provision of such information when they draw up weather service contracts but the general public have, at present, no mechanism with which to seek such feedback.

Acknowledgements

The ideas for this paper were stimulated when working on a research report for the European Space Agency entitled a ‘Study of the direct economic effects of the Meteosat programme’ (ESA, 1993). I would like to thank Paul Robinson of Bramshill Consultancy Limited and Professor Peter Stubbs of the University of Manchester for their interest and support during that study. I would also like to thank the UK Met Office for supplying a copy of their ‘in-house’ verification statistics (UK Met. Office, 1995). Last, but by no means least, I would like to thank Damian Fleming and Claire Redfern for their enthusiasm and time consuming data collection.

References


Quality and accuracy of weather forecasts in the UK


