UK index documentation

Introduction

The UK Index (henceforth called the Index) is a measure of the forecasting skill of limited area NWP (Numerical Weather Prediction) models over the UK.

It is *currently* based on the 12 km North Atlantic European (NAE) model forecasts. However, other UK-specific higher-resolution model forecasts may replace some of the NAE forecasts in the Index in future.

The Index is based on forecasts of selected parameters out to 48 hours ahead for a selected set of positions verified by comparison with station observations and is based on 36 months of data. The UK 1.5 km model (UKV) however, produces forecasts only out to T+36h. Subject to operational acceptance tests, a future revision of the Index will be assessing "Day 1" UKV forecast and "Day 2" NAE forecasts.

A score is calculated for each parameter included in the Index. The individual scores are then combined in a weighted average to form a single value.

The Index is compiled from the following parameters

- Near-surface (1.5m) temperature
- Near-surface (10m) wind speed & direction
- Precipitation yes/no (equal to or greater than 0.5, 1.0 and 4.0 mm over the preceding 6 hours)
- Total cloud amount yes/no (equal to or greater than 2.5, 4.5 and 6.5 oktas)
- Cloud base height given at least 2.5 oktas yes/no (equal to or less than 100, 500 and 1000 m above ground)
- Near-surface visibility yes/no (equal to or less than 200, 1000 and 4000 m)

verified at

- quality-controlled station positions across the UK, and
- at the 3-6 hourly intervals up to 48 hours.

Observations and forecasts use available WMO block 03 station positions across the UK, excluding the Republic of Ireland. The station list is reviewed once a year (at the start of the financial year). This ensures that any changes in the land-observing network (in terms of new stations and closures) have been taken into account. Stations and observations can be excluded from calculations via two independent processes (from verification):

1. if observations have been flagged to be erroneous by the real-time quality-control process provided by the data assimilation system and/or
2. by the quality control processes applied to the MIDAS climatological database by the Observations quality team in Edinburgh.

From 1st April 2008 a change was introduced to the way cloud base height observations are used in the Index. Automatic reports of cloud base height are now included in the Index, whereas manual reports of cloud base height are excluded. The treatment of total cloud amount observations has not changed so automatic observations are excluded, whereas manual observations are included. For all other parameters both automatic and manual observations are included. Note that some stations are fully automatic, whilst other stations are semi-automatic (eg automatic at night-time or at the weekend, manual at other times).

Note also that observations of wind speed are reported in knots, but are converted to metres per second before use in the Index calculation. The model forecasts of wind are in metres per second.

All model forecast runs are verified. There are four runs per day. Forecasts are verified at 00, 06, 12 and 18 UTC to coincide with observations availability and standard reporting times.

- The NAE runs at 00, 06, 12 and 18 UTC. For each run eight forecasts covering a 48-hour period can be verified viz T+6, 12, 18, 24, 30, 36, 42 and 48.
- The higher-resolution models (e.g. UKV) run at staggered times 03, 09, 15 and 21 UTC. For each run a maximum of 6 forecasts covering a 36-hour period can be verified viz T+3, 9, 15, 21, 27, 33 and 36.

As described above, once the UKV has passed operational acceptance tests, the Index will be constructed from the following forecast lead times:

- T+9, 15, 21 and 27 from the UKV
- T+30, 36, 42 and 48 from the NAE

This document gives a detailed description of the mechanisms used to calculate the Index. It describes all of the stages from interpolation to the observation locations, through the calculation of scores, to the final compilation of the Index.

Collection of raw UK Index data

The data used by the calculation is produced within the non-critical suite. This is part of the operational schedule that runs on the supercomputer. Each time the non-critical suite runs (4 times a day), it works through a series of steps. One of these is the Station-Based Verification (SBV) step for the model. When this step executes it first extracts observations from the MetDB before running separate verification steps against sonde observations and surface observations. Data required for the UK Index calculation is produced in the surface observations step as the parameters used are with respect to surface or near-surface observations. By splitting the verification steps if one particular sub-step has problems (e.g. if there are no sonde observations), then this does not affect the other calculations. Therefore data for the UK Index is only missing if no surface observations are found within the MetDB.

If on a particular run there are no valid surface observations extracted from the MetDB then all parameters used within the Index will be missing. If the following run is OK all parameters will be present except for rainfall as this needs the previous run's observations in order to calculate accumulations.

Once a day a step within the non-critical suite runs to extract a particular days data from the generated files and places it into another set of files. These new files are then transferred to the verify package (which is an Oracle database) and the data is ingested into station-based tables for the model (one for surface data and the other for upper-air data)

Once the data has been stored on verify, a job is executed to update a further table with the persistence values.

Occasionally there can be problems with the transfer of the data. If this is the case then jobs can be manually run to recover the missing data and produce the persistence values.

Interpolation from model grid to observation positions

UK-specific models have a horizontal grid length of the order of kilometres across the UK. Typically the locations of the verifying observations will lie between grid points, so interpolation is necessary in order to calculate appropriate forecast values at station positions.

In the case of the continuous variables (temperature and wind) a simple bilinear interpolation is employed, taking a distance-weighted average of the four surrounding grid point values. For wind, the u- and v- components are interpolated separately.

However, in the case of the categorical variables (precipitation, total cloud amount, cloud base height and visibility) bilinear interpolation can result in unrepresentative values due to sharp gradients in the fields, so in this case it is appropriate to use the value at the nearest grid point.

Verifying observations are obtained from the reported synoptic (SYNOP) observations at 00, 06, 12 and 18 UTC. The reporting practice for accumulated precipitation means that some manipulation has to be performed. At 00 and 12 UTC, the precipitation over the last 6 hours is reported, and therefore no further calculation is required. However, at 06 and 18 UTC the 12-hour accumulation is reported, so the 6-hour accumulation reported 6 hours earlier has to be subtracted to obtain the required value. A separate document detailing the rules for site lists and observations quality control now exists.

Note: The model produces forecast fields of convective and dynamic accumulated rain and snow, which are added together at each forecast time. The value at the previous forecast time is then subtracted to obtain the total precipitation over the 6-hour period.

Verification of 1.5m temperature

The basic verification measure for temperature forecasts is the monthly root mean square error (*rmse*). Once the observed and forecast values are available at the station positions for a given forecast range, the monthly *rmse* may be calculated for the composite of stations according to the standard formula:

$$\text{rmse} = \sqrt{\frac{1}{n}\sum_{n}\left(t_f - t_o\right)^2}$$

where:

n is the number of matching pairs of forecasts & observations in the month
t_f is the forecast temperature
t_o is the observed temperature.

Note that if any observations are missing, then the forecasts verifying at that time cannot be included in the verification.

The Index components for temperature (and also for wind) are based on persistence skill scores. A persistence forecast is one where the forecast value remains the same as the initial value throughout the forecast period. T+0 (analysis) fields (rather than observations) are used as the persistence forecast. However, because of the diurnal nature of near-surface temperature, a 24-hour persistence forecast is used for all forecast ranges. This means, for example, that the equivalent persistence forecast for a T+12 temperature forecast verifying at midday is the T+0 temperature at midday yesterday (rather than the T+0 temperature at midnight, the start of the forecast). Values of *rmse* for persistence forecasts are calculated in exactly the same way as described above.

To produce the skill score the samples of forecast data and persistence data are equalised. Thus, if a forecast is missing then the corresponding persistence forecast is discarded. Similarly, if a persistence forecast is missing then the corresponding forecast is discarded. If an observation is missing then the corresponding forecast and persistence forecast are not used in the calculations.

Verification of 10m wind

The basic verification measure for wind forecasts is the monthly root mean square vector wind error (*rmsvwe*). This may be calculated for the composite of stations according to the standard formula:

$$\mathrm{rmsvwe} = \sqrt{\frac{1}{n}\sum_{n}\left|V_f - V_o\right|^2}$$

where:

n is the number of matching pairs of forecasts & observations in the month
V_f is the forecast wind vector
V_o is the observed wind vector.

Note that if any observations are missing, then the forecasts verifying at that time cannot be included in the verification.

Once again verification of a persistence forecast is required in order to calculate a skill score. For the same reasons given for temperature in the previous section, a 24-hour persistence forecast of wind is used for all forecast ranges. Values of *rmsvwe* for persistence forecasts are calculated in the same way as described above.

Calculation of skill scores for temperature and wind

Once the forecast and persistence *rms* errors have been determined for a particular forecast range, the monthly skill score may be calculated. This is defined in terms of Reduction of Variance, i.e.

$$SS = 1 - \frac{r_f^2}{r_p^2}$$

where:

r_f is the *rms* forecast error
r_p is the *rms* persistence error.

The smaller the ratio between forecast and persistence errors, the closer the skill score will be to 1 (perfection). If the forecast error is greater (worse) than the persistence error, then the skill score will be negative and is in fact unbounded.

The 36-month skill scores for both parameters, calculated as simple means of the monthly values, are used as components of the final Index.

Verification of precipitation, total cloud amount, cloud base height and visibility

Forecasts (and observations) of 6-hour accumulated precipitation, total cloud amount, cloud base height and visibility are classified separately for each of three specified thresholds as 'yes' (events) or 'no' (non-events). A 'yes' means that the parameter was forecast (observed) to be equal to or greater/less than the threshold, i.e. precipitation equal to or greater than the threshold, total cloud amount equal to or greater than the threshold, cloud base height equal to or less than the threshold, and visibility equal to or less than the threshold. The threshold values are given in section 1 above.

Over a month a 2x2 contingency table for each forecast range, for each parameter and for each threshold is created from matching pairs of forecasts and observations for all stations. 36-month contingency tables are then formed by simple addition of the monthly tables, and used to compute the Equitable Threat Score (ETS) for precipitation, total cloud amount, cloud base height and visibility, which form components in the Index calculation.

Calculation of equitable threat score for precipitation, total cloud amount, cloud base height and visibility

The Equitable Threat Score is used as the basis for the precipitation, total cloud amount, cloud base height and visibility components of the Index, the definition being:

$$ETS = \frac{R - \text{"chance"}}{T - \text{"chance"}}$$

where:

R is the number of observed events which were correctly forecast
T is the total number of events which were either observed or forecast.

Subtraction of "chance" from the numerator and denominator removes those observed events which are expected to be correctly forecast by chance. It is given by:

$$\text{"chance"} = \frac{F*O}{N}$$

where:

F is the number of events forecast
O is the number of events observed
N is the total number of events plus non-events.

This score has properties similar to the rms-based score described in section 5, i.e. it takes the value of 1 for a perfect forecast, and is negative if the forecast is worse than chance, although in this case there is a lower limit of -1/3.

Compilation of the Index

The processes above produce 112 scores for the 36-month period, representing the verification of 6 parameters (temperature, wind, precipitation, total cloud amount, cloud base height and visibility), 4 of which have 3 thresholds, at 8 forecast ranges. There are some further stages in the compilation of the Index from these scores.

Firstly a simple weighted average, S, of the 112 individual scores is calculated. The weighting factors are chosen to give equal overall weighting to each of the 5 components (wind, temperature, precipitation, cloud and visibility), so total cloud amount and cloud base height get lower individual weighting than precipitation and visibility.

$$S = \frac{1}{\sum_i w_i}\left(\sum_i (w_i SS_i)\right)$$

where:

w_i is the weight for the i-th parameter (see table below)
SS_i is the 36-month score for the i-th parameter.

The following table defines the weights, w_i , the same for each forecast range:

Parameter	**Threshold**			
	No	**1st**	**2nd**	**3rd**
Temperature	**6**			
Wind	**6**			
Precipitation		**2**	**2**	**2**
Total Cloud Amount		**1**	**1**	**1**
Cloud Base Height		**1**	**1**	**1**

Visibility		**2**	**2**	**2**

Thus, the Index is defined as:

$$I = \{S/S_0\}*I_{old}$$

where:

S_0 is the value of S on 31st March 2007
I_{old} is the value of the old Index on 31st March 2007.

Precision

All calculations are carried out 'exactly' using floating point arithmetic in Oracle SQL, and using the following units: temperature in degrees Kelvin, wind in metres per second, precipitation in millimetres, total cloud amount in fraction, cloud base height in metres and visibility in metres.

Owner: verification team
Last updated: 30/04/2010 12:19:58