P3.10A

DEVELOPMENTS IN NUMERICAL CLEAR AIR TURBULENCE FORECASTING AT THE UK MET OFFICE

D. Turp* and P. Gill
Met Office, Exeter, Devon, UK.

1. INTRODUCTION

Turbulence is one of the main meteorological hazards to en-route air traffic. Regional forecasts of areas where turbulence is likely are currently provided by the World Area Forecast Centres (WAFCs) in the graphical format of medium level significant weather charts. The international aviation community has expressed a desire for the WAFCs to introduce new icing and turbulence products that better satisfy the requirements of navigation and air traffic management systems (ICAO and WMO, 2002).

In response the Met Office (which houses WAFC London) has developed a system that will produce global forecasts of clear air turbulence (CAT) in numerical format. These forecasts extend up to 36 hours ahead and are used primarily for flight planning. They are now routinely distributed by WAFC London on a trial basis.

The system uses Ellrod's TI1 index (Ellrod and Knapp, 1992) to predict wind shear induced CAT and an algorithm to forecast turbulence caused by breaking mountain waves based on gravity wave stresses (described in Turp et al. (2006) and Turner (1999)).

This paper describes the verification work conducted on the forecasts of wind shear CAT which lead to the selection of the TI1 index, and current work investigating the potential of the UK Met Office's "WAFTAGE" wind and temperature nowcasting tool to produce forecasts of wind shear CAT.

2. COMPARISON OF CAT FORECASTS AND AIRCRAFT DATA

2.1 *CAT Indices Investigated*

Shear induced CAT has been the subject of much research over recent decades. Many predictors of shear induced CAT have been proposed and a number of studies have attempted to verify these predictors. Some of these studies indicate that the Ellrod Indices proposed by Ellrod and Knapp (1991) tend to perform better than the others investigated (McCann, 1993 and Brown et al., 2000 for example). There are two Ellrod Indices, TI1 and TI2. Both utilise vertical wind shear (VWS) and deformation, but TI2 also considers the effect of convergence. TI1 is used by WAFC Washington for forecasting shear induced turbulence, and TI2 by the Air Force Global Weather Central in Nebraska. TI1 and TI2 are defined as follows:

$$TI1 = VWS \times Deformation \quad (1)$$
$$TI2 = VWS \times (Deformation + Convergence) \quad (2)$$

The performance of these two predictors was compared with that of the Met Office's existing CAT algorithm. This algorithm calculates a form of the Dutton Index (Dutton, 1980), E, calculated from horizontal wind shear (HWS) in $10^{-5}s^{-1}$ and vertical wind shear (VWS) in $10^{-3}s^{-1}$, as follows:

$$E = ((5 \times HWS) + (VWS)^2 + 42)/4 \quad (3)$$

This is then converted to the *CAT Index* (referred to in some literature as the CAT Probability) by mapping the values of E in a non-linear manner to a value between 0 and 7.5 (Bysouth, 1998).

2.2 *Approach*

Two methods were used to verify the CAT algorithms. The first method was to conduct a qualitative verification study where the forecasts were compared with significant weather charts. The second method was to conduct a quantitative verification study where the forecasts were compared directly with observations of turbulence or no turbulence using a set of statistical measures. Forecasts with a lead time of 0 and 24 hours (i.e. T+00 and T+24) were investigated in both cases.

There are a number of statistical measures which are commonly used to assess the performance of a set of forecasts. Each measure assesses a different aspect of forecast performance, e.g. how good it is at predicting an event to occur, or if it is likely to produce many false alarms. To gain an impression of the general performance of the forecasts several statistical measures need to be used. However in this study the measures used need to be chosen with care, as the observations used are

Corresponding author address: D. Turp, Met Office, FitzRoy Road, Exeter, Devon. EX1 3PB. UK.
Email: debi.turp@metoffice.gov.uk

Statistic	Range	Description
Probability of Detection of “yes” observations (PODy)	0 to 1 Best score=1 Worst score=0	Fraction of “yes” observations correctly forecast
Probability of Detection of “No” observations (PODn)	0 to 1 Best score=1 Worst score=0	Fraction of “no” forecasts correctly forecast. 1-PODn=false alarm rate
True Skill Statistic (TSS)	-1 to 1 Perfect forecasts=1 Random forecasts=0 Forecasts inferior to random forecasts < 0	Measures ability of forecast system to separate “yes” from “no” cases. Also measures performance of forecasts compared with random unbiased forecasts.
Relative Operating Characteristic (ROC) curve	Area: 0 to 1 Best score: Area=1 System has no skill if area=0.5	Plot of PODy versus 1-PODn for a variety of thresholds. Area under curve is a measure of skill. The ideal curve would lie through top left hand corner of the plot.

TABLE 1. Verification statistics used in this study.

irregular in space and time and some statistical measures are sensitive to the distribution of observations (Brown and Young 2000). The measures used are listed in Table 1.

2.3 *Aircraft Observational Data Used*

For the verification study observations of turbulence from the “turbulent events database” were used. This database consists of turbulent encounter data from the Global Aircraft Data Set (GADS data) (Jerrett and Turp 2005). In this database the events are categorised according to the likely cause of the turbulence. A set of events that had been categorised as likely cases of shear induced CAT and which also occurred within an hour of the model forecast validity times (00Z, 06Z, 12Z and 18Z) were used as observations of turbulence for verification of the algorithms.

For the quantitative verification a set of “no turbulence” observations was obtained from the original GADS data. This was achieved by selecting observations where derived equivalent gust velocity (a measure of turbulence: see Hoad and Ogden (2004) for more details) was less than 0.5, indicating “nil” turbulence. To constrain the size of the set of “no turbulence” observations it was limited to those observations taken at 00Z, 06Z, 12Z or 18Z on dates where a turbulent encounter occurred elsewhere.

A total of 300 “turbulence” observations and 460 “no turbulence” observations recorded between August 2004 and March 2005 were used for the quantitative analysis.

3. ANALYSIS OF THE RESULTS

3.1 *Qualitative Verification*

For this analysis forecasts with a lead time of 0 and 24 hours were produced daily and compared with significant weather charts. When conducting this verification it must be considered that significant weather charts are forecasts and also that there are small differences between the significant weather charts produced by WAFC Washington and WAFC London.

To assess the suitability of significant weather charts for use in this verification study, a selection of turbulence observations taken from the “turbulence events database” were plotted on the corresponding significant weather charts. It was found that whilst many turbulent observations were located in CAT areas marked on WAFC London significant weather charts, a large portion were located on or close to jets which had no associated CAT areas indicated, or near (but not in) marked CAT areas. This was considered whilst comparing the forecasts with the significant weather charts.

Figure 1 shows an example comparison between T+24 forecasts produced by the three algorithms and the corresponding significant weather chart. In this case there was an observation of turbulence which occurred in the mid Atlantic. This was close to jets and a CAT area on the significant weather chart, and all three algorithms also predicted it. The CAT Index forecast the CAT regions indicated over the south-eastern US and the Mediterranean but missed the CAT region to the north of Scotland. The TI1and TI2 forecast the CAT regions over south-eastern US, Canada and the Mediterranean and also some of the region to the north of Scotland.

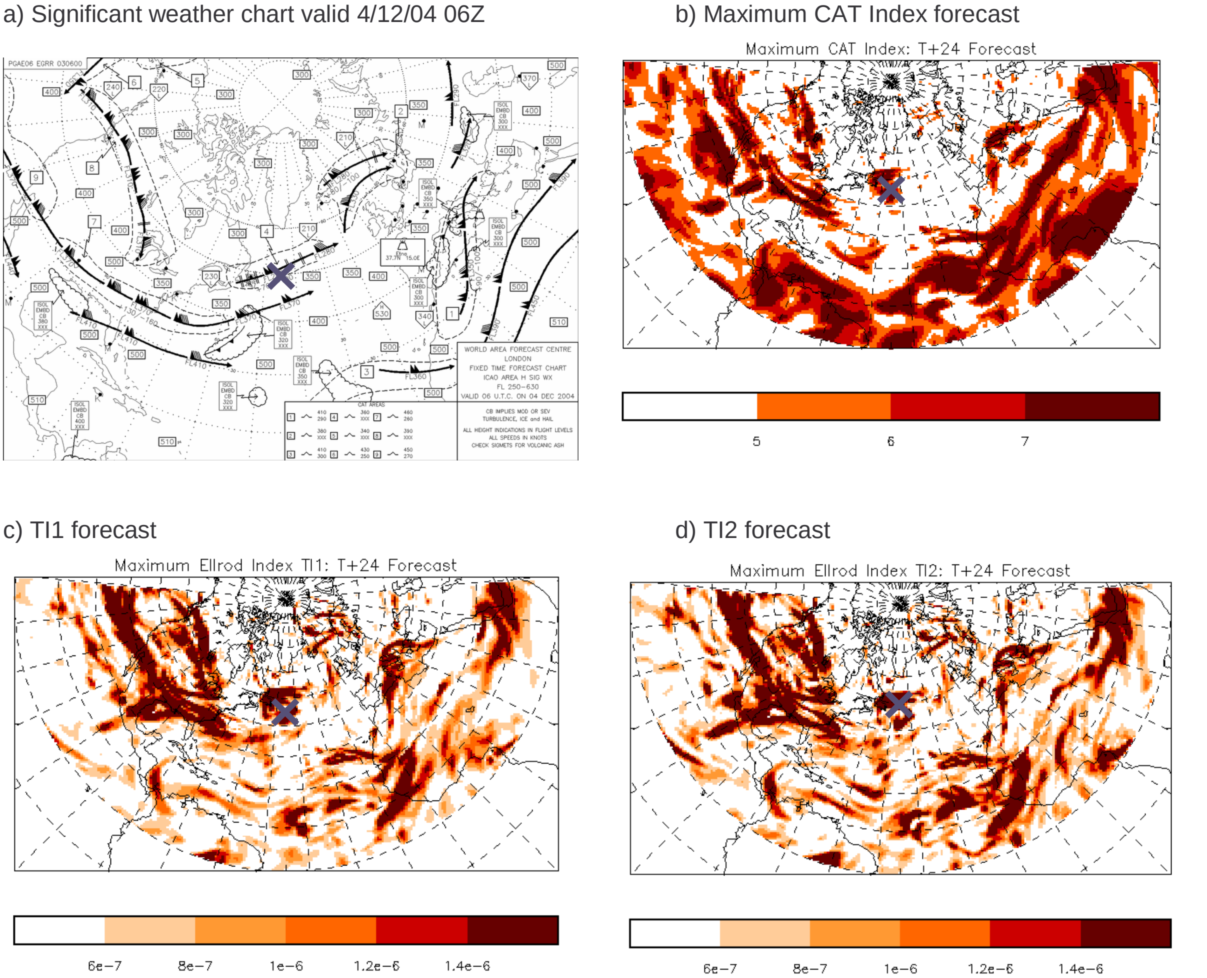


Figure 1. Example comparison of significant weather chart and forecasts produced by the CAT Index, TI1 and T12 algorithms, for 06Z 4/12/04. Plots of the forecast fields are of maximum value over all levels (400-150hPa), so all forecast CAT is shown. The blue cross indicates the location of the observation of turbulence.

In general, it was found that:

- there was good agreement of CAT Index forecasts with significant weather chart CAT areas and jets in some cases, but not with others
- CAT Index tends to overpredict, especially in the tropics at 150 hPa
- there was good agreement of TI1 forecasts with significant weather chart CAT areas and jets in most cases
- TI1 also tends to overpredict, especially in the tropics at 150 hPa
- TI2 was very similar to TI1.

3.2 *Quantitative Verification*

The performance of the forecasts from each algorithm was examined by matching the observations of turbulence and no turbulence with the algorithm values. The algorithm values were then converted to "turbulence"/"no turbulence" forecasts by application of various thresholds for the occurrence of turbulence. From these values of PODy, PODn and TSS for each threshold were calculated.

Figures 2 and 3 show the variation of PODy, PODn and TSS with threshold, for each algorithm considered, for both the T+00 (Figure 2) and T+24 forecast fields (Figure 3). In general, it was found that:

- If the thresholds were low, PODy was high but PODn was low (i.e. most turbulent observations were forecast correctly and "no turbulence" observations were incorrectly forecast as turbulence, as expected);
- If the thresholds were high, PODy was low and PODn high (i.e. few turbulence observations were correctly forecast but most "no turbulence"

observations were correctly forecast, as expected);

- The optimum threshold indicated by TSS was low for all algorithms; if these thresholds were used to determine whether turbulence was forecast or not forecast, the algorithms would overforecast most of the time.
- The CAT Index has little skill when if the threshold is less than 2.
- TI1 and TI2 perform similarly, but TI1 performs slightly better than TI2.

In general, the T+00 fields perform marginally better than the T+24 fields.

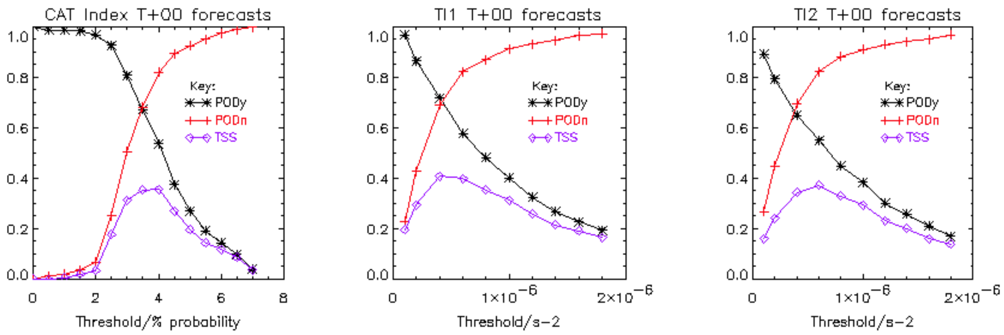


FIGURE 2. Graphs showing the variation of PODy, PODn and TSS for T+00 forecasts for the three algorithms CAT Index, TI1 and TI2.

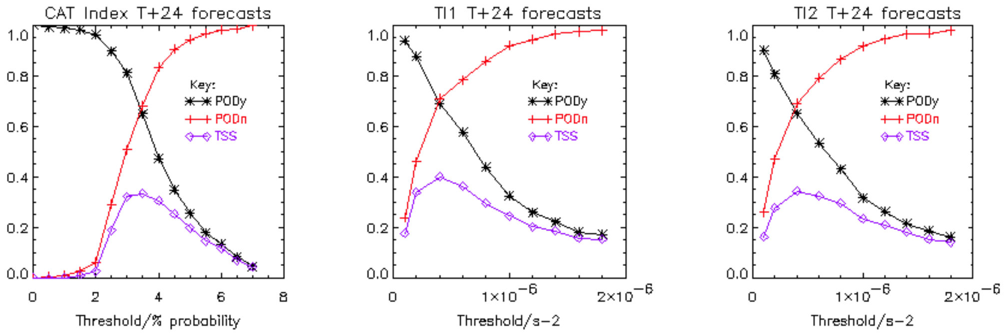


FIGURE 3. Graphs showing the variation of PODy, PODn and TSS for T+24 forecasts for the three algorithms CAT Index, TI1 and TI2.

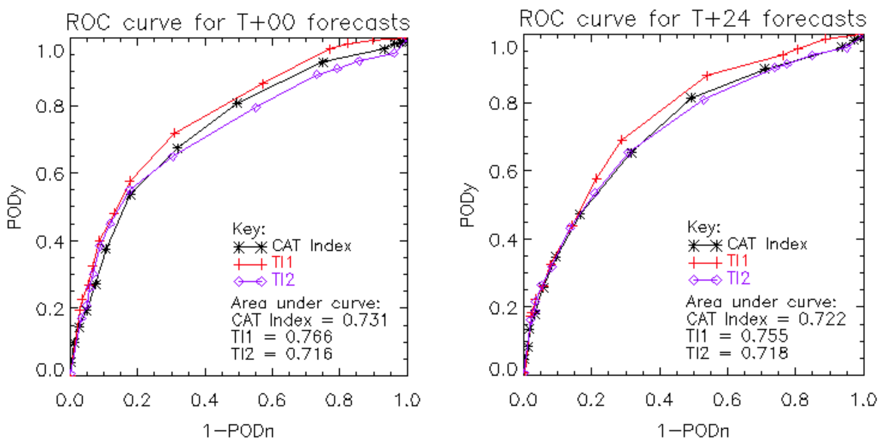


FIGURE 4. ROC curves CAT Index, TI1 and TI2 algorithms, for T+00 and T+24 forecasts.

Figure 4 shows ROC curves for the three algorithms, for the T+00 and T+24 fields. These indicate that the TI1 algorithm performs the best. The CAT Index and TI2 algorithms perform about the same (these graphs show CAT Index as performing slightly better than TI2; ROC curves plotted earlier in the project with less observations indicated TI2 being marginally better than CAT Index).

On the basis of these results TI1 was chosen as the algorithm for the shear-induced turbulence part of the new numerical format turbulence product, with a threshold of $0.6 \times 10^{-6}\ s^{-2}$ for the prediction of CAT.

4. INVESTIGATING THE USE OF "WAFTAGE" TO PRODUCE WIND SHEAR CAT FORECASTS

As part of the European Commission funded "FLYSAFE" project (Lunnon et al. 2006) the benefit of using Met Office's WAFTAGE (Winds Analysed and Forecast for Tactical Aircraft Guidance over Europe) wind and temperature nowcasting tool to produce forecasts of shear induced CAT is currently being investigated.

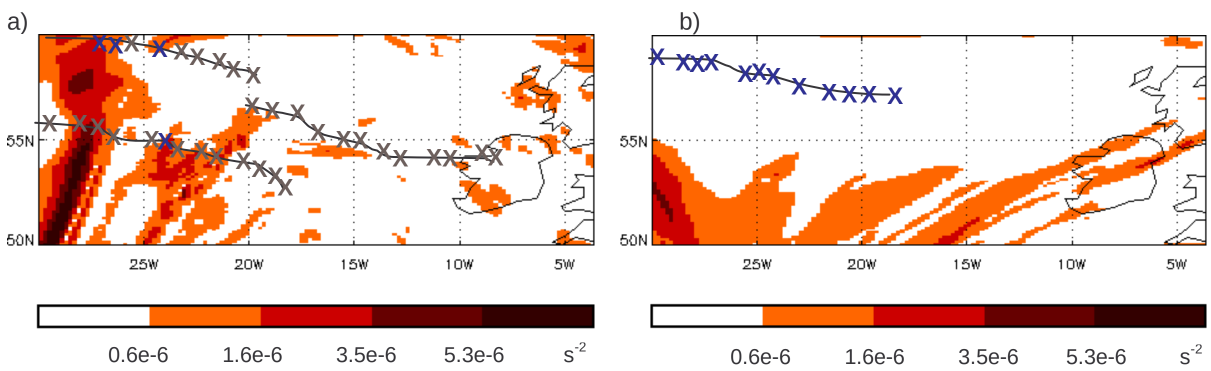
4.1 *Producing CAT Forecasts Using WAFTAGE*

The WAFTAGE nowcasting tool works by adjusting model forecasts of wind and temperature according to recent observations to provide a new forecast (for details see e.g. Sharpe (2005)). Since there is a significant time lag between the model run and the forecast validity time, there is an advantage of using WAFTAGE to forecast wind and temperature as it uses recent observations to update forecasts shortly before the forecast validity time.

As CAT predictors are often based on wind and temperature gradients, it is possible to produce a CAT forecast from WAFTAGE forecasts. These forecasts should show improved skill in predicting CAT as they have been produced using recent observations.

Input model fields	Wind and temperature forecasts for between 5 and 11 hours ahead, from the NAE model
Input observations	AMDAR aircraft observations of wind and temperature
Area covered	50° to 60°N, 30°W to 3.6°W
Levels on which wind and temperature forecasts were produced (main levels)	425, 400, 375, 350, 325, 300, 275, 250, 225, 200, 175, 150, 125 hPa
Levels on which CAT forecasts were produced (intermediate levels)	412.5, 387.5, 362.5, 337.5, 312.5, 287.5, 262.5, 237.5, 212.5, 187.5, 162.5, 137.5 hPa
Time WAFTAGE was run	Run 1 hour before the forecast validity time

TABLE 2. Details of the WAFTAGE trial.



X = CAT observation   X = "No CAT" observation   _ = aircraft tracks

FIGURE 5. WAFTAGE CAT forecasts (using Ellrod's TI1 index) for two cases in January 2007: a) for 17Z, 2nd January 2007 at 262hPa and b) for 08Z, 9th January 2007 at 237hPa.

For this investigation the WAFTAGE system was set up over the North Atlantic and to use input forecasts from the Met Office North Atlantic and European (NAE) model. It was run during the 3 month period December 2006 - February 2007. Table 2 gives further details of this trial.

Forecasts of shear induced CAT during this 3 month period were calculated from the resulting wind and temperature forecasts using a number of predictors (Browns Index (Brown 1973), CAT Index, and Ellrods TI1 and TI2 indices). These were then compared to observations of shear CAT and no CAT on a case-by-case basis and using statistical measures. The observations of shear CAT were taken from the "turbulent events database" and the set of "no CAT" observations was obtained from the original GADS data using the method described in section 2.3.

4.2 *Preliminary Results*

The results have only been processed for January 2007 so far. These indicate that there is little difference between CAT forecasts produced from WAFTAGE output and the input NAE model fields, probably because of the limited number of input observations at the levels of interest. There was also little difference in the performance of the different predictors.

There were some cases where the WAFTAGE CAT forecasts performed well (such as on the 2^nd^ January 2007, illustrated in Figure 5a) and some cases where it performed poorly (such as on the 9^th^ January 2007, as illustrated in Figure 5b). Further investigation of the meteorological situation on dates when the WAFTAGE forecasts performed poorly using satellite imagery has indicated that there were showers in the vicinity of the observations on some of these dates. This means that the observed turbulence in these cases may actually be caused by convection rather than wind shear.

5. FUTURE WORK

Further investigations with WAFTAGE CAT forecasts will focus on the following:

- re-analysis of the data including analysis of the data for December 2006 and February 2007;
- using NAE model input fields with a shorter lead time;
- investigate increasing the influence of input observations within the WAFTAGE tool;
- investigate whether using more input observations would be beneficial.

If the WAFTAGE forecasts eventually show sufficient skill in predicting CAT over the North Atlantic, a system to produce these forecasts routinely will be developed.

REFERENCES

Brown, R., 1973: New Indices to locate clear-air turbulence. *Meteorological Magazine*, 102, 347-361.

Brown, B. G., Mahoney, J. L., Henderson, J., Kane, T. L., Bullock, R. and Hart, J. E., 2000: The turbulence algorithm intercomparison exercise: Statistical verification results. *Preprints, 9th Conference on Aviation, Range and Aerospace Meteorology*. 11-15 Sept 2000, Orlando.

Brown, B. G. and Young, G. S., 2000: Verification of Icing and Turbulence Forecasts: Why Some Verification Statistics Can't Be Computed Using PIREPS. *Preprints, 9th Conference on Aviation, Range and Aerospace Meteorology*. 11-15 Sept 2000, Orlando.

Bysouth, C. E., 1998: A Comparison of Clear Air Turbulence Predictors. *Met Office Forecasting Research Technical Report No. 242.*

Dutton, M. J. O., 1980: Probability Forecasts of clear air turbulence based on numerical model output. *Meteorological Magazine*, 109, 293-310.

Ellrod, G. P. and Knapp, D. I., 1992: An Objective Clear Air Turbulence Forecasting Technique: Verification and Operational Use. *Weather Forecasting*, 7, 150-165.

Jerrett, D. and Turp, D. J., 2005: Characterisation of Turbulent Events Identified in the GADS Aircraft Dataset. *Met Office Forecasting Research Technical Report.*

Hoad, D. J. and Ogden, D. J. 2004: Characterisation of aircraft turbulence reports. *Met Office Forecasting Research Technical Report.*

Lunnon, R. W., Hauf, T., Gerz, T. and Josse, P., 2006: FLYSAFE – meteorological hazard nowcasting driven by the needs of the pilot. *Preprints, 12th Conference on Aviation, Range and Aerospace Meteorology.* 29 Jan-2 Feb 2006, Atlanta.

McCann, D. W., 1993: An evaluation of clear-air turbulence indices. *Preprints, 5th International Conference on Aviation Weather Systems*. 2-6 Aug. 1993, Vienna, Virginia.

Turner, J., 1999: Development of a mountain wave turbulence prediction scheme for civil aviation. *Met*

Office Forecasting Research Technical Report No. 265.

Sharpe, M. A. 2005: Investigation of the performance of the WAFTAGE Nowcasting system and the effect of covariance statistic function modification. *Met Office Forecasting Research Technical Report.*