Forecaster Assessment of Turbulence Algorithms: A Summary of Results for the Winter 2000 Study

Jennifer Luppens Mahoney[1] and Barbara G. Brown[2]

Submitted to:
FAA Aviation Weather Research Program Turbulence Product Development Team

25 September 2000

---

[1] NOAA Forecast Systems Laboratory, 325 Broadway R/FS5, Boulder, CO 80305-3328
[2] National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80303

1. Introduction

The FAA Aviation Weather Research Program (AWRP) has funded the Turbulence Product Development Team (PDT) to develop model-based forecasts of clear-air turbulence (CAT). In support of this development, the quality of various CAT forecasts were objectively evaluated over two winter seasons: 1 January–31 March 1999 (Brown et al. 1999, 2000) and 10 January–31 March 2000 (Brown et al. 2000). The winter 2000 evaluation was modified to also include a subjective assessment of the CAT algorithms by Aviation Weather Center (AWC) forecasters covering the period from 18 February-31 March 2000. The goals of the subjective assessment were to: 1) supplement the objective assessment with a meteorological classification of turbulence events, 2) identify the frequency of the meteorological factors leading to turbulence, 3) obtain a subjective evaluation of algorithm performance, and 4) compare the differences/similarities between the objective and subjective assessments of the quality of the turbulence forecasts.

This document summarizes the results of the winter 2000 subjective assessment of the turbulence forecasts by the AWC forecasters, addressing goals 1, 2, and 3. The comparison of differences/similarities between the objective and subjective assessments, goal 4, will be done at a later date.

The information presented in this document should be used to supplement other objective information to help identify the strengths and weakness of each of the algorithms.

2. Description of Assessment

2.1 Approach

Five forecasters at the AWC participated in assessing the performance of 12 turbulence algorithms during the period 18 February–31 March 2000. The algorithms were applied to data from the RUC-2 (Rapid Update Cycle, Version 2) model (Benjamin et al. 1998), with model output obtained from the National Centers for Environmental Prediction NCO (NCEP Computer Operations). Model forecasts issued at 1200, 1500, and 1800 UTC, with lead times of 3, 6, and 9 hours were included in the evaluation. The evaluation was limited to the region of the atmosphere at 20,000 ft and above.

Displays of the various CAT algorithms were created at NCAR and made available to the forecasters through an easy-to-use Web-based graphical user interface [http://www-ad.fsl.noaa.gov/afra/rtvs/turb_eval_2000] on the Real-Time Verification System (RTVS; Mahoney et al. 1997) of the Forecast Systems Laboratory (FSL) of NOAA. The forecasters were asked to view the displays each day and compare the output from these model-based forecasts to their assessment of the location of CAT, its strength, and its source (e.g., mountain waves). Forecasters were allowed to use all available sources of data and observations [e.g., pilot reports (PIREPs), satellite data, model forecasts] to evaluate these CAT features.

During the evaluation period, forecasters were asked, but not required, to fill out a questionnaire each day during their shift describing the weather situation at a specific time period and the performance of the algorithms in capturing the character of the turbulence at that time. Since this process was voluntary, only a subset of the total number of turbulence cases was classified. At the end of the evaluation, the questionnaires were returned to the authors at FSL and NCAR for analysis.

2.2 Tools

2.2.1 Web-based interface

Web-based tools were developed on the RTVS, which allowed forecasters to view displays of the various CAT forecasts, as well as the corresponding verification statistics [http://www-ad.fsl.noaa.gov/afra/rtvs/turb_eval_2000]. Figure 1 shows an example of a display of an ITFA algorithm forecast; each of the four panels represents a different flight level. These displays allowed the forecasters to quickly assess the output from each algorithm. The forecasters were able to obtain larger views of each panel by double-clicking on the selected image. In addition to the algorithm displays, the forecasters were able to access time series displays of statistical verification results, which summarized the past performance of the individual algorithms. These displays included time series of the True Skill Statistic and the average % Volume covered by the forecasts.

2.2.2 Questionnaire

The questionnaire addressed three main topics: 1) weather classification, 2) algorithm assessment, and 3) statistical verification (Appendix A). In Section I of the questionnaire, forecasters addressed the severity of turbulence, causes of the turbulence, location, and time of day the turbulence event occurred. In Section II of the questionnaire, forecasters evaluated whether the CAT forecasts captured the turbulence well, or did not capture it well, and whether the turbulence was overforecasted. In Section III, forecasters were asked to evaluate the statistical results and determine whether the results were consistent with their views of algorithm quality. The questionnaire was created with guidance from other members of the Turbulence PDT, and then enhanced based on feedback from the forecasters and other AWC staff prior to the start of the evaluation.

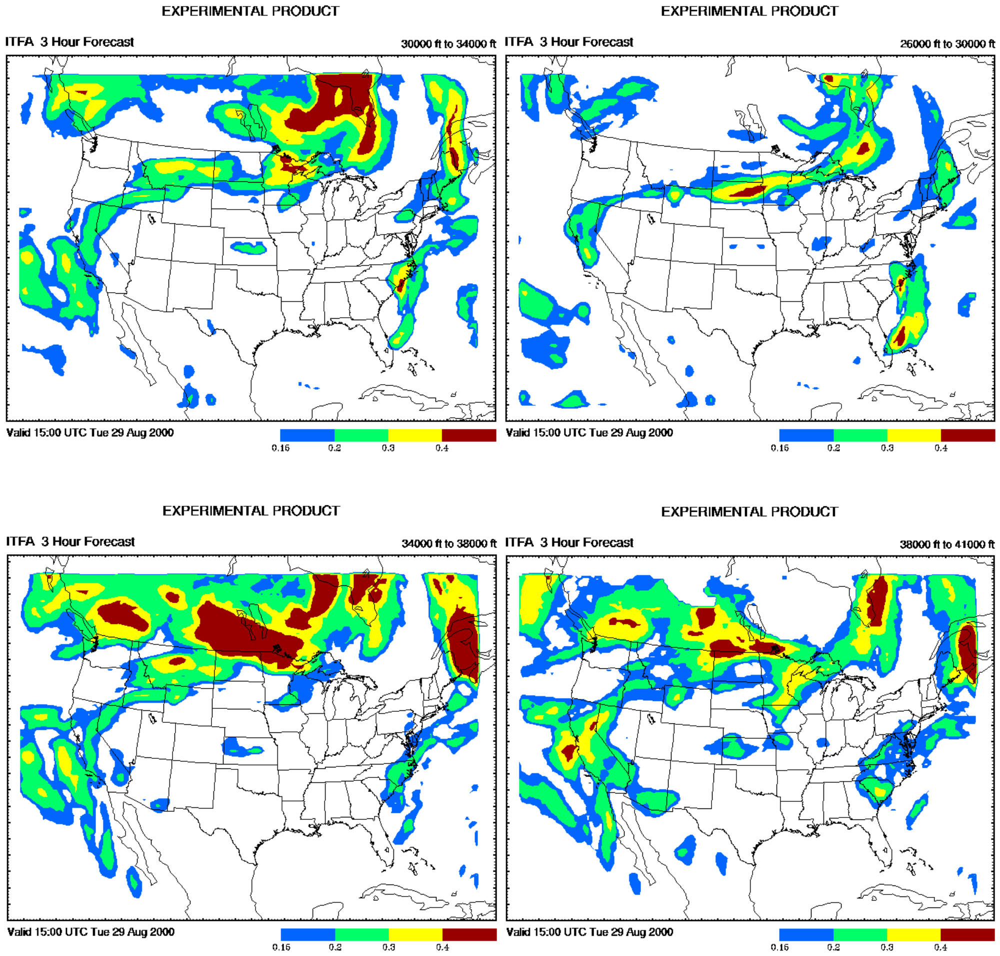


Fig. 1. Experimental ITFA, 3-h forecasts, each panel represents a different flight levels, colors indicate turbulence intensity.

2.3 *Description of the model-based turbulence algorithms*

The CAT algorithms used in this assessment are described below. Due to apparent errors that were identified in the Ulturb algorithm output, the results from that algorithm were excluded from this evaluation.

Brown-1: This index is a simplification of the Ri tendency equation originally derived by Roach (1970). The simplifications involve use of the thermal wind relation, the gradient wind as an approximation to the horizontal wind, and finally some empiricism (Brown 1973).

CCAT: The CCAT (Clark's Clear Air Turbulence) index has been used on a semi-operational basis by the U.S. Navy's FNMOC for at least 2 decades. CCAT was developed by Leo Clark in consultation with Hans Panofsky, by applying aerodynamicist

Theodore Theodorsen's theory for the generation of vortices to clear-air turbulence. There is no direct documentation on this index other than a definition and evaluation in an Naval Research Laboratory verification study document (Vogel and Sampson 1996).

DTF3, 5: The DTF ("Diagnostic Turbulence Formulation") algorithms were developed to take into account several sources of turbulent kinetic energy (tke) in the atmosphere (e.g., upper fronts), with the output in terms of tke (Marroquin 1995, 1998). These algorithms are related to one another, with the algorithm associated with each larger algorithm number incorporating more complexity.

Dutton: This index is based on linear regression analyses of a pilot survey of turbulence reports over the North Atlantic and North West Europe during 1976 and various synoptic scale turbulence indices produced by the then-operational UK Met Office forecast model (Dutton 1980). The result of the analyses was the "best fit" of the turbulence reports to meteorological outputs for a combination of horizontal and vertical wind shears.

Ellrod-2: This index was derived from simplifications to the frontogenetic function. As such it depends mainly on the magnitudes of the potential temperature gradient, deformation and convergence (Ellrod and Knapp 1992).

Endlich: The Endlich "index" is based on a paper by R. Endlich (1964). He compared jet stream structures to turbulence measurements and found best agreement of observations was obtained for an empirical parameter that is simply the product of the local wind speed and the wind direction vertical gradient.

ITFA: The Integrated Turbulence Forecasting Algorithm (ITFA) uses fuzzy logic techniques to integrate available turbulence observations [in the form of PIREPs and vertical accelerometer (AVAR) data] together with a suite of turbulence diagnostic algorithms (a superset of algorithms used in the verification exercise and others) to obtain a turbulence forecast (Sharman et al. 1999).

Richardson Number: Theory and observations have shown that at least in some situations patches of CAT are produced by what is known as Kelvin-Helmholtz (KH) instabilities. This occurs when the Richardson number (Ri), the ratio of the local static stability to the local shears, becomes small. Therefore, theoretically, regions of small Ri should be favored regions of turbulence (Drazin and Reid 1981; Dutton and Panofsky 1970; Kronebach 1964).

SCATR: This index is based on attempts by several investigators to forecast turbulence by using a time tendency (i.e., prognostic) equation for the Richardson number (Roach 1970). The version used in this study was based on a formulation of this equation in isentropic coordinates by John Keller, who dubbed the algorithm "SCATR" (Specific CAT Risk; Keller 1990).

Vertical wind shear: Wind shear is known to be a destabilizing force from the time of Helmholtz. This can be seen from its inverse relation to Richardson's number: large values favor small Ri, which in turn produce turbulence in stratified fluids (Drazin and Reid 1981; Dutton and Panofsky 1970).

UlTurb: The UlTurb (Upper-Level Turbulence) forecasting index was developed by Don McCann at AWC (McCann 1997). It attempts to correlate unbalanced (i.e., nongeostrophic) flow to regions of CAT. Three different measures of this imbalance are computed, and the maximum of these measures relates to turbulence potential. The correlation between unbalanced flows and turbulence is supported at least qualitatively from numerous field experiments, both over the continental U.S. and the North Pacific (Knox 1997).

3. Results

3.1 General comments

At the end of the evaluation period, 57 questionnaires were collected, with five forecasters participating in the evaluation. These questionnaires represented a total of 40 days. In some cases, more than one forecaster completed a questionnaire for the same day. Table 1 lists the cases that were evaluated, according to category, region, and date of turbulence. As shown in Table 1, the western region was the focus with 27 turbulence cases, with 19 in the central region, and 5 in the eastern region (note; in some instances, some turbulence cases were grouped into several categories). Five days during the evaluation period were considered *big* days with a lot of turbulence activity, 23 were

Table 1. A summary of days with turbulence categorized by turbulence activity and forecast region.

<table>
<tr><th colspan="3"><i>Big Days</i></th></tr>
<tr><td>West</td><td>Central</td><td>East</td></tr>
<tr><td>7 and 14 March</td><td>26 February
19 and 20 March</td><td></td></tr>
<tr><th colspan="3"><i>Moderate Days</i></th></tr>
<tr><td>West</td><td>Central</td><td>East</td></tr>
<tr><td>20, 21, 22, 23,
26, 27 February
1, 2, 3, 4, 8, 10, 15, 17, 21
22, 26, 27, 29, 30 March</td><td>22, 27 February
8, 11, 15, 19, 21, 22
27, 29 March</td><td>3, 11, 15, 29 March</td></tr>
<tr><th colspan="3"><i>Small Days</i></th></tr>
<tr><td>West</td><td>Central</td><td>East</td></tr>
<tr><td>18, 20 February
9, 12, 25, 28 March</td><td>18, 19, 20, 29 February
1, 28 March</td><td>25 March</td></tr>
<tr><th colspan="3"><i>No Turbulence or not Classified</i></th></tr>
<tr><td>West</td><td>Central</td><td>East</td></tr>
<tr><td>1 March</td><td>12 and 24 March</td><td></td></tr>
</table>

considered to be *moderate* days, 9 were considered to be *small* days, 1 day had *no* turbulence activity, and 2 days were not classified. In general, when more than one forecaster evaluated the turbulence on a given day, their assessments of the turbulence category (i.e., “big day,” “moderate day”) were in agreement. However, in some cases the forecasters disagreed on the level of activity for a particular day, as is the case for 19 March, as shown in Table 1 indicating the difficulties with evaluating turbulence.

3.2 Weather classification

The results presented in this section focus on the moderate and severe turbulence cases that fell within the three categories of turbulence activity presented in Table 1 (e.g., big, moderate or small days).

3.2.1 Severity and causes of turbulence

The forecasters classified 19 days with moderate turbulence and 20 days as having severe turbulence above a flight level of 20,000 ft. These cases are listed in Table 2. Cases that appear in both columns in Table 2 are the result of two or more entries with differing turbulence intensities, or turbulence cases that fell in different regions. If a case was identified by two or more forecasters, both indicating the same severity, it was entered only once in Table 2 under that particular severity.

Table 2. Summary of the moderate and severe turbulence cases

Moderate	Severe
18, 19, 20, 22, 23, 27, 29 February 1, 8, 9, 11, 15, 19, 22, 24, 25, 26, 27, 28 March	22, 26 February 1, 2, 3, 4, 7, 10, 12, 14, 15, 17, 19 20, 21, 25, 26, 27, 29, 30 March

The major meteorological causes of turbulence that were identified by the forecasters included jet streams, mountain waves, convection, and upper-level troughs. For all of the moderate and severe cases (including those cases identified by different forecasters to have the same severity), 16 were considered to be due to a combination of an upper-level trough and the jet stream; 11 were believed to be the result of an upper-level trough; 10 were believed to be due to a jet stream; 7 were considered to be associated with a combination of convection, jet stream, and upper-level trough; and 4 were believed to be due to a combination of mountain waves, jet streams, and upper-level troughs. However, when the inferred causes of the turbulence are categorized by turbulence severity, as shown in Tables 3 and 4, turbulence of moderate intensities were more often associated with convection than severe reports of turbulence.

3.2.2 Location of the turbulence

The forecasters were asked to graphically depict the location of the turbulence directly on the questionnaire. These results are presented in Appendix B.

Table 3. Summary of the major meteorological causes of severe turbulence

Causes	Severe Cases	Number of Cases
Upper trough and jet stream	22 February 2, 3, 7, 10, 17, 20, 21, 30 March	9
Jet stream	26 February 26, 29 March	3
Upper trough	1, 3, 4, 12, 19, 21, 25, 27, 29 March	9
Mountain waves, upper trough, and jet	14, 17 March	2
Convection, upper trough, and jet	15 March	1

Table 4. Summary of the major meteorological cases of moderate turbulence

Causes	Moderate Cases	Number of Cases
Upper trough and jet stream	27 and 29 February 9, 11, 19, 22, 25 March	7
Jet stream	18, and 27 February 15, 26, 27, 28 March	6
Upper trough	29 February and 1 March	2
Mountain waves, upper trough	23 February and 8 March	2
Convection, upper trough, and jet	18, 20, 22 February 8, 15, 24 March	6

3.2.3 Time

Tables 5 and 6 summarize the time of day in which the forecasters noted the moderate and/or severe turbulence cases. Interestingly, the majority of severe cases shown in Table 5 occurred between 1800 and 0000 UTC, while the majority of moderate cases shown in Table 6 occurred earlier, between 1200 and 1800 UTC. These results could be a function of the time in which the evaluations were completed or a function of the forecaster's work schedule.

3.2.4 Mountain wave boxes

Generally, throughout the assessment period, turbulence due to mountain wave activity was at a minimum, with only 4 cases being identified as mountain wave cases. However, in 9 cases, forecasters were able to assess the mountain wave boxes and did affirm that in 8 out of the 9 cases, the boxes correctly captured the mountain wave turbulence events. In one of the cases, the mountain wave event was just east of the boxes.

Table 5. Time of day when the severe turbulence cases occurred

Time of Day	Severe Cases	Number of cases
1200 - 1500 UTC		0
1500 - 1800 UTC	26 February and 3 March	2
1500 - 1800 and 1800 - 2100 UTC	7, 20, and 29 March	3
1800 – 2100 UTC	26 February and 26 March	2
1800 – 2100 and 2100 – 0000 UTC	4, 17, 19, 21, 27, 30 March	6
2100 – 0000 UTC	22 February 1, 2, 4, 10, 12, 14, 15, 21, 25, 29 March	11

Table 6. Time of day when the moderate turbulence cases occurred

Time of Day	Moderate Cases	Number of cases
1200 - 1500 UTC	8 and 9 March	2
1500 - 1800 UTC	18, 19, 27, 28 and 29 February 1, 8, and 25 March	8
1500 - 1800 and 1800 - 2100 UTC	27 February 11, 26, and 27 March	4
1800 - 2100 UTC	20 February and 15 March	3
1800 - 2100 and 2100 - 0000 UTC	19 and 24 March	2
2100 - 0000 UTC	18, 22, and 23 February 22 March	4

3.3 Algorithm assessment

The results presented in this section pertain to the quality of the turbulence algorithms, as considered in Part II (questions 6-12) of the questionnaire. Ulturb is excluded from this evaluation since an error was detected in the displays. In considering these results, note that the forecasters were able to choose more than one response to each of the questions.

3.3.1 Which algorithms were considered and when?

Nearly all algorithms were considered by the forecasters during each assessment session with a few exceptions. The forecast issue and lead times evaluated most often are listed in Table 7, in order of frequency. In general, the 6-h lead times were used most often, with the 1200, 1500, and 1800 UTC issue times being the most widely used. On the other hand, the analyses were not used, mainly because they are not available to the forecasters early enough to be useful for formulating their forecasts.

3.3.2. How did the algorithms perform?

The results of the forecasters' assessments of algorithm performance are presented in Table 8. In general, the algorithms that were believed to capture the turbulence and produce the correct amount of turbulence include Brown-1, ITFA, Ellrod-2, and Richardson Number. On the other hand, the forecasters did not generally think that the Endlich and CCAT algorithms captured the turbulence well and, in the majority of the cases, the consensus was that these algorithms produced too much turbulence.

Table 7. Forecast issue and lead times most often used to evaluate the algorithms

Forecast Issue and Lead Time	Frequency
1200 UTC, 6-h	9
1500 UTC, 6-h	9
1800 UTC, 6-h	9
1200 UTC, 3-h	8
1800 UTC, 3-h	7
1200 UTC, 9-h	5
1500 UTC, 3-h	3
2100 UTC, 3-h	2
1500 UTC, 0-h	1
1200 UTC, 12-h	1
1500 UTC, 9-h	1
1200 UTC, 0-h	0
1800 UTC, 0-h	0
2100 UTC, 0-h	0

If the algorithm performance is separated by intensity as summarized in Table 9, the cases are nearly evenly split between moderate and severe turbulence with a few exceptions. DTF5, Dutton, and Vertical Wind Shear were more likely to capture the turbulence when during moderate events as compared to severe events.

The forecasters identified the problems with the forecasts as directly related to the biases in the algorithms and not with the numerical models.

3.4 Forecaster comments on algorithms that captured turbulence

This section summarizes forecaster comments concerning the algorithms that seemed to correctly capture the turbulence regions. The date the turbulence occurred is listed for reference. For more detail, see Appendix B.

Table 8. Summary of turbulence algorithm performance for all cases

Algorithms	Number times the algorithm was considered	Captured the turbulence well	Did not capture the turbulence well	Forecast too much turbulence
Brown -1	56	27	24	9
CCAT	56	11	35	33
DTF3	53	17	24	24
DTF5	54	11	25	24
Dutton	54	10	28	29
Ellrod-2	54	23	21	19
Endlich	52	0	43	50
ITFA	54	27	17	5
Richardson Number	52	21	24	1
Vertical Wind Shear	48	13	26	16

Table 9. Summary of turbulence algorithm performance categorized by intensity

Algorithms	Number times the algorithm was considered	Captured the turbulence well		Did not capture the turbulence well		Forecast too much turbulence	
		Mod	Sev	Mod	Sev	Mod	Sev
Brown -1	56	17	10	9	15	4	5
CCAT	56	7	4	13	22	21	12
DTF3	53	8	9	12	12	15	9
DTF5	54	8	3	8	17	18	6
Dutton	54	8	2	12	16	16	13
Ellrod-2	54	13	10	10	11	11	8
Endlich	52	0	0	19	24	28	22
ITFA	54	16	11	7	10	4	1
Richardson Number	52	13	8	9	15	1	0
Vertical Wind Shear	48	9	4	9	17	10	6

3.4.1 Severe turbulence cases

22 Feb. - Ellrod-2 did a little better in the Pacific northwest with anticyclonic flow.
26 Feb. - Best were Brown-1 and Ellrod-2. Both slightly overforecast areas, but pattern was very good. Could not determine where algorithms suggest severe.
3 Mar. - ITFA performed best and did not forecast too much turbulence in east as did most others.

4 Mar. - DTF3 and Richardson Number captured turbulence well over central California between flight level 220 and 260 where most moderate-severe reports occurred.
4 Mar. - All algorithms forecast too much turbulence in southeast with eastern trough.
7 Mar. - DTF3, DTF5, Richardson Number, and Vertical Wind Shear showed similar patterns, but Richardson Number was not strong enough.
10 Mar. - None did well in the west. A sharp southwest trough spawned many moderate reports. A combination of low Richardson Number and the anomalous gradient instability were the best aspects of Ulturb.
12 Mar. - Too much turbulence forecast in the NY-PA-VA area. Richardson Number was better in Colorado than in the east.
14 Mar. - Dutton was better in the northwest and over mountains in Wyoming and Colorado.
15 Mar. - ITFA was not bad, but too little turbulence in the 20-30K range. Richardson Number not too bad, but too much turbulence up north.
19 Mar. - Brown-1, ITFA, and Vertical Wind Shear captured turbulence well without overforecasting. Several algorithms from Brown-1 to Ellrod-2 did well between FL260-300.
20 Mar. - Brown-1 and ITFA captured western turbulence well and showed similarities. CCAT, DTF3, and DTF5 not enough turbulence forecast over northern intermountain region.
21 Mar. - ITFA, Vertical Wind Shear, and Richardson Number did the best job of capturing turbulence from VT to western SD and over central and southern Rocky Mountains.
21 Mar. - DTF3 did better over VT on west side of the upper trough.
25 Mar. - All algorithms were poor. Richardson Number did better than others by not forecasting as much turbulence.
27 Mar. - Turbulence over ridge in southwest was missed by all. ITFA worst with too little turbulence and Endlich bad with too much turbulence.
29 Mar. - For some reason most algorithms forecast too much turbulence in the East Coast trough, but not enough in the west.
29 Mar. - ITFA, DTF3, and Richarson Number seemed to capture area around western KY/TN best of all, though none of the algorithms were strong enough across lower Mississippi Valley Gulf states.
30 Mar. - DTF3 and Ellrod-2 forecast the turbulence in Pacific Northwest, west of the trough in Negative Vorticity Advection and anticyclonic shear.

3.4.2 Moderate turbulence cases

18 Feb. - Most way over forecast turbulence in the northeast. Brown-1 had the mean area of western turbulence too far north.
18 Feb. - DTF3 was way too much in the west.
19 Feb. - Endlich consistently did not do very well. CCAT did better than expected.
22 Feb. - Endlich way too much turbulence.
23 Feb. - Endlich predicted too much turbulence everywhere. Richardson Number showed too little turbulence. Vertical Wind Shear and Brown area showed moderate turbulence over Idaho and eastern Oregon, and Brown-1 showed some severe turbulence reported over eastern IA.
27 Feb. - Several did pretty well, question was if turbulence would be moderate or severe. Algorithms not much help with the forecast decision.
29 Feb. - Brown-1 captured the upper deformation zone turbulence best from north-central plains through upper Mississippi Valley at FL260-FL330.
8 Mar. - Brown-1, DTF3, DTF5, and Dutton all forecast moderate turbulence at FL240-360 over northern plains; few PIREPs in this area. ITFA did reasonably well in north-central plains. Ellrod-2 did best over central Rocky Mountains where most PIREPs reported at FL280-390.
8 Mar. - Brown-1 and Richardson Number were consistently better than other indices.
9 Mar. - Many of the indices seemed to have a reasonable area at lower altitudes (220-240)… then the area expanded and became stronger at higher altitudes (310-360).
11 Mar. - Once again, forecasts were better at low altitudes than at high altitudes.
15 Mar. - DTF3 and DTF5 seemed best over the west and from the lower Mississippi Valley to Ohio Valley.
19 Mar. - Most of the algorithms showed too much turbulence over the northeastern U.S.
22 Mar. - CCAT too far south. Several algorithms did well with turbulence pattern from southern U.S. to central plains and FL200 to 308 southern U.S. and mid to upper FL308 plains. Maybe ITFA did best overall.
24 Mar. - ITFA and Richardson Number not too bad, but did not capture the turbulence well.
25 Mar. - Most turbulence reports were between FL240-350. Several algorithms captured the turbulence well. Brown-1 did best between FL240-350.
26 Mar. - CCAT had considerable amount of turbulence forecast for central and eastern U.S. Most of the algorithms seem to forecast less turbulence in the lower FL (20s) than in the upper ranges (FL 30 and lower 40s).
27 Mar. - Several algorithms did fairly well over central and southern plains, but not so well over the Rocky Mountains and southern U.S. Ellrod-2 seemed best overall.
28 Mar. - Several algorithms did well in the southern U.S. and southern plains. DTF3, DTF5, Dutton, Ellrod-2, Endlich, ITFA, and Vertical Wind Shear all were similar. Brown-1 and CCAT were not strong enough.

3.5 Statistical verification

Forecasters compared the statistical verification results to the relative ability of the various algorithms to predict turbulence. In 27 cases, the forecasters agreed that the statistical results were consistent with their perception of algorithm performance, while in 10 of the cases the statistical verification results did not agree with their perception of algorithm performance.

4. Summary

The results of this study are two fold: 1) they provide a great deal of information regarding the important sources of turbulence associated with particular cases and types of turbulence and 2) identify particular situations in which various algorithms work better than others. These results can lead to improvements in turbulence forecasts; for example, the information could be applied to develop dynamic weighting approaches in ITFA that would change the weighting scheme based on the synoptic situation.

The results also will aid in interpretation of the objective verification results. One of the next steps in this study is to correlate specific questionnaires with the verification results for the corresponding day. In addition, a similar evaluation will be undertaken during winter 2000-2001 to increase the evaluation sample size.

Acknowledgments

This research is in response to requirements and funding by the Federal Aviation Administration. The views expressed are those of the authors and do not necessarily represent the official policy and position of the U.S. Government.

We would like to thank the Aviation Weather Center forecasters who participated in this evaluation. Without them, this work could not have been performed. We also would like to thank Clinton Wallace of the AWC for being the AWC focal point for this evaluation and for providing guidance to the forecasters when needed. We would also like to thank the RTVS Team (Judy Henderson, Joan Hart, Andy Loughe, Beth Sigren, and Chris Fischer who are the developers of the RTVS, and Nita Fullerton for her helpful review of this work.

References

Benjamin, S.G., J.M. Brown, K.J. Brundage, B.E. Schwartz, T.G. Smirnova, and T.L. Smith, 1998: The operational RUC-2. *Preprints, 16th Conference on Weather Analysis and Forecasting*, Phoenix, AZ, American Meteorological Society, 249-252.

Brown, B.G., J.L. Mahoney, R. Bullock, J. Henderson, and T.L. Kane, 1999: Turbulence algorithm intercomparison: 1998-1999 Initial Results. FAA Turbulence Product

Development Team Report to the FAA Aviation Weather Research Program (Available from B. Brown, Research Applications Program, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000 or from the WWW at http://www-ad.fsl.noaa.gov/afra/rtvs/reference_mats.html).

Brown, B.G, J.L. Mahoney, J. Henderson, Tressa L. Kane, Randy Bullock, and Joan E. Hart, 2000: The turbulence algorithm intercomparison exercise: Statistical verification results. *Preprints, 9th Conference on Aviation, Range, and Aerospace Meteorology*, Orlando, Fl., American Meteorological Society.

Brown, R., 1973: New indices to locate clear-air turbulence. *Meterol. Mag*., **102**, 347-361.

Drazin, P.G. and W.H. Reid, 1981: **Hydrodynamic Stability**. Cambridge, 527 pp.

Dutton, J. and H. A. Panofsky, 1970: Clear Air Turbulence: A mystery may be unfolding. *Science*, **167**, 937-944.

Dutton, M.J.O., 1980: Probability forecasts of clear-air turbulence based on numerical model output. *Meterol. Mag.*, **109**, 293-310.

Ellrod, G.P. and D.I.Knapp, 1992: An objective clear-air turbulence forecasting technique: verification and operational use. *Wea. Forecasting*, **7**, 150-165.

Endlich, R. M., 1964: The mesoscale structure of some regions of clear-air turbulence. *Mon. Wea. Rev.*, **92**, 261-276.

Keller, J. L., 1990: Clear Air Turbulence as a response to meso- and synoptic-scale dynamic processes. *Mon. Wea. Rev.*, **118**, 2228-2242.

Knox, J. A., 1997: Possible mechanism of clear-air turbulence in strongly anticyclonic flows. *Mon. Wea. Rev.*, **125**, 1251-1259.

Kronebach, G. W., 1964: An automated procedure for forecasting clear-air turbulence. *J. App. Met.*, **3,** 119-125.

Mahoney, J.L., Judy Henderson, P.A. Miller, 1997: A description of the Forecast Systems Laboratory's Real-Time Verification System (RTVS). *Preprints, 7th Conference on Aviation, Range, and Aerospace Meteorology*, Long Beach, CA, American Meteorological Society, J26-J31.

Marroquin, A., 1995: An integrated algorithm to forecast CAT from gravity wave breaking, upper fronts and other atmospheric deformation regions. *Preprints, 6th Conference on Aviation Weather Systems*, Dallas, TX, American Meteorological Society, 509-514.

Marroquin, A., 1998: An advanced algorithm to diagnose atmospheric turbulence using numerical model output. *Preprints, 16th Conference on Weather Analysis and Forecasting*, Phoenix, AZ, American Meteorological Society.

McCann, D. W., 1997: A "novel" approach to turbulence forecasting. Preprints, *Seventh Conf. On Aviation, Range and Aerospace Meteorology*, Long Beach, CA, American Meteorological Society, 158-163.

Roach, W.T., 1970. On the influence of synoptic development on the production of high level turbulence. *Quart. J. R. Met. Soc*., **96**, 413-429.

Sharman, R, C. Tebaldi, and B. Brown, 1999: An integrated approach to clear-air turbulence forecasting. Preprints, *Eighth Conf. On Aviation, Range, and Aerospace Meteorology*, Dallas, TX, American Meteorological Society, 68-71.

Vogel, G.N. and C.R. Sampson, 1996: Clear air tubulence indices derived form U.S. Navy numerical model data: a verification study. Naval Research Laboratory, Monterey, NRL/MR/7543-96-7223, 30 pp.

APPENDIX A

Turbulence Algorithm Questionnaire

Please fill out one questionnaire every day, if possible, considering only turbulence at 20,000 ft and above. If you have problems getting data or have questions regarding this questionnaire, please contact Jennifer Mahoney at mahoney@fsl.noaa.gov (303-497-6514) or Barbara Brown at bgb@ucar.edu (303-497-8468).
NOTE: All displays (including mountain wave boxes) are on the web at
http://www-ad.fsl.noaa.gov/afra/rtvs/turb_eval_2000
(User name: RTVS Password: turb2000)

Name:_______________________________________

Current Date (UTC):____________________ **Current Time (UTC):**___________

Date of turbulence (UTC) if different from above: _____________________

On my shift, I focused on the following region(s):
___ **West** ___**Central** ___**East**

How would you categorize this day, in terms of turbulence activity? (We are most interested in turbulence events during the hours 1200 to 0000 UTC.)

___ A “big” day ___A “moderate” day ___A “small” day

Comments:

I. Weather Classification

***Note**: Using only PIREPs to identify the moderate to severe turbulence events or turbulence outbreaks above 20,000 ft, please answer the following questions:*

1. What was the maximum severity of turbulence today?
___Chop ___Light ___Mod ___Severe
Comments:

2. What is the major cause of the turbulence?
___Jet Stream ___Mountain Waves ___Convection
___Upper level trough

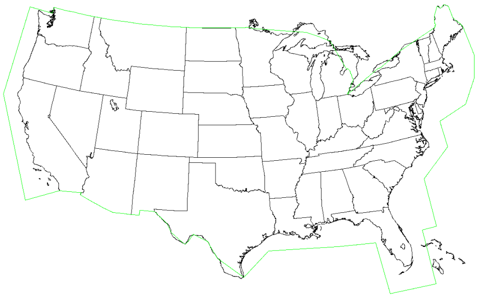
Other______________________

3. Where is the turbulence located? ***(Draw on map, if desired)***
___Over entire U.S. ___East Other:_________
___Northern half of U.S. ___Central _____________
___Southern half of U.S. ___West

4. At what time of day did the major events occur?
___12-15 UTC ___15-18 UTC ___18-21 UTC ___21-00 UTC

5a. Do the mountain wave boxes shown on the web site (listed above) accurately capture today’s mountain wave events?
___Yes ___No

5b. If No, how should the mountain wave boxes be defined (use map below to draw new boxes)?



II. Algorithm Assessment

Still considering turbulence at levels of 20,000 ft and above, please review the algorithm output provided on the web site. <u>Check all responses that apply</u>.

6. Which forecast issue and lead times did you consider?
___12Z, 0-h ___12Z, 3-h ___12Z, 6-h ___12Z, 9-h ___12Z, 12-h
___15Z, 0-h ___15Z, 3-h ___15Z, 6-h ___15Z, 9-h
___18Z, 0-h ___18Z, 3-h ___18Z, 6-h
___21Z, 0-h ___21Z, 3-h

7. Which algorithms did you consider?
___Brown-1 ___CCAT ___DTF3 ___DTF5 ___Dutton ___Ellrod-2
___Endlich ___ITFA ___Rich # ___Vert. Wind Shear ___Ulturb

8. Which algorithms seemed to capture the turbulence well?
___Brown-1 ___CCAT ___DTF3 ___DTF5 ___Dutton ___Ellrod-2
___Endlich ___ITFA ___Rich # ___Vert. Wind Shear ___Ulturb

9. Which algorithms did *not* seem to capture the turbulence well?
___Brown-1 ___CCAT ___DTF3 ___DTF5 ___Dutton ___Ellrod-2
___Endlich ___ITFA ___Rich # ___Vert. Wind Shear ___Ulturb

10. Which algorithms seemed to forecast too much turbulence?
___Brown-1 ___CCAT ___DTF3 ___DTF5 ___Dutton ___Ellrod-2
___Endlich ___ITFA ___Rich # ___Vert. Wind Shear ___Ulturb

11. Do the problems with the forecasts seem to be due to biases in the numerical model or to an inaccurate forecast by the turbulence algorithms?
___Numerical Model ___Algorithms ___Can't Determine

Comments:

12. For the algorithms that captured the turbulence, please list any trends (e.g., varying regional capabilities) that you noticed.

III. Statistical Verification

Please review the time series plots of TSS and % Volume provided on the web site

13. Are the verification statistics consistent with your views regarding the relative ability of the various algorithms to predict turbulence?

___Yes ___No

Comments:

APPENDIX B

Forecaster Evaluation Forms