International Civil Aviation Organization

SIXTEENTH MEETING OF THE COMMUNICATIONS/NAVIGATION/SURVEILLANCE AND METEOROLOGY SUB-GROUP (CNS/MET SG/16) OF APANPIRG

Bangkok, Thailand, 23 – 27 July 2012

Agenda Item 8: **World Area Forecast System (WAFS)**

4) Gridded forecasts (icing, turbulence and CB) visualisation

VERIFICATION OF THE WAFS CAT POTENTIAL AND THE TRIAL HARMONIZED VERSION OF GRIDDED FORECAST PRODUCTS

(Presented by Hong Kong, China)

SUMMARY

This paper presents an update on the verification of WAFS gridded clear air turbulence forecast using flight data from flights landed/departed from Hong Kong, China.

This paper relates to –

Strategic Objectives:

A: Safety - *Enhance global civil aviation safety*
C: Environmental Protection and Sustainable Development of Air Transport - *Foster harmonized and economically viable development of international civil aviation that does not unduly harm the environment*

Global Plan Initiatives:
GPI-19 Meteorological Systems

1. INTRODUCTION

1.1 Verification of the WAFS trial gridded forecasts of the clear-air turbulence (CAT) potential was reported in the Sixth Meeting of WAFSOPSG in March 2011. Using WAFS trial gridded CAT potential products between October 2008 and September 2009, and the "sky truth" based on the eddy dissipation rate (EDR) calculated from the flight data recorder (FDR) data, it was observed that the gridded CAT potential products of WAFC London and WAFC Washington differed in forecast values of en-route turbulence in short-range (up to 24 hours of forecast) and over different geographical areas. Discussion was raised on the harmonization of the products from the two WAFCs to facilitate interpretation and applications of forecast products by the users.

1.2 In this paper, the verification of the CAT potential is extended to cover forecast runs up to November 2010. The number of verifying data points increased by about 40% with the number of available flights increased from about 1400 to near 2000. The same verifying area covering East and Southeast Asia is considered to facilitate the comparison of results of the previous study, in order to obtain a more comprehensive review on the general performance and characteristics of the gridded CAT potential products over the Asian regions.

1.3 Starting from 29 November 2011 (1200 UTC model run), the "harmonized" version of gridded turbulence forecast product has been made available from the WAFC Provider States for "trial and evaluation". That is in response to the recommendation by WAFSOPSG on the operation of WAFS forecasts for cumulonimbus cloud, icing and turbulence tentatively in November 2013, subject to success of harmonized products[1]. In essence, the harmonized mean forecasts are derived by taking the "mean of the mean" values while the maximum forecasts are derived by taking the higher value of the maximum values from each of the WAFC forecasts. As the harmonization might impact on the performance of the gridded forecasts, the CAT potential forecasts subsequent to harmonization is verified separately.

2. DISCUSSION

2.1 Following the methodology in the previous study, two categories of turbulence intensity were considered for verification of CAT potential along the flight path within 1.25 × 1.25 degree latitude/longitude grid box. The vertical levels range from 400 hPa to 150 hPa at an interval of 50 hPa. The thresholds of turbulence intensity are mean EDR $\geq$ 0.1 $m^{2/3}$ s^{-1} and maximum EDR $\geq$ 0.4 $m^{2/3}$ s^{-1}. They represent light turbulence or above on average, and moderate turbulence level or above for the peak turbulence respectively (Section 2.6 of Appendix 4 in ICAO Annex 3).

2.2 Figure 1 shows the relative operating characteristic (ROC) curves for the CAT potential products of WAFC London and WAFC Washington at forecast time of 6 hour (red), 12 hour (green), 18 hour (blue) and 24 hour (purple). For light turbulence or above on average (Figure 2 (a) and (b)), the observed frequency of the turbulence is above 1.7%. Similar to the previous study, the two WAFCs continue to show comparable skills. ROC curves for the four forecast hours are similar. For a false alarm rate of about 0.2, the hit rate of WAFC London is between 0.33 – 0.38 while that of WAFC Washington is close to 0.4. For the ROC curves of WAFC London (Washington), the area under the curves (AUC; the larger the better in performance) are about 0.58 – 0.59 (0.62 – 0.63). In case of moderate turbulence or above for the peak turbulence, the observed frequencies are reduced to about 0.2%. WAFC Washington has better skills that the hit rates of the four forecast ranges are 0.35 – 0.45 when the false alarm rate is 0.2, and the hit rates of WAFC London are around 0.3 or slightly lower. The higher skills of the WAFC Washington forecast are also revealed from the AUC, in which the values are 0.60 – 0.65 while those of WAFC London are in the range of 0.53 and 0.57.

2.3 The values of the above AUC are slightly smaller than those in the verification study against GADS data between 20 N and 50 N (Gill, 2012[2]). In that study, the hit rate was about 0.48 for light or greater turbulence (DEVG $\geq$ 2) (Figure 6 in Gill, 2012; not reproduced here) and the value of AUC of WAFC London (Washington) is 0.708 (0.720).

2.4 It is also worth to note that for the gridded WAFC Washington forecasts, discontinuities over the top-right-hand corner of the ROC curves are more significant (Figure 2(b) and (d)) compared with those in the previous study. More pronounced discontinuities are found in the ROC curves of

[1] Guidance on the Harmonized WAFS Forecasts for Cumulonimbus Cloud, Icing and Turbulence forecasts; Version 1.0; 28 November 2011.
[2] Gill, P.G., 2012: Objective verification of World Area Forecast Centre clear air turbulence forecasts. Submitted to Met. Appl.

WAFC London in the present and previous studies. It is further noted that the number of grid-point forecasts with small CAT potential over the extended period (i.e. October 2009 to November 2010) from both WAFCs is smaller compared with the period in the previous study (i.e. October 2008 to September 2009). A possible reason of this change might attribute to the on-going development effort in the two WAFCs to reduce their difference in the forecast gridded CAT potential, at least in terms of spatial coverage and frequency of occurrence for the small values of gridded CAT potential.

2.5 Performance of CAT potential for flight routes over various geographical coverage is also investigated. The numbers of flights over the 4 quadrants (N/E/S/W) used in the verification are shown in Figure 2. If the flight route is short and found near the boundary between quadrants, the flight data will be excluded in the quadrant-based verification. The number of such cases is about 10% of the total flight routes.

2.6 In the regional verification results (Figure 3(a)-(d)) for 24-hour forecast, limited / no skill (i.e. below the light grey diagonal line) is seen over the W quadrant in the two WAFC products and is again similar to the previous study. A possible reason for the lack of skill for this quadrant is the difficulty in forecasting clear air turbulence caused by the complex mountainous terrain within this quadrant. For the 24-h forecast of average turbulence at light turbulence or above, WAFC London forecasts for E and N quadrants have AUC of 0.59 and 0.66 respectively but the forecast for S quadrant only has limited / no skill. WAFC Washington show skills with AUC between 0.61 and 0.65 in E, S and N quadrants. For the moderate turbulence level or above for the peak turbulence (maximum EDR ≥ 0.4 $m^{2/3}$ s^{-1}), the AUC of WAFC London for E, N and W are ranged between 0.56 and 0.59 whereas the AUC of WAFC Washington forecast for the four quadrants ranges from 0.58 to 0.63, and that again indicates a better skill from the WAFC Washington forecasts in general.

2.7 As a first attempt to understand the performance of the "harmonized" version of gridded turbulence forecasts, verification has also been made to the flight data collected during December 2011 to April 2012. Totally, there are about 180 flights available providing 3000 data points to verify the harmonized gridded CAT potential for 6, 12, 18 and 24 hour of forecasts.

2.8 Figure 4(a) and 4(b) shows the respective ROC curves for the mean EDR ≥ 0.1 $m^{2/3}s^{-1}$ and maximum EDR ≥ 0.4 $m^{2/3}$ s^{-1}. For the former, i.e. on light turbulence or above on average, the AUC generally decreases from 0.63 at T+6 h to 0.57 at T+24 h of forecasts. Though the frequency of occurrence of turbulence (0.6 – 0.7 %) is only one-third of that discussed in the paragraph 2.2 above (i.e. 1.7% shown in Figure 1(a) and 1(b)), the AUC of the harmonized product in the 6-24 hours of forecasts are within the ranges of AUCs for the two WAFC products for 2008-2010. Meanwhile, in the case of moderate turbulence or above for the peak turbulence, the AUC of 24-h forecast drops below 0.5 indicating limited / no skill. However, only three turbulence cases are found for each of the forecast hours. The verification shall be extended to obtain more comprehensive results on the performance of the harmonized products.

2.9 Considering (a) the lack of skill of the gridded CAT forecast from the two WAFC over the W quadrant for the period from October 2008 to November 2010; (b) the limited period after harmonization of the WAFCs forecasts for a full assessment of its performance and (c) the result of the preliminary verification of the "harmonized" forecasts being not particularly promising, it is considered that the proposed deletion of "trial" nature of the gridded forecast products in Amendment 76 to ICAO Annex 3 might be pre-mature.

3. SUMMARY

3.1 From the above discussions, it is observed that:

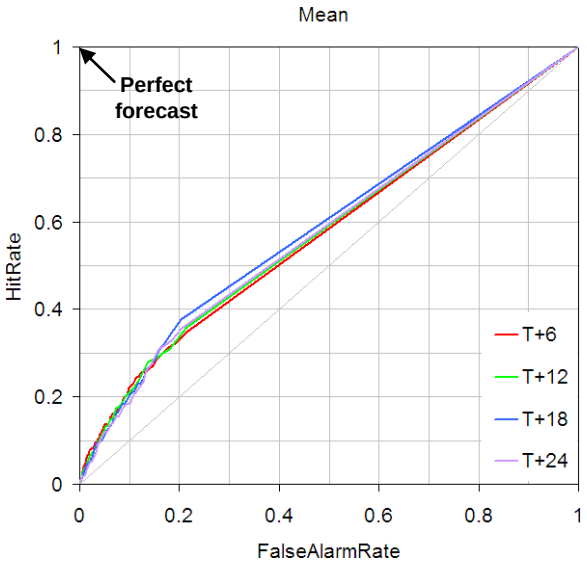
a) The gridded CAT potential forecast products from WAFC London and WAFC Washington show difference in prediction of the average or peak turbulence in short-range (6-24 h) and over different geographical regions in East and Southeast Asia before harmonization;

b) Initial verification of the harmonized products suggests little or no skill for T+24h of forecast. More data to cover a longer verification period is undoubtedly needed for a more comprehensive verification of the performance since the number of turbulence events is rather limited in the present study. The performance of the harmonized product over different geographical region is yet to be determined to quantify the performance of products for different flight routes over diverse topography underneath.

4. ACTION BY THE MEETING

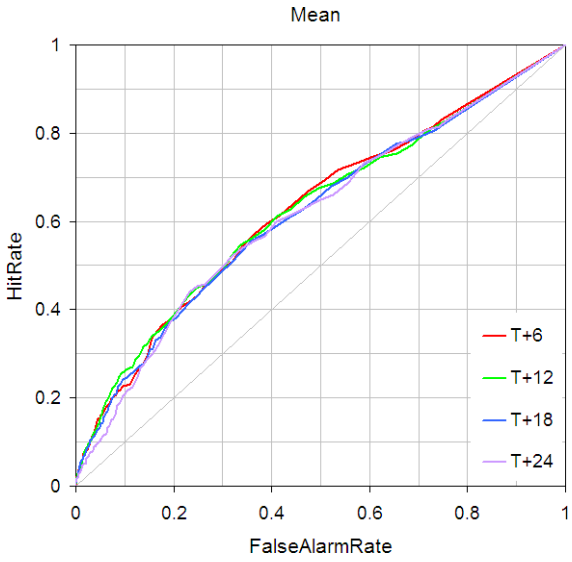
4.1 The meeting is invited to note the information contained in this paper.

– – – – – – – – – – – – –



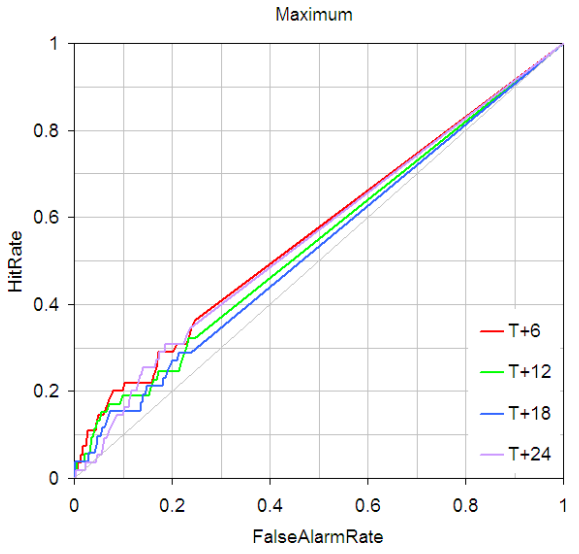
	T+6	T+12	T+18	T+24
D	27646	27414	27810	27744
T	471	466	467	475
F	1.7	1.7	1.7	1.7
A	0.58	0.58	0.59	0.58

(a) UK



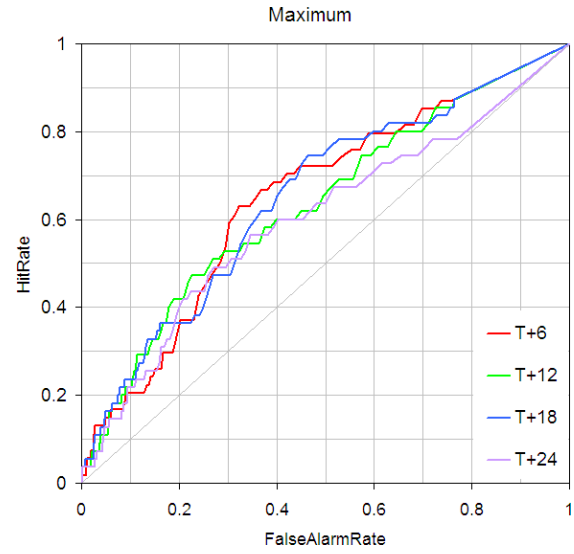
	T+6	T+12	T+18	T+24
D	28307	28403	28079	28295
T	469	469	468	469
F	1.7	1.7	1.7	1.7
A	0.63	0.63	0.62	0.62

(b) US



	T+6	T+12	T+18	T+24
D	27646	27414	27810	27744
T	55	53	52	55
F	0.20	0.19	0.19	0.20
A	0.57	0.54	0.53	0.56

(c) UK



	T+6	T+12	T+18	T+24
D	28307	28403	28079	28295
T	54	55	55	55
F	0.19	0.19	0.20	0.19
A	0.65	0.64	0.65	0.60

(d) US

Fig. 1 ROC curves of mean EDR ≥ 0.1 $m^{2/3}s^{-1}$ (upper) and maximum $EDR^{1/3} \geq 0.4$ $m^{2/3}s^{-1}$ (lower) for UK and US CAT potential forecasts at T+6, 12, 18 and 24 hours. Number of data points, number of turbulent events exceeding the respective $EDR^{1/3}$ thresholds, the observed frequency of turbulence (in %) and the area under the curve are denoted by D, T, F and A respectively.

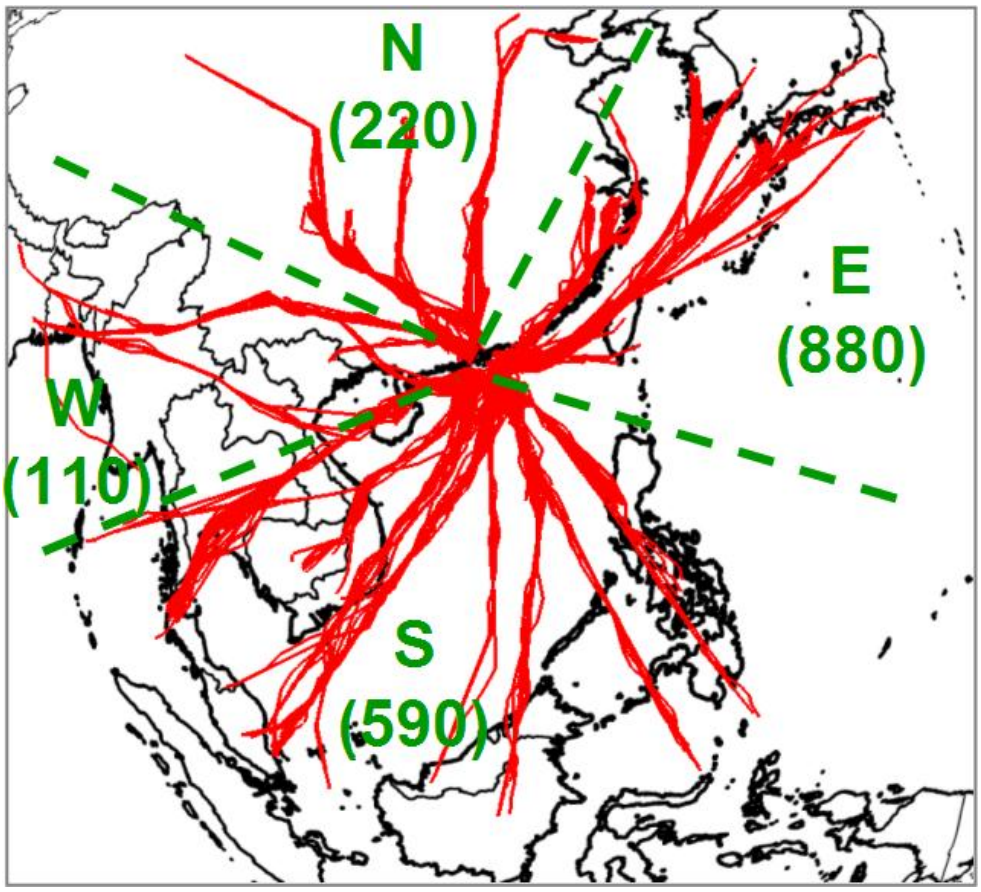
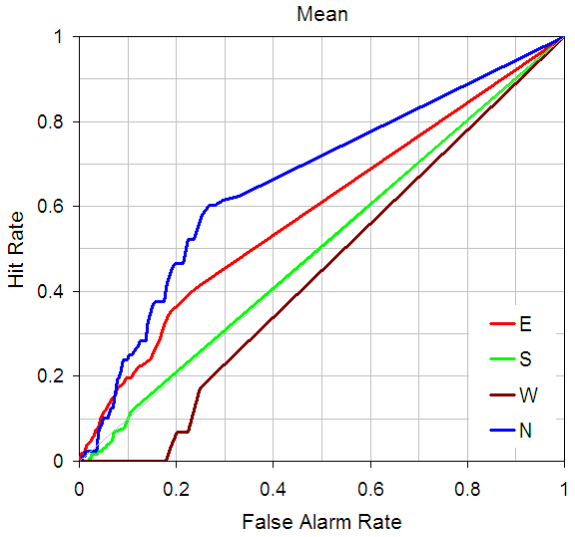
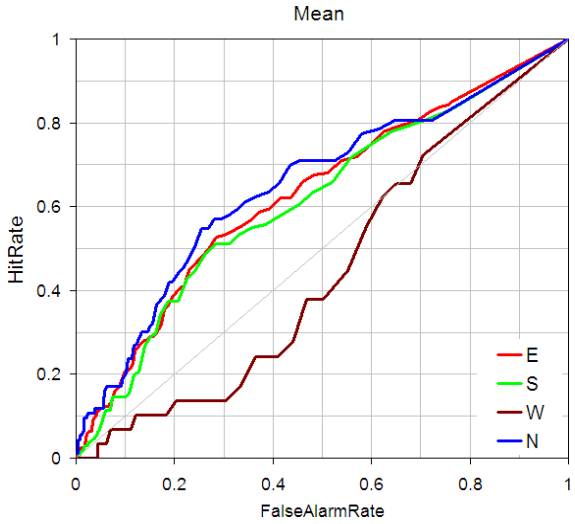


Fig. 2 Distribution of flight routes in the four quadrants N, E, S and W. Number of flights used to calculate the verification statistics (see Fig. 3) are given in brackets.



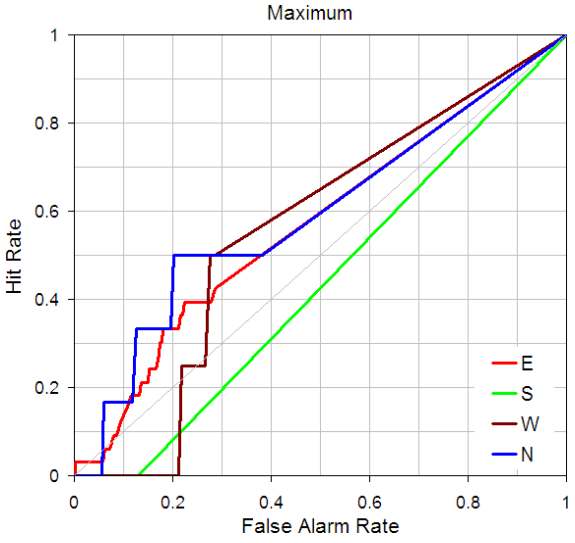
	E	S	W	N
D	11504	8958	1630	3356
T	219	127	29	88
F	1.9	1.4	1.8	2.6
A	0.59	0.50	0.45	0.66

(a) UK



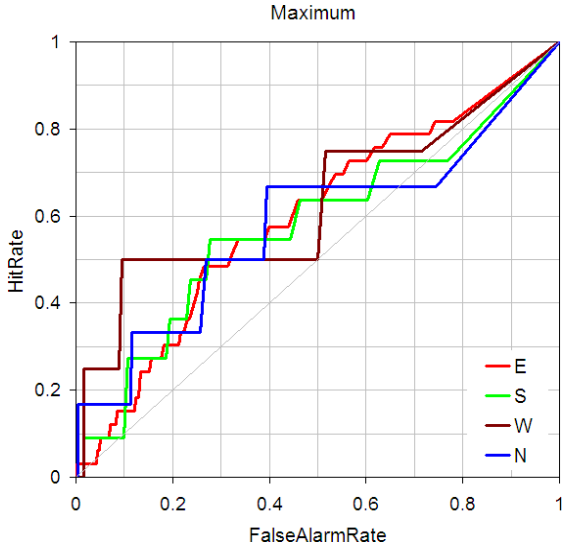
	E	S	W	N
D	11782	9082	1661	3430
T	203	131	29	93
F	1.7	1.4	1.7	2.7
A	0.63	0.61	0.45	0.65

(b) US



	E	S	W	N
D	11504	8958	1630	3356
T	33	11	4	6
F	0.29	0.12	0.25	0.18
A	0.57	0.44	0.56	0.59

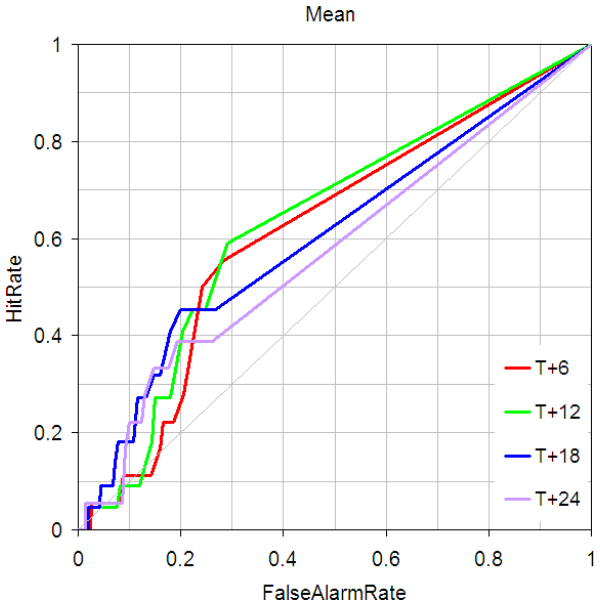
(c) UK



	E	S	W	N
D	11782	9082	1661	3430
T	33	11	4	6
F	0.28	0.12	0.24	0.17
A	0.60	0.58	0.63	0.58

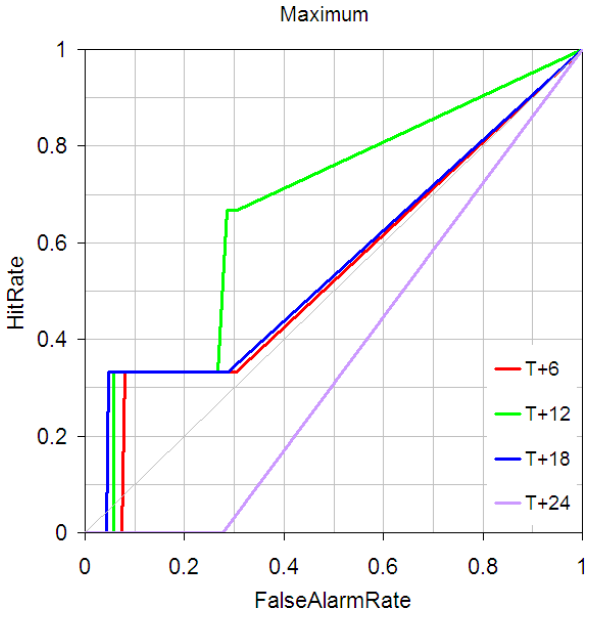
(d) US

Fig. 3 ROC curves of mean EDR $\geq$ 0.1 $m^{2/3}s^{-1}$ (upper) and maximum $EDR^{1/3} \geq 0.4$ $m^{2/3}s^{-1}$ (lower) for the T+24h forecast over the four quadrants of flight routes. Number of data points, number of turbulent events exceeding the respective $EDR^{1/3}$ thresholds, the observed frequency of turbulence (in %) and the area under the curve are denoted by D, T, F and A respectively.



	T+6	T+12	T+18	T+24
D	2908	3137	3114	3011
T	18	22	18	18
F	0.62	0.70	0.58	0.60
A	0.62	0.63	0.60	0.57

(a)



	T+6	T+12	T+18	T+24
D	2908	3137	3114	3011
T	3	3	3	3
F	0.10	0.10	0.10	0.10
A	0.54	0.67	0.56	0.36

(b)

Fig. 4. ROC curves of (a) mean EDR $\geq$ 0.1 $m^{2/3}s^{-1}$ and (b) maximum $EDR^{1/3} \geq 0.4$ $m^{2/3}s^{-1}$ for the harmonized CAT potential forecasts at T+6, 12, 18 and 24 hours. Number of data points, number of turbulent events exceeding the respective $EDR^{1/3}$ thresholds, the observed frequency of turbulence (in %) and the area under the curve are denoted by D, T, F and A respectively.

— END —