A complementary measure to assess temporal uncertainty within Terminal Aerodrome Forecasts

Michael Sharpe and Andre Lanyon, Operational Verification Systems and Products, Met Office UK

IVMW-O, 18th November 2020

Aerodrome Reports and Forecasts Part C (WMO):

Aerodrome forecasts (TAFs) are complete descriptions of the meteorological elements expected at and over the aerodrome throughout the whole of the forecast period, including any changes considered to be significant to aircraft operations

Example:

EGCC 122300Z 1300/1406 30006KT 5000 -RA BKN005
BECMG 1301/1304 9999 NSW SCT012
PROB30 TEMPO 1304/1309 3000 BR BKN003.

Met Office

Change group between 01Z & 04Z:
- vis BECoMinG ≥10,000m (denoted by 9999)
- cloud BECoMinG SCaTtered@1,200ft
- No Significant Weather

30% PROBability of a TEMPOrary change between 04Z & 09Z:
- in visibility to 3,000m
- mist (BRume in French!)
- BRokeN cloud @300ft

Issue day/time

Wind dir & speed

Weather type

Airport ID

Validity period

Initial vis (m)

Cloud amount/ht(ft)

EGCC 122300Z 1300/1406 30006KT 5000 -RA BKN005

BECMG 1301/1304 9999 NSW SCT012

PROB30 TEMPO 1304/1309 3000 BR BKN003.

TAF Term	Forecast Definition
BEGMG	
TEMPO	
PROB *p*	

PROB *p* TEMPO

*when there are no intermediate categories

Met Office

Appendix H. TAF coding quick reference guide.

	ALL TAFs – Direction	ALL TAFs – Speed
WIND	30 degrees or more, mean before or after 20KT or more	10KT change in MEAN speed
	60 degrees or more, mean before or after 10KT or more	10KT change in GUST, the mean speed before or after 15KT or more

	GENERAL TAFs	OFFSHORE OPERATIONS TAFs*
VISIBILITY	10km or more (9999)	10km or more (9999)
	5000m to 9000m	7000m to 9000m
	1500m to 4900m	5000m to 6000m
	800m to 1400m	3000m to 4900m
	350m to 750m	1500m to 2900m
	300m or less	800m to 1400m
		350m to 750m
		300m or less

	ALL TAFs – Begin or end of…	ALL TAFs - Begin, end or change in intensity of…
WEATHER	Thunderstorm (with or without precipitation) (TS)	Freezing precipitation (FZDZ, FZRA)
	Squall (SQ)	Moderate or heavy precipitation, including showers
	Funnel Cloud (tornado or waterspout) (FC)	Duststorm (DS)
	Low drifting snow, dust or sand (DRSN, DRDU, DRSA)	Sandstorm (SS)
	Blowing snow, dust or sand (BLSN, BLDU, BLSA)	Other weather if associated with significant change in visibility or cloud (note, HZ, BR, FU and DU not to be use with visibility in excess of 5000M)
	Freezing fog (FZFG)	
	Ice crystals (IC)	
	CAVOK	

	GENERAL TAFs	OFFSHORE OPERATIONS TAFs*
CLOUD BASE	5000ft or more (NSC/CAVOK)	5000ft or more (NSC/CAVOK)
	1500ft to 4900ft	1500ft to 4900ft
	1000ft to 1400ft	1000ft to 1400ft
	500ft to 900ft	700ft to 900ft
	200ft to 400ft	500ft to 600ft
	100ft or less (including VV///)	200ft to 400ft
	Changes apply to lowest BKN/OVC cloud layer.	100ft or less (including VV///)

	ALL TAFs
CLOUD AMOUNT	When amount of lowest cloud below 1500ft changes from half cover or less (NIL, FEW, SCT) to more than half cover (BKN, OVC), or vice versa.
	The development or dissipation of CB does not, of itself, trigger a cloud change group. Changes must be triggered by application of height/amount thresholds, although CB can then be included in addition.

*OFFSHORE HELICOPTER OPERATIONS TAFs			
	Blackpool EGNH	North Denes EGSD	Kirkwall EGPA
	Humberside EGNJ	Aberdeen EGPD	Scatsta EGPM
	Liverpool EGGP	Benbecula EGPL	Sumburgh EGPB
	Norwich EGSH	Inverness EGPE	Wick EGPC

Forecasters use a categorical quick-reference threshold guide reflecting:

- ICAO categories
- WMO Manual

Met Office

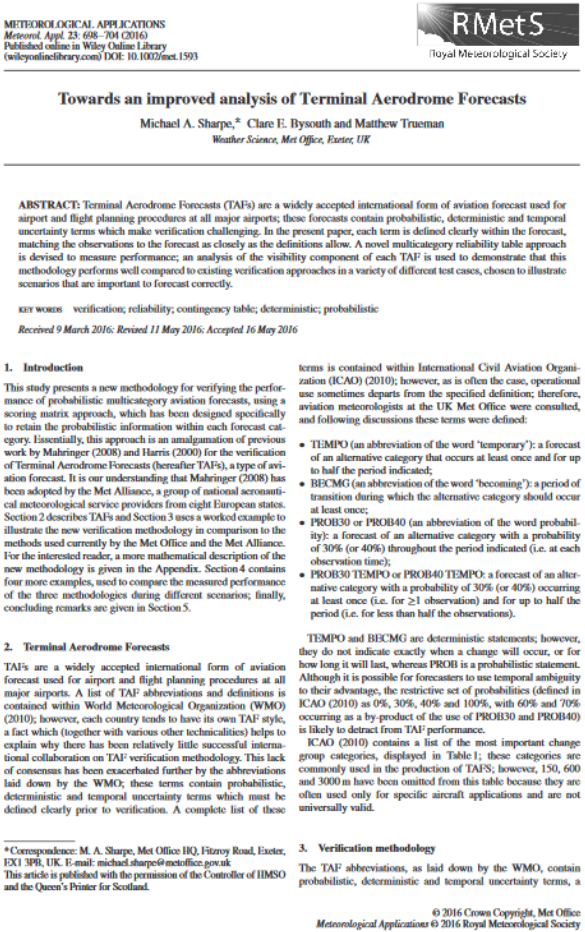
New verification methodology

- Published 2016
- Presented at IVMW 2017
- Now operational at Met Office

Uses:

- WMO definitions to interpret TAF terms
- Multiple categories
- Probability information
- Variety of scores

Why do we need another verification methodology?



METEOROLOGICAL APPLICATIONS
Meteorol. Appl. 23: 698–704 (2016)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/met.1593

RMetS
Royal Meteorological Society

Towards an improved analysis of Terminal Aerodrome Forecasts

Michael A. Sharpe,* Clare E. Bysouth and Matthew Trueman
Weather Science, Met Office, Exeter, UK

ABSTRACT: Terminal Aerodrome Forecasts (TAFs) are a widely accepted international form of aviation forecast used for airport and flight planning procedures at all major airports; these forecasts contain probabilistic, deterministic and temporal uncertainty terms which make verification challenging. In the present paper, each term is defined clearly within the forecast, matching the observations to the forecast as closely as the definitions allow. A novel multicategory reliability table approach is devised to measure performance; an analysis of the visibility component of each TAF is used to demonstrate that this methodology performs well compared to existing verification approaches in a variety of different test cases, chosen to illustrate scenarios that are important to forecast correctly.

KEY WORDS verification; reliability; contingency table; deterministic; probabilistic

Received 9 March 2016; Revised 11 May 2016; Accepted 16 May 2016

1. Introduction

This study presents a new methodology for verifying the performance of probabilistic multicategory aviation forecasts, using a scoring matrix approach, which has been designed specifically to retain the probabilistic information within each forecast category. Essentially, this approach is an amalgamation of previous work by Mahringer (2008) and Harris (2000) for the verification of Terminal Aerodrome Forecasts (hereafter TAFs), a type of aviation forecast. It is our understanding that Mahringer (2008) has been adopted by the Met Alliance, a group of national aeronautical meteorological service providers from eight European states. Section 2 describes TAFs and Section 3 uses a worked example to illustrate the new verification methodology in comparison to the methods used currently by the Met Office and the Met Alliance. For the interested reader, a more mathematical description of the new methodology is given in the Appendix. Section 4 contains four more examples, used to compare the measured performance of the three methodologies during different scenarios; finally, concluding remarks are given in Section 5.

2. Terminal Aerodrome Forecasts

TAFs are a widely accepted international form of aviation forecast used for airport and flight planning procedures at all major airports. A list of TAF abbreviations and definitions is contained within World Meteorological Organization (WMO) (2010); however, each country tends to have its own TAF style, a fact which (together with various other technicalities) helps to explain why there has been relatively little successful international collaboration on TAF verification methodology. This lack of consensus has been exacerbated further by the abbreviations laid down by the WMO; these terms contain probabilistic, deterministic and temporal uncertainty terms which must be defined clearly prior to verification. A complete list of these terms is contained within International Civil Aviation Organization (ICAO) (2010); however, as is often the case, operational use sometimes departs from the specified definition; therefore, aviation meteorologists at the UK Met Office were consulted, and following discussions these terms were defined:

- TEMPO (an abbreviation of the word 'temporary'): a forecast of an alternative category that occurs at least once and for up to half the period indicated;
- BECMG (an abbreviation of the word "becoming"): a period of transition during which the alternative category should occur at least once;
- PROB30 or PROB40 (an abbreviation of the word probability): a forecast of an alternative category with a probability of 30% (or 40%) throughout the period indicated (i.e. at each observation time);
- PROB30 TEMPO or PROB40 TEMPO: a forecast of an alternative category with a probability of 30% (or 40%) occurring at least once (i.e. for ≥1 observation) and for up to half the period (i.e. for less than half the observations).

TEMPO and BECMG are deterministic statements; however, they do not indicate exactly when a change will occur, or for how long it will last, whereas PROB is a probabilistic statement. Although it is possible for forecasters to use temporal ambiguity to their advantage, the restrictive set of probabilities (defined in ICAO (2010) as 0%, 30%, 40% and 100%, with 60% and 70% occurring as a by-product of the use of PROB30 and PROB40) is likely to detract from TAF performance.

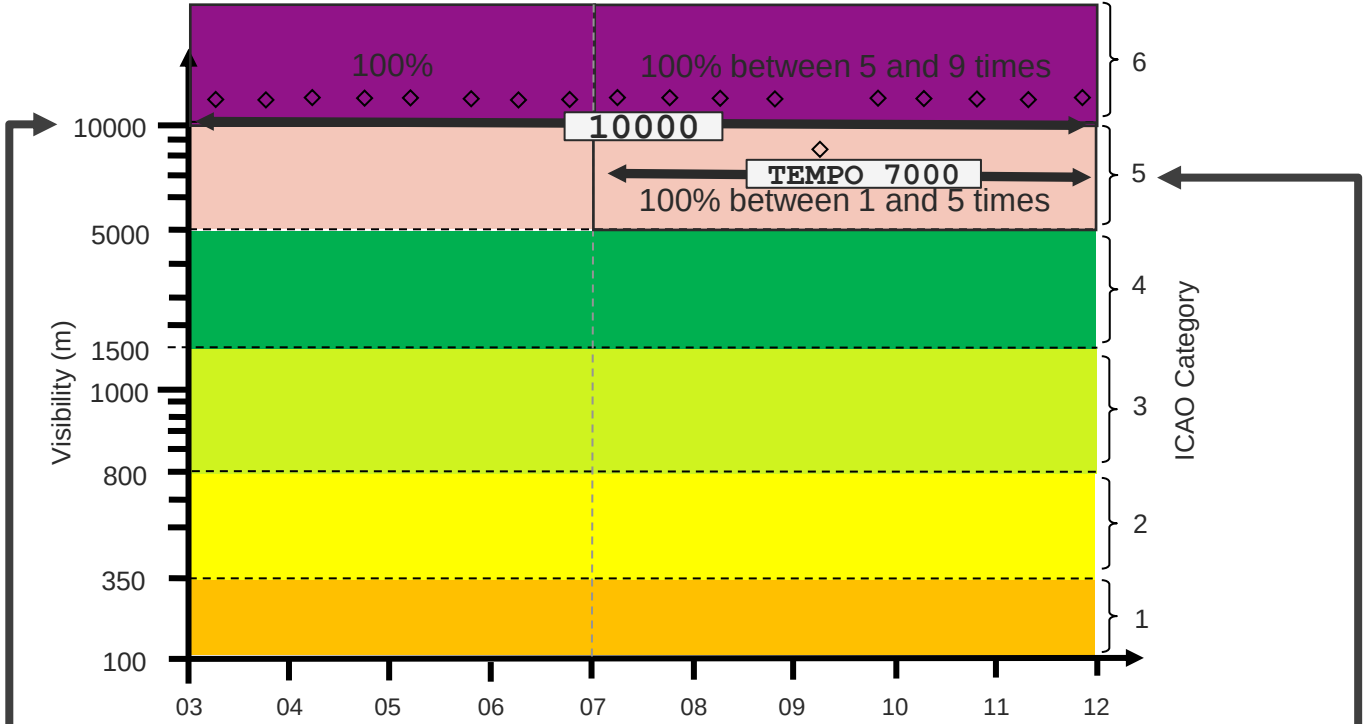
ICAO (2010) contains a list of the most important change group categories, displayed in Table 1; these categories are commonly used in the production of TAFS; however, 150, 600 and 3000 m have been omitted from this table because they are often used only for specific aircraft applications and are not universally valid.

*Correspondence: M. A. Sharpe, Met Office HQ, Fitzroy Road, Exeter, EX1 3PB, UK. E-mail: michael.sharpe@metoffice.gov.uk

3. Verification methodology

The TAF abbreviations, as laid down by the WMO, contain probabilistic, deterministic and temporal uncertainty terms, a

WMO TAFs rules can be hedged!

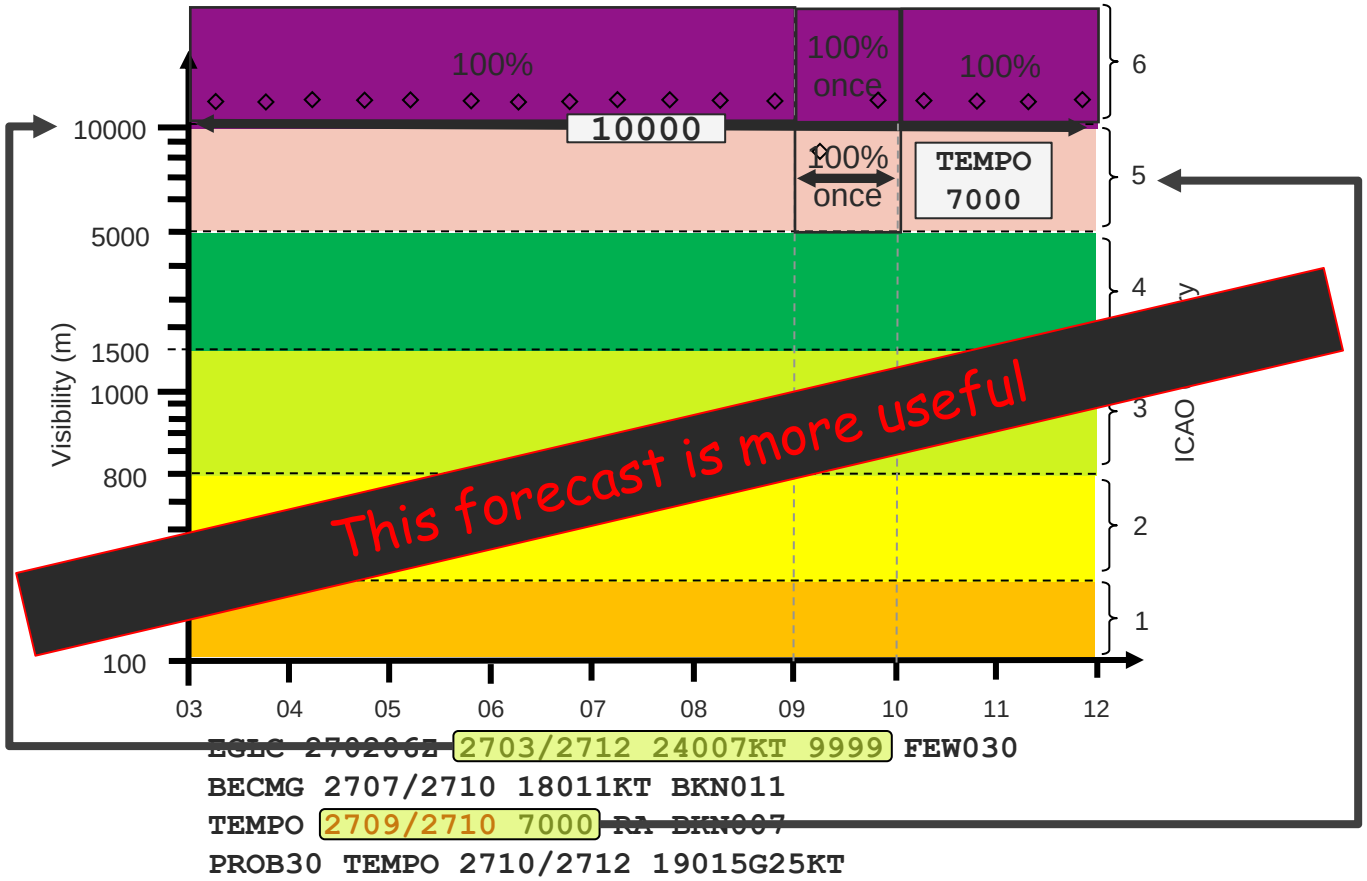


~~EGLC 2702062~~ 2703/2712 24007KT 9999 FEW030
BECMG 2707/2710 18011KT BKN011
TEMPO 2707/2712 7000 ~~RA BKN007~~
PROB30 TEMPO 2710/2712 19015G25KT

A PERFECT SCORE!

Met Office

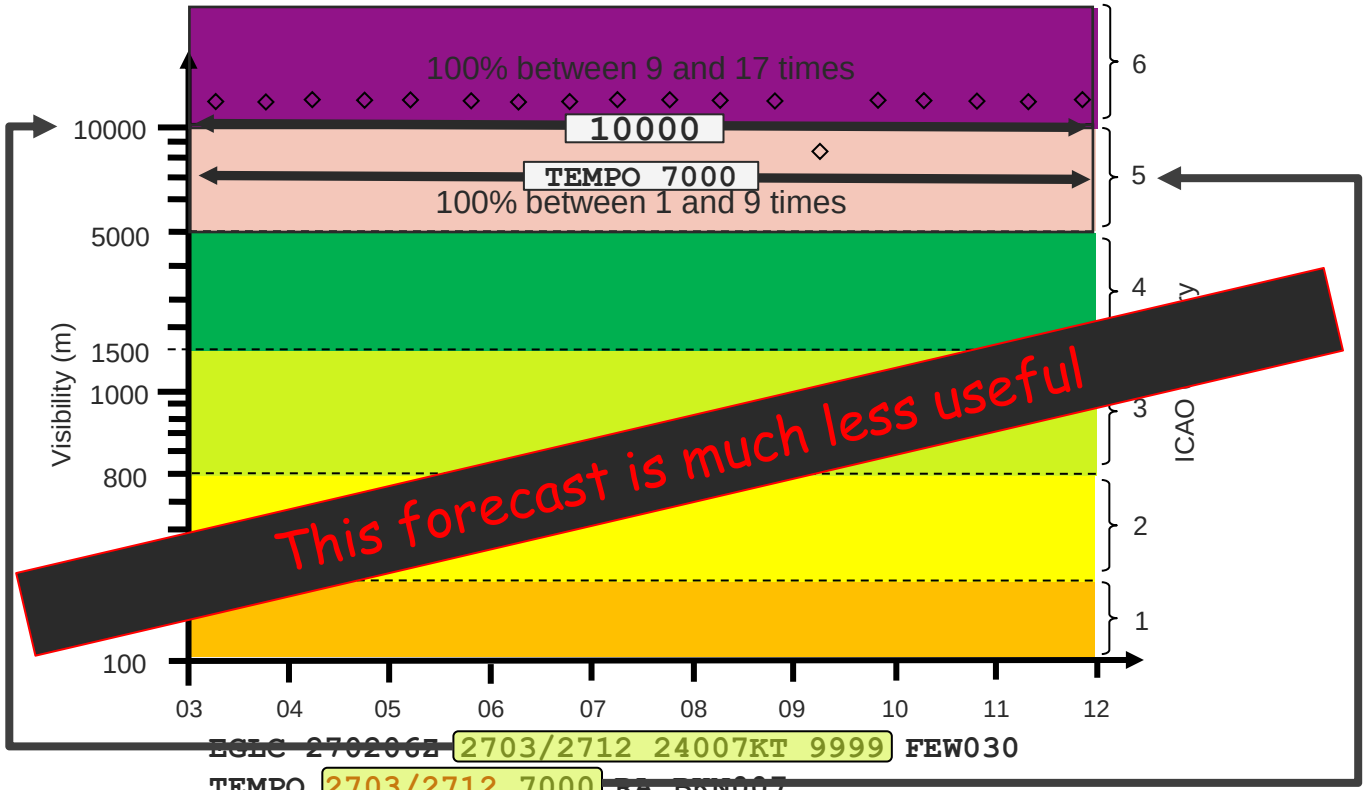
WMO TAFs rules can be hedged!



~~EGLC 2702062~~ 2703/2712 24007KT 9999 FEW030
BECMG 2707/2710 18011KT BKN011
TEMPO 2709/2710 7000 ~~RA BKN007~~
PROB30 TEMPO 2710/2712 19015G25KT

A PERFECT SCORE!

WMO TAFs rules can be hedged!

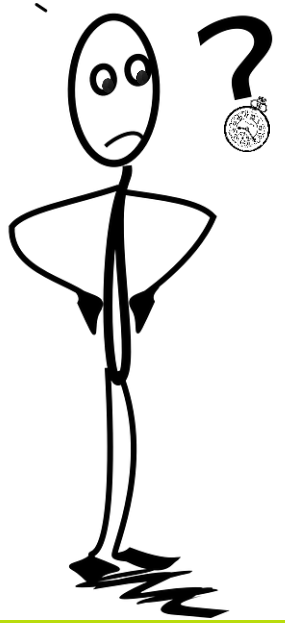


~~EGLC 2702062~~ 2703/2712 24007KT 9999 FEW030
TEMPO 2703/2712 7000 ~~RA BKN007~~
~~BECMG 2707/2710 18011KT BKN011~~
~~PROB30 TEMPO 2710/2712 19015G25KT~~

A PERFECT SCORE!

WMO TAFs rules can be hedged!

- WMO definitions do not penalise temporal uncertainty
- Affects:
  - all **`(PROB)TEMPO`** groups
  - all **`BECMG`** groups
  - some **`PROB`** ***`p`*** groups
- These terms are **NOT WRONG**
  - using them should **NOT** be penalised
  - but ***excessive*** use can indicate hedging
- So a complementary *uncertainty-penalising* score is required
  - to measure uncertainty within TAFs
  - achieved by interpreting all TAF terms as probabilities



Met Office

TEMPO (‘temporary’):

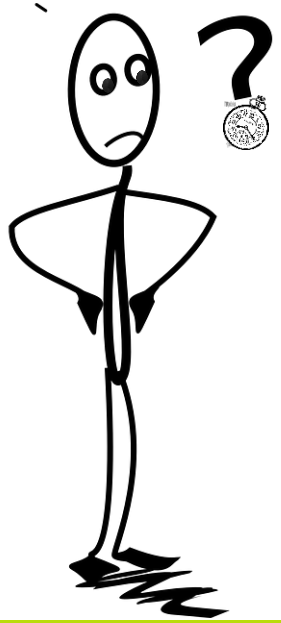
- alternative category is forecast *at least* once and *up to half* the period

If **TEMPO** period is n-hours

- there are $2n$ observations in total
- $[n, 2n)$ are forecast $\in M$ (main category)
- $[1, n]$ are forecast $\in A$ (alternative category)

If $P(O_j^i \in M)$ = probability j^{th} observation $\in M \mid M$ observed i times:

- $P\left(O_j^n \in M\right) = n/2n$
- $P\left(O_j^{n+1} \in M\right) = (n+1)/2n$
- …
- $P\left(O_j^{2n-1} \in M\right) = (2n-1)/2n$



Met Office

TEMPO (‘temporary’):

- alternative category is forecast *at least* once and *up to half* the period

If $P(O_j^i \in M)$ = probability j^{th} observation $\in M \mid M$ observed i times:

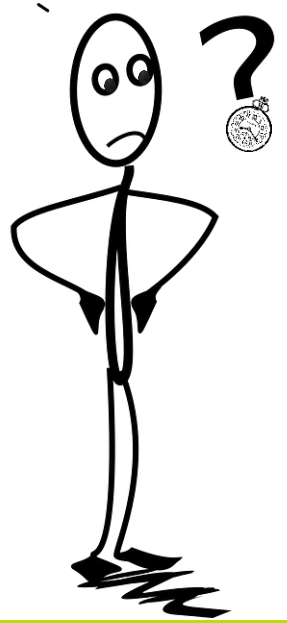
- $P(O_j^n \in M) = n/2n$
- $P(O_j^{n+1} \in M) = (n+1)/2n$
- …
- $P(O_j^{2n-1} \in M) = (2n-1)/2n$

Assuming all are equaly likely

$$P(O_j \in M) = 1/n\left[P(O_j^n \in M) + P(O_j^{n+1} \in M) + \cdots + P(O_j^{2n-1} \in M)\right]$$
$$= (3n-1)/4n$$

An *arithemic progression* − remember those?!

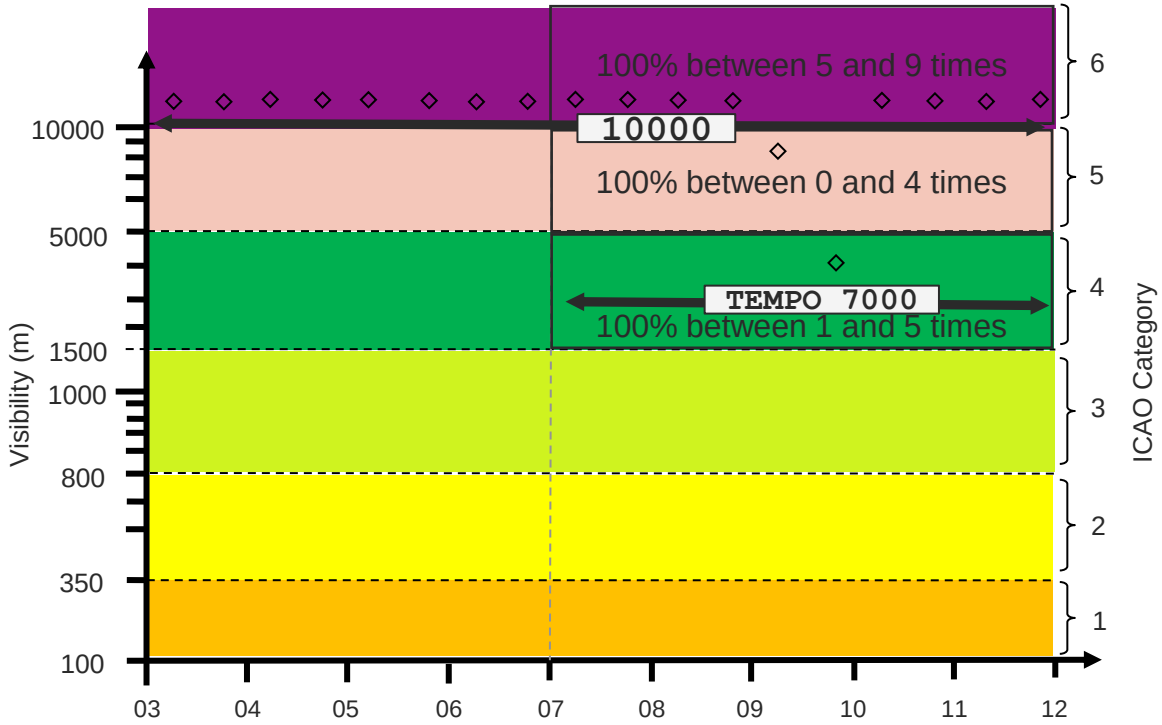
So $P(O_j \notin M) = (n+1)/4n$

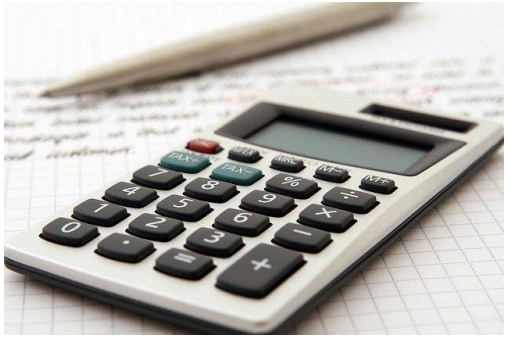


Met Office

TEMPO (‘temporary’):

- alternative category is forecast *at least* once and *up to half* the period





But, observations are continuous so:

- if A (alternative category) occurs
- intermediate categories (I) should not be penalised

So, from 07Z to 15Z the forecast is:

- Purple: 5 to 9 times
- Pink: 0 to 4 times
- Green: 1 to 5 times

TEMPO (‘temporary’):

- alternative category is forecast *at least* once and *up to half* the period

$$P(O_j \in M) = (3n-1)/4n$$

$$P(O_j \in A) = \begin{cases} (n+1)/4n & if\ I = \phi \\ (n+3)/8n & if\ I \neq \phi \end{cases}$$

$$P(O_j \in I) = (n-1)/8n$$

PROB *p* TEMPO (‘temporary’):

- alternative category is forecast with *p%* probability *at least* once and *up to half* the period

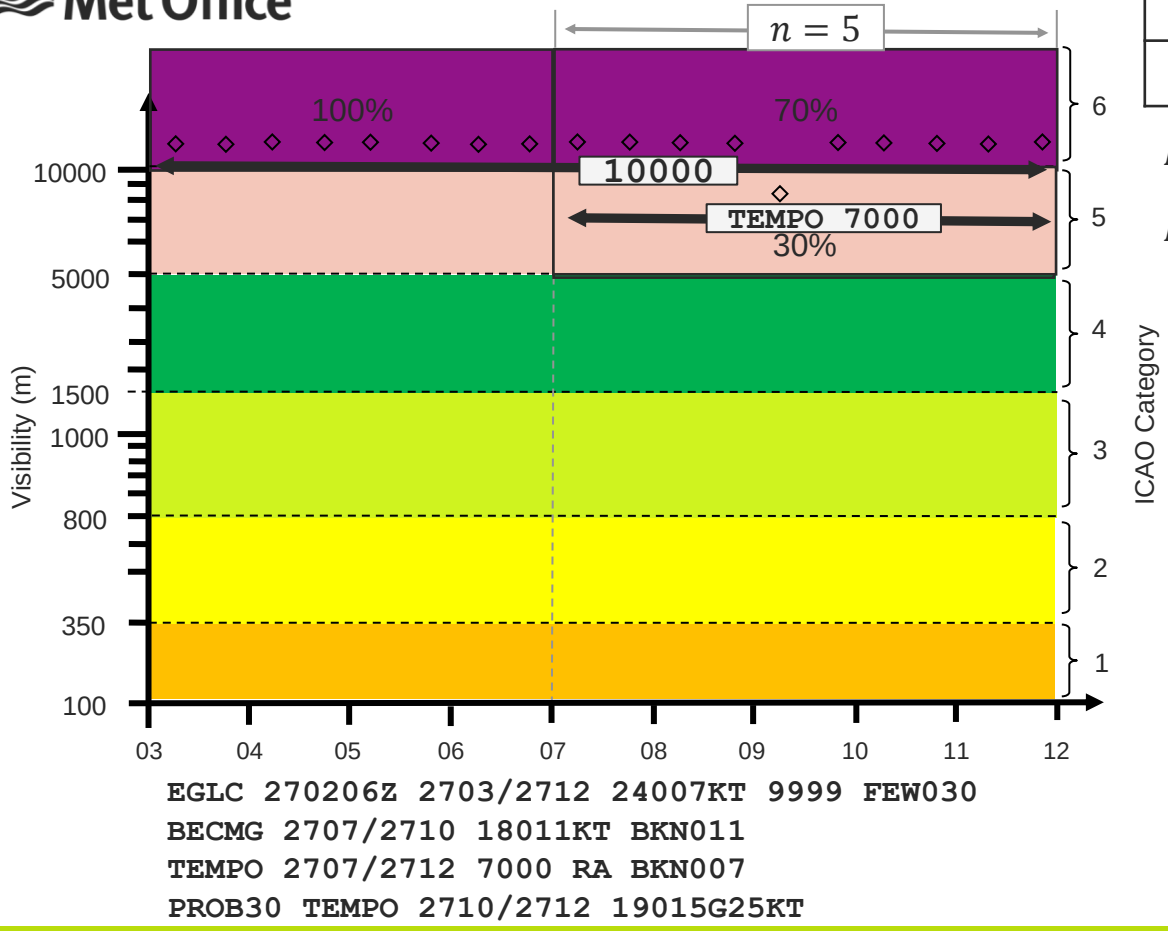
$$P(O_j \in M) = 1 - p(n+1)/4n$$

$$P(O_j \in A) = \begin{cases} p(n+1)/4n & if\ I = \phi \\ p(n+3)/8n & if\ I \neq \phi \end{cases}$$

$$P(O_j \in I) = p(n-1)/8n$$

Similar method used to derive expressions for:

- **BECMG** (‘becoming’)
- **PROB *p***



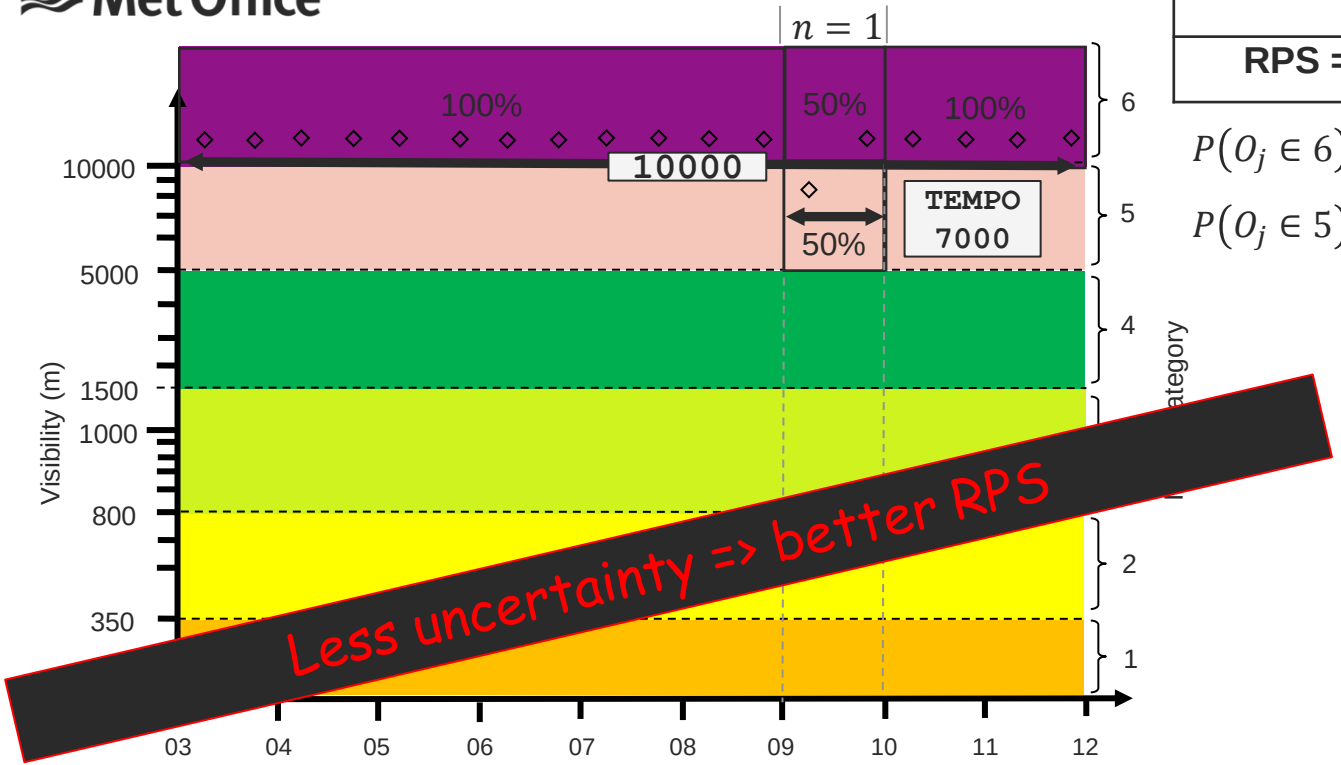
EGLC 270206Z 2703/2712 24007KT 9999 FEW030
BECMG 2707/2710 18011KT BKN011
TEMPO 2707/2712 7000 RA BKN007
PROB30 TEMPO 2710/2712 19015G25KT

Main measure	Uncertainty – penalising measure
RPS = 0.0	RPS = 0.07

$$P(O_j \in 6) = (3 \times 5 - 1)/4 \times 5 = 0.7$$

$$P(O_j \in 5) = (5 + 3)/8 \times 5 \quad = 0.3$$

Met Office

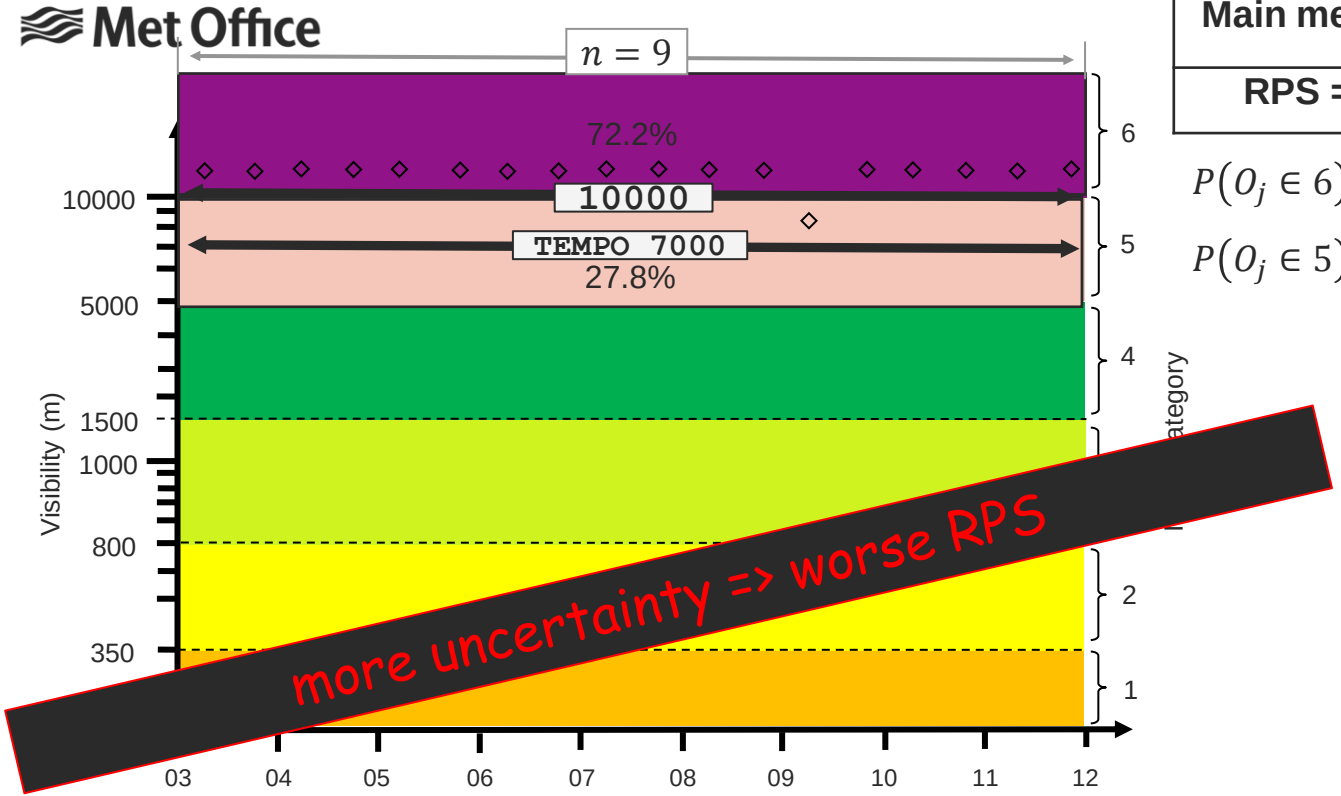


Main measure	Uncertainty – penalising measure
RPS = 0.0	RPS = 0.03

$$P(O_j \in 6) = (3 \times 1 - 1)/4 \times 1 = 0.5$$

$$P(O_j \in 5) = (1 + 3)/8 \times 1 \quad = 0.5$$

EGLC 270206Z 2703/2712 24007KT 9999 FEW030
BECMG 2707/2710 18011KT BKN011
TEMPO 2709/2710 7000 RA BKN007
PROB30 TEMPO 2710/2712 19015G25KT



Main measure	Uncertainty – penalising measure
RPS = 0.0	RPS = 0.1

$$P(O_j \in 6) = (3 \times 9 - 1)/4 \times 9 = 0.722$$

$$P(O_j \in 5) = (9 + 3)/8 \times 9 \quad = 0.278$$

```
EGLC 270206Z 2703/2712 24007KT 9999 FEW030
TEMPO 2703/2712 7000 RA BKN007
BECMG 2707/2710 18011KT BKN011
PROB30 TEMPO 2710/2712 19015G25KT
```

Summary

- New methodology* reflects TAF performance according to WMO rules
- WMO rules allow a canny forecaster to hedge without penalty
- Cannot modify the methodology to account for this
- Need a secondary score:
  - **Uncertainty-penalising measure**
  - Monitors uncertainty within TAFs
  - Identifies airports where TAFs are particularly uncertain
  - Mis-interprets the TAF
  - By interpreting TAF terms as categorical probabilities

} So, it should not be the main measure

*Sharpe *et al* (2016)

Thank you

Michael Sharpe and Andre Lanyon, Operational Verification Systems and Products, Met Office UK

IVMW-O, 18th November 2020