The Role of Meteorological Forecast Verification in Aviation

Günter Mahringer, November 2012

Introduction

Aviation Meteorology is internationally regulated. Services are standardized and harmonized by the International Civil Aviation Organization (ICAO).

Core Standards and recommended practices for MET Services are described in Annex 3 to the Convention on International Civil Aviation: Meteorological Service for International Air Navigation.

Annex 3, 2.2.3.: From 15 November 2012, each contracting state shall ensure that the designated meteorological authority establishes and implements a properly organized quality system …

Quality management is also required by European legislation. Certification according to Single European Sky (SES) regulations is required for each organization providing air navigation services including meteorological services for aviation.

Introduction

ICAO Met Group for Europe and North Atlantic (METG) has defined recommendations for Key Performance Indicators (KPIs).

KPIs should be used by regulators, National Supervisory Authorities (NSA) and service providers to ensure:

(a) Aviation forecasts are skillful … and comply with ICAO standards and recommended procedures,

(b) Aviation observations comply with ICAO standards and recommended procedures,

(c) Production resilience,

(d) Resource is effectively directed at improving the process,

(e) Changes to observation, forecast and dissemination processes are justified, and

(f) Maintain customer confidence in aviation MET services.

Meteorological Forecasts for Aviation

Standard meteorological forecasts for international aviation, provided by designated meteorological authorities:

- **Aerodrome forecasts (TAF, 9 – 30 hrs)**
- **Landing forecasts (TREND, 2 hrs)**
- Forecasts for take-off
- Area forecasts for low-level flights

Warnings for international aviation:

- **SIGMET, AIRMET (en-route)**
- **Aerodrome warnings**
- **Wind shear warnings**

World Area Forecast System (provided by designated centres):

- **wind and temperature charts**
- **Significant weather charts**

Other products in agreement with users (e.g. Air Traffic Management)

Verification of TAFs

The METG identifies the Terminal Aerodrome Forecast (TAF) as the key aerodrome-related Met product for aviation.

TAFS are

- **Produced for all international airports**
- **Widely used for flight planning by operators → economic impact**
- **(Fairly) standardized worldwide**
- **Quality of TAF is probably linked to other products for low levels like TREND, Aerodrome warnings, and low level area forecasts**

The TAF was probably the first specific aviation Met product for which verification projects were undertaken.

Objective verification is possible because

- **TAFs contain defined values for defined times**
- **Observational reference is easily available (METARs)**

TAF Verification in the Met Alliance

The Met Alliance is a cooperation between the Aviation Met Services of: CH, B, NL, IRL, A, D, F, L

As old TAF verification methods were not accepted by forecasters, TAF Verification was redesigned.

A common standard has been created which is currently used in 7 Met Alliance countries and by additional customers.

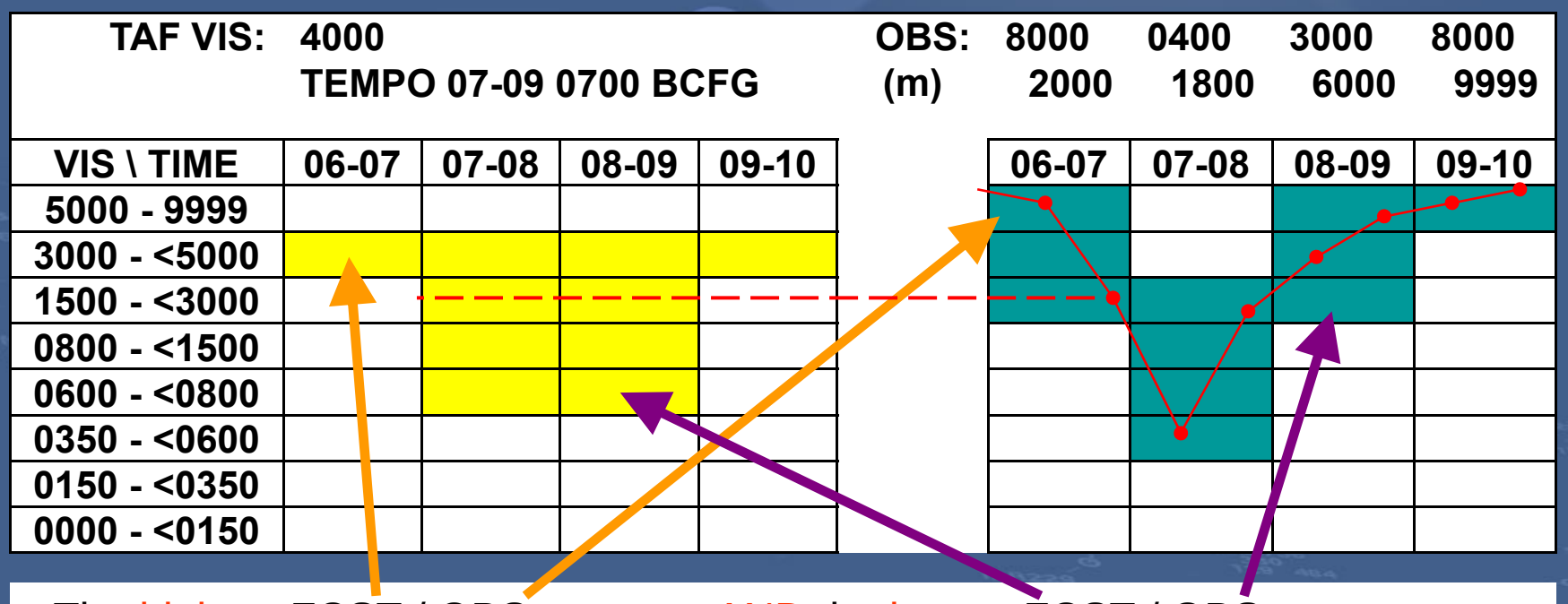
The Verification method is based on observed / forecast ranges and time intervals.

This approach avoids problems with assigning probabilities to simultaneously valid forecast conditions.

TAF Verification Method

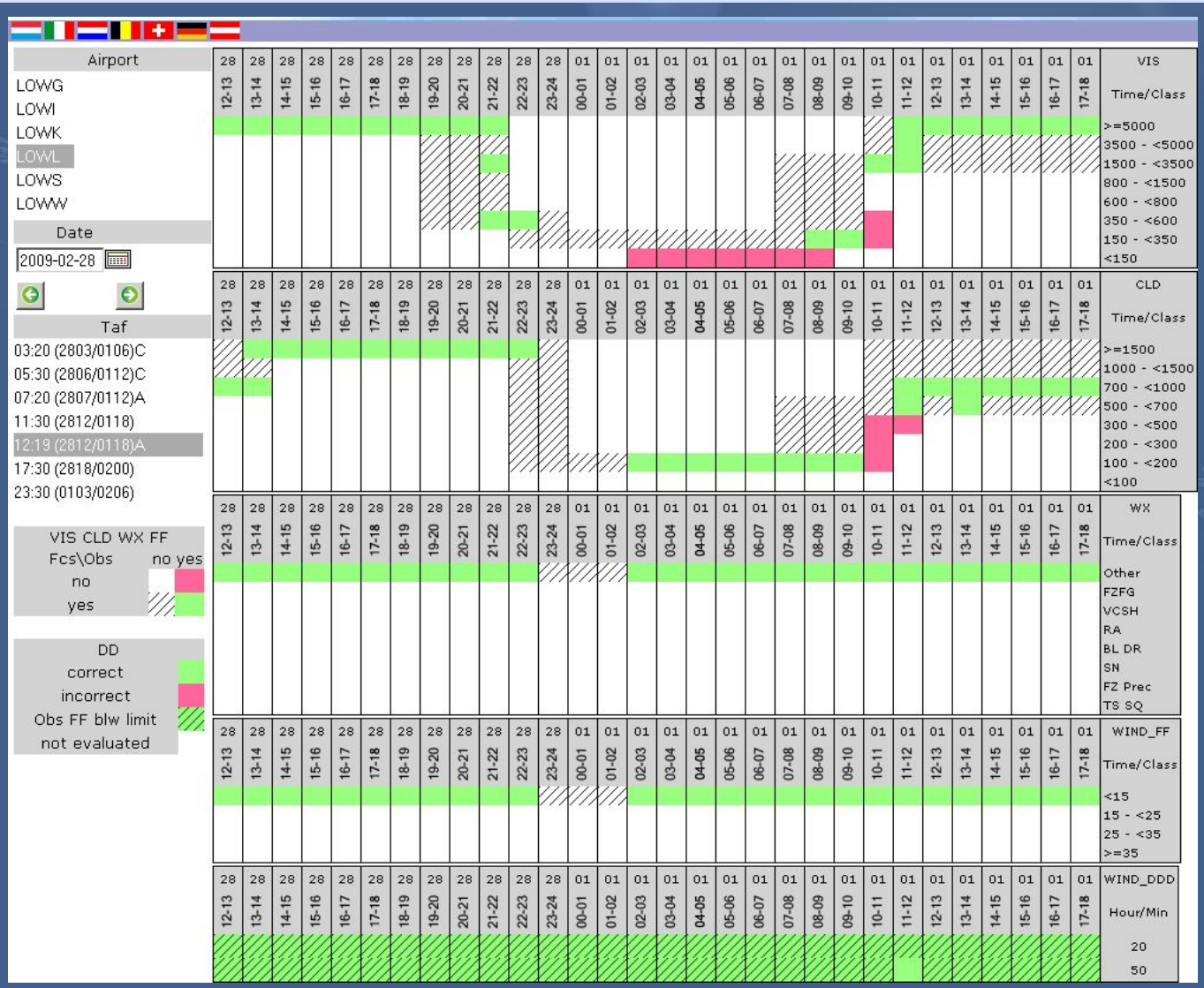
TAF VIS: 4000 TEMPO 07-09 0700 BCFG	06-07	07-08	08-09	09-10
5000 - 9999				
3000 - <5000				
1500 - <3000				
0800 - <1500				
0600 - <0800				
0350 - <0600				
0150 - <0350				
0000 - <0150				

OBS: 8000 0400 3000 8000 (m) 2000 1800 6000 9999	06-07	07-08	08-09	09-10
5000 - 9999				
3000 - <5000				
1500 - <3000				
0800 - <1500				
0600 - <0800				
0350 - <0600				
0150 - <0350				
0000 - <0150				



- The highest FCST / OBS category AND the lowest FCST / OBS category are verified for each hour.
- The FORECAST RANGE is verified based on the OBSERVED RANGE (instead of a SINGLE OBSERVATION value)
- PROB.. groups may be included or excluded.

Result displays (1): Online Verification (→ forecasters)



Result displays (2): Contingency Tables (→ Forecasters)

LOWG VISIBILITY - Maxima over all FCST hours, Period 2008 11 10 – 2009 03 31

FCST \ OBS	<150	150 - <350	350 - <600	600 - <800	800 - <1500	1500 - <3500	3500 - <5000	≥5000	SUM
<150	7	7	2	3	3	10	6	25	63
150 - <350	20	41	11	14	4	32	24	79	225
350 - <600	5	42	8	5	3	7	7	49	126
600 - <800	0	0	0	0	0	0	0	0	0
800 - <1500	2	7	5	6	2	9	6	74	111
1500 - <3500	3	35	24	10	26	207	107	188	600
3500 - <5000	6	81	40	8	45	422	250	637	1489
≥5000	15	140	75	27	92	522	754	9606	11231
SUM	58	353	165	73	175	1209	1154	10658	13845

LOWG VISIBILITY - Minima over all FCST hours, Period 2008 11 10 – 2009 03 31

FCST \ OBS	<150	150 - <350	350 - <600	600 - <800	800 - <1500	1500 - <3500	3500 - <5000	≥5000	SUM
<150	79	127	72	10	17	83	50	146	584
150 - <350	25	192	86	23	52	180	90	459	1107
350 - <600	0	23	38	4	15	63	60	149	352
600 - <800	0	0	0	2	4	3	2	17	28
800 - <1500	1	13	23	4	62	216	74	215	608
1500 - <3500	3	49	53	19	86	550	336	790	1885
3500 - <5000	11	50	36	10	56	256	241	1015	1675
≥5000	14	77	53	9	67	244	216	6926	7606
SUM	133	531	361	81	359	1595	1068	9717	13845

Many misses in Maxima, moderate rate of misses in Minima for low visibilities

Hit rate ~60% for low visibilities (Minima)

Many false alarms for low visibilities (Max, Min)

Maximum visibility is regarded less important for flight operations than minimum visibility.

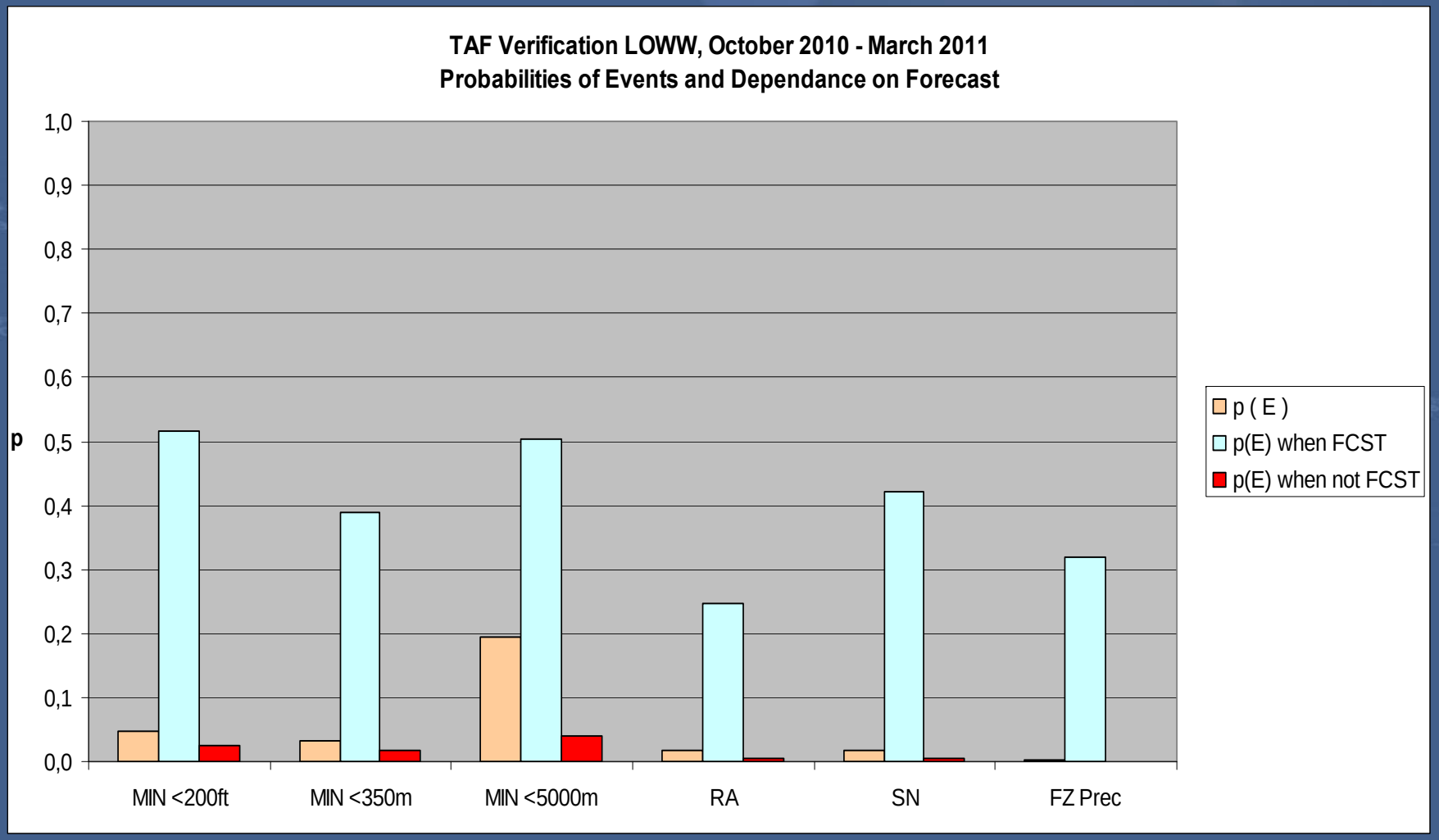
Forecasters are cautious not to miss events of visibility reductions.

At long forecast ranges (up to 30 hours), missed events are hard to avoid.

Bias: low Min visibilities are forecasted too often, mostly by TEMPO, PROB

Many false forecasts of "slight" visibility reductions

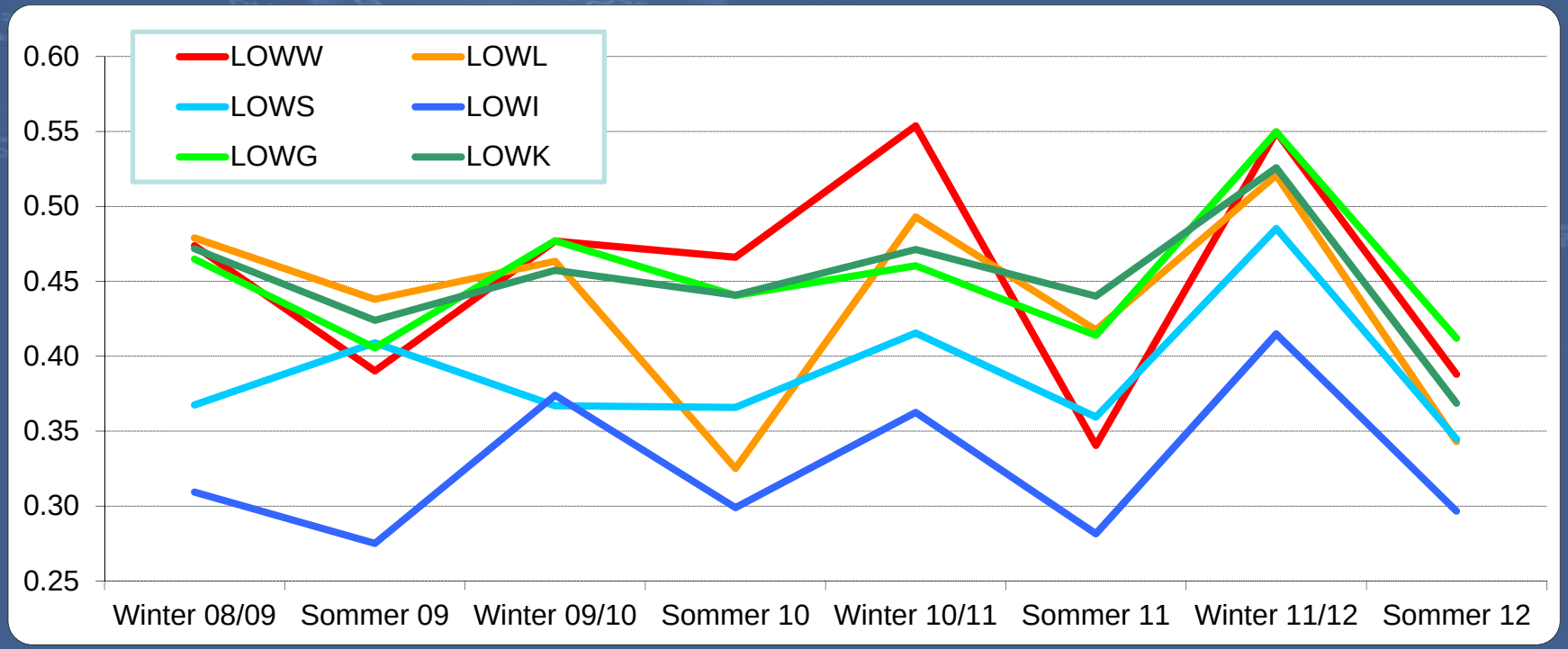
Result displays (3): Reliability of TAFs (→ customers)



Result displays (4): Summary Scores (→ management)

Important:

- **Score should be proper, and**
- **increasing the score should be desirable.**



Problems with “% correct” type of scores which are asked for by ICAO Annex 3 and therefore also by management (WMO knows ;-)

Comparison TAF - AUTOTAF

FCST \ OBS	<150	150- <350	350- <600	600- <800	800- <1500	1500- <3500	3500-<5000	>=5000	SUM
<150	16	14	2	0	0	9	6	23	70
150 - <350	2	17	3	7	3	9	5	157	203
350 - <600	0	0	0	0	1	0	0	9	10
600 - <800	0	0	0	0	0	0	0	2	2
800 - <1500	0	0	0	0	2	7	5	65	79
1500 - <3500	0	1	0	0	4	4	19	152	180
3500 - <5000	0	1	0	3	0	21	25	554	604
>=5000	0	1	0	5	11	19	51	9788	9875
SUM	18	34	5	15	21	69	111	10750	11023

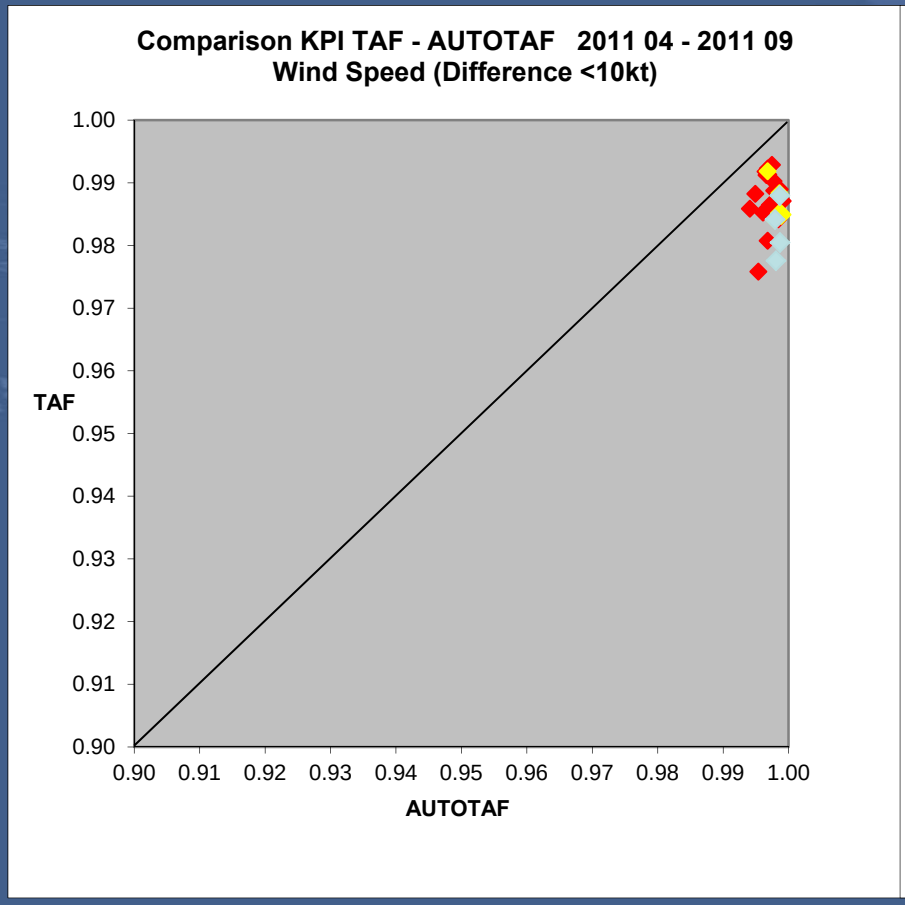
TAF:
+ Very high hit rate especially in low VIS
+ Less missed events
- More false alarms
- Less "correct negatives"

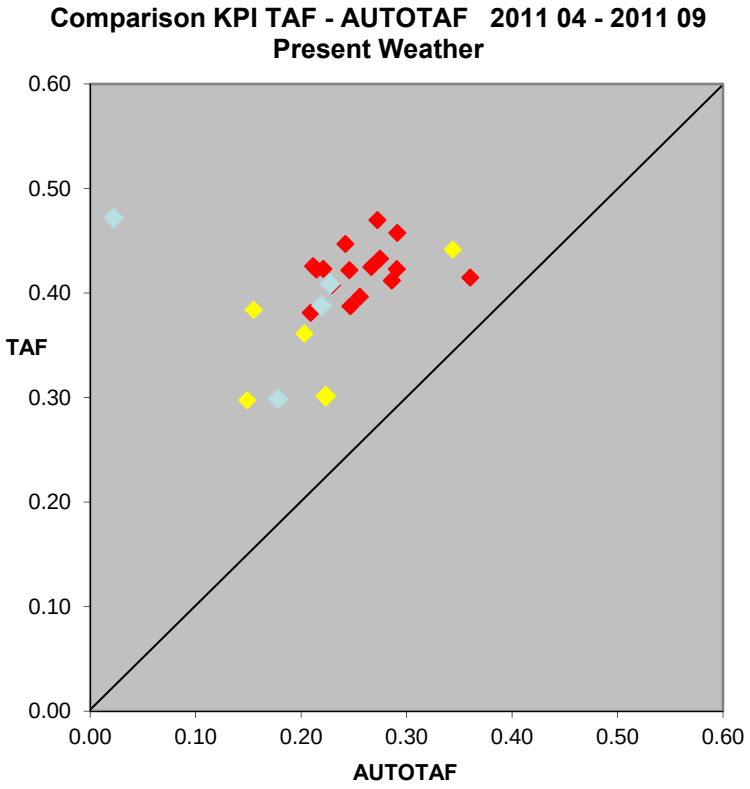
Above: TAF LOWX Below: AUTOTAF LOWX

FCST \ OBS	<150	150- <350	350- <600	600- <800	800- <1500	1500- <3500	3500- <5000	>=5000	SUM
<150	2	0	0	0	1	1	1	1	6
150 - <350	6	13	1	1	2	6	5	49	83
350 - <600	2	3	1	0	1	3	1	35	46
600 - <800	0	1	0	0	0	2	3	14	20
800 - <1500	4	2	1	1	0	3	5	39	55
1500 - <3500	0	2	1	0	0	4	3	113	123
3500 - <5000	1	1	0	1	0	7	27	281	318
>=5000	3	14	2	12	20	52	69	10299	10471
SUM	18	36	6	15	24	78	114	10831	11122

AUTOTAF:
+ Better Bias
+ Less false alarms
+ More "correct negatives"
- Lower hit rates
- Many "by far" missed events

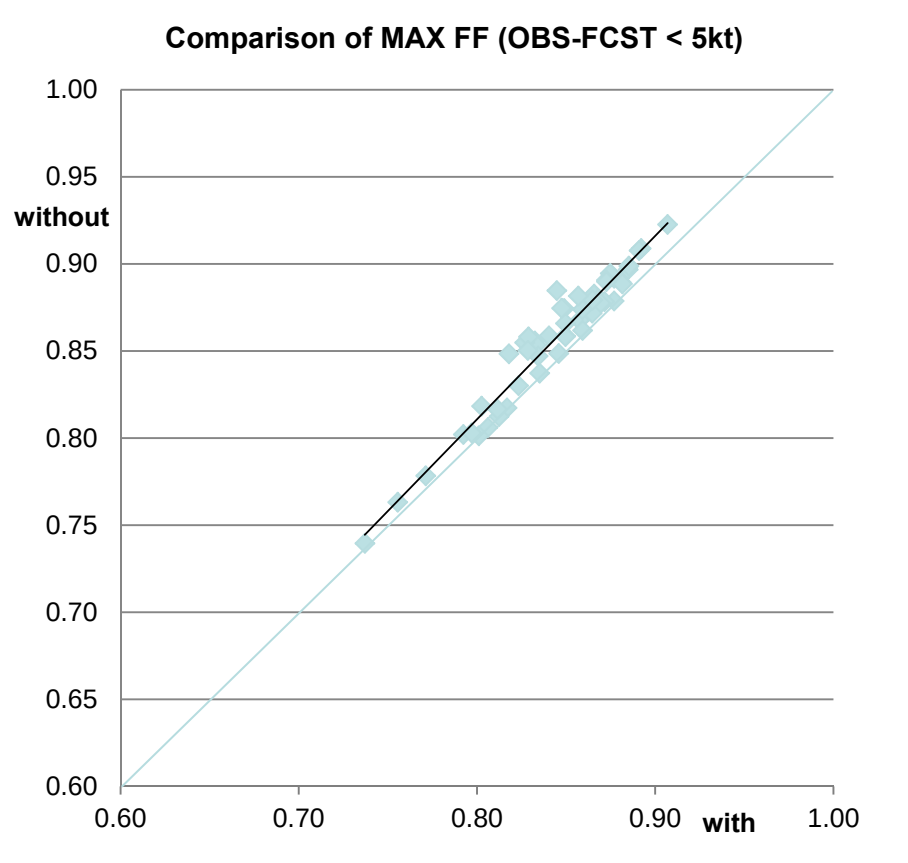
Comparison TAF - AUTOTAF

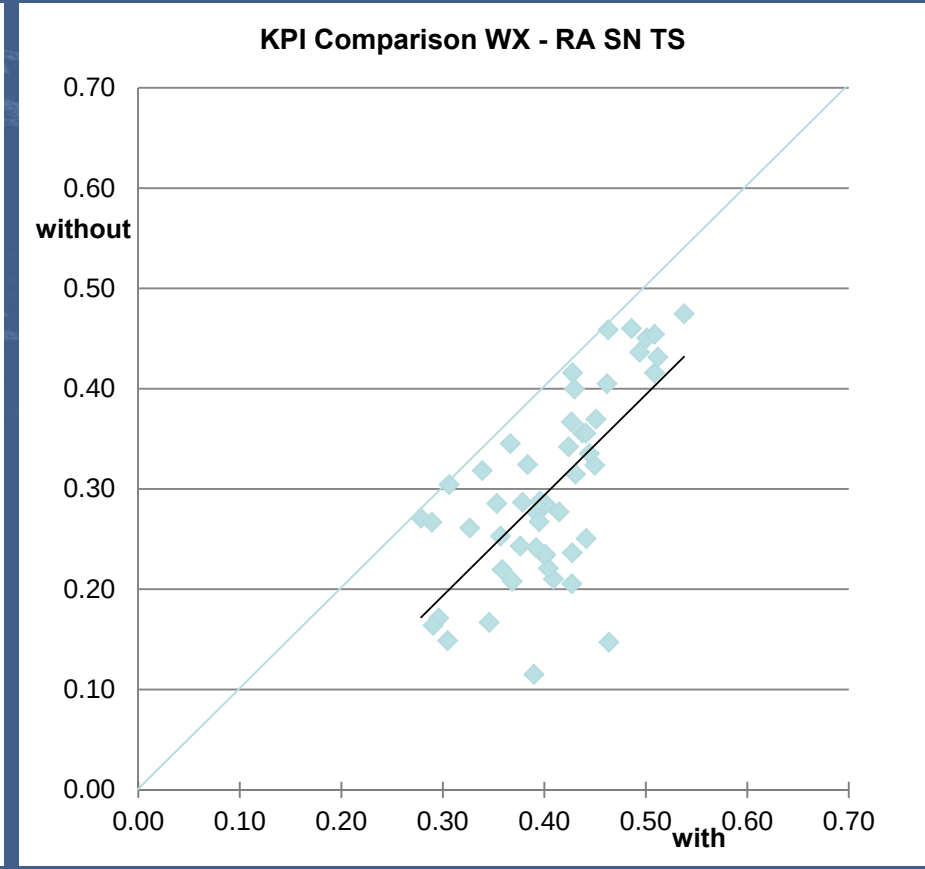




Probabilistic Information in TAFs (PROB TEMPO Groups) → Customers

PROB TEMPO groups are often disregarded for flight planning.
For wind: PROB TEMPO groups bring no quality improvement.
For WX, many hits are in PROB TEMPO groups → disregarding reduces score





Economic value of TAFs
(MeteoSwiss, 2011)

Costs	extra fuel	no extra fuel
Weather < min	C1	L
Weather > min	C2	0

C1	C2 + (1-p) x Div
C2	Carrying extra fuel
L	(1-q) x Div

OBS \ FCST	Below min	Above min
Below min	f11	f21
Above min	f12	f22

From TAF verification

Costs with TAF = f11 x C1 + f12 x C2 + f21 x L
Costs without TAF = (f11+f21) x C1 + (f12+f22) x C2 + A

Difference = economic value of TAF

C1, C2, L: dependent on aircraft type and flight duration
A: Costs for carrying extra fuel to reach 2nd alternate aerodrome

Costs are the “ultimate verification measure”.

Economic value of TAFs
(MeteoSwiss, 2011)

Reference:
MeteoSchweiz: Der volkswirtschaftliche Nutzen von Meteorologie in der Schweiz - Verkehr und Energie. Schlussbericht, 15. Juni 2011. econcept AG, Zürich

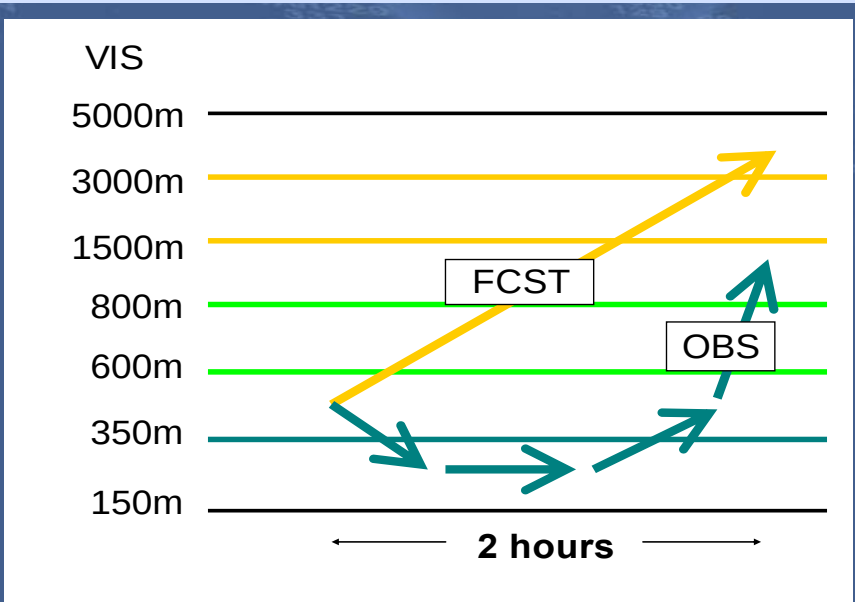
Remark:
Airlines often use a „planning“ threshold (e. g. for VIS / Ceiling) that is higher than actual approach minima.
TAF verification results could easily be used to evaluate and optimize planning thresholds.

Verification of TREND Forecasts

METAR CCCC 06005KT 0500 FG VV001 …

BECMG 4000 BR BKN006=

The TRENDs are verified by investigating if significant changes were:



- **OBS and FCST**
- **FCST (but not OBS)**
- **OBS (but not FCST)**
- **not OBS, not FCST**

A common activity for TREND verification is envisaged within the Met Alliance.

Verification of Aerodrome Warnings

Wind warnings:

- **Hourly verification**
- **Monthly results and details.**

Day	Warning			No warning		
	No event	Near event	Event	No event	Near event	Event
2012-10-27	0	8	5	5	5	1
2012-10-28	0	5	2	5	12	0
2012-10-31	0	1	1	20	2	0
Summe	1	16	19	655	46	7

Verification of Aerodrome Warnings

Example: Wind Warnings:

ISSUE TIME: 270332
AD WRNG 1 VALID 270400/270700
SFC WIND SE 20KT MAX 30 FCST WKN=

→ Too late!

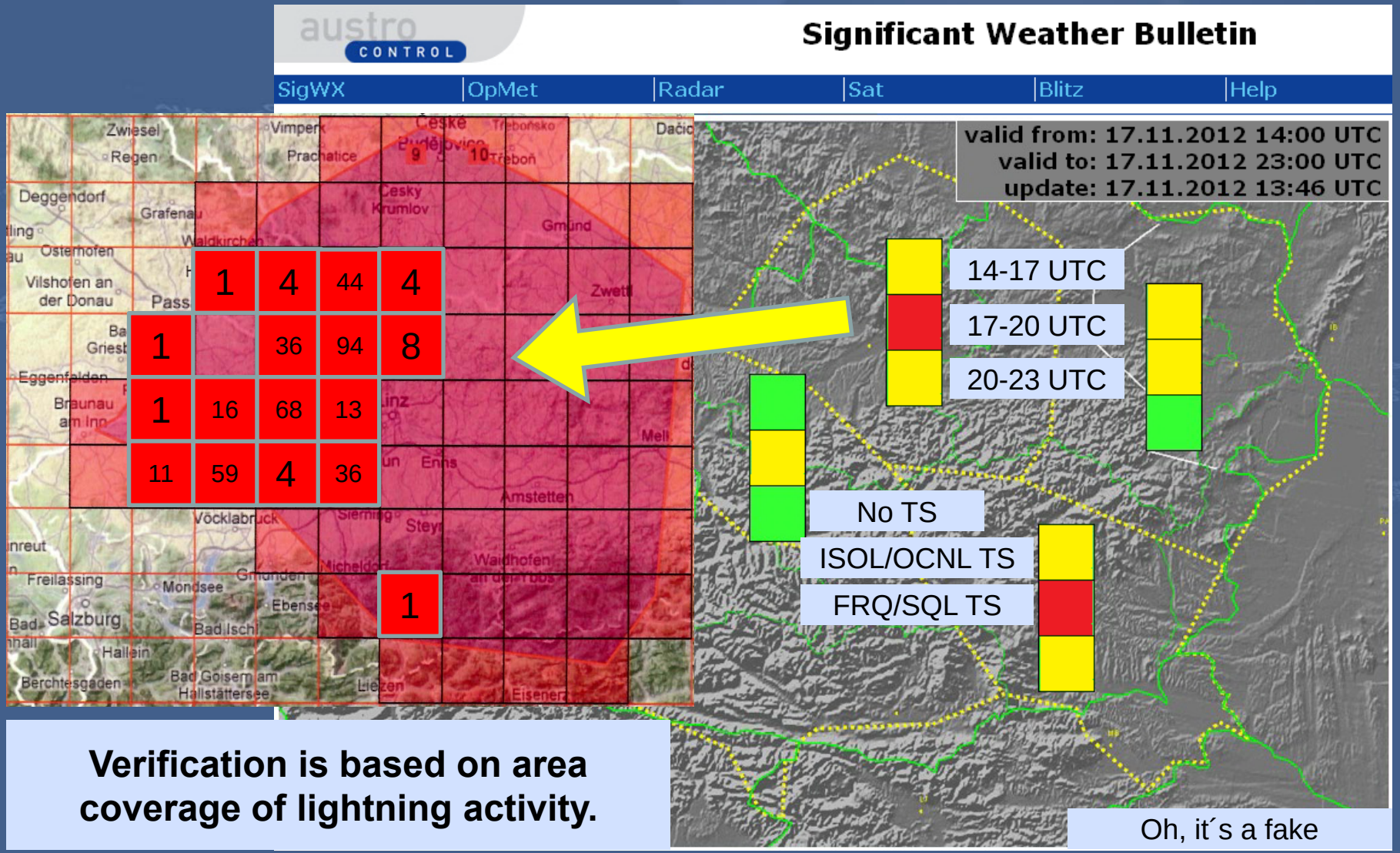
Verification is based on sensor data.
„Near event“: ≤ 5 kt below threshold

Problems:
Pre-warning time is difficult to define
- **Before strongest (hourly) gust?**
- **Before first gust > threshold?**

What is the actual time of issue?

HH-HH	Maxwind	fast Event	Event	gewarnt
00-01	27	x		nein
01-02	27	x		nein
02-03	27	x		nein
03-04	30		x	nein
04-05	30		x	30
05-06	26	x		30
06-07	25	x		30
07-08	23			
08-09	21			
09-10	21			
10-11	17			
11-12	20			

Verification of Forecasts for Air Traffic Management



Verification is based on area coverage of lightning activity.

Final Remarks

- Verification: “need to have” in aviation meteorology.
- Verification results should be fair to be accepted by Met forecasters.

  For this, attention has to be given to the properties and regulations for forecast production.
- Verification results should help in developments, e. g. by evaluating new methods and products.
- Verification results should be understandable to be used by management.

  Tracking the trend of a score should be easy enough
- Verification results should be specific for different aviation user requirements.

  Not all users are equally happy with the same forecast.
- Verification results should be designed to be used for estimating the economic value of forecasts.