to be published at the Seventh Conference on Aviation, Range, and Aerospace Meteorology 77th AMS Annual Meeting, 2-7 February 1997, Long Beach, CA

9.19 AUTOMATION OF AVIATION FORECASTS - THE PROJECTS AUTO-TAF AND AUTO-GAFOR

Klaus Knüpffer *
Meteo Service GmbH Berlin, Germany

1. INTRODUCTION

Meteo Service develops and implements statistical weather forecasting systems for interested weather services (e.g. Knüpffer 1993, 1995). The company has developed a comprehensive software system aiming at the automation of aviation weather forecasts in close cooperation with the German Weather Service. It has been implemented in 1996 and consists of the components TAF-guidance, Auto-TAF and Auto-GAFOR which are described in sections 2, 3 and 4. The TAF-guidance is a Model Output Statistics (MOS) system which produces forecasts for a specialized TAF oriented predictand set using a specialized predictor set. The Auto-TAF and Auto-GAFOR components translate TAF-guidance matrix outputs into TAFs and GAFORs complying with WMO and ICAO regulations.

2. TAF-GUIDANCE

2.1 General Aspects

TAF-guidances are MOS forecasts for a special set of elements to be forecasted in TAFs and GAFORs. The MOS regression equations have been derived based on more than 4 years of data of the European Model (EM, developed by the German Weather Service, horizontal resolution about 50 km) and observations (SYNOP) for about 100 stations separately for four seasons.

The TAF-guidance predictand set contains the categorical and probabilistic information needed for the automatic generation of TAFs and GAFORs. Table 1 shows an example of a TAF-guidance. Much care has been dedicated to the proper definition of the predictands (left column) in order to meet the special TAF and GAFOR needs and to make the encoding as easy as possible. The predictand name P_BKN<1500ft is to read as "Probability of ceiling below 1500 feet". Predictands are the result of a linear combination of predictors. The regression algorithm selects and weightens them. Table 2 shows the operational scheme of TAF-guidance production.

* *Corresponding author address:* Klaus Knüpffer, METEO SERVICE weather research GmbH, Teltower Damm 25, D-14169 Berlin, Germany; e-mail <k.knuepffer@meteo-service.met.fu-berlin.de>.

TAF-Guidance, Ausgabe Mo 27.05.96, 9.13z
St 10637 Frankfurt/M 50:03 N 8:36 E 111m

	OBS	VORHERSAGEN MOS				
PREDIKTAND	08z	09z	12z	15z	18z	21z
DD	22	22	23	25	25	29
FF /kt	10	11	10	10	9	7
FF_MAX_GUST		17	20	26	23	18
P_MAX_GST>25	0	12	29	52	38	17
P_MAX_GST>40	0	0	1	5	2	0
Tl /C	13	14	15	16	14	13
Td	13	12	11	11	11	12
VIS / 100m	OK	OK	OK	OK	OK	OK
P_Vis< 8 km	0	0	1	1	0	0
P_Vis< 5 km	0	0	1	1	0	0
P_Vis< 3 km	0	0	1	1	0	0
P_Vis<1.5km	0	0	0	0	0	0
P_Vis<800m	0	0	0	0	0	0
P_Vis<400m	0	0	0	0	0	0
P_Vis<200m	0	0	0	0	0	0
P_ww_STRATIF	0	23	14	5	2	11
P_ww_SHOWER	0	5	16	30	30	11
P_Cb %	0	2	11	39	36	10
P_ww_THSTORM	0	0	2	17	15	1
P_ww_REC_VCY	0	9	24	23	19	16
P_ANY_PRCPTN	0	28	30	35	32	22
P_ww_LIQUID	0	28	30	35	32	21
P_ww_FREZING	0	0	0	0	0	0
P_ww_SOLID	0	0	0	0	0	1
RR/6hr mm/10			11		10	
P_RR/6h>.1mm			69		65	
P_RR/6h>5 mm			12		11	
N	7	7	6	6	4	4
Okta<5000ft	5	5	5	3	1	1
Okta<1500ft	1	1	1	0	0	1
Okta<1000ft	0	0	0	0	0	0
Okta< 500ft	0	0	0	0	0	0
Okta< 200ft	0	0	0	0	0	0
P_BKN<5000ft	100	55	20	0	0	0
P_BKN<1500ft	0	0	2	0	0	0
P_BKN<1000ft	0	0	2	0	0	0
P_BKN< 500ft	0	0	0	0	0	0
P_BKN< 200ft	0	0	0	0	0	0
Ceiling		45	OK	OK	OK	OK
P_OVC<5000ft	0	0	0	0	0	0
..	..	..	..	..	..	..
P_OVC< 200ft	0	0	0	0	0	0
Overcast		OK	OK	OK	OK	OK

Table 1: TAF-guidance. From left to right: Predictand name ('P_' means 'Probability of'), last observation, forecasts.

Issue hh.mm	Last obs	EM-run	Valid Time TAF	GAFOR
00.13	23	12	01-10	
01.30	00	12		03-12
03.13	02	12	04-13	
04.41	04	12	12-06	06-15
06.13	05	12	07-16	
07.30	07	00		09-18
09.13	08	00	10-19	
10.41	10	00	18-12	12-21
12.13	11	00	13-22	
13.30	13	00		15-00
15.13	14	00	16-01	
16.41	16	00	00-18	
18.13	17	00	19-04	
21.13	20	12	22-07	
22.41	22	12	06-00	

Table 2: TAF-guidance: Operational scheme

2.2 Predictors

Predictor sources are model forecasts and observations. From the EM a set of direct model output (DMO) parameters like T2m, precipitation, cloud cover (different types, different layers), short and long wave radiation and others are used together with the upper air geopotential heights, temperatures and relative humidities at 6 levels (1000, 950, 850, 700, 500 and 300 hPa).

Many predictors are transformed and defined according to synoptical reasoning, for example:

- Geostrophic winds and voriticity indices are derived from geopotential heights.
- Wet bulb potential temperatures are calculated from the upper air predictors.
- Gradients and advections of several parameters are calculated.
- Non-linearily transformed and vertically integrated relative humidity values (vertical maximum, average and certain products) have been added for cloud cover (ceiling) and precipitation forecast. In most cases the regression prefers these predictors instead of using DMO cloud cover predictors.
- Thunderstorm indices (TotalTotals, Ko, S, Steinbeck, Darkov and one synthetical index) are derived and offered to the statistics as predictors for convective predictands. Most of these indices are defined between 850 and 500 hPa. Research has shown that a re-definition of these traditional thunderstorm indices for higher (700 up to 300 hPa) and lower (925 to 700 hPa) levels produces helpful predictors for the forecasts of convective phenomena.
- Fog predictors. Non-linear transformations of spread, wind speed and cloud cover are used for the description of potential radiation fog situations. Obstructions to visibility by precipitation are also expressed as predictors.

Markov chain predictors (=conditional climatologies of the predictands) have been provided by the German Military Aviation Service for this project. These are probabilities of certain (NATO-) colour states for visibility and ceiling in dependence on the last observation, the month, the weather type (cyclonal or anticyclonal) and the orography class of a station based on 30 years of observations. They are widely used by the regression. Other predictors are: Persistency of the statistical forecast, statistical forecasts of other elements, predictors of valid time +/- 3 and +/- 6 hours, sine and cosine of the day in year, step functions for handling changes of the numerical model.

2.3 MOS Equations

The regression algorithm has been optimized with the goal to minimize the RMSE of the predictands on independent data. This optimization is based on results of research on statistical overfitting. Details about this can be found in Knüpffer (1996).

Tables 3a and 3b show examples of MOS equations for the predictands "Probability of the occurence of thunderstorm in the ww code" (P(Ths)) in summer and "Maximum gust of the recent 3 hours" (FX3) in spring. Both equations have been derived for forecasts issued at 10.41z. They are valid at 18z at the airport of Frankfurt/Main. The following explanations shall help in understanding how the prognostic information present at issue time is combined to a forecast of these convective phenomena.

Out of a set of about 150 potential predictors the regression algorithm first selects the predictor with the highest linear correlation to the predictand. For P(Ths) this is the Ko-Index, defined between 700 and 300 hPa (index H) with a correlation coefficient of R(Pd) = 0.37. Next, the regression algorithm selects the predictor which has the highest correlation to the residual (=error) of the one-predictor-equation. It is the S-index, defined between 850 and 500 hPa (index M). This thunderstorm index inhers independent prognostic information from the first one. This can be seen from the fact that the correlation of this predictor to the residual (R(Res)) is only a bit smaller than its correlation to the predictand. The next predictor which enters the regression equation is the statistical forecast of any convective precipitation which has been calculated earlier and is used now as a potential predictor. Finally, the rotor of the geopotential height field at 1000 hPa, normalized by the wind speed at 1000 hPa has been selected by the regression. The

St=10637 Issue=06z Lead Time=+012
EM-Run=12z Predictand=P(Ths) Summer

R(Pd)	R(Res)	Name	dRVI	Co	Wgt
0.37	0.37	ThsKoH	13.2	2.48	32
0.36	0.30	ThsSM	7.2	4.76	30
0.28	0.18	StF(ww8)	1.5	0.22	19
0.29	0.15	Rot1/FF1	0.5	0.10	19
		Constant		-0.6	

#Cases= 422 #PotPr=165
MV(Pd)= 3.4 RV = 27 RMSE =11.5
SD(Pd)=13.5 E(RVI)= 21 E(RMSI)=12.0

Table 3a: MOS Equation for Probability of Thunderstorm

St=10637 Issue=06z Lead Time=+012
EM-Run=12z Predictand=FX3 Spring

R(Pd)	R(Res)	Name	dRVI	Co	Wgt
0.73	0.73	StF(FF)	53.2	1.96	62
-0.28	-0.19	ThWAdv85	1.5	-0.10	-15
0.35	0.18	ThsSL	1.3	0.59	13
		Constant		-0.7	

#Cases= 379 #PotPr=156
MV(Pd)=14.2 RV = 58 RMSE = 5.7
SD(Pd)= 8.8 E(RVI)= 55 E(RMSI)= 5.9

Table 3b: MOS Equation for Maximum Gust

rotor is defined as Laplacian applied to the 1000 hPa geopotential height field.

To the right of the predictor names the expected reduction of variance on independent data (dRVI) due to inclusion of the predictor is shown. The regression algorithm stops after the inclusion of the 4th predictor because the correlation of the 5th predictor to the residual of the 4-predictor-equation is lower than a critical correlation coefficient R_{crit}. R_{crit} is a function of the number of cases and the number of potential predictors. Next to the right the regression coefficients of the final regression equation are shown. The regression equation for forecasting the probability of thunderstorm is:

```
P(Ths) = -0.6 + 2.48*ThsKoH + 4.76*ThsSM
         + 0.22*StF(ww8) + 0.10*Rot1/FF1
```

The weight in the last column gives an indication on the importance of a certain predictor in the regression equation. The sum of the absolute values of all weights of all predictors is 100%, the sign is that of the regression coefficient. Weights are regression coefficients normalized by their standard deviation. They are helpful for the synoptic interpretation of a regression equation: It can be seen that the final forecast for P(Ths) is to 62% a combination of the two thunderstorm indices with about equal weight (although not equal coefficients). The remaining 38% are shared by the last two predictors. Summary statistics of this equation can be found in the lower part of the table. It can be seen that the reduction of variance of the equation at the development sample is 27%. The expected reduction of variance on independent data is only 21%.

The regression equation for FX3 can be interpreted synoptically in an analogous way: The best first guess for FX3 is to multiply the statistical forecast for FF by a factor of about 2. The following predictors can be interpreted as corrections to this simple approach: If the advection of wet bulb temperature at 850 hPa is positive (typical warm front situation) then gusts are suppressed and/or if it is negative (typical cold front situation) then it is positive. Finally, if the S-Index, defined between 925 and 700 hPa (index L), indicates readiness of the lower troposphere for convective activity then statistics adds further knots to the forecast.

The complete set of MOS equations can be interpreted as a synoptical knowledge base containing the following basic aspects of making a forecast for a specific element at a specific location in a specific season for a specific lead time:

- usefulness of the predictors and combinations of them for the local weather forecast;
- forecast accuracy of these predictors (systematical and random errors of the numerical model forecast);
- weighten the prognostic information in an optimum way.

This knowledge, fixed in about 2 million regression equations, is applied operationally by using the MOS equations for forecast production.

3. AUTO-TAF

3.1 Introduction and Example

TAFs are automatically generated based on the TAF-guidance. The TAF encoding algorithm and verification results are described in the next sections. An example of the Auto-TAF derived from the guidance (table 1) is shown in table 4.

```
EDDF 271019 24010KT 9999 SCT045
   PROB30 1319 25015G25KT 3000 TSRA
   BKN013CB
```

Table 4: Auto-TAF derived from the TAF-guidance (Table 1)

3.2 Auto-TAF Encoding Algorithm

The encoding of the TAF-Guidance into a TAF complying with ICAO and WMO regulations is controlled by two general criteria:

a) The loss of relevant information due to the encoding shall be minimized.

b) The code should be as short as possible.

These criteria are conflicting. A bonus/penalty system shall control the optimization in the process of Auto-TAF generation. This system works as follows:

1. For each of the four weather groups (wind, visibility, clouds and significant weather) an average forecast for the valid time of the whole Auto-TAF is produced.

2. Penalty points for this TAF are calculated as follows: For each hour the difference between TAF-guidance and Auto-TAF is squared and multiplied by a weighting factor. The weighting factor is an expression for the importance of the element and can be configurated by the user (weather service).

3. A trigger starts to try to separate the weather groups into two (or three, if the weather element has its extremum somewhere in the middle of the valid Time of the TAF) using BECMG.

4. Penalty points are re-calculated including penalty points for creating new groups which can also be configurated by the user. This allows the user to control the average and maximum length of the Auto-TAF.

5. The most effective change (with the highest reduction of penalty points) is executed and the system goes back to point 2.

6. If there is no useful BECMG to add anymore then the system tries to put together BECMGs into FM. If the reduction of penalty points due to the reduction of the number of groups is higher than the addition of penalty points due to loss of fit between Auto-TAF and TAF-guidance then FM is used.

7. The system continues with including PROB and TEMPO. TEMPO is interpreted as probability of 40%, and there are also possibilities to express lower probabilities than 30% by using TEMPO PROB or PROB TEMPO. The lowest probability which can be expressed by using TEMPO PROB30 is 12%. This approach is used in the TAF verification scheme of the German Weather Service which has been developed and described by Balzer (1994). The penalty point system is also applied to the probabilities. The algorithm always finds out the most efficient PROB or TEMPO group to add to the TAF. Efficiency is defined by the user in a configuration file.

8. Finally, the most efficient addition to the TAF saves less penalty points than the creation of additional TAF code would add. Then the TAF of optimum length has been found and the process of TAF generation is stopped.

The points 1 to 8 only describe the main processing. There are many details which require a lot of consideration and effort but can be handled using the penalty point system. For example, inconsistencies between the weather elements may occur at certain time steps in certain situations. These can be detected and then be punished with penalty points such that they will not occur any further.

In a first attempt to put these principles into practice Meteo Service has implemented an encoding program at the German Weather Service which produces syntactically correct TAFs. As far as it was to implement easily the principles above have been considered in this program.

3.3 Verification

The Auto-TAFs have been recepted surprisingly positive by the forecasting community. However, first verification results (Balzer, 1996) showed substantial deficiencies of this implementation. Comparative verification has been undertaken using the TAF verification scheme of the German Weather Service (Balzer, 1994) for more than 2000 TAFs issued in July 1996 for 17 places in Germany. The three methods compared are the forecaster, Auto-TAF and persistency.

In figure 1 the RMSE curves for wind speed are shown: The forecasters and persistency curves show normal behaviour. The RMSE increases with increasing lead time. However, Auto-TAF has its optimum 5 hours ahead and the RMSE of the 1 hour forecast is higher than the RMSE of the 5 hours and even the 9 hours forecast. The reason for this is obvious: The trigger of the encoding algorithm works too conservative and produces rarely more than one wind group for the whole 9 hours lead time. Thus, the wind speed is the averaged wind speed of these 9 hours, and the value in the centre of these 9 hours (5 hours) has the lowest deviation from this average, on the average. On the other hand it can be seen that beyond 2 hours ahead Auto-TAF produces the best wind speed forecasts.

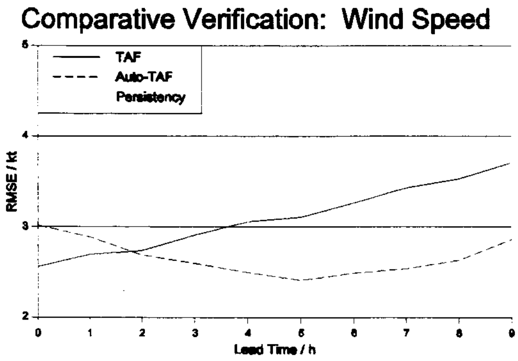

Figure 1

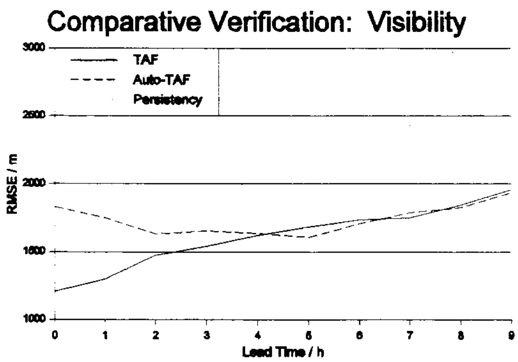

Figure 2

Figure 2 shows the same verification for visibility. A relative minimum at a lead time of 5 hours can also be found. The forecaster and persistency are better than Auto-TAF on the average. The TAF-guidance produces better visibility forecasts than persistency because persistency is one of the predictors for the guidance. Therefore it is likely that - especially in the first four hours - still too much information is lost in the process of encoding.

Figures 1 and 2 show better performance of persistency (lead time = -1 h) than the forecaster in the very short range. This is the effect of focussing the TAFs to lead times beyond 1 hour: The forecaster expects the pilot to be able to interpolate between the last METAR and the second TAF hour and keeps the TAF simpler this way.

4. AUTO-GAFOR

GAFORs based on the TAF-guidances of representative stations for a GAFOR area are automatically generated. For each area single GAFOR codes based on only visibility, ceiling and the combined code (Index) are produced using four different definitions:

- Prevailing (50%):= Code belonging to the "expected" flight conditions with 50% chance of better and 50% chance of worse conditions, averaged over the representative non-mountain stations.

- Variable (X%): = Code belonging to flight conditions with X% chance of worse conditions, averaged over the representative non-mountain stations. X is variable and can be changed by the user (weather service).

- Minimum (50%): = Code belonging to the "expected" flight conditions of a fictive station with the worst conditions of all representative non-mountain stations.

- Mountain (50%): = Optionally, if there are mountain stations in the GAFOR area, for these stations a GAFOR is produced separately. Mountain stations are considered not to be representative for the GAFORs.

The GAFORs of the area in which the example airport Frankfurt/Main is situated are shown in table 5.

Issue-Time	GA		Visibility				Ceiling				Index			
1996052707	42	PREVAILING(50%)	C	C	C	C	O	O	C	C	O	O	C	C
1996052707	42	VARIABLE__(30%)	C	C	C	C	D	D	O	O	D1	D1	O	O
1996052707	42	MINIMUM___(50%)	C	C	C	C	O	O	O	O	O	O	O	O
1996052707	42	BERG_10635(50%)	X	C	C	C	X	X	C	C	X	X	C	C

Table 5: Auto-GAFOR generated from TAF-guidances issued for different stations in GAFOR Area 42. The columns are: Obs(07z) and forecasts valid from 9-11, 11-13 and 13-15z.

5. OUTLOOK

The Auto-TAF and Auto-GAFOR systems have a high potential for automation of the production of aviation forecasts. The automatic forecasts are expected to be of equal or better accuracy than the human forecasts on the average. Completely automation of TAF and GAFOR production is not the goal, but 80% of all forecast situations may not need human correction anymore when the deficiencies of this system are removed. Especially in complicated situations the forecaster should be able to add value to the automatic forecasts. The weak point of the whole system is currently the TAF encoding algorithm which can be improved considerably within the near future. Attempts for improving the TAF-guidance will be directed towards the extended use of conditional climatologies of the predictands in dependence of predictor values. This shall help to improve the capability of the linear regression algorithm to consider non-linear facts. The inclusion of the output of a boundary layer model which produces categorical forecasts of the elements needed in Auto-TAFs would also be desirable.

In view of the possibility to produce high quality statistical forecasts of the elements forecasted in a TAF automatically, the aviation community might think over the current regulations for TAF production. Instead of loosing information in the process of translating the TAF-guidance TAF syntax it might be better to produce graphics with time series of the elements forecasted in the guidance, perhaps with most severe weather in red.

Acknowledgements

This work was done under contract of the German Weather Service. Thanks to the engagement of Bernd Richter and the contributions of Konrad Balzer in many fruitful discussions it was possible to reach this state of the art within relatively short time. Many thanks are finally also directed to my colleague Dik Haalman who has written the TAF-encoding software and who was involved in many helpful discussions during this work.

6. REFERENCES

Balzer, K., 1994: TAF-Verifikation - Eine Dokumentation der Methodik, *DWD,* Potsdam

Balzer, K., 1996: First Comparative Verification of Auto-TAFs, personal communication, *DWD,* Potsdam

Knüpffer, K., 1993: Recent Developments in Statistical Weather Forecasting - Implications on Institutionalization, Operational Services and Marketing. Preprint Volume: *First European Conference on Applications of Meteorology*, 27th September - 1st October, Oxford (UK)

Knüpffer, K., 1995: Statistical Weather Forecasting Systems for the German Weather Service and the Swedish Meteorological and Hydrological Institute. Preprint Volume: *2nd European Conference on Applications of Meteorology*, September 25-29, Toulouse, 132-135

Knüpffer, K., 1996: Methodical and Predictability Aspects of MOS Systems. Preprint Volume: *13th Conference on Probability and Statistics in Atmospheric Sciences*. February 21-23, San Francisco, California, 190-197

In view of the possibility to produce high quality statistical forecasts of the elements forecasted in a TAF automatically, the aviation community might think over the current regulations for TAF production. Instead of loosing information in the process of translating the TAF-guidance TAF syntax it might be better to produce graphics with time series of the elements forecasted in the guidance, perhaps with most severe weather in red.

Acknowledgements

This work was done under contract of the German Weather Service. Thanks to the engagement of Bernd Richter and the contributions of Konrad Balzer in many fruitful discussions it was possible to reach this state of the art within relatively short time. Many thanks are finally also directed to my colleague Dik Haalman who has written the TAF-encoding software and who was involved in many helpful discussions during this work.

6. REFERENCES

Balzer, K., 1994: TAF-Verifikation - Eine Dokumentation der Methodik, *DWD,* Potsdam

Balzer, K., 1996: First Comparative Verification of Auto-TAFs, personal communication, *DWD,* Potsdam

Knüpffer, K., 1993: Recent Developments in Statistical Weather Forecasting - Implications on Institutionalization, Operational Services and Marketing. Preprint Volume: *First European Conference on Applications of Meteorology*, 27th September - 1st October, Oxford (UK)

Knüpffer, K., 1995: Statistical Weather Forecasting Systems for the German Weather Service and the Swedish Meteorological and Hydrological Institute. Preprint Volume: *2nd European Conference on Applications of Meteorology*, September 25-29, Toulouse, 132-135

Knüpffer, K., 1996: Methodical and Predictability Aspects of MOS Systems. Preprint Volume: *13th Conference on Probability and Statistics in Atmospheric Sciences*. February 21-23, San Francisco, California, 190-197