Improved Low visibility and Ceiling Forecasts at Schiphol Airport

Final report, part 1

KDC - LVP Project team

Photo: Peter de Vries (KNMI)

July 2008

KDC/2008/0089

KNMI publication 222

Status page

Authors			
Name	Organization (Function)	Approval	Date
C.J. de Rover	KLM (Support Manager ATC Desk)		
D. Vogelezang	KNMI (Scientist)		
H. van Bruggen	KNMI (Forecaster)		
N. Maat	KNMI (Scientist)		
L. Smit	LVNL (Researcher)		
J. Heijstek	NLR (Sr. R&D Manager)		
M. Keet	Schiphol Group (Advisor)		
J. Wijngaard	KNMI (Scientist, project manager KNMI)	[signature]	10 July 2008
R. ten Hove	Schiphol Group (Sr. Advisor) Chairman KDC - LVP Project team	[signature]	10 July 2008

Acceptance			
Name	Organization (Function)	Approval	Date
E. Westerveld	LVNL (Manager R&D) Chairman Management Team KDC	[signature]	15.07.08

Document information	
Title	Improved Low Visibility and Ceiling Forecasts at Schiphol Airport, Final report, part 1
Document number	KDC/2008/0089
Version number	1.0
Version date	8 July 2008
Status	Final

Executive summary

Autumn 2006 the management team of the KDC (Knowledge & Development Centre) decided to start a project to improve forecasting of low visibility conditions at Schiphol Airport. A project team was formed with representatives of LVNL, KLM, Schiphol Group, NLR and KNMI as the meteorological consultant and provider.

This KDC - LVP project focuses on increased reliability and accuracy of the low visibility forecasts as well as on the development and deployment of a tool to present this forecast information to the appropriate operational management for support in their decision making process. It is not the intention of the project to change the flow restrictions (aircraft per hour) itself. Other KDC-projects are focused on initiatives to increase the runway capacity during low visibility conditions.

In order to have results available as soon as possible, the project has been split in two parts. Part 1 concentrates on improvements in the current forecast products to obtain the so-called quick wins. Part 2 will focus on new sensor technology, high resolution models and other innovative developments. This report marks the end of part 1, where part 2 is yet to be started and will probably be included in the project scope of the KvK/KBS (Kennis voor Klimaat / Klimaat Bestendig Schiphol - Dutch for "Knowledge for Climate / Climate Resistant Schiphol").

Airport capacity reduces due to low visibility, resulting in delays, diversions and cancellations leading to increased workload and additional operational costs/expenses. Based on the forecast of low visibility conditions, flow restrictions are enforced to reduce this operational impact. This is acceptable as long as the forecasts are accurate and reliable, that is, the forecast equals correctly predicted the observed conditions (Hits). Visibility conditions worse than forecasted (Misses) or better than forecasted (False Alarms), increase the operational costs unnecessarily and such erroneous forecasts should therefore be limited as much as possible.

An analysis has been made on the occurrence of low visibility and ceiling. It showed the very low occurrence of conditions with Low Visibility Procedures (LVP-phases A, B, C and D occur in total less than 5% of the ATM operational time per year). More climatological information on visibility and ceiling is presented in the KNMI report "Low Visibility and Ceiling Forecasts at Schiphol, Part 1 - Assessment of the current system".

Several improvements to the forecasts have been proposed, developed en verified. They include general improvement to both the SKV (Schiphol Kans Verwachting, - Dutch for Probability Forecast Schiphol) and the tool for the low visibility forecast (TAFG = Terminal Aerodrome Forecast Guidance) due to:

- New and closer up-stream observing sites used in the statistical model,
- Inclusion of RVR (Runway Visual Range) in addition to MOR (Meteorological Optical Range),

- Joint probabilities for visibility and ceiling.

This results in more accurate and reliable probabilities for LVP conditions.

Specific changes to the SKV are:

- LVP-phase C added,
- 6 hours extension of forecast period,
- Long and short forecast period combined.

The new TAFG and SKV show a significant improvement in accuracy and reliability. The number of Hits has almost doubled, and the number of False Alarms reduced with 25-50% in comparison with the old TAFG and SKV. The number of Misses shows almost no change.

The reduction in avoidable expenses is strongly dependent on the cost sensitivity of False Alarms versus Misses, as well as the decision threshold. The project team analysed a method to determine the decision threshold aimed at the lowest expenses. Although no detailed cost information was available, the analysis showed that the improved forecast can lead to a substantial cost reduction.

The results of the verification were such that the project team decided to implement the improvements to the TAFG and SKV as soon as possible. The improved TAFG and SKV have seen their operational implementation on May 26th 2008 (formalized June 2nd 2008). From that day onward, the forecasts for Schiphol Airport benefit from the improvements gained in the KDC - LVP project. Yet the CPS tool (Capacity Prognosis Schiphol) of KLM-OCC does not fully benefit from the improvements because it still works with its old input format.

Based on the findings of the KDC - LVP project part 1, the project team recommends:

- To initiate part 2 of the KDC - LVP project, wherever practicable in connection with the KvK/KBS (Kennis van Klimaat / Klimaat Bestendig Schiphol) project starting September 2008,
- To adapt the CPS tool to make best use of the improved low visibility forecasts,
- To improve the decision making process by optimizing the benefit of the probabilistic forecast by determining the decision threshold(s),
- To investigate whether the TAFG and SKV should be extended to more than one location at the airport,
- To determine an optimal update frequency and temporal output resolution of the visibility forecast, in agreement with users from LVNL, KLM and AAS, but also meteorologically meaningful.

Content

1. Introduction

1.1 Background

Low visibility conditions have a direct negative influence on the available operational capacity of Schiphol Airport. Accurate, reliable and unambiguous information concerning actual and expected low visibility conditions is crucial to the decision-making process during conditions where Low Visibility Procedures (LVP) must be applied.

If a limited runway capacity at Schiphol Airport is known in advance, delaying or even cancelling flights could be considered. On the other hand, a more accurate and reliable forecast of the duration of periods with fog will reduce the time that the ATM (Air Traffic Management) system operates at less capacity. This may be achieved by timely issuing or discontinuing a flow control measure.

The Dutch Meteorological Institute (KNMI) introduced in 2003 the Schiphol Kans Verwachting (SKV, Dutch for Probability Forecast Schiphol) to give probabilities for the operational LVP-phases. Although with the introduction of the SKV a significant improvement was achieved, the forecast suite still did not met all needs of the airport services and operators.

In the second half of 2006 the board of KDC (Knowledge and Development Centre) decided to start a project to improve forecasting of low visibility conditions at Schiphol Airport. As a result of an improved prediction of fog and other low visibility situations, it is expected that measures (like flow restrictions) can be taken more adequately to minimize the disturbance and delay of operations.

1.2 Outline of this Document

This document describes part 1 of the KDC - LVP project "Improved Low visibility and Ceiling forecasts at Schiphol Airport". The aim of this document is to explain the study and its goals and to present the results of the developed improved tool for the low visibility forecast (TAFG, Terminal Aerodrome Forecast Guidance).

Chapter 2 gives the project objective and requirements. Chapter 3 describes the operational implications of low visibility for today's commercial aviation, the effect on airport (runway) capacity as well as the effect of inaccurate forecasts. In chapter 4 the different meteorological products and the developed and implemented changes in the TAFG are described.

Verification of the revised visibility forecast is described in chapter 5, where in chapter 6 a method to use the probabilistic forecast in the decision making process is discussed. In chapter 7 the project deliverables are summarised and finally in chapter 8 concluding remarks and recommendations are given.

2. Project Description

2.1 Objective

Airport capacity is limited due to low visibility. If no precautionary measures are taken this capacity shortfall will lead to an increased workload for Air Traffic Control (ATC). To deal with the increased workload flow restrictions will be enforced when reduced runway capacity is expected.

Many flights (mainly European) remain at their airport of origin and will experience large delays and often cancellation. This leads to a significant increase of operational cost (re-booking passengers or hotel accommodations, crew scheduling, increased fuel burn and emissions etc) due to delay and cancellation of flights,

Inaccurate low visibility forecasts have a negative impact on the operational costs as related flow restrictions may be unnecessary or issued too late. The prime objective is to reduce the number of unnecessary flow restrictions due to incorrect visibility prediction as well as to reduce the number of situations where flow restrictions were lacking but visibility conditions would have justified it.

The ultimate goal is to use the existing capacity of Schiphol more efficiently. The number of flow restrictions due to reduced visibility must be as small as possible. Increase of runway and ground movement capacity are also promising solutions but they are part of other KDC projects and studies and therefore not included in this KDC - LVP project.

On short term the improvement of the accuracy and reliability of the low visibility forecast has the highest potential. Prevention of low visibility conditions will have potentially a higher (positive) operational impact but may in its best be a mid to long term solution. Fundamental research on causes of low visibility conditions as well as possibilities to prevent and to dissolve low visibility will take several years. For this reason the project has been divided into two separate parts.

Part 1: Short term improvement of the low visibility forecast model. Focus on the development of a more accurate and reliable Low Visibility forecast tool.

Part 2: Medium/long term improvements. To gain a better insight in the microclimate at Schiphol with regard to LVP conditions.

Part 1 is subject of the KDC project described in this document. It focused on the development of a low visibility tool to be used not only within the operational environment of LVNL but also within those of KLM (Operations Control Centre) and Schiphol Group (Airside Operations).

2.2 Project Requirements

It is assumed that in the former situation the low visibility forecasts had an accuracy / reliability of approximately 30%. This may be translated in such a way that only in 3 out of 10 occasions the actual low visibility condition is equal to the predicted conditions ("hit"). The 30% value is only a subjective indication based on gut feeling.

For the project a set of requirements have been made for which below a short list is given.

- "Hit"- rate of 60% on forecasts 3 hours in advance,
- Forecast must be made 24 hours in advance (preferably 36 hours),
- To be easily integrated into operational systems of LVNL, KLM and AAS.

In chapter 7 the compliance of the project part-1 result with the above requirements is described.

3. Low Visibility and Ceiling in Aviation

3.1 Low Visibility Definitions

At Schiphol Airport a set of thresholds for visibility and ceiling (cloud base at least 5 octas coverage) are defined for operational use. Table 1 shows the different visibility classifications and the related visibility and ceiling ranges. Each visibility class corresponds to its own specific operational procedures. Different visibility parameters are distinguished: MOR (Meteorological Optical Range), VIS (VISibility) and RVR (Runway Visible Range). MOR is a parameter more or less directly available from the visibility sensors, where both VIS (Visibility for Aeronautical purposes) and RVR are calculated values and dependent on background luminance and lamp (approach lights) settings.

Visibility classification	Visibility [VIS / RVR]		Ceiling
Good	> 5 km [VIS]	and	> 1000 ft
Marginal	1.5 – 5 km [VIS]	or	300 – 1000 ft
LVP phase A	550 – 1500 m [RVR]	or	200 – 300 ft
LVP phase B	350 – 550 m [RVR]	or	< 200 ft
LVP phase C	200 – 350 m [RVR]		
LVP phase D	< 200 m [RVR]		

Table 1. Low Visibility Procedure categories.

3.2 Climatology

To get a first impression of the occurrence of Low Visibility events the numbers of days per year with at least one hour of LVP conditions are presented in figure 1. For the period 1977 - 2006 a decrease in LVP conditions can be seen. This is not unique for the Schiphol Airport location as it is also seen at other locations in the Netherlands.

The occurrence of Low Visibility is also dependent on season and time of the day (see figure 2 for dependency on time of day). The dependency on season and time of the day is an important factor when forecasting LVP conditions.

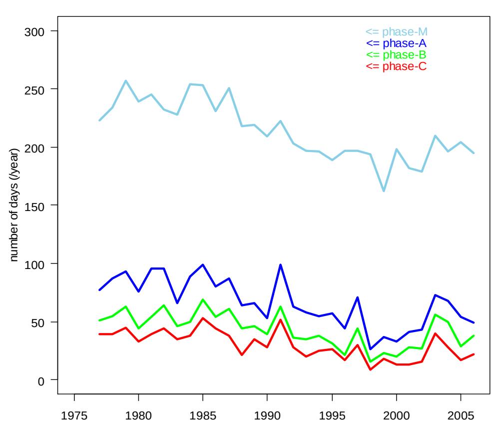


Figure 1. Occurrences of LVP, annually Visibility 1977-2006 - Schiphol.

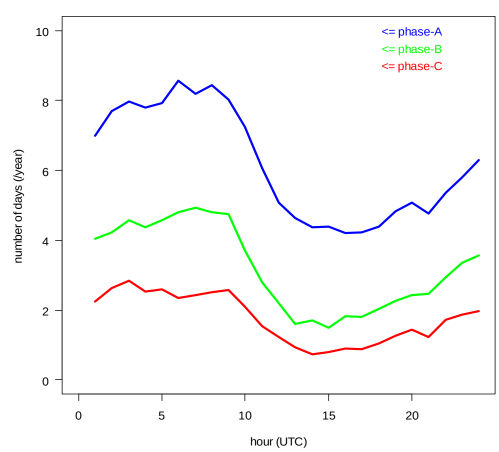


Figure 2. Occurrences of LVP, daily Visibility 1977-2006 - Schiphol.

In figure 3 the percentage of time (average per year) is given for the different visibility conditions. It is clear that low visibility conditions phase A to D have a low probability. More detailed information on climatology can be obtained from the interim report "Low Visibility and Ceiling Forecasts at Schiphol, Part 1 - Assessment of the current system"[1].

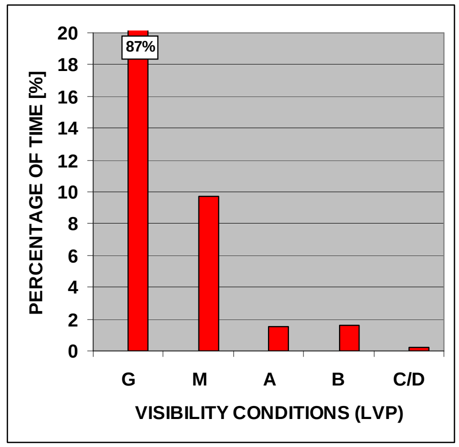


Figure 3. Occurrences of LVP in percentage of time. Period: 2003/05 - 2007/04.

3.3 Operational Implications due to Low Visibility

Reduced visibility will limit the runway capacity. At Schiphol Airport this limitation is twofold:

- Requirement to operate parallel arrival runways only,

  The converging runway lay-out of Schiphol Airport requires limitations in runway use and runway combinations during operations under LVP conditions.

- Increased separations on final approach,

  Also larger in-trail separations on final approach as well as a reduced number of aircraft moving on the airport's taxiways are required to prevent unacceptable disturbance of the Instrument Landing System (ILS). With decreasing visibility a clear reception of the ILS signals becomes more important.

Table 2 gives an indication of the capacity restrictions as a function of the visibility and ceiling classification. The application of a flow restriction is not only dependent on the visibility (RVR and ceiling) conditions but also on the expected duration of the low visibility situation. If low visibility conditions are expected to last only for a short period (for example 1 hour) the flow restriction could be less stringent or even omitted.

Visibility classification	Capacity restrictions due to visibility (movements/hour)
Good	68 arrivals or 74 departures, max 104/108 movements No flow restrictions
Marginal	68 arrivals or 74 departures, max 104/108 movements Use of independent parallel runways required No flow restrictions
LVP phase A	56 arrivals or 52 departures, max 80 movements In general no flow restrictions
LVP phase B	44 arrivals or 52 departures, max 74 movements Flow restrictions in force
LVP phase C	30 arrivals and 17 departures, max 47 movements Flow restrictions in force
LVP phase D	16 arrivals and 20 departures, max 36 movements Flow restrictions in force

Table 2. Low visibility conditions and flow restrictions.

[1] "Low Visibility and Ceiling Forecasts at Schiphol, Part 1 - Assessment of the current system", written by KNMI in October 2007.

3.4 *Inaccurate Visibility Forecasts*

More accurate and reliable forecasts of the occurrence and the duration of low visibility conditions reduces the number of times that ATC needs to issue flow restrictions and operate with less capacity. Three situations with respect to reduced visibility (actual/observed and forecast) are distinguished (figure 4):

1. Actual visibility in accordance with forecast (Hit). Justified flow restrictions, high but acceptable and unavoidable costs.
2. Actual visibility worse than the forecast (Miss). Lack of adequate/timely flow restriction, high operational costs due to holding, diversions and cancellations.
3. Actual visibility better than forecast (False Alarm). Unnecessary flow restrictions, high operational costs due to cancellations and aircraft remaining at outstations/origin.

Based on the forecast of low visibility conditions, flow restrictions can and will be enforced as during low visibility conditions runway capacity is limited. Flow restrictions reduce the number of aircraft flying to Schiphol Airport with the goal to lower the pressure on the ATM system.

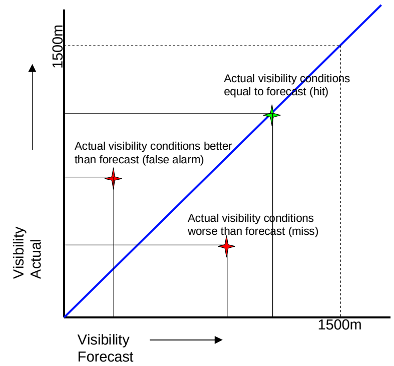


Figure 4. Definition of Hit, Miss and False Alarm.

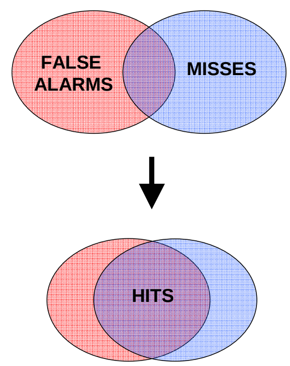


Figure 5. Less Misses and False Alarms implies more correct forecasts (Hits).

More accurate and reliable visibility forecasts will lead to a decrease in False Alarms and Misses resulting in more correct forecasts (Hits) as is schematically shown in figure 5.

4. Improvement to the Low Visibility Forecast

4.1 Forecast Products

At KNMI it was recognized already some years ago that the standard TAF (Terminal Aerodrome Forecast) bulletin was not appropriate to supply all relevant forecast information to the users at the airport. This resulted in the introduction of the Schiphol Kans Verwachting (SKV, Probability Forecast Schiphol) in 2003. By using probabilistic forecasts within this SKV, the uncertainty that is inherent in weather forecasts could be dealt with. A schematic construction of the aeronautical forecasts is depicted in figure 6. Main source of this forecast is the output of the numerical weather prediction model HIRLAM (High Resolution Limited Area Model). Real-time observations and HIRLAM data are the main data sources for the TAF Guidance (TAFG), an automatic statistical post processing module that provides essential information to the forecaster. At the back end of the cascade the two bulletins including forecasts related to visibility and ceiling are produced:

- TAF Schiphol: Prescribed ICAO format,
- SKV Schiphol: User tailored forecast.

Until the recent upgrade, both TAF and SKV were provided in a short and a long version. The short version having a lead time of 9 hours starting 3 hours after issue time with an update frequency of 3 hours, where for the long version the lead time was 24 hours starting 7 hours after issue time with an update frequency of 6 hours.

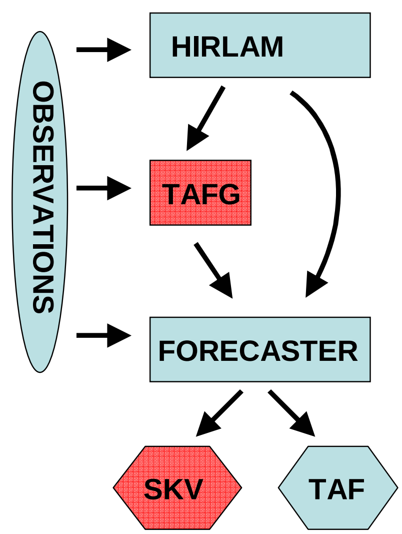


Figure 6. Schematic overview of the forecast cascade.

With the upgrade the TAF system remained unaltered because of the internationally prescribed format. For the SKV however the short and long version have now been combined into a single bulletin covering a lead time of 3 to 31 hours with an update frequency of 3 hours.

For the very short lead times (0 to 3 hours ahead) the skills of the human forecaster (MAS - Meteorological Advisor Schiphol) appear very valuable. He/she is situated at the KNMI head office or at the ATC facilities when weather conditions require. The current procedure with the MAS is requested by the operational departments of ATC Netherlands (LVNL) and KLM-OCC (Operations Control Centre). It has been selected for its ability to react on short term and sometimes unpredicted or complex meteorological changes/events.

The focus of the project is on the improvement of the TAFG and the SKV. Where the format and update rate of the TAF is regulated by strict international (ICAO) rules, the SKV is a local product and therefore more flexible. Changes as required by the project definition can therefore only be implemented in the short and long SKV. Both short and long TAF formats will remain unchanged.

4.2 Developed Changes to Forecast Products

Several changes and improvements to the visibility forecast have been developed and implemented in the product suite:

- New -closer- upstream sensor sites are used in the statistical model (TAFG),
- In addition to MOR, RVR is given as an output parameter,
- Joint probabilities for visibility and ceiling have been derived.

In the next sub paragraphs these changes will be shortly described.

4.2.1 New - Closer - Upstream Sensor Sites

Nearby upstream weather station data has been introduced for the construction of the statistical probability distribution for visibility. These stations are the so-called ‘Fog detection sites’ situated around Schiphol Airport in the neighbourhood of Nieuw-Vennep, Nieuwkoop, Muiden and Assendelft (see figure 7). The automated observations from these sites are used to forecast the probability distribution of visibility at Schiphol airport. In the old system the stations upstream were about 50 km away from Schiphol. The new sites are much closer (approximately 20 km) and have been proven to be important for short term forecasting of small-scale phenomena like low visibility.

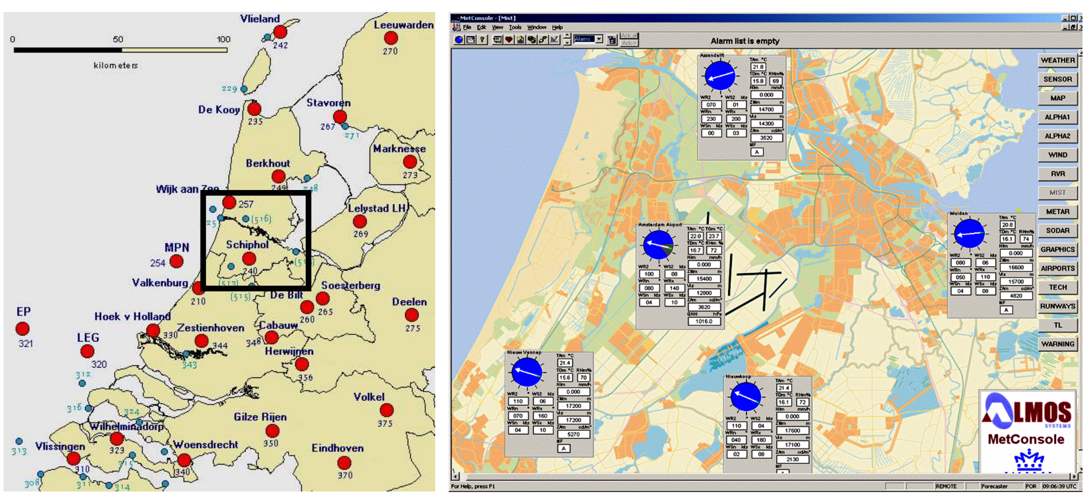


Figure 7. Additional up-stream sensors in the close vicinity of Schiphol Airport.

4.2.2 MOR to RVR Translation

The statistical model, the TAFG, contains (joint) probabilities for visibility and ceiling. The probabilities for visibility have so far always been in terms of MOR. The new extended version of the statistical post processing model contains probabilities for RVR as well.

Combination of MOR and BackGround Luminance (BGL) results in the RVR which is extensively described in vd Meulen (1993)[2]. In table 3 the relationship between MOR and RVR is presented. There is a significant difference between MOR and RVR during situations with reduced BGL (i.e. night time). Especially in winter the visibility forecast for the early morning, based on MOR, could give a low visibility value where in practice RVR could be much higher. It therefore could happen that enforced flow restrictions, based on MOR forecasts, were in practice not necessary (i.e. false alarm).

RVR as a function of MOR and background luminance (cd/m2)

MOR	RVR					
	NA	SD	SL	DD	DG	DH
	<50	50-300	300-1000	1000-4000	4000-12000	>=12000
50	175	150	125	125	100	100
100	325	275	250	200	175	150
150	500	375	325	275	250	200
200	650	500	400	350	300	250
250	800	600	550	450	350	275
300	900	700	650	550	400	325
350	1000	800	700	600	500	375
400	1100	900	800	650	550	400
450	1200	1000	800	750	600	450
500	1400	1000	900	800	650	500
550	1500	1100	1000	800	700	550
600	1600	1200	1000	900	750	600
700	1800	1300	1200	1000	800	700
800	2000	1500	1300	1100	900	800
900	P2000	1600	1400	1100	900	900
1000		1800	1500	1200	1000	1000
1100		1900	1600	1300	1100	1100
1200		2000	1700	1400	1200	1200
1300		P2000	1800	1500	1300	1300
1400			1900	1500	1400	1400
1500			2000	1600	1500	1500
1600			P2000	1700	1600	1600
1700				1700	1700	1700
1800				1800	1800	1800
1900				1900	1900	1900
2000				2000	2000	2000
2100				P2000	P2000	P2000
	> phase A	phase A	phase B	phase C	phase D	

Table 3. MOR related to RVR, values for different background luminance conditions:
Bright day (DH), Normal day (DG), Dark day (DD), Twilight-light (SL), Twilight-dark (SD) and Night (NA).

The RVR calculation needs as input MOR, background luminance and lamp setting of the approach lights.

2 J.P. van der Meulen (ed), 1993. Runway Visual Range - Observing and reporting practices in the Netherlands. KNMI publication, 100 pp.

So far, only actual measurements of the background luminance were available, not forecasts. A statistical forecast for the background luminance has therefore been developed. In this forecast the solar elevation angle appeared an important predictor (as expected), but also parameters like cloud amount, cloud base / ceiling, humidity, precipitation and MOR proved to be useful for the estimate of the future background luminance values.

With these new forecasts for the background luminance, it is possible to translate MOR based values to RVR based values for every hour in the output series of the TAF guidance. For example, a probability of MOR<400 m can be translated into a probability of RVR<x, where x can range from 400 to 1100 m dependent on the background luminance (see table 3). This value for x may be different for every hour in the forecast range, because the forecasted background luminance may vary from hour to hour. Probabilities for all other MOR thresholds are used for the translation into RVR probabilities at every forecast hour. From these series of RVR probabilities the desired probabilities for RVR<1500 m, RVR<550 m and RVR<350 m are calculated by interpolation.

4.2.3 From Separate to Joint Probabilities for Visibility and Ceiling

The old TAFG system produced separate probability forecasts for visibility (MOR) and ceiling. Mathematical combination of two probabilities, which is necessary since the LVP categories depend on combined thresholds for visibility and ceiling, was done afterwards. With the upgrade to the new TAFG system it was decided to produce the operationally relevant combinations of visibility (MOR) and ceiling directly.

The new statistical TAF guidance model contains forecasts for joint probabilities of visibility and ceiling. The visibility component of these combinations is still in terms of MOR. The procedure described in section 4.2.2 is used to translate these probabilities to probabilities of a combination of RVR and ceiling. As a result, these probabilities directly represent the probabilities of LVP-phases. The model provides the probabilities for the following situations:

- LVP-phase A or worse,
- LVP-phase B or worse,
- LVP-phase C or worse.

5. Verification of the Probabilistic Visibility Forecast

Verification (validation) of the new model / algorithm has been be done by comparing the results with the past. Meteorological data, on which the forecasts in the past have been based, were collected and stored. From this data selection "new" LVP probability forecasts have been calculated using the new and improved model / algorithm (see Appendix A).

A good probabilistic forecast should at least be reliable: For a large number of events the forecasted probabilities of an event should correspond with its observed frequency (see appendix B). Graphically this correspondence can be presented in a "reliability diagram". Figure 8 shows the reliability for the short term forecasts for LVP-phase A or worse of the old TAFG (blue), the new TAFG (red) and the old SKV (green), combined for all 8 issue times and 3 lead times (+3, +6, +9).

As the old SKV is based on the forecasters own interpretation of the old TAFG, there is no such data available yet for a "new" SKV based on the new TAFG. This data can only be generated and collected during day to day operations.

Numbers in the diagram depict the number of cases within the bins of fixed width (10%). Most cases fall in the first bin (0-10%), their number exceeds 20000. For layout purposes these numbers have been removed from figure 8 and figure 9.

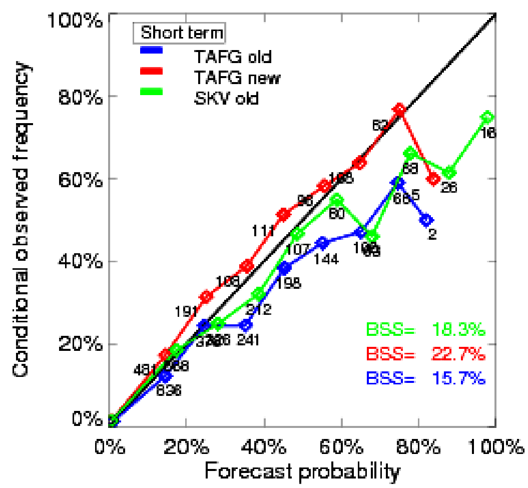


Figure 8. Reliability diagram for short term forecast of LVP <= phase A (3 - 9 hours ahead).

As the diagonal in the diagram represents the theoretical optimal reliable forecast it is clear that the new TAFG is much more reliable than the old TAFG. Except for the highest bin (80-90%), for the new TAFG all points are almost perfectly on the diagonal.

The old TAFG shows overforecasting which can partly be explained by the fact that the probabilities in the old TAFG are in terms of MOR where the actual observations are in terms of RVR. This overforecasting is only partly compensated by the forecaster: The SKV is slightly more reliable than the old TAFG. On the other hand, both TAFGs hardly produce probabilities exceeding 80%. The forecasters are though able to issue reliable higher probabilities than the old TAFG, which indicates that they add value to the old TAFG by issuing more distinct forecasts. Even stronger overforecasting is seen for phases B and C/D (see appendix D) and in general we may conclude that forecasting probabilities for short term phase C/D is at the edge of skillful forecasting at the moment. Also note that overconfident probability forecasts may result in False Alarms and too often declaration of flow restrictions.

In the reliability diagrams also the Brier Skill Scores (BSS, see appendix B) for the forecast systems are shown. The BSS represents the quality of a forecast compared to a reference forecast where 100% is a perfect forecast and 0% means not better than climatology. For LVP-phase A or worse the BSS for all 3 systems is positive, where the BSS for the new TAFG is the highest (22.7%, but on dependent data). Note that the BSS for the SKV is higher than the BSS for the old TAFG, again an objective confirmation of added value by the human forecaster compared with the automatic TAFG system.

Figure 9 shows the reliability diagram for the long term forecasts of the SKV and TAFGs. Compared with the short term, overforecasting by both TAFG and SKV is stronger; forecasted probabilities are too high. All forecasted probabilities of the TAFG above 30% are affected by this overforecasting. Again this may largely be explained by the MOR versus RVR difference between the old TAFG forecast and the observations. This effect is not seen for Moderate and Good visibilities since MOR and VIS observations are in general the same for these visibility conditions.

Again the improvement in reliability by the new TAFG is seen. The red curve stays near the diagonal whereas the old TAFG and SKV again show heavy overforecasting. As with the short term forecasts, the SKV is slightly better than the old TAFG; but both show very low Brier Skill Scores.

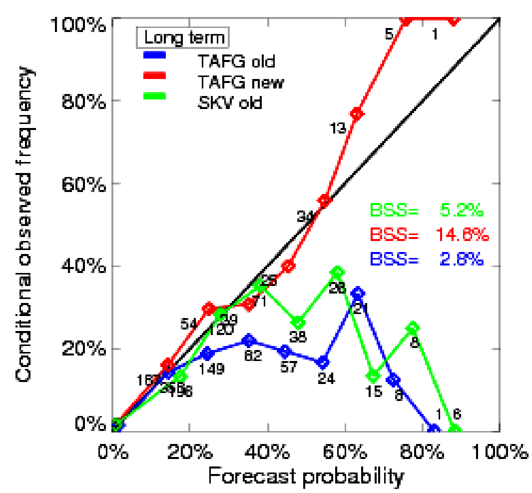


Figure 9. Reliability diagram for long term forecast of LVP <= phase A (10 - 31 hours ahead).

6. Decision Support Application

6.1 Construction and Verification of a Deterministic Forecast

In the operational practice at ATC, at a certain moment a choice has to be made between the four different categories from the probability forecast for LVP. Figure 10 shows an example of such a probability distribution in a graphical form. Note that no distinction is made here neither between "Good" / "Moderate", nor between "C" / "D". The former because the G-M boundary has no operational implications with respect to flow restrictions, the latter because the meteorological forecast of "D" has hardly any skill.

An objective strategy to choose from the probability forecast is the following:

- Take the category that coincides with a fixed percentage P. In the example P=50 would lead to LVP-phase A, while P=25 would lead to LVP-phase B.

The forecast of an LVP category, based on a P=x value, is matched with the actual observed LVP category. For all forecasts in the database (see appendix A) this results in pairs of forecast and observation of deterministic LVP categories which can be summarized in the 4x4 verification/performance matrix below (tables 4 and 5). This procedure has been applied to both the old and the new TAFG. It is important to remember that the construction of the probabilities for LVP differs for the old and the new TAFG (as explained in section 4.2.3).

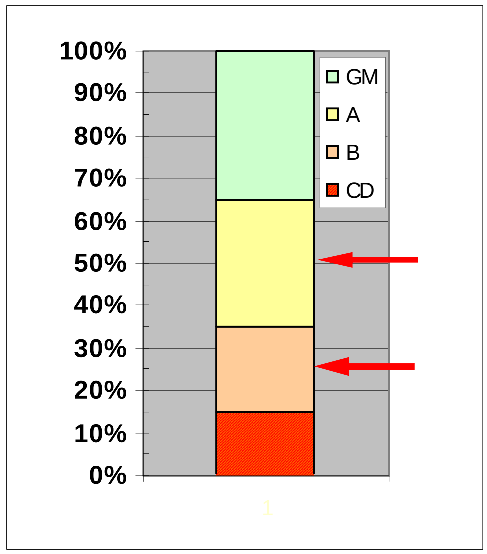


Figure 10: Example of a forecasted cumulative probability distribution for LVP from the TAFG/SKV. Both red arrows indicate example choices of 25 and 50%.

Table 4 and 5 show the verification matrices with the FC-OBS (Forecast - Observation) pairs for all forecasts from the TAFGs of 02 UTC valid at 06 UTC. In total 1449 forecasts are used (long data set). Left and right from the slashes denote results from the old/new TAFG. Table 4 shows the results for P=50 from the probability distribution. Table 5 for P=25. Comparing tables 4 and 5, events tend to shift to the left in the matrix with P=25; overall resulting in fewer Misses but more False Alarms.

Forecast old/new TAFG (p=50)						
		C	B	A	GM	
Observation	C	0/0	2/3	1/0	1/1	4
	B	8/0	9/21	10/5	19/20	46
	A	2/0	7/4	2/4	23/26	34
	GM	3/0	9/6	21/3	1332/1356	1365
		13/0	27/34	34/12	1375/1403	1449

Table 4. Performance matrix of deterministic LVP categories (old / new TAFG 02 +4; P50.
Purple: FC and OBS in same class → Hit
Blue: FC better than OBS → Miss
Red: FC worse than OBS → False Alarm

Forecast old/new TAFG (p=25)						
		C	B	A	GM	
Observation	C	3/2	0/1	1/1	0/0	4
	B	14/0	11/29	11/6	10/11	46
	A	7/0	7/10	7/4	13/20	34
	GM	12/1	10/13	59/18	1284/1333	1365
		36/3	28/53	78/29	1307/1364	1449

Table 5. Performance matrix of deterministic LVP categories (old / new TAFG 02 +4; P25.
Purple: FC and OBS in same class → Hit
Blue: FC better than OBS → Miss
Red: FC worse than OBS → False Alarm

An alternative presentation of tables 4 and 5 is given in figures 11 and 12. It shows the change in the number of False Alarms, Misses and Hits for the old versus the new TAFG. Note that the Hits in the bottom right cells (GM-GM) are not included in figures 11 and 12, but since for both situations the number of GM-GM cases increases in the new TAFG this can be regarded as an increase in Hits. Most prominent difference is a sharp decrease in the number of False Alarms in the new TAFG, visible in both the P=50 and P=25 graphs. This is in-line with the conclusion from the reliability diagrams that the old TAFG showed too much overforecasting, where the new TAFG is much more reliable. The decrease in False Alarms is accompanied by an increase in Hits while the number of Misses does not change much.

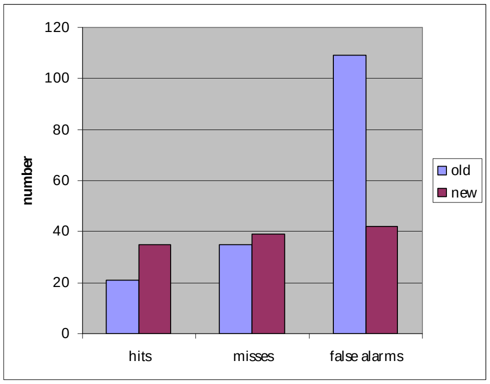


Figure 11. Comparison between old and new TAFG. Expressed in Hits, Misses and False Alarms. (TAFG 02 +4; **P25**) GM-GM cases not included.

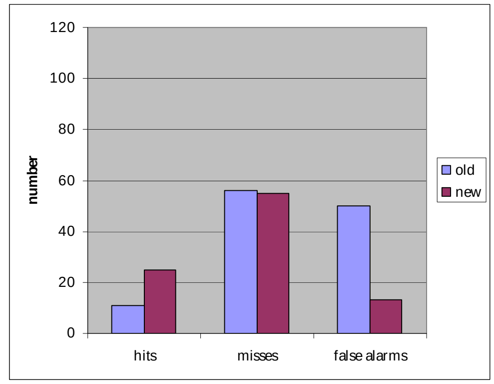


Figure 12. Comparison between old and new TAFG. Expressed in Hits, Misses and False Alarms. (TAFG 02 +4; **P50**) GM-GM cases not included.

As is clear from the differences between figures 11 and 12, the balance between Hits, Misses and False Alarms can be optimized through the choice of the value of P. It is therefore important to carefully determine the value(s) of P to be used in daily operations. Complications may well occur here since different stakeholders may opt for different balances, which implies that for each one a different value for P may be more practical.

6.2 Expense Analysis

Although not part of the project definition, the project team sees it as their obligation to explicitly point out the meaning and the usefulness of the percentile threshold (P) for the decision making process related to low visibility conditions. This chapter describes how probabilistic forecasts may support the decision making process. Although presented for LVP forecast here, the approach may also be applied to probabilistic forecasts on wind, de-icing etc.

The probabilistic forecast gives information on scenarios to be expected. However in the daily operation these probabilities have to be transformed into a specific decision (see previous sections). The question is; what is the optimal percentile threshold to be used in daily operations?

Fa/Mi = 1

Observed LVP category	Forecast LVP category			
	CD	B	A	GM
CD	0	1	10	100
B	1	0	1	10
A	10	1	0	1
GM	100	10	1	0

Fa/Mi = 0.5

Observed LVP category	Forecast LVP category			
	CD	B	A	GM
CD	0	1	10	100
B	0.5	0	1	10
A	5	0.5	0	1
GM	50	5	0.5	0

Table 6. Examples of expense matrices.

To determine the optimum percentage threshold for decision making with the lowest expenses, more information about the cost of Misses and False Alarms must be obtained. The matrices in table 6 show example fictitious expenses for every pair of forecasted and observed LVP category.

As no detailed cost information was available the expenses of Misses and False Alarms have been indexed. In practice, pairs in the outer (upper right and lower left) corners are often more expensive events than those closer to the diagonal. The "extra costs or damage due to wrong forecast" should be assigned to every cell in the matrix, assuming that cells on the diagonal are perfect forecasts which give no additional cost (unavoidable expenses).

It is assumed that the cost of Misses is equal to the cost of False Alarms (Fa/Mi ratio = 1) however it is possible that Misses are more expensive than False Alarms (Fa/Mi ratio < 1) or the other way around (Fa/Mi ratio > 1). Table 6 shows examples of both an expense matrix where Misses and False Alarms are equally expensive (left matrix) and an expense matrix where Misses are twice as expensive as False Alarms.

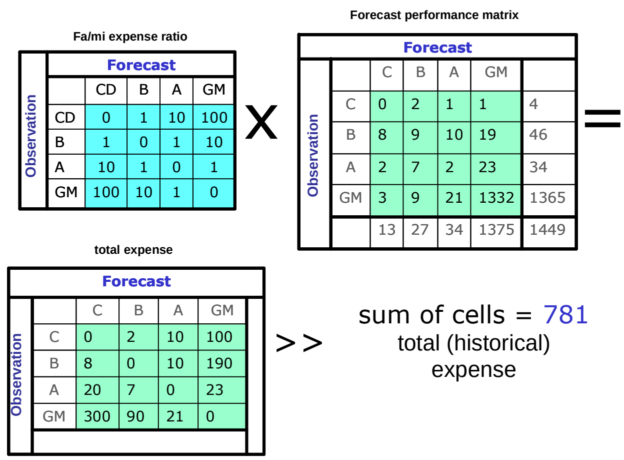


Fig 13. Calculation of total expense for percentage threshold P50, Fa/Mi ratio = 1 and assumed / indexed expense matrix.

Multiplying cell by cell the 4x4 verification/performance matrix (see also section 6.1) by an expense matrix and summing up all cells of this multiplication gives an indication of the (user) specific extra costs/expenses. In figure 13 this process is graphically presented. Consider e.g. the lower left corner, which represents cases where the forecast was LVP-phase C but the observed condition was only M(arginal) or G(ood), i.e. a strong false alarm. This situation is assigned an expense of 100. In the total dataset it happened 3 times. So for the whole period these forecasts contribute 300 to the total historical expense of 781, which is clearly the largest contribution to the total as can be seen from the lower diagram in figure 13.

The total expense can be determined for any value of P, for both the old and new TAFG. As an example, table 7 shows the total expense for the old and the new TAFG for P=50 and P=25 values, assuming Fa/Mi ratio = 1.

From this table the benefit of the new TAFG is clear. The expense reduces with 25 - 50% in comparison with the old TAFG, dependent on the threshold percentage.

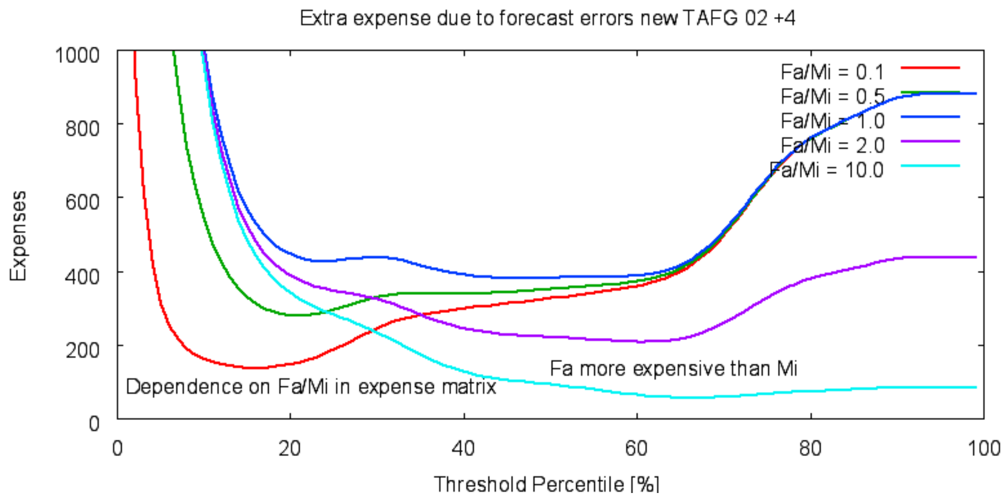
Expense for TAFG02 +4

choice	Old TAFG	New TAFG
P=50	781	401
P=25	1584	405

Absolute "cost" much lower for new cf. current

Table 7. Total expense for old and new TAFG and for P=25 and P=50. Assuming Fa/Mi ratio = 1.

Where there is a significant expense difference between P25 and P50 for the 'old' TAFG, this difference is hardly visible for the 'new' TAFG. This is dependent on the Fa/Mi ratio. If the Fa/Mi ratio changes also the benefit of the new TAFG will change. In case a false alarm is more expensive than a miss, the difference between the old and the new TAFG will increase. However if a miss is more expensive the



Red:	**Fa/Mi=0.1**	**Px min 15-20%**
Green:	**Fa/Mi=0.5**	**Px min 20-25%**
Dark blue:	**Fa/Mi=1**	**Px min 25-70%**
Purple:	**Fa/Mi=2**	**Px min 55-65%**
LightBlue:	**Fa/Mi=10,**	**Px min 60-80%**

Figure 14. Expense as a function of threshold percentage for different Fa/Mi ratios.

difference will decrease. The improved TAFG shows an increased number of Hits and a reduced number of False Alarms. The number of Misses remains approximately the same.

For every user (ATC, Airlines, airport operator,…) the values in the expense matrix are different. This means that for the same forecast each user (or group of users) should make their decisions on their own estimated expense matrix and resulting threshold percentage (P) to get their optimal results.

In figure 14 the effect on the total expense of some example Fa/Mi ratios (expense matrices) and threshold percentages (P-values) is given. From this figure it becomes clear that for a user with Fa/Mi = 2 the optimum is found near P=60%. Note that the calculations in this section are done for the TAFG LVP forecast of 02 UTC valid at 06 UTC. Other combinations of issue time and lead time may give different optimal P-values since:

1) The sensitivity of the airport operation to the LVP forecast (i.e. the numbers in the expense matrix) varies during the day and,
2) The TAFG performance matrices depend on issue time and lead time.

Although filling the expense matrix will not be an easy task, it is a good way to optimize the information from the probabilistic forecast to the user's specific needs.

7. Results and Project Deliverables

7.1 Summary of results

An update for the Probabilistic Forecast Schiphol (SKV) has been developed. This update has improved the accuracy and reliability of the low visibility and ceiling forecast and has brought it in line with the operational criteria on the use of Low Visibility Procedures (LVP) at Schiphol Airport. Also the number of Hits has increased where the number of False Alarms has seen a large reduction.

For the SKV improved visibility forecasts are achieved due to:

- Joint probabilities for visibility and ceiling,
- New and closer up-stream sites,
- In addition to MOR: RVR.

In addition, modifications to the SKV are:

- LVP-phase C added,
- Extension of the forecast period with 6 hours,
- Long and short term forecast period combined.

An expense analysis shows how a probabilistic forecast can support the decision making process. Although no detailed information was available on avoidable operational cost, it could be shown that a clear reduction in cost, due to imprecise forecasts, can be achieved. The improvement to the TAFG and the SKV may well reduce this cost to less than half the operational cost related to the current TAFG and SKV. This benefit is strongly dependent on the actual cost of inaccurate low visibility forecast. It is assumed that the cost of a false alarm is equal to the cost of a miss. If this ratio changes also the benefit will change.

7.2 Compliance with project requirements

The results of the changes to the low visibility forecast tool has been compared with the project requirements as described in the project proposal/definition (see also paragraph 2.2).

- "Hit"- rate of 60% on forecasts 3 hours in advance.

  The Hit rate as formulated in the proposal, defined as Hits / (Hits + Misses) (see appendix C), can be reached easily. However, the "hit"- rate as intended in the project definition is equal to the Critical Success Index (CSI). Where for the hit-rate this 60% requirement is reached, for the CSI this is not. For the CSI a 60% score is almost impossible because of the very low occurrence of LVP conditions (LVP-phase A, B, C and D less than 5%). In appendix C the dependency of the CSI on the occurrence of the forecasted condition is discussed in more detail. Verification of the new TAFG shows a significant improvement in CSI value compared to the old TAFG; e.g. the CSI for LVP-phase A or worse improved from 0.27 to 0.46 (P25) (see also Appendix E). The project team judged this to be a success.

- Forecast must be made 24 hours in advance (preferably 36 hours).

  The forecast period is extended with 6 hours in the improved TAFG and presented in the new SKV resulting in a forecast period ranging from 3-31 hours ahead.

- To be easily integrated into operational systems of LVNL, KLM and AAS.

  The new and improved forecast is successfully integrated into the systems of LVNL and AAS. For the CPS tool of the KLM it is only partly achieved: To assist the decision making process within day to day operations, the CPS tool (Capacity Prognosis Schiphol, see also http://cdm.klm.com/cps/default.asp) has been developed by KLM. The effectiveness of this tool is improved by implementing the more accurate and reliable information with respect to the visibility forecast. However the extensions of the improved low visibility forecast are not available in the CPS tool. Changes to the CPS tool will have to be made to get the maximum benefit out of the improved TAFG and new SKV.

7.3 Implementation

The improved TAFG and SKV have been implemented within the operational departments of KNMI, LVNL and KLM-OCC on May 26th 2008. Starting at that day, the forecasts for Schiphol Airport are based on the improvements achieved within the KDC - LVP project. On June 2nd 2008 the implementation has been formalised.

However the CPS tool of KLM-OCC still works with its old input format. Adaptation of the CPS tool will take place on a later date.

7.4 Revised format for Probability Forecast Schiphol

The SKV can be seen as the operational presentation of the final output of the forecast cascade (see also figure 6 in chapter 4)

With respect to the old SKV, the SKV incorporates the following changes:

- Combined short and long term probability forecasts; one SKV for Schiphol. The old SKV has two parts (a short and a long one) the new SKV consists of one part,
- Addition of LVP-phase C. Probabilities for RVR less than 350 meters (LVP-phase C) are given,
- Extension of the forecast range with 6 hours,
- More accurate and more reliable probabilities for LVP conditions.

Saturday 15 January 03 UTC till Sunday 16 January 06 UTC

utc:	03	04	05	06	07	08	09	12	15	18	21	24	03	06
Visibility < 5 km and/or ceiling < 1000 ft (%)	60	70	80	90	90	80	40	20	5	5	5	10	30	50
RVR < 1500 m and/or ceiling < 300 ft (%)	30	40	50	50	50	40	10	5	0	0	0	0	0	10
RVR < 550 m and/or ceiling < 200 ft (%)	15	20	25	30	30	20	5	0	0	0	0	0	0	0
RVR < 350 m (%)	5	10	15	20	20	10	0	0	0	0	0	0	0	0
Wind direction (deg)	160	160	160	160	160	160	160	170	180	190	230	240	240	240
Wind speed (kt)	5	4	5	5	5	6	7	9	9	9	10	12	12	13
Gusts (kt)											15	17	18	19
Standard deviation wind direction (deg)	30	30	30	30	30	25	25	20	20	20	20	15	15	15
Standard deviation wind speed (kt)	2	2	2	2	2	2	2	2	2	2	2	3	3	3
Temperature (°C)	1	0	-1	-1	-1	-1	0	2	5			5	4	4
Dewpoint (°C)	1	0	-1	-1	-1	-1	0	0					2	2
Snow (%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Moderate or heavy snow (%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Freezing rain (%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Remarks	Short term	Long term
Visibility and ceiling		
Wind		
Temperature/dewpoint		
Precipitation		

Last update: short term 00.10 utc, long term 22.50 utc

Table 8. Revised format of Probability Forecast Schiphol (SKV).

7.5 Project Deliverables

Next to the updated and implemented changes to the TAFG and changed presentation in the SKV, the project also has some other related deliverables. The first is this final report which has been written by the project team. Once more independent verification material is available, KNMI will produce a scientific report in which the improvements are described in more detail.

To support the implementation of the new SKV and the transmission of information towards the operational departments of LVNL, KLM-OCC, AAS and KNMI a leaflet has been produced. This information leaflet is available for distribution among other stakeholders.

A tool to present information on conditional climatology in an accessible way was developed. Under low visibility circumstances, an important question for further operations is: "when does this situation change and how does it change?". Conditional climatology can give a first and quick answer but does not replace the information given in the TAFG and SKV (see also appendix F).

Summary project deliverables:

- Improved and extended TAFG and SKV,
- Final report on part 1 of the KDC - LVP project,
- Information leaflet,
- Interim report on Climatology,
- Conditional Climatology tool.

8. Concluding Remarks

8.1 Lessons Learned

The project evolved towards an interesting and informative study on meteorological forecasting, operational practice and scientific foundations. Next to the achieved results and deliverables the project also gained interesting information related to:

- Mutual understanding on process and terminology,
- The significance and planning horizon of the different aviation forecasts,
- Availability and refresh rate of meteorological data,
- The difference between MOR and RVR.

The KDC - LVP project itself has resulted in a significant improvement of the visibility forecast but also the gained understanding on the above mentioned issues and the propagation of this knowledge will have a positive influence on the day to day operation/process.

8.2 Recommendations

Based on the findings of the KDC - LVP project part 1, it is recommended:

- To initiate part 2 of the KDC - LVP project, wherever practicable, in connection with the KvK/KBS (Kennis van Klimaat / Klimaat Bestendig Schiphol) project in September 2008,
- To adapt the CPS tool to make best use of the improved low visibility forecasts,
- To improve the decision making process by optimizing the benefit of the probabilistic forecast by determining the decision threshold(s),
- To investigate whether the TAFG and SKV should be extended to more than one location at the airport,
- To determine an optimal update frequency and temporal output resolution of the visibility forecast, in agreement with users from LVNL, KLM and AAS, but also meteorological meaningful.

8.3 Part 2 of the KDC - LVP Project

For part 2 of the LVP project a project definition has to be made. Part 2 of the KDC - LVP project will focus on medium/long term improvements including fundamental research towards causes of low visibility as well as possible prevention.

- The use of high resolution 3D models,
- Introduction of new sensor technology,

  Examples are Satellite and IR images, LIDAR (Laser Imaging Detection And Ranging).

- New algorithms/models,

  To improve the physical modelling of fog, e.g. through application of a 1 Column model.

- Fog sensitivity chart, aerial planning guidelines (water, forest, municipalities, etc.),

  The creation of a low visibility sensitivity chart for Schiphol Airport as well as aerial planning guidelines with respect to the development, existence and dissipation of low visibility (fog) will also be part of the project.

- Possible Adaptations to the SKV will also be examined.

  This follows from the recommendations of part 1. Dedicated SKV's for specific parts of the airport (i.e. the Polder runway 18R-36L) can also be one of the practical results from the use of high resolution models in combination with (more) local observations and will be analysed on their applicability. In addition, investigate an appropriate update rate, forecast resolution and presentation format. However this can only be useful if other meteorological elements are also taken into account.

First an inventory will be made of all known possibilities and technologies. The applicability for the location Schiphol will be analysed as well as a cost/benefit and realisation time frame. A priority list will be made with recommendations for the next step(s) to be taken.

Including part 2 within project scope of the KvK/KBS (Kennis voor Klimaat / Klimaat Bestendig Schiphol - Dutch for "Knowledge of Climate / Climate Resistant Schiphol"), may provide the necessary exchange as projects like "Improved Wind Forecasts" and "Meteo Server" will be executed under the supervision of KvK/KBS.

8.4 Acknowledgement

Although not part of the initial project plan the project team is very pleased to see the improved TAFG and SKV be implemented within the meteorological forecast suit and being operational from may 26th 2008 (formalized June 2nd 2008), well before the formal end of the KDC - LVP project part 1.

A word of thanks goes to everyone involved in this project, not only the members of the project team but also those at the KNMI who worked behind the scenes, and those at KLM, LVNL, AAS and NLR who provided the project team with valuable information and support.

Acronyms

AAS	Amsterdam Airport Schiphol (Schiphol Group)
ATC	Air Traffic Control
ATM	Air Traffic Management
BGL	BackGround Luminance
BZO	Beperkt Zicht Operaties (Reduced Visibility Operations)
CLB	Cloud Base (ceiling)
CPS	Capacity Prognosis Schiphol
CSI	Critical Success Index
FAR	False Alarm Ratio
FC	Forecast
HIRLAM	High Resolution Limited Area Model
HR	Hit Rate
ICAO	International Civil Aviation Organization
ILS	Instrument Landing System
KDC	Knowledge & Development Centre
KLM	Koninklijke Luchtvaart Maatschappij
KNMI	Koninklijk Nederlands Meteorologisch Instituut (Royal Netherlands Meteorological Institute)
LIDAR	Laser Imaging Detection And Ranging
LVNL	Luchtverkeersleiding Nederland (Air Traffic Control the Netherlands)
LVP	Low Visibility Procedure
MAS	Meteorological Advisor Schiphol
METAR	Meteorological Aviation Routine weather report
MOR	Meteorological Optical Range
NLR	Nationaal Lucht & Ruimtevaart laboratorium (Dutch National Aerospace Laboratory)
OBS	Observation
OCC	(KLM) - Operations Control Centre
RVR	Runway Visual Range
SKV	Schiphol Kans Verwachting
SYNOP	Report of surface observations of a land station
TAF	Terminal Aerodrome Forecast
TAFG	TAF Guidance
VIS	Visibility

Appendices

Appendix A Verification data

In the comparison between the old TAFG, the new TAFG and the SKV we have used a short and a long dataset with forecasts and observations, where the long one completely overlaps the short one. For both TAFG systems and the observations we have data from 2003/05/01 until 2007/04/30. For the SKV we have data between 2004/07/07 and 2007/04/30. In the verification, in principle we used the second dataset when the SKV is considered or compared with the other data. The first (long) dataset is used in chapter 6 concerning the decision-making.

The current TAFG data is from the automatic operational TAFG system, running at KNMI. Note that in normal circumstances, every nine months an update of the regression equations in this system is performed. However, the updates to the TAFG system under the flag of the KDC - LVP project are such that we call it a new system. The regressions of this new system have been applied to historic data (so-called re-computations) which gives a total of 4 years of (dependent) data. "Dependent" here means that data is taken that is used in the development of the system.

The observational data is from the autosynop position near runway 27. Visibility (MOR) measurements are done with a "Forward Scatter" instrument. The same instrument measures the Background luminance (BGL). Combination of MOR and BGL results in the Runway Visual Range (RVR) which is extensively described in vd Meulen (1993). Ceiling height, the height where the cloud cover coverage is 5/8 or more, is measured with a LIDAR instrument. Combining visibility and ceiling height, applied to the thresholds in table 3 (chapter 3) results in the observed LVP-phase.

Reference

- J.P. van der Meulen (ed), 1993. Runway Visual Range - Observing and reporting practices in the Netherlands. KNMI publication, 100 pp.

Appendix B Probabilistic Verification Measures

Reliability

Reliability of the forecast means that when an event has e.g. a forecasted probability of 30%, in fact in 30% of these cases it really happens and in 70% of these cases it does not. Of course, the contribution to the error is large when a forecast that an event will happen is 99%, but it does not occur (or vice versa). On the other hand, given a forecast of 99%, nothing is wrong with 99 occurrences out of a set of 100 forecasts.

In a reliability diagram the forecasted probabilities are plotted (in bins) against the observed frequency. In the most ideal case, all points should lay on the x=y diagonal. Points below the diagonal suffer from over-forecasting (probabilities too high) and points above the diagonal indicate under-forecasting (probabilities too low).

Resolution

Conditional on reliability, a forecast should also have resolution. Forecasts with a good resolution are as close as possible to 0% or 100%. In the reliability diagram this leads to many data points in the lower left corner and/or in the upper right corner.

The Brier Score (BS)

is very similar to the mean-square-error which is a common verification measure. For a certain dichotomous (yes/no) event it assesses the (squared) difference between the forecasted probability of the event and the observation (event happened = 1, not happened = 0).

$BS = 1/N \sum_{i=1}^{N} (P_f - P_o)^2$ where P_f is the forecasted probability, P_o the observed value (0 or 1) and N the number of cases. BS=0 is perfect and BS=1 the worst.

The Brier Skill Score (BSS)

expresses the quality of the forecast of an event relative to a reference forecast. In this project the sample climatological probabilities are used for the reference forecast. Maximum BSS = 100%. When BSS drops below 10% the value added by the forecast relative to climatology becomes marginal.

Appendix C Scores for Yes-No forecasts

With a 2-category forecast (yes/no) an event happens or does not happen. It can be evaluated using a 2 by 2 contingency table (see below).

		Forecast	
Observation		Yes	No
	Yes	A	B
	No	C	D

Different scores can be derived from this table:

Hit Rate or Probability of Detection (HR/POD; A/A+B) The HR gives the fraction of the observed "yes" events which were correctly forecasted.

False Alarm Ratio (FAR; C/A+C) The FAR gives the fraction of predicted "yes" events that actually did not occur (i.e. were a false alarm).

Probability of False Detection (POFD; C/C+D) The POFD gives the fraction of wrong forecasts given the event did not occur.

Critical Success Index (CSI; A/A+B+C) The CSI measures the fraction of observed and/or forecasted events that were correctly predicted.

Dependency of verification scores on the observed frequency

The value of a score at itself says almost nothing about the often desired qualification "good" or "bad". Many verification measures depend very much on the observed frequency of the event (see Kok, 2000). An example for the CSI is given below.

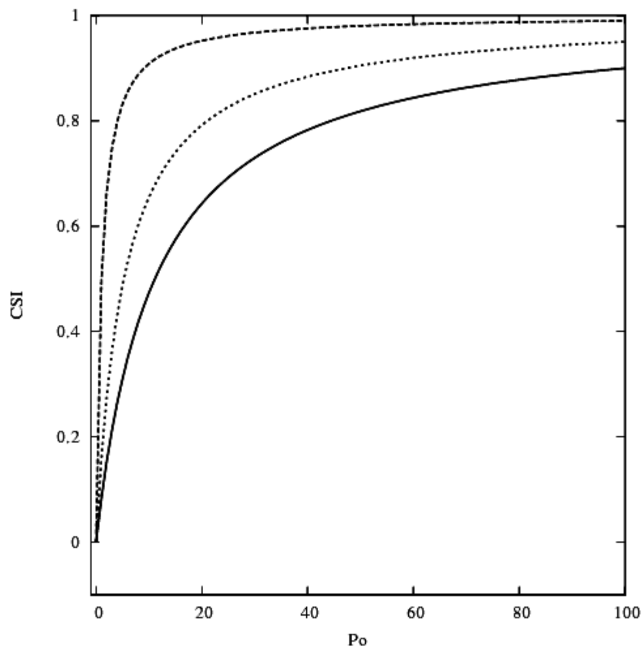


Figure C-1. Dependency of the CSI on observed frequency (Po) and accuracy (99% (dashed), 95% (dotted) and 90% (solid)).

In the figure above the relation between the observed frequency (Po) and the CSI for different accuracies is given (99%, 95% en 90%). 95% accuracy means in this case that 95 percent of the events and the non-events are predicted correctly. It is clear that with decreasing Po – given e.g. an accuracy of 95% - the maximal attainable CSI reduces. In this perspective a forecast for LVP-phase A or worse with a CSI of 0.46 is a relatively good forecast. On the other hand a CSI around 0.5 for an event with an observed frequency of about 1 (the event almost always happens) is only considered "extremely bad", since it can easily be obtained by mere chance.

References:

- C.J. Kok, 2000. On the behaviour of a few popular verification scores in yes/no forecasting. KNMI scientific report: WR 2000-04, 73 pp.
- D.S. Wilks, Statistical methods in the atmospheric sciences. Academic Press, 464 pp.

Appendix D Probability forecast verification

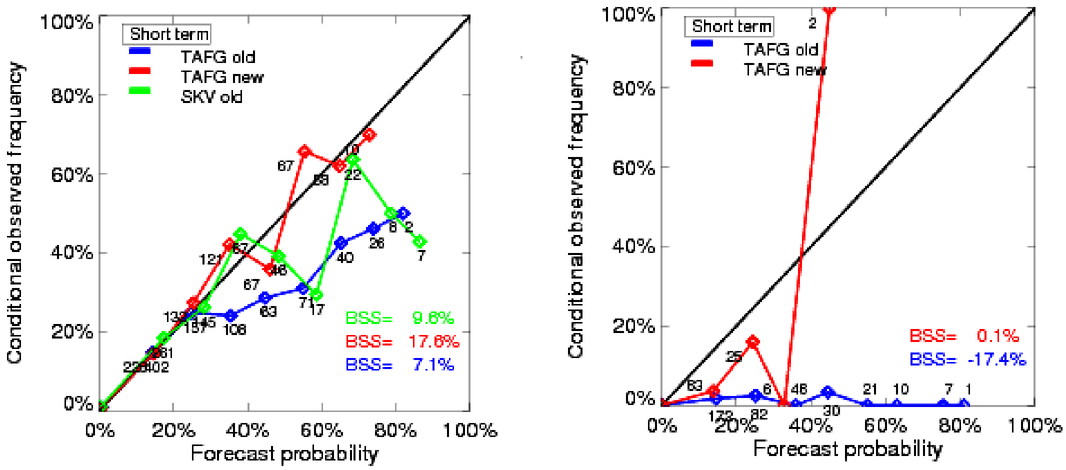


Figure D-1. Reliability diagrams for the events LVP-phase B or worse (left panel) and LVP-phase C or worse (right panel) for the short term forecasts (+4,+7,+10).

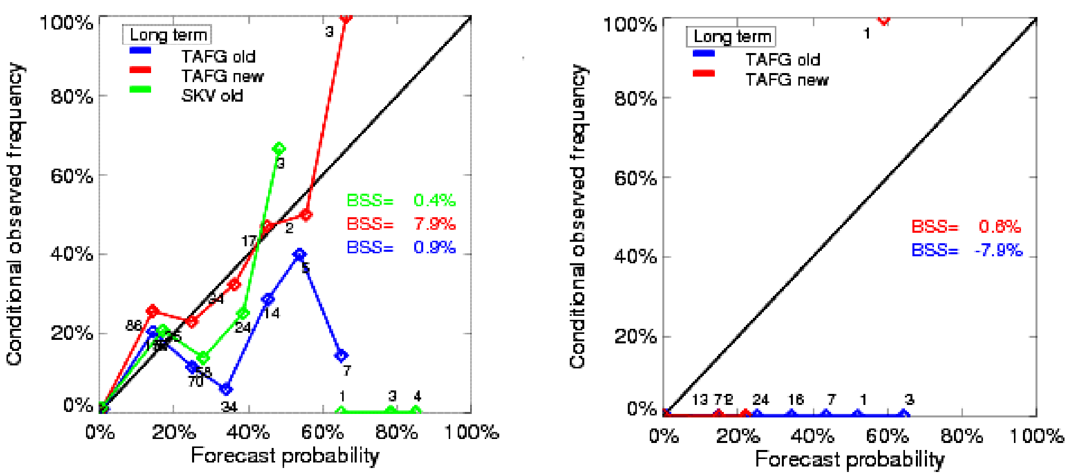


Figure D-2. Reliability diagrams for the events LVP-phase B or worse (left panel) and LVP-phase C or worse (right panel) for the long term forecasts (+9,+12,+15,+18,+21,+24.).

Appendix E Categorical Verification of LVP

Concerning the airport capacity, flow restrictions come in force when LVP-phase A or worse is observed / forecasted. We therefore now concentrate on the forecast that the LVP conditions will be worse or equal to A. This is a yes/no type forecast (for more details see appendix C). Then table 4 can be summarized into table E.1 (where cells are added). Note that in table E.1 the values still depend on the P=50 threshold to construct the deterministic forecast for visibility and ceiling. Also, several verification scores for the events in table E.1 are calculated. Appendix C gives an explanation of these categorical verification scores. The hit rate (HR) of 0.49 means that for all situations that LVP-phase A or worse conditions were observed at 06 UTC; the TAFG of 02 UTC also forecasted these conditions at 06 UTC in 49% of the cases. Of course then the other 51% of the forecasts would Moderate or Good visibility. The false alarm ratio (FAR) of 0.45 on the other hand says that in 45% of the situations that LVP-A or worse conditions were forecasted for 06 UTC it did not occur, and in fact Moderate or Good visibility was observed. Finally a CSI (Critical Success Index, see appendix C) of 0.35 can be interpreted such that of all situations that LVP-A or worse conditions were forecasted and/or observed 35% was right forecasted. Note that when comparing old and new TAFG, the HR gets worse while both FAR and CSI improve. This illustrates that drawing conclusions about model improvements should in general not be based on a single measure alone.

		Forecast old/new TAFG		
		yes	no	
Observation	yes	41/37	43/47	84
	no	33/9	1332/1356	1365
		74/46	1375/1403	1449

HR = 0.49 / 0.44
FAR = 0.44 / 0.20
CSI = 0.35 / 0.40

Table E-1. Verification of deterministic LVP categories. "LVP-phase A or worse" conditions as taken from the probability distributions at (P=50). TAF guidance +4 (lead time: 4 hours) issued 02 UTC. Blue values are Misses, red are False Alarms and violet are the correct forecasts. Left values:old TAFG, right values: new TAFG.

The same exercise as with P=50 has been done with P=25 (see table E-2). The HR and CSI are then clearly improved compared to the P=50 scores, on the other hand at the same time the FAR is somewhat higher.

Forecast old/new TAFG				
Observation		yes	no	
	yes	61/53	23/31	84
	no	81/32	1284/1333	1365
		142/85	1307/1364	1449

HR = 0.73 / 0.63
FAR = 0.57 / 0.38
CSI = 0.27 / 0.46

Table E-2. As table E.1 but for P=25.

Dependence of the scores on percentile threshold choice

As shown before, the verification scores depend on the threshold-percentiles which are applied to determine the forecasted LVP-class. We can explore this further by calculating the scores for a varying threshold-percentile (between 1 and 99). As can be seen from Figure E-1, for all scores maxima/minima appear. A clear feature is a strong coupling between hit rate and false alarm rate that both run from high values for low thresholds to low values for higher ones. CSI, which takes into account both the Hits, False Alarms and the Misses, shows a maximum at a threshold of about 25%, although in the range of 20% to 50% the CSI is rather insensitive to the applied threshold percentile. Since the different measures have their optimal value at different percentiles, this suggests that the "optimal percentile" depends on the chosen measure which is not a desirable property.

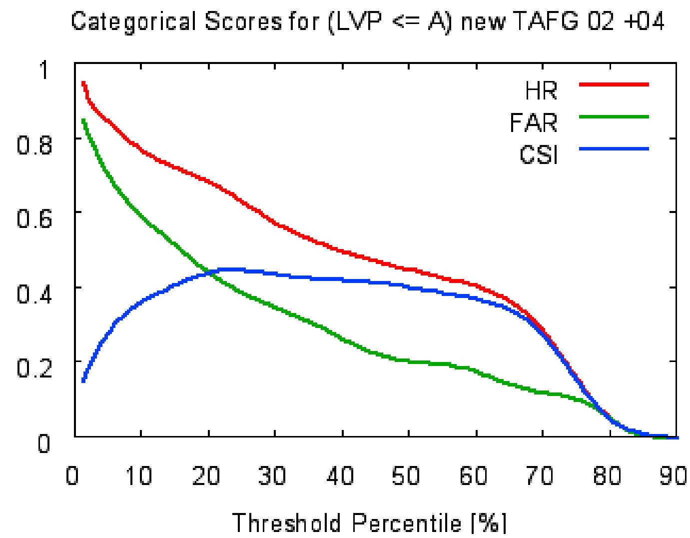


Figure E-1. Hit Rate (HR), False Alarm Ratio (FAR) and Critical Success Index (CSI) as a function of threshold percentage for the event LVP-phase A or worse for the new TAFG 02 UTC +4 forecasts.

In general verification scores deteriorate with increasing lead time. But, as is discussed in appendix C, beside on lead time, most verification scores also depend strongly on the climatological occurrence of the event. Since a forecast can do nothing about the climatological occurrence of the event it is essential to know this dependence for the correct interpretation (and comparison) of the forecast scores. This feature is illustrated in Figure E-2. The CSI for the event LVP-phase C or worse is much lower than that for the other two events. This difference can largely be assigned to the difference in climatological occurrence for the three events which implies different "degrees of difficulty" regarding the forecasts for the three events. This creates a perspective for the (original) project requirements where targets of 60% and 30% "Hit rate" were set for short and long term forecasts respectively.

The improvement in CSI for LVP-phase C or worse is remarkable though: Around P=25 a distinct optimum is found. For the long term forecasts (see figure E-3), as expected, CSI scores are much lower compared to the short term forecasts. This difference can fully be assigned to a usually found decrease in forecast skill with higher lead times.

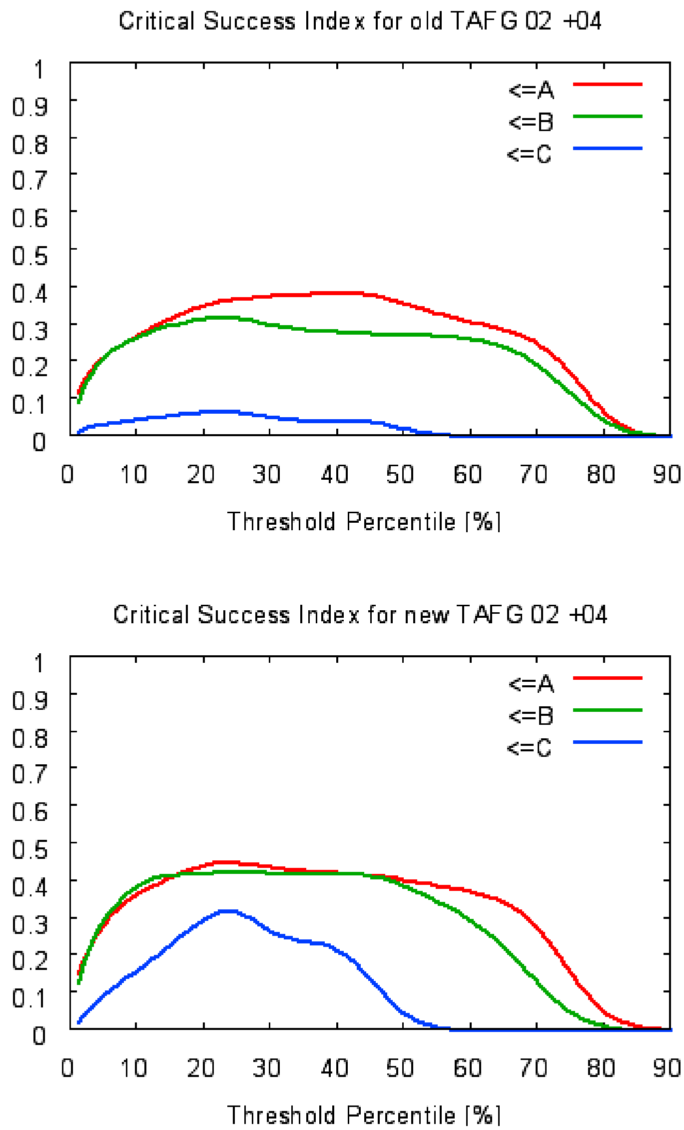


Figure E-2. Critical Success Index (CSI) as a function of threshold percentage for the events LVP-phase A or worse (red), LVP-phase B or worse (green) and LVP-phase C or worse (blue) for the old (upper panel) and new (lower panel) short term TAFG 02 UTC +4 forecasts.

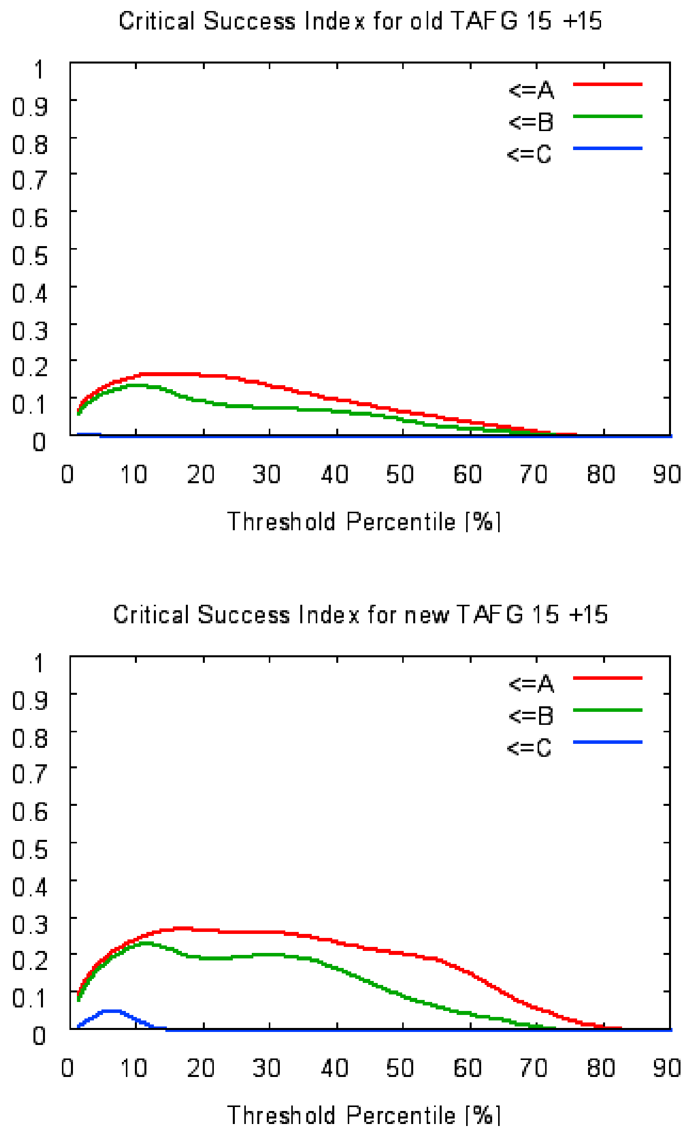


Figure E-3. As figure E-2 but for the long term TAFG 15 UTC +15 forecasts.

Appendix F Conditional Climatology

For conditional climatology location, time of year, hour of the day and of course the starting conditions are relevant. An example is shown if figure F.1, valid for January at the location Schiphol Airport. Starting condition is LVP-phase B at 6 UTC. The chance is about 50% that after one or two hours the condition improves. There is also a chance, be it significantly smaller, for worse conditions. More detailed information on climatology can be obtained from the interim report "Low Visibility and Ceiling Forecasts at Schiphol, Part 1 - Assessment of the current system".

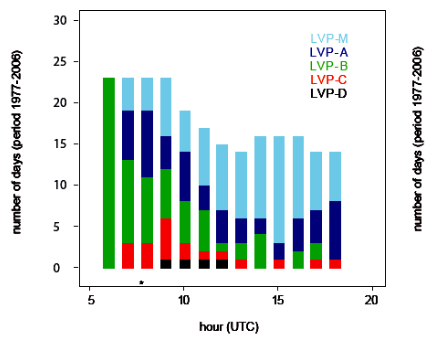


Figure F-1. Conditional Climatology for Schiphol starting at 06 UTC in January. Initial condition LVP-B.

The Conditional Climatology tool is currently available on http://www.knmi.nl/samenw/kdclvp/mist but will in the near future also become accessible through www.eham.aero. Within this tool, location (i.e. other airports in the Netherlands), time of year, hour of the day and starting conditions can be easily altered.

Reference

- Low Visibility and Ceiling Forecasts at Schiphol, Part 1 - Assessment of the current system, KNMI, October 2007.