Numerical Guidance Methods for Decision Support in Aviation Meteorological Forecasting

A. J. M. JACOBS AND N. MAAT

Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands

(Manuscript received 7 October 2003, in final form 25 June 2004)

ABSTRACT

Numerical guidance methods for decision making support of aviation meteorological forecasters are presented. The methods have been developed to enhance the usefulness of numerical weather prediction (NWP) model data and local and upstream observations in the production of terminal aerodrome forecasts (TAFs) and trend-type forecasts (TRENDs) for airports. In this paper two newly developed methods are described and it is shown how they are used to derive numerical guidance products for aviation. The first is a combination of statistical and physical postprocessing of NWP model data and in situ observations. This method is used to derive forecasts for all aviation-related meteorological parameters at the airport. The second is a high-resolution wind transformation method, a technique used to derive local wind at airports from grid-box-averaged NWP model winds.

For operational use of the numerical guidance products encoding software is provided for automatic production of an alphanumeric TAF and TREND code. A graphical user interface with an integrated code editor enables the forecaster to modify the suggested automatic codes. For aviation, the most important parameters in the numerical guidance are visibility and cloud-base height. Both have been subjected to a statistical verification analysis, together with their automatically produced codes. The results in terms of skill score are compared to the skill of the forecasters' TAF and TREND code. The statistical measures suggest that the guidance has the best skill at lead times of +4 h and more. For the short term, mainly trend-type forecasts, the persistence forecast based on recent observations is difficult to beat. Verification has also shown that the wind transformation method, which has been applied to generate 10-m winds at Amsterdam Airport Schiphol, reduces the mean error in the (grid box averaged) NWP model wind significantly.

Among the potential benefits of these numerical guidance methods is increasing forecast accuracy. As a result, the related numerical guidance products and encoding software have been integrated in the operational environment for the production of TAFs and TRENDs.

1. Introduction

Aviation meteorological forecasters at the civil airports in the Netherlands, in cooperation with weather observers at the airports, are responsible for accurate weather observations and forecasts. The meteorological observations and forecasts are provided on a routine basis to the appropriate air traffic service authorities. Timely information on sudden changes in weather conditions is needed to ensure that safety and efficiency are guaranteed at all times during flight operations. For reasons of efficiency these meteorological observations and forecasts are produced and distributed in an alphanumeric coded form according to regulations from the International Civil Aviation Organization (ICAO 1998) and the World Meteorological Organization (WMO 1998). Airport observations are provided every half hour as "meteorological aviation routine weather reports" (METARs). Weather forecasts are provided for the short term, up to 2 h in advance, as trend-type forecasts (TRENDs), and for today and tomorrow as terminal aerodrome forecasts (TAFs).

TAFs and TRENDs contain detailed forecast information on wind, visibility, cloud amount, cloud-base height, and precipitation in the vicinity of the airport. The TAF is produced 12 times a day: 8 short TAFs with lead times from 1 to 10 h, and 4 long TAFs with lead times from 8 to 26 h. The lead times are with respect to the issue times of forecasters' codes. The TREND is added to the METAR and therefore produced every half hour. The TREND is valid for the next 2 h. TAFs and TRENDs are produced for all airports in the Netherlands. However, forecasters at the Royal Netherlands Meteorological Institute (KNMI) only produce them for the four civil airports.

The TAF and TREND codes should be relatively short, but at the same time should contain all the rel-

Corresponding author address: A. J. M. Jacobs, Royal Netherlands Meteorological Institute (KNMI), P.O. Box 201, 3730 AE De Bilt, Netherlands.
E-mail: Albert.Jacobs@knmi.nl

evant meteorological information on future changes in weather conditions. Changing weather conditions are included in the TAF code by the weather change indicators BECMG, TEMPO, and FM, and the probability of occurrence indicator PROB. In the TREND code, only the weather change indicators BECMG and TEMPO are used; PROB indicators are not allowed. The change indicators describe definite (BECMG), temporary (TEMPO), or abrupt (FM) weather changes. The probability of occurrence indicator gives the likelihood (in tens of percent) that a weather change will occur. Only probabilities of 30% (PROB30) and 40% (PROB40) are used. Lower probabilities are not considered to be sufficiently significant, while higher probabilities should be indicated by use of a change indicator. When no significant changes are expected to occur in the TREND, the indicator NOSIG is used.

The production of aviation meteorological forecasts is based mainly on forecasters using numerical weather prediction (NWP) model data in combination with available observations. Although advancements in NWP modeling have been substantial over the last decade, these models have not reached a state where clouds and precipitation can be resolved at the spatial and temporal resolutions needed for airport weather forecasts. Various physical processes, such as those associated with fog and low stratus clouds, are not adequately described in the NWP model due to the complexity of those processes and the lack of sufficient vertical resolution in the lower planetary boundary layer (PBL). As a result, sudden weather changes on small spatial scales can only be analyzed and predicted if the forecaster has access to detailed observation information regarding recent and current weather developments. In particular, the quality of short-term forecasts, up to 6 h, depends mainly on the availability of local and upstream observations. Besides these observations of atmospheric conditions, high-resolution topographical information, such as land-use characterizations and detailed positioning of various geographic objects, contributes to a better understanding of near-surface processes affecting the fluctuations in the weather conditions. For the forecaster, it is difficult to objectively integrate the data originating from these different, often independent, sources. Nevertheless, his or her efforts finally must result in a detailed analysis and forecast of the evolution of the atmosphere in the vicinity of the airport.

Research and development activities have been initiated with the aim of developing numerical guidance methods to compensate for the deficiencies of the NWP model by integrating more objectively the available observations with existing model data. The main purpose of these activities was to develop and implement methods to increase efficiency in the production of aviation weather forecasts in an operational environment while increasing the accuracy of the final product. Basically, two methods have been developed that can be used to derive numerical guidance products with estimates of the wind, visibility, significant weather, cloud-base height, and cloud amount, on the local airport scale. The purpose of this paper is to describe the numerical guidance methods and their related numerical guidance products, and to show their potential in aviation meteorological forecasting.

The first numerical guidance method developed is based on statistical and physical postprocessing of NWP model data and observations. This method is used to derive numerical TREND guidance for the short term and numerical TAF guidance for the longer term. The guidance products, containing all the site-specific forecast information needed for the production of TAFs and TRENDs, are produced to support the forecaster. In the development phase of this method, the pattern followed is similar to that of most of the worldwide developments in the (semi-) automation of aviation meteorological TAF production (Lynn 1997; Kilpinen 1994). It is a combination of direct model output (DMO) from a NWP model, physical postprocessing of DMO, and model output statistics (MOS; Glahn and Lowry 1972). However, instead of using only local observations from the forecast site, as most MOS-based systems do, the approach presented in this paper uses observations from upstream locations as well. A combined physical–statistical approach is often used to forecast adverse weather conditions related to poor visibility and low clouds. In order to encourage European cooperation on the development of forecast techniques in this working area, in 2001 the European Union (EU) has established a 5-yr program on short-range forecasting methods of fog, visibility, and low clouds [European Cooperation in the Field of Scientific and Technical Research Action 722 (COST-722); information available online at http://cost.cordis.lu/src/home.cfm]. For practical use, encoding software has been provided that translates the TAF and TREND guidance into an automatically produced alphanumeric code. The TAF and TREND guidance and the AUTOTAF encoding software have been developed and implemented in close cooperation with the German company Meteo Service Weather Research (Knüppfer 1997). A graphical module for presentation of the guidance parameters, a graphical editor for control and modification of the automatically produced codes, TREND encoding software, and some sophisticated integral parts of the TREND guidance system have been developed at KNMI.

NWP model data forecasts are grid-box-averaged values, valid at the airport and vicinity. However, local differences in land use and surface roughness at the airport can cause local meteorological parameters, particularly wind, to deviate significantly from the grid-box-averaged value. This deviation is called the representativeness error of the model. Recently, a method has been developed that increases the representativeness of local wind forecasts. This, second, numerical guidance method can be used to refine NWP grid-box-averaged model winds to local values. The method is called downscaling and is described in more detail in Verkaik and Smits (2001).

In the remainder of this paper the two numerical guidance methods and related products are described in more detail. In section 2 the automated TAF and TREND production system and the downscaling method are described. An example of a forecast product derived from these numerical guidance methods and arising from a case study is presented in section 3. Besides details on forecasting techniques, the example also demonstrates the value of the numerical TREND guidance in the TREND writing process. In section 4, the opportunities and difficulties of the automatic encoding of the TAFs and TRENDs, such as the linguistic component, are discussed. Verification results for visibility and cloud-base-height forecasts from the numerical guidance and the automatically produced codes (in relation to the quality of the TAF and TREND codes traditionally produced by aviation forecasters) are presented in section 5. A discussion of the potential of these numerical guidance methods for aviation meteorological forecasting, together with some of the possibilities for future improvements, is given in section 6.

2. Numerical guidance methods and products

a. *The TAF and TREND guidance*

1) Model data and observations

For the development of the TAF and TREND guidance, the operational mesoscale (10–20-km horizontal resolution) NWP model HIRLAM is used as background (Undén et al. 2002). The NWP model has 31 vertical levels, of which 10 are in the lowest 3000 m. The NWP model is run every 6 h and produces forecasts with lead times through 48 h on an output time interval varying from 1 to 3 h. The model data used for post-processing consist of upper-air winds, humidity, and temperature at various pressure levels, and the geopotential height, total cloud cover, and precipitation. Furthermore, near-surface parameters of the wind, temperature, and pressure (available at heights of 0, 2, and 10 m above the surface) are used. The development of a MOS-based system requires a huge amount of model data, as much as 3–4 yr of historical field data. For the development of the statistical forecast equations, the model field data are used with a limited output time step of 6 h. For the production of the TREND guidance, in addition to model field data, model time series data (for various near-surface elements in the neighborhood of the airport) on an output time interval of 30 min also are used.

The observation information for the TAF guidance is based on hourly synoptic observations from a network of 24 stations in the Netherlands and nine adjacent foreign stations. The average distance between the stations in this network is approximately 30 km. Figure 1 shows a selection of observation stations from this network surrounding Amsterdam's Schiphol airport. The TREND guidance is based primarily on the half-hourly METAR observations from the airport. Missing weather elements in the METAR observations are supplemented with synoptic observations from the nearest station in the selected network. Observed near-surface variables used in the development of the MOS forecast equations consist of temperature, humidity, precipitation, wind, visibility, total cloud cover, and cloud-base height and cloud cover at multiple layers (with a maximum of four layers). The site-specific observations at the airport are used in the system to monitor the current weather at the airport and to derive local climatology. Neighboring observations are used to account for the upstream weather conditions. Finally, the airport climatology, persistence of the observations at the site, and the conditions observed upstream of the site are factored into the forecasts.

2) Predictands

All the essential aviation-related weather parameters are contained in the TAF and TREND guidance. Besides the standard deterministic information on, for example, air temperature and surface wind, categorical and probabilistic forecasts related to clouds, visibility, and significant weather conditions in the vicinity of the airport are included as well. Table 1 shows a selection of some of the forecast meteorological variables, called predictands, which are present in the TAF guidance.

The categorical predictands for visibility and cloud-base height, recognized by the probability abbreviation "P(. . . < . . .)" in the first column of Table 1, are related to threshold values for visibility and cloud-base height that have been established as international standards by ICAO (ICAO 1998). The ICAO thresholds of these ranked classes, including additional national thresholds, are presented in Table 2. The predictand P (ceiling height <1000 ft) in Table 1, for example, is defined as the probability of the cloud ceiling below 1000 ft, where the ceiling is defined as the first cloud layer with coverage of at least 5 oktas. The guidance also contains deterministic information on visibility and ceiling. The deterministic forecasts are derived from the probabilistic forecasts as the 50% probability of occurrence. In precipitation situations, conditional forecasts for visibility and ceiling can be used, which are also available in the guidance. Furthermore, the guidance contains information on the probability of precipitation with various intensities (light, moderate, or heavy) and the probability of stratiform versus convective precipitation including possible phase types (solid, liquid, or freezing).

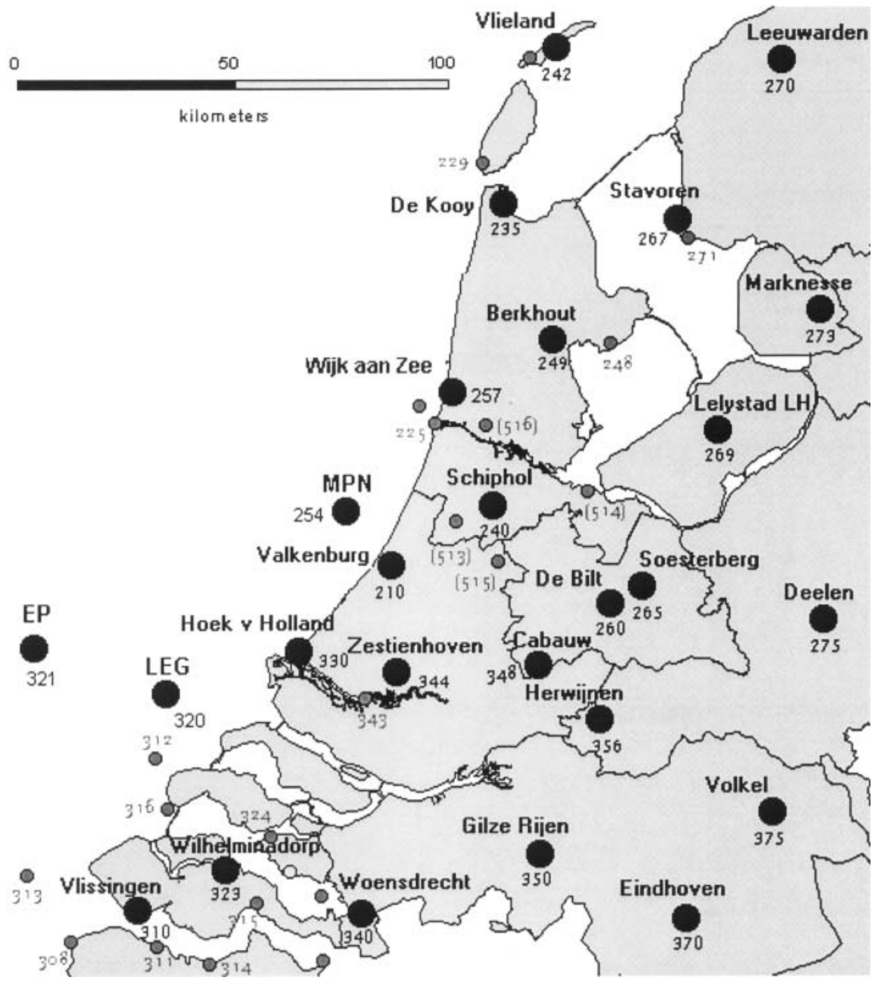

FIG. 1. Map of the synoptic observation network in the western Netherlands surrounding Amsterdam's, Schiphol International Airport (middle) of the figure. The large filled circles indicate the (upstream) synoptic observing stations used for the computation of the guidance. The small dots refer to wind measurement locations.

3) MOS FORECAST EQUATION DEVELOPMENT AND POTENTIAL PREDICTORS

For each predictand, MOS forecast equations are developed depending on season, forecast issue time, and forecast lead time. For this purpose, 4 yr of HIRLAM model forecasts and observations, acquired from KNMI's mass storage archive, are used. The DMO parameters and observed variables (including station climatology) mentioned in section 2a(1) serve as primary predictors in the equation development. Furthermore, additional predictors have been derived from them by linear and nonlinear transformation. Table 3 presents an overview of the most important potential predictors, derived from DMO parameters and observations, which are used for visibility and cloud ceiling forecasts in the TAF and TREND guidance. A more extensive list of predictors, for all guidance parameters, can be obtained from Meteo Service Weather Research who developed the statistical interpretation methods for the TAF and TREND guidance (Knüppfer 1997 and references therein).

For each of the predictands in the guidance, a separate forecast equation is developed by multiple linear regression techniques. The regression algorithm has been optimized to minimize the root-mean-squared error (rmse), defined in section 2a(4), of the predictands. In the regression algorithm, predictors from the potential predictor set are added one by one to the forecast equation, by selecting the one with the highest linear correlation to the predictand in the first step and to the residual (error) in the following steps. The algorithm stops when the next predictor has a correlation to the residual lower than a predefined critical correlation coefficient (Carr 1988).

Additional relative humidity (RH) predictors have been defined for cloud cover and precipitation forecasts on five atmospheric pressure levels. In many cases

TABLE 1. Examples of meteorological variables in the TAF guidance.

Predictand	Observations 21h	Forecast 1 Mar 2001 (UTC) 0600	0900	1200	1500	1800	2100	0000
Temperatures								
Air temp (°C)	1	0	1	2	2	1	1	0
Dewpoint temp (°C)	0	0	1	0	0	0	−1	−1
Visibility (OK = visibility >= 10 km)								
Visibility (100 m)	OK	24	24	37	63	90	OK	88
Visibility in precipitation (100 m)		24	15	14	26	20	27	29
P (visibility < 5 km)	0	69	70	61	42	35	26	30
P (visibility < 3 km)	0	58	58	44	30	22	14	16
P (visibility < 1.5 km)	0	36	38	21	12	6	6	8
P (visibility < 800 m)	0	18	17	5	1	1	1	2
Clouds (cloud cover and cloud-base height)								
Total cloud cover (oktas)	6	7	7	7	7	7	7	7
Ceiling height (ft)	25	12	8	8	18	30	32	34
Ceiling height in precipitation (ft)		11	8	8	11	16	19	24
P (ceiling height < 5000 ft)	100	79	80	79	77	74	70	69
P (ceiling height < 1500 ft)	0	62	75	60	47	32	31	26
P (ceiling height < 1000 ft)	0	39	60	59	44	23	23	12
P (ceiling height < 500 ft)	0	31	30	25	13	14	7	4
P (ceiling height < 200 ft)	0	7	7	1	1	1	1	2
P (ceiling height < 100 ft)	0	7	3	0	1	0	1	2
Precipitation (type and phase)								
P (any precipitation)	0	67	77	65	70	58	60	52
P (stratiform precipitation)	0	61	66	44	47	36	39	30
P (convective precipitation)	0	6	11	21	23	22	21	22
P (thunderstorm)	0	2	1	1	4	1	2	0
P (solid precipitation)	0	33	46	36	32	38	39	40
P (liquid precipitation)	0	34	31	29	38	20	21	12
P (freezing precipitation)	0	2	1	1	4	1	2	0

these predictors are preferred over the DMO cloud cover and precipitation. The smoothed binary predictor RH_Bin in Table 3 mainly supports cloud cover forecasts. The binary predictor is equal to 0 or 1, depending upon whether RH exceeds the predefined cutoff threshold. The predictor is smoothed resulting in a logistic, S-shaped, curve. The value at the threshold is always 0.5 (or 50%) while the slope at the threshold depends on the predictand and the forecast lead time for which the predictor is used. The power function transformation RH_Exp is used mainly for precipitation forecasts. From the original DMO RH and the nonlinear transformed RH_Bin and RH_Exp predictors, vertically aggregated values such as the maximum (RH_Max), average (RH_Ave) and product (RH_Prd) over the five pressure levels are derived.

FogRadTd and FogRadFF in Table 3 are potential radiation fog predictors that are used for visibility and low cloud forecasts. A decrease in dewpoint depression ($T - T_d$) and wind speed (FF) leads to increased predictor values often resulting in an increased probability of poor visibility and low ceilings.

The contribution of upstream synoptic observations for short-term forecasts is included in the MOS system by defining additional advection predictors, denoted as Adv_Trj in Table 3. The advection predictors account for the deviation of the observed predictand at the forecast site, Obs(site), from the observations of that predictand at the five nearest upstream locations, Obs[upstr (i)], $i = 1, \ldots, 5$. A trajectory model is used to compute the path of the airmass displacements. In the trajectory model, HIRLAM model winds at levels of 925 or 1000 hPa are used to compute the trajectory path. Locations nearest to the trajectories starting point are identified as the upstream locations. The advection predictor is defined as

$$\text{Adv\_Trj} = \text{Obs(site)} + \sum_{i=1}^{5} \{\text{Obs[upstr}(i)] - \text{Obs(site)}\} \times W(i), \quad (1)$$

where $W(i)$ is the normalized relative station weight (in percent) for station i. The station weight depends on

TABLE 2. Ranked classes of visibility and ceiling, established by ICAO and additional national requirements. The thresholds used for TAF and TREND verification are shown in bold face.

Visibility threshold values (range, m)	Ceiling threshold values (range, ft)
150, 350, 600, **800, 1500, 3000, 5000,** and **8000**	100, **200, 500, 1000** and **1500**

TABLE 3. Potential predictors in the TAF and TREND guidance, derived from DMO parameters and observations.

Predictor group name	Input variables	Atmospheric input levels (hPa), or surface	Transformation	Threshold (T)/ exponent (E)
RH_Bin	DMO relative humidity	1000, 925, 850, 700, 500	Smoothed binary	T = 70%, 90%
RH_Exp	DMO relative humidity	1000, 925, 850, 700, 500	Power function	E = 2, 6
RH_Avr	DMO relative humidity	1000, 925, 850, 700, 500	Vertical average	—
RH_Prd	DMO relative humidity	1000, 925, 850, 700, 500	Vertical maximum	—
RH_Max	DMO relative humidity	1000, 925, 850, 700, 500	Vertical product	—
Rotation	DMO geopotential height	1000, 500	Laplacian operator	—
Grad_ThetaW	DMO wet-bulb potential temperature (thetaW)	850	Gradient	—
Sun_Alt_Sin	Sun altitude	Surface	Harmonic	—
FogRadTd	DMO air temperature (T), DMO dewpoint (Td)	Surface Surface	$100/\max(2, T - T_d)$	—
FogRadFF	DMO wind speed (FF)	Surface	100/max(2, FF)	—
Adv_Trj	Observed predictands, DMO wind	Surface, 1000, 925	Linear	—

the geographical distance from the trajectory starting point and the difference in elevation with respect to the forecast site.

4) GENERAL FEATURES OF THE TAF GUIDANCE

The TAF guidance is produced for 20 locations in the Netherlands and issued 1 h prior to the issuance of the forecasters' TAF. Each TAF guidance is based on the latest available HIRLAM model output and recent observations. The quality of the deterministic parameters in the TAF guidance, such as air temperature, surface wind speed, and total cloud cover, is regularly validated in relation to its background (reference) HIRLAM model. For this purpose, the guidance forecasts and HIRLAM forecasts are verified against local synoptic observations. For the deterministic variables, the forecast error (E) is defined as the metric difference between the forecast (F) and the corresponding observation (O):

$$E = F - O. \tag{2}$$

The accuracy of the forecasts is expressed in terms of statistical quantities of the forecast errors: the mean error or bias (ME), the root-mean-square error (rmse), and the variance of the error [var(E)]:

$$\begin{cases} \text{ME} = \dfrac{1}{N}\displaystyle\sum_i E_i, \\ \text{Rmse} = \sqrt{\dfrac{1}{N}\displaystyle\sum_i E_i^2}, \\ \text{var}(E) = \dfrac{N}{N-1}[(\text{rmse})^2 - (\text{ME})^2], \end{cases} \tag{3}$$

where N denotes the total number of pairs of forecasts and observations. The square root of the variance is also known as the standard deviation in the error (SD):

$$\text{SD} = \sqrt{\text{var}(E)}. \tag{4}$$

Based on these verification quantities, the forecast skill of two models can be compared by computing the skill score in terms of reduction of variance (RV). The RV is defined as

$$\text{RV} = \left(\frac{\text{MSE}_H - \text{MSE}_G}{\text{MSE}_H}\right) \times 100\%, \tag{5}$$

where MSE_H and MSE_G are the mean-square errors (square of rmse) of the HIRLAM and TAF guidance forecasts, respectively. In Fig. 2, the RV of the total cloud cover in the TAF guidance is presented in relation to the HIRLAM model for Schiphol airport. The RV values are computed over the period October 1998–May 1999. The resulting RV values in the figure can be interpreted as percentage improvement, in terms of MSE, over HIRLAM. Clearly the guidance has improved skill related to HIRLAM. This is expected, as HIRLAM forecasts are grid-box-averaged values. The deviation of the locally observed cloud

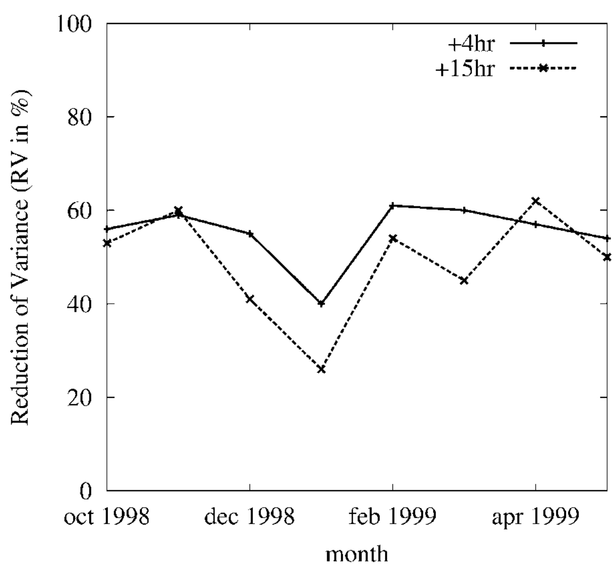

FIG. 2. The RV in the TAF guidance total cloud cover related to the HIRLAM total cloud cover, for lead times +4 and +15 h. Results are for Schiphol airport.

cover from the grid-box average may be large. The purpose of the MOS technique that is applied to develop the forecast equation for total cloud cover is twofold: (a) to increase the local representativeness of the forecast values, and (b) to compensate for systematic errors in the HIRLAM model. Figure 2 shows the combined effect leads to a significant improvement in forecast skill.

5) Advanced features of the TREND guidance

After a first evaluation proved the usefulness of the TAF guidance for decision-making support to the aviation forecaster, it was decided to develop the TREND guidance according to the same principles. Here the key issue was to generate meteorological forecasts of high accuracy for the short term, 0–6 h, on an output time interval of 30 min and issued every 30 min. The METAR served as primary observation input. The TREND guidance is produced for four civil airports in the Netherlands.

The meteorological information content of the TREND guidance is the same as for the TAF guidance. Furthermore, the central guidance component is also based on the MOS methodology. For the development of the TAF guidance, the MOS system was originally single station based. NWP model data are combined with only local observations from the forecast site. Forecasting TRENDs, however, depends more on the monitoring and extrapolation of local weather at and in the vicinity of the airport. In order to integrate the information from neighboring weather observations into the MOS system, the scheme was extended from a single-station approach to a multistation approach. For that purpose, for each forecast site, a surrounding network of observing stations has been selected. The observation network that has been selected for Schiphol airport is shown in Fig. 1. In the multistation MOS system, the contribution of upstream synoptic observations is included by defining the additional advection predictors. The idea of the multistation approach is based on the observation-based forecast system developed by Vislocky and Fritsch (1997) for short-term prediction of visibility and cloud ceiling. They showed that single-station MOS, climatology, and persistence forecasts could be significantly improved for the short term, up to +6 h, by using weather observations from an entire network of stations surrounding the forecast site. Recently, Leyton and Fritsch (2003) showed that an increase in the spatial density of observation networks leads to further improvement in short-term (0–6 h) forecasts of ceiling and visibility.

In relation to standard MOS techniques, the inclusion of advection predictors has further reduced the variance for visibility and cloud base by another 10%–20%. In Fig. 3, the additional RV, estimated for independent data, derived from a development sample of 4 yr (1996–99), is shown for visibility and low clouds (below 1000 ft) for all TREND guidances with an issue time of 0300 UTC for Schiphol airport. In the figure the lead times vary from 30 up to 180 min. Very short-term cloud and visibility forecasts are dominated by persistence of observations from the forecast site. For longer lead times, the impact of persistence decreases and advection predictors become more important. This results in an increase of the RV with time as seen in Fig. 3.

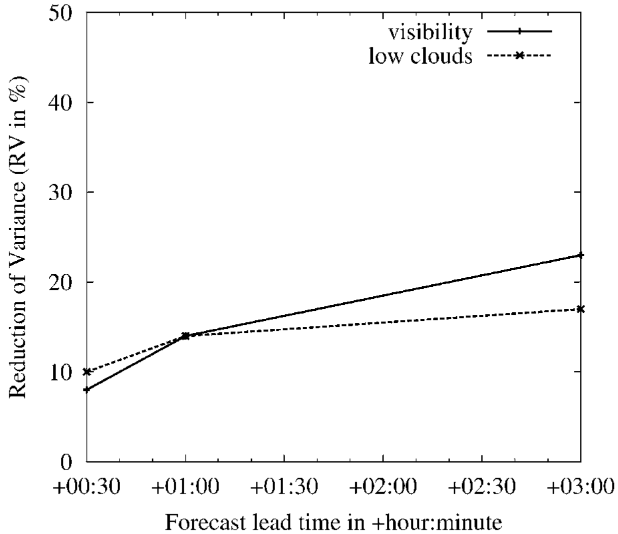

FIG. 3. Additional RV for visibility and low clouds (below 1000 ft) caused by advection predictors. Results are for TREND guidances with issue time 0300 UTC for Schiphol airport. The RV values are estimates for independent data, derived from a 4-yr developmental dataset.

The combined use of these multiple observation datasets proved to be very powerful. Therefore, the multistation approach has been integrated into the TAF guidance as well.

6) Automatic encoding and forecaster interaction

The TAF and TREND guidance contain all the meteorological information necessary to produce the TAF and TREND codes. For both guidance products an encoding system has been developed that selects the most relevant meteorological information from the guidance, groups the different meteorological phenomena, and transforms the forecast information into the actual code. Based on the information in the guidance, an automatically derived formulation for the aeronautical code is suggested. The forecaster, however, makes the final decision on the code. For this, a graphical user interface with an integrated code editor has been developed as a fundamental part of the TAF and TREND production chain. The editor enables the forecaster to control and possibly modify the suggested automatic codes. Finally a syntax program checks the modifications made by the forecaster before the code is disseminated to the user community.

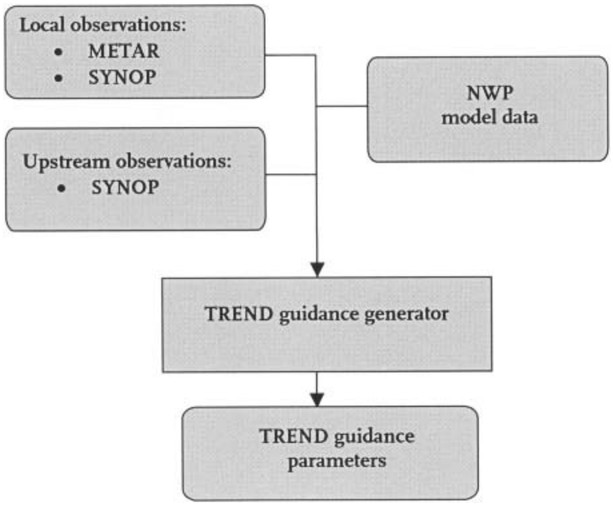

FIG. 4. Automatic generation of the numerical TREND guidance.

The numerical guidance products including its linguistic encoding components have been integrated into the operational production chain to support the human forecaster. The operational systems for TAF and TREND production follow a similar pattern, which basically consists of two modules: (a) the numerical guidance generator, which automatically produces the guidance parameters, and (b) the semiautomatic user interface where the guidance parameters and the encoding component are integrated into the graphical editor. In Fig. 4, the numerical guidance generator for the TREND code is presented. Figure 5 gives an overview of the TREND user interface with the built-in code generator. According to the implemented scheme, forecasters are only allowed to intervene on an alphanumeric code level. Upstream interaction, for example modification of TREND parameters, is not implemented. For TAF production a similar user interface has been put into use.

b. *Downscaling—A method for the computation of the local wind*

Airport capacity can be reduced significantly due to strong surface winds at airport takeoff and touchdown areas. The impact of surface winds on the aircraft depends on the angle between the wind direction and the geographical orientation of the runway. In general, aircraft cannot take off and land if the crosswind and tailwind components exceed certain threshold values. Schiphol airport in the Netherlands has several runways placed at different geographical orientations to cope with different wind directions. Switching runways at the proper time can maximize headwinds and thus minimize the reduction in airport capacity. To support air traffic controllers in regulating the incoming and outgoing air traffic, accurate and timely local wind forecasts at the airport are needed. In particular, information on the average wind speed and wind direction, including possible wind gusts, is required.

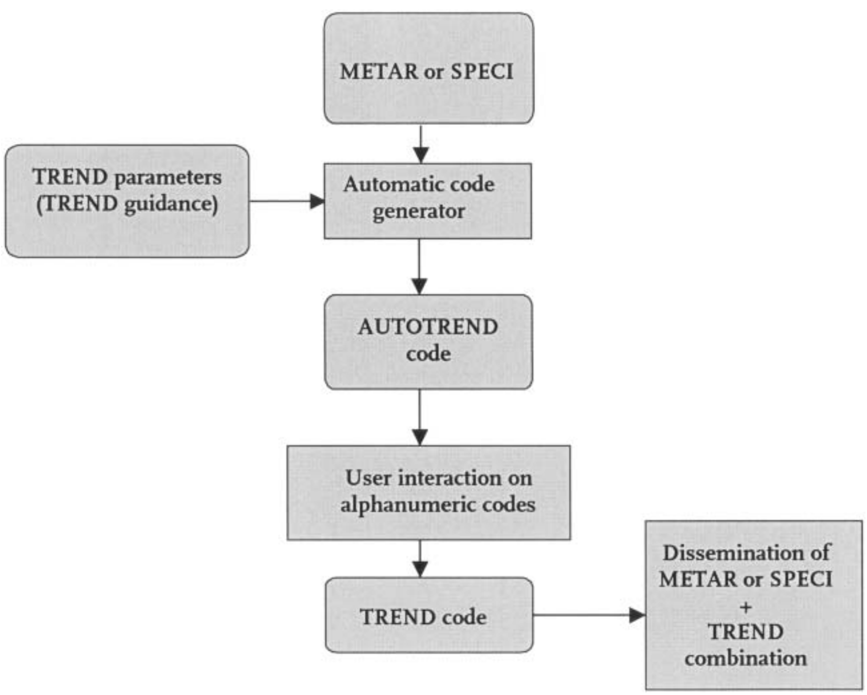

FIG. 5. The TREND code integrated in the user interface.

A high-resolution wind transformation method called downscaling has been developed and tested for the computation of the local average wind on spatially small scales such as airports (Verkaik and Smits 2001; De Rooy and Kok 2004). The method is based on a physical two-layer model of the PBL where the upper boundary condition is provided by NWP model data and roughness information of the surface is derived from high-resolution land-use maps. The downscaling refines NWP model winds near the earth's surface. The method has been validated for the computation of the +3 h forecast of the 10-m wind speed and wind direction at various locations at Schiphol airport. Only moderate and high wind speeds, 3 m s^{-1} and up, have been validated as lower winds are not relevant to aviation. For this validation, model wind fields from the operational mesoscale NWP model HIRLAM were used as input for the downscaling.

Figure 6 shows the verification results of the downscaling method and HIRLAM for the synoptic observation location at Schiphol airport, for different atmospheric stability conditions (unstable, neutral, and stable). In the figure, the ME and SD [from (2)–(4)] in the wind speed errors are presented for each wind direction. Figure 6 clearly shows that downscaling of NWP model winds reduces the bias in the surface winds significantly while the standard deviation in the error reduces slightly. NWP models such as HIRLAM use a grid-box-averaged surface roughness to derive the grid-box-averaged surface winds. For Schiphol airport this grid-box-averaged surface roughness includes the high roughness of the city of Amsterdam and its surroundings. The resulting grid-box-averaged roughness for most locations in the grid box is too high. This leads to a large underestimation of the wind speed for a wide range of wind directions. The deviation of the local wind from the grid-box-averaged value is called the representativeness error of the NWP model. By downscaling the model winds with a more realistic wind-direction-dependent local roughness, the representativeness error can be reduced as is shown by Fig. 6.

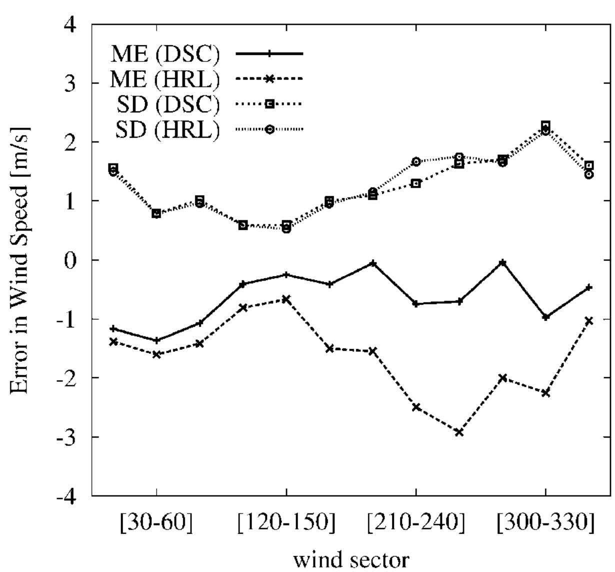

FIG. 6. Wind speed error statistics in the computation of the +03 h forecast average wind speed for Schiphol airport. Bias (ME) and SD of the downscaling method (DSC) and HIRLAM (HRL) are compared. The errors are sampled per wind sector of 30°. The verification period is Nov 2001–Feb 2002.

Currently, aviation forecasters at Schiphol airport are evaluating downscaling of NWP model wind fields. For this purpose, downscaling wind time series for eight wind measurement positions at the airport are produced. First experiences with the downscaling procedure are promising and clearly show potential for the use of this new product in support of operational forecasting. For practical use at the airport, the downscaling wind forecasts are tailored to several more runway specific products. One of these products is the crosswind and tailwind component at the touchdown positions at Schiphol airport. Figure 7 gives an example of a possible crosswind (perpendicular) and tailwind (parallel) forecast, up to +48 h, for one of the touchdown positions (36R) at the airport.

3. The use of advection predictors for short-term forecasts

In this section we describe how advection predictors are applied in the computation of short-term numerical guidance forecasts. In the example presented in section 3a, advection predictors are applied to forecast low stratus clouds at Schiphol airport a few hours in advance. The value of the numerical guidance is demonstrated in section 3b by discussing its role in the TREND writing process.

a. Advection of low stratus clouds: A case study

The general weather situation on 12–13 June 2000 for the Netherlands, as described by the forecaster (from KNMI's FBNL40 EHWX weather bulletin), was as follows:

> "In the evening of 12 June 2000, moist and stable air is advected with a southwesterly airflow. The advected air mass, which mainly affects the western part of the Netherlands, contains cloud fields in the lower levels that can be thick enough to produce some light drizzle. During the night visibilities will gradually deteriorate in mist and locally in fog."

In Table 4, the observed wind, cloud-base height, and cloud cover are given for the nocturnal period 0100–0300 UTC for Schiphol airport and some of its neighboring synoptic observing stations to the southwest. According to the wind observations, the airflow at Schiphol airport was from the southwest. This results in the advection of the low clouds, observed at the up-

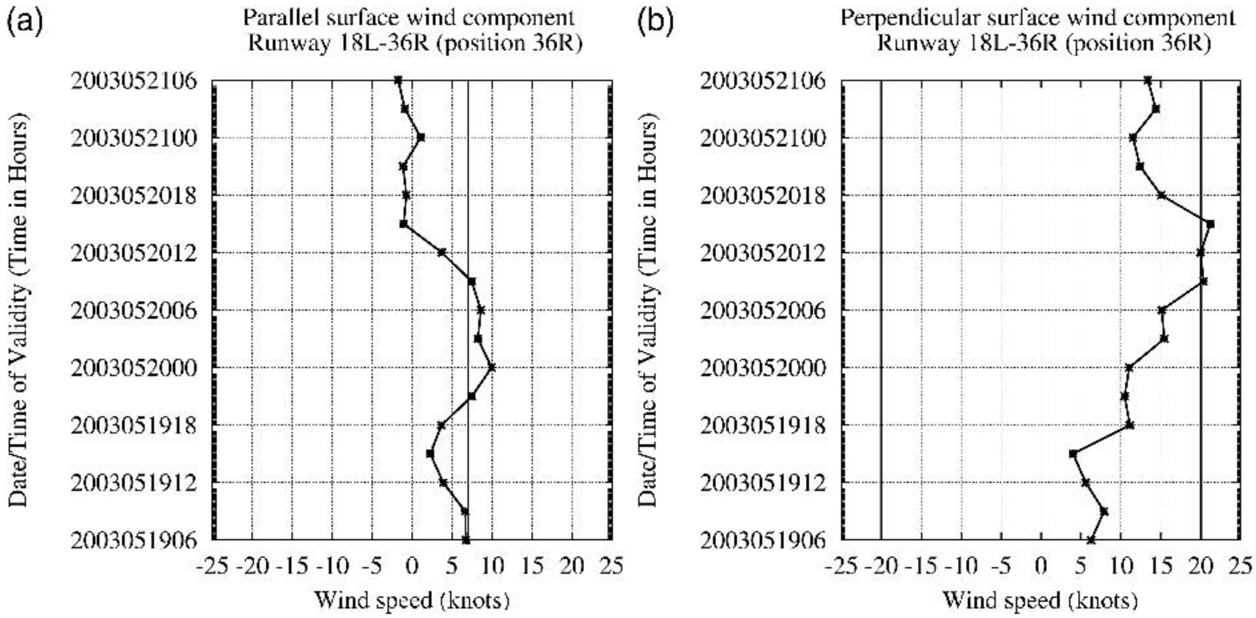

FIG. 7. Forecast (a) tailwind and (b) crosswind at touchdown position 36R at Schiphol airport. The threshold values for tailwind (7 kt) and crosswind (20 kt) are shown by straight vertical lines.

stream locations Rotterdam (airport Zestienhoven), Valkenburg, and Vlissingen, toward Schiphol airport.

The low stratus clouds were already present at the upstream locations at least 1 h before they were observed at Schiphol airport. Figure 8 graphically presents the mechanism of advection of air particles for this example. On the left in Fig. 8, the airmass trajectories for Schiphol airport are presented for the heights 1000 and 925 hPa. In this example, characterized by a low-level wind maximum in a stable (nocturnal) boundary layer, the wind speed at 1000 hPa is higher than the wind speed at 925 hPa. The trajectories are computed by displacing the air particles with NWP model winds, from HIRLAM, at 1000 and 925 hPa, respectively. The source locations of the airmass trajectories arriving at Schiphol airport in 1 and 2 h are marked by squares and triangles, respectively. Note that, according to the figure, the air particles along the 1000-hPa trajectory arriving at Schiphol airport in 2 h come from a source area that is close to Vlissingen. As a result, the observed low clouds at Vlissingen will have a high impact on the clouds at Schiphol airport within 2 h. On the right in Fig. 8, cloud amounts below 1500 ft at Schiphol airport and the upstream area are shown. In the figure, cloud amounts for all locations are summed, but for each location they are depicted by different shadings. Note the low clouds are present at all upstream locations several hours before they are advected to Schiphol airport.

In order to demonstrate the impact of the advection predictors on the short-term forecasts for the low stratus clouds, we have presented a typical MOS forecast equation, in tabular form, in Table 5. The table gives detailed predictor information on forecasting cloud amounts below 1500 ft at Schiphol airport 2 h in advance. In the table cloud amounts are specified in percentage, where 100% equals 8 oktas. In this summer equation only the 925-hPa advection predictor is selected. Upstream locations are presented in order of relevance, according to their relative station weights.

TABLE 4. Observed wind and clouds upstream from Schiphol airport during 0100–0300 UTC at 13 Jun 2000. Wind direction (in °) and wind speed (in m s^{-1}) are separated by the solidus (/). Cloud-base height is given in units of 100 ft with corresponding cloud cover in oktas. Additional cloud layers are separated by the solidus (/).

	Wind direction (°), wind speed (m s^{-1})			Cloud-base height (100 ft)			Cloud cover (oktas)		
Location	1 h	2 h	3 h	1 h	2 h	3 h	1 h	2 h	3 h
Schiphol	210/6	220/5	220/6	200	7	5/6	1	8	4/8
Rotterdam	230/6	240/6	240/6	9/11	7	7	2/6	8	8
Valkenburg	220/5	230/5	230/6	11	7	5/6	6	8	4/8
Vlissingen	220/7	230/7	250/9	7/10	6/10	8/10	3/7	3/8	3/8

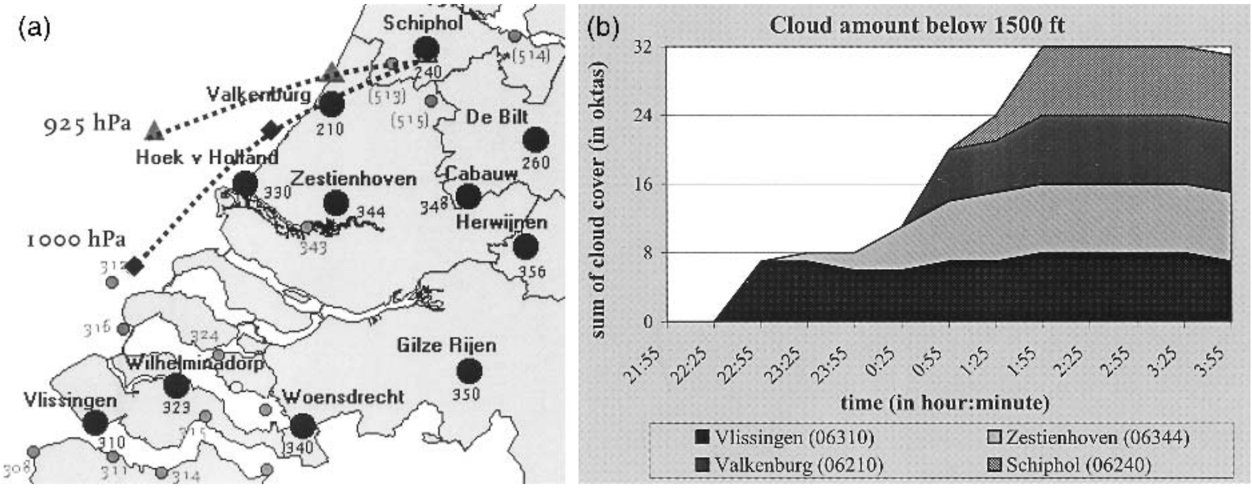

FIG. 8. Trajectories (a) upstream from Schiphol airport and (b) upstream clouds below 1500 ft.

The value of the advection predictor is determined by (1). This predictor, among several others, is finally weighted into the MOS forecast equation. According to the table, the forecast value is mainly determined by the relative humidity predictor RH_1000_90%_Bin (RH_Bin at 1000 hPa and with threshold 90%) and by the advection predictor. Persistence of the latest observation at the forecast site [okta < 1500 ft(−2)Obs] plays only a minor role in the forecast, due to its relative low coefficient. In this case, the contribution is zero due to the observed value.

Figure 9 shows cloud amount forecasts for each ranked class (see Table 2) according to the TREND guidance issued for Schiphol airport at midnight 2355 UTC (0000 in the figure). In the upper part of the figure, the total cloud amount, N, also is presented. In Fig. 10, ceiling forecasts for Schiphol airport are presented for several consecutive TREND guidance runs. Note, the 0055 UTC TREND guidance was the first forecast to indicate the low stratus clouds at Schiphol airport overnight. At the production time of the guidance (0100 UTC), no low clouds, below 1500 ft, were present at Schiphol airport. The guidance forecast 5 oktas of clouds below 1500 ft at 0255 UTC, 2 h in advance (see also Table 5). Almost 80% of the forecast cloud amount was due to airmass advection from Rotterdam, Vlissingen, and Valkenburg.

b. *The TREND writing process*

The TREND guidance is produced periodically every 30 min with the METAR as the primary observation input. The TREND code indicates whether significant changes are forecast to occur with respect to the actual observed weather in the METAR. Significant changes occur if the weather conditions are forecast to change through the threshold values (see Table 2) during the

TABLE 5. Example of a MOS forecast equation, in tabular form, for the short-term prediction of cloud amounts below 1500 ft, in summer.

Location: Amsterdam's Schiphol airport Issue: 0100 UTC 13 Jun 2000	Forecast lead time: +2 h Season: Summer	Predictand: okta < 1500 ft	
Trajectory: Adv_Trj_925	Trajectory start: 3.42 = lon	51.54 = lat	Obs(site) = 0.0 (in %)
Upstream location	Obs(upstr) (in %)	Weight (in %)	Value (in %)
Rotterdam	75.0	35.1	26.4
Vlissingen	87.5	29.7	26.0
Valkenburg	75.0	21.6	16.2
Schiphol	0.0	8.1	0.0
Gilze Rijen	0.0	5.4	0.0
Predictor	Value	Coefficient	Product
RH_1000_90%_Bin	75.7	0.1859	14.0656
Rotation_1000	−12.5	0.1442	−1.8025
Sun_Alt_Sin	−4.9	0.1584	−0.7826
Okta<1500ft(−2)Obs	0.0	0.0907	0.0000
Adv_Trj_925	68.6	0.6910	47.3895
Constant		0.9565	0.9565
		Forecast:	59.8265%

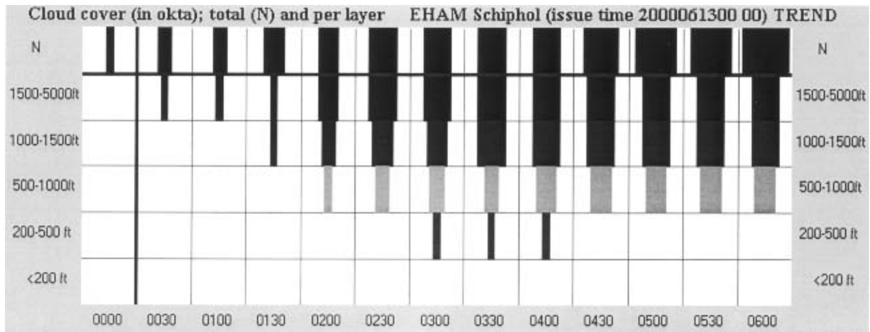

FIG. 9. TREND guidance total cloud cover (N) and cloud cover by layer for Schiphol airport.

2-h TREND period. For practical use, the TREND code should be added instantaneously to the METAR. When a new TREND code must be made, which is at the METAR observation time, the TREND guidance based on this actual observation is not yet available. Therefore the TREND guidance of 30 min prior to the actual observation time must be used. Figure 11 gives an overview of the time schedule that has been implemented for production and usage of the TREND guidance.

Operationally the forecaster has access to each of the consecutive TREND guidance forecasts. Although the 2355 UTC guidance, in the example presented in Fig. 9, shows cloud amounts at lower altitudes are forecast to increase, these changes have no significant impact on the ceiling. Nevertheless, this guidance and the actual weather conditions observed upstream of the airport are a signal for the forecaster to stay alert during the night. The forecaster also has access to a user interface where he or she can consult the corresponding AUTOTREND codes. In Table 6, the observed METAR clouds (cloud amount and cloud-base height) and corresponding encoded AUTOTREND clouds are presented for the consecutive TREND guidances in the example. The 30-min difference between the production time of the guidance and the time of the actual METAR is accounted for in the encoding. Due to this time lag, no significant cloud ceiling changes are indicated before 0125 UTC. Reduction of the time lag would improve the value of the numerical guidance.

According to this example, and generally recalling the additional reduction of variance due to the advection predictors (see Fig. 3), inclusion of advection predictors based on upstream observations is obviously a very successful new technique in short-term cloud cover and cloud-base-height forecasting. Most parameters in the TREND guidance benefit from this new technique, but the greatest gain occurs for visibility and cloud-related elements.

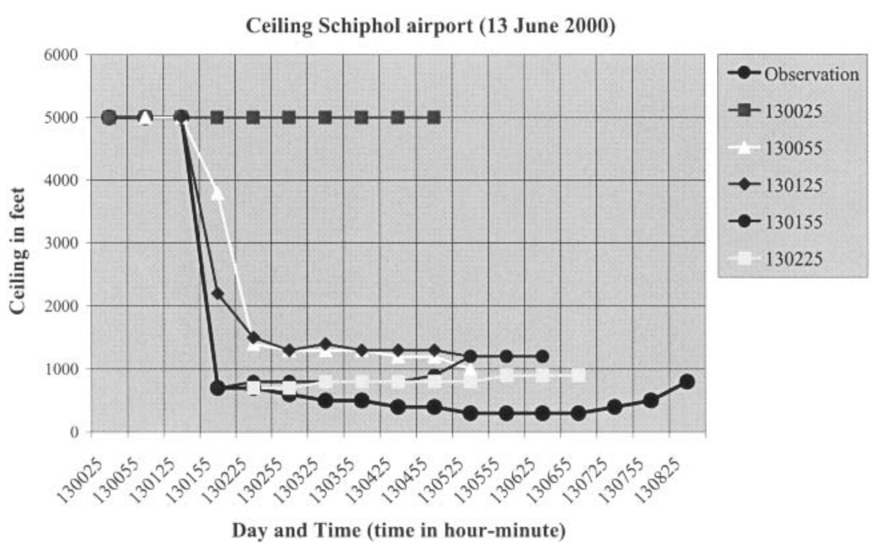

FIG. 10. Deterministic cloud ceiling forecasts in the TREND guidance for Schiphol airport, 13 Jun 2000.

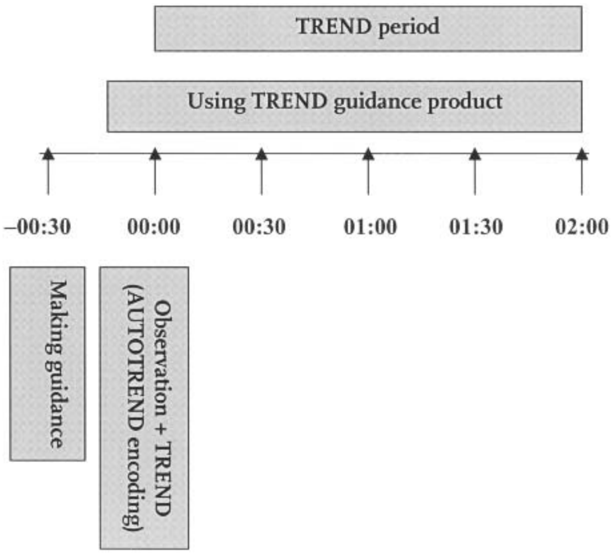

FIG. 11. Schedule for making and using the TREND guidance and AUTOTREND code.

4. Automatic encoding of TAFs and TRENDs

a. TAF encoding

Although the quality of the TAF guidance parameters was satisfactory, the forecasters' initial experiences with the AUTOTAF encoding system were disappointing. The first version of the encoding system complied with ICAO regulations; yet, some national requirements were not fulfilled. Furthermore, due to the encoding of the AUTOTAF, some of the content quality of the guidance was lost. Nevertheless, the TAF guidance and automatic encoding have been successfully implemented in the operational production chain. In 2002 a new version of the encoding system has been implemented and this version complies with ICAO and national regulations. The first verification results of the new AUTOTAF encoding system are presented in section 5.

The German company Meteo Service Weather Research (Knüppfer 1997) developed the first version of the TAF encoder. In recent years this encoder has been improved in close cooperation with KNMI. The process of converting a numerical TAF guidance into an encoded TAF, according to the ICAO regulations (ICAO 1998; WMO 1998), is very complex. The encoded TAF should be relatively short, while the TAF guidance contains a large amount of data. This means substantial data reduction must take place under the condition that the most relevant meteorological phenomena appear in the encoded TAF. One challenge is how to select the most relevant features; another is how to group different phenomena. Aviation forecasters with many years of operational experience are very efficient in grouping relevant phenomena. This has resulted in the development of certain techniques for writing a TAF. For instance, when a frontal system with precipitation passes, the usual technique is to write a general BECMG group where the deterioration conditions for wind, visibility, weather, and clouds are combined, followed by a TEMPO group covering the worst conditions. In some situations, however, PROB groups or PROB TEMPO combinations are used instead. It seems that every synoptic situation has its own personal TAF style.

TABLE 6. METARs and corresponding AUTOTREND codes. CAVOK stands for "cloud and visibility OK," which means visibility >= 10 km, no clouds (any amount) below 5000 ft, no cumulonimbus clouds (any height), and no significant weather. Cloud amounts in the table are specified by FEW (1 or 2 oktas), SCT (3 or 4 oktas), BKN (5–7 oktas), or OVC (8 oktas). Cloud-base height is in 100 ft. BKN013 means 5–7 oktas at 1300 ft.

METAR clouds (valid 13 Jun 2000)		AUTOTREND clouds
Issue time (UTC)	Observed clouds	Encoded clouds
0025	CAVOK	NOSIG
0055	CAVOK	NOSIG
0125	FEW009 SCT011	BECMG BKN013
0155	OVC007	BECMG BKN013
0225	OVC007	NOSIG

Underlying the process of composing a TAF code is the recognition and combination of relevant weather patterns. This process is difficult to automate. Moreover, the international TAF code standard itself leaves room for many encoding possibilities. Sometimes it is useful to use a TEMPO group while a PROB group may be chosen as well. It is very difficult to build an effective decision tree using realistic semantics for the TAF that matches the most relevant information in the guidance. The TAF encoding software that has been developed is partly based on the operational experience of the forecaster. The weather situation, however, is not always easy to determine from the numbers in the TAF guidance, and the problem of multiple solutions also remains.

The TAF encoder is designed to make the TAF code as short as possible. In the configuration settings of the encoder the maximum number of words allowed in the TAF code is specified. Until recently the default encoder settings were used. These settings, that caused significant deviations in conditions that were only valid during the first or the last hour of the TAF, were ignored. In the second half of 2002, the encoder settings were modified, allowing for more words in the TAF code. Significant changes at the edges of the TAF period now appear in the code. Despite all efforts to optimize the settings in the automatic code generator, verification results in the next section show that much skill is lost by converting a numerical TAF guidance into a TAF code.

b. TREND encoding

The encoding of TRENDs is much easier than the encoding of TAFs because PROB groups are not allowed in the TREND code and the lead time, which is only 2 h, is much shorter. Therefore, only TRENDs with NOSIG, BECMG, and TEMPO groups can be produced. The only difficulty is the combination of elements such as wind, visibility, weather, and clouds. The TREND encoder was been built at KNMI and was designed to keep the encoding relatively simple. The use of time statements has not been implemented, resulting in TREND codes with only TEMPO or BECMG groups without time specification. This choice was made as forecasters rarely use time specifications in their TRENDs. The relative simplicity of TREND encoding has resulted in the development of an AUTOTREND encoder, which, contrary to the AUTOTAF encoder, performs quite satisfactorily.

The TREND encoder analyzes transitions between visibility and ceiling classes. These classes are defined by the threshold values established either by ICAO or by national agreements with aviation authorities (see Table 2). The class transition analysis is based on the deterministic values from the numerical TREND guidance. From the guidance, a 2-h period is analyzed. If no changes in classes are forecast, a TREND NOSIG is generated. In other cases, if the transition toward another class is sustained during the 2-h period, a TREND with BECMG is generated. A TREND with TEMPO is produced when a higher or lower class is expected temporarily within the 2-h period. In cases where transitions in both directions occur (improvement followed by deterioration or vice versa), a TREND for the deterioration is generated. For winds of at least 10 kt, a 30° directional change has been established as the threshold value for wind shifts. The threshold for a change in wind speed is also 10 kt. Relevant weather phenomena will also be reported in the TREND following ICAO regulations (ICAO 1998). The TREND for weather phenomena is again based on the numerical TREND guidance.

5. Verification of TAF and TREND forecasts

a. Introduction

In this paper two new methods have been introduced to support the forecaster in the production of TAFs and TRENDs. In order to assess the potential of these methods for aviation meteorological forecasting, an objective verification is necessary. Verification results of the downscaling method, and in particular for application in the computation of the local average wind speed at Schiphol airport, are presented in section 2b. The potential for downscaling NWP model winds in producing TAF and TREND codes has not yet been assessed.

In this section we focus on the assessment of the automated TAF and TREND production system. In this system the numerical TAF and TREND guidance are produced as intermediary products. The final product consists of the automatically produced TAF and TREND code. In order to assess the additional value of the numerical guidance and the automated code in the operational TAF and TREND production, all intermediary and final products are verified. The verification results of these automated products are presented in comparison to the verification results of the forecasters' codes. Because visibility and cloud ceiling are the most important parameters in the TAF and TREND, we will restrict the verification results to those parameters only. Furthermore, in this paper the verification results are presented only for Schiphol international airport.

b. TAF and TREND verification method

The TAF guidance and the TAF code have been verified for visibility and cloud ceiling forecasts using the method proposed by Gordon (1989). In this method, the performance of the forecasts is assessed by calculation of the ranked probability skill score (RPS). A similar approach is suggested by Reid (1978). Aviation forecasts are categorical forecasts, where forecast values are probabilities assigned to ranked categories; therefore, Gordon's verification method is a suitable choice. The boundaries between the categories are threshold values (see Table 2) that are significant to all users of aviation forecasts. Similar to the classical rmse, the RPS score measures the distance between forecasts and observations. In the RPS context forecasts are given a better score, a lower RPS value, if they are closer in ranking to the observed category. RPS values can vary between 0, which represents a perfect forecast, and 1, representing no skill. For TREND verification the same method has been used. The observation data used in the verification are extracted from the METAR bulletins that are provided half-hourly for each civil airport.

c. Verification results for the TAF guidance and AUTOTAF code

The proposed TAF verification method can be used directly to assess the performance of visibility and cloud ceiling parameters in the TAF guidance, because these are essentially probabilistic. The TAF code, however, must first be converted to a probabilistic forecast. This implies that probability values have to be assigned to the weather change groups and the probability of occurrence groups. By definition, probability values of 30% and 40% are assigned to the probability of occurrence groups PROB30 and PROB40, respectively. Clearly, the probability value assigned to the TEMPO group must be less than 50%; otherwise, the use of a BECMG group in the code would be more appropriate, and higher than 0%, because for very low probability values, the TEMPO group has insufficient impact on the forecast. A pilot verification study has shown that the RPS score is insensitive to variations of the assigned probability value in the range 10%–40%. As a result, in the TAF verification method, a 25% probability is assigned to TEMPO groups. The probability values for group combinations can now be computed straightforwardly; for example, PROB40 TEMPO implies a 10% probability. For BECMG groups, equal probabilities are assigned to all categories involved in the weather change that is forecast to occur.

The verification for visibility and cloud ceiling covers the period January 1999–July 2002 for the TAF guidance and January 2001–July 2002 for the AUTOTAF code. In Fig. 12 the RPS values for visibility forecasts at Schiphol airport are presented as moving 12-month averages for the period January 2001–July 2002. In the figure, the results of the encoded AUTOTAF are presented and compared to the results of the intermediary TAF guidance and the final TAF from the forecaster. In the figure, the forecast lead time for all TAF-related products is +5 h relative to the start time of the TAF period. This is at the center of the short TAF valid time window. By this definition, the lead time is the same for each product. However, the effective lead time, the lead time relative to the issue time of each of the products, is different. In the operational production chain the forecaster issues the TAF 1 h before its valid time. The numerical TAF guidance and AUTOTAF code are issued 1 h before the forecasters' TAF is issued. This gives the forecaster sufficient time to consult the guidance and the suggested AUTOTAF code. According to Fig. 12, the TAF guidance has the lowest RPS value of all products and, thus, the best skill. The difference in RPS between the TAF guidance and the TAF of the forecaster seems small, but depending on the use of the products, this difference can be of practical significance. Especially for users of weather-related products in aviation, this small improvement could lead to substantial economic savings. Moreover, the guidance contains a complete forecast probability distribution, which, supplementary to the limited TAF code, has been received by the user community as very useful. This means that at a lead time of +5 h, the numerical guidance is a powerful instrument having the potential to (a) support the aviation forecaster and improve the human skill, and (b) provide the user community with additional, and more detailed, information on forecast weather changes. The RPS differences between the TAF guidance and the forecasters' TAF typically increase with increasing lead time, as will be shown next.

The amount of information reduction due to encoding leads to skill loss. In Fig. 12 this skill loss is shown by the increased RPS values of the encoded AUTOTAFs. According to the figure, the increase is about 20% of the guidance RPS. In general the increase in the RPS value is small compared to the RPS value of the guidance itself. Therefore, the skill loss (SL) can be defined as

$$SL = \left(\frac{RPS_A - RPS_G}{RPS_A}\right) \times 100\%, \qquad (6)$$

where RPS_A and RPS_G are the RPS of the AUTOTAF code and TAF guidance, respectively. Note that by this definition, SL is a strongly nonsymmetric measure. In the case of (nearly) "perfect" encoding, the skill loss approaches 0%. However, if the codes being evaluated have lower RPS value than the guidance, then SL < 0%. Negative SL values should be interpreted with care.

Figure 13 shows the skill loss due to automatic encoding of visibility forecasts in the short TAF guidance

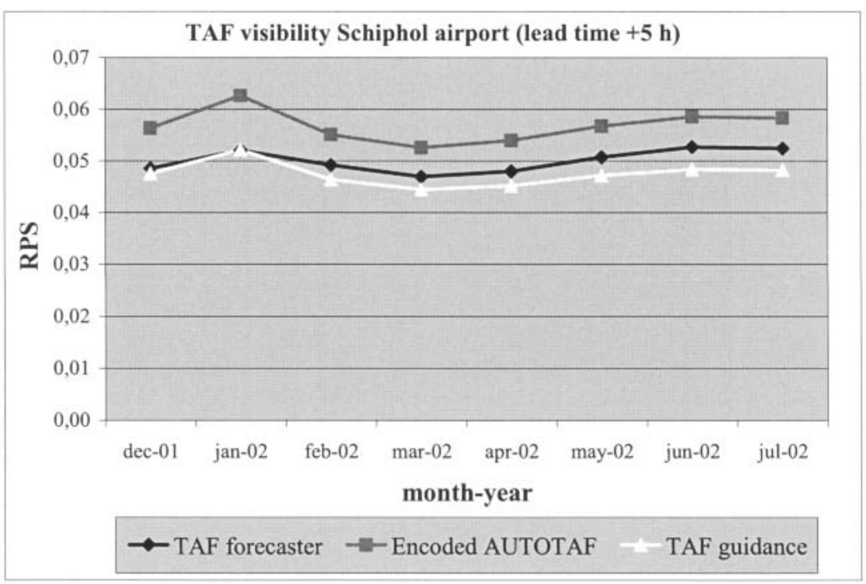

FIG. 12. Moving 12-month average of RPS for TAF visibility forecasts at Schiphol airport, ending at the indicated month. Results are valid for a lead time of +5 h, relative to the start time of the TAF period.

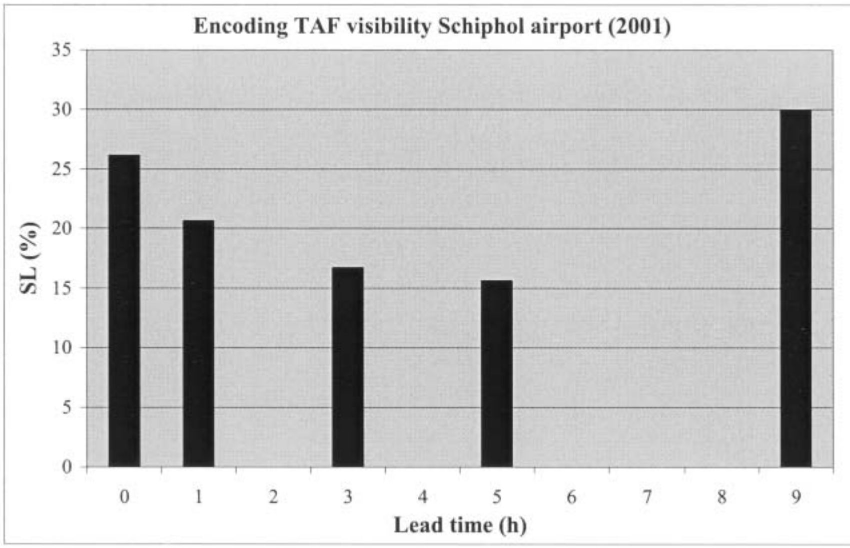

FIG. 13. Skill loss (SL) for visibility forecasts at Schiphol airport (2001) due to encoding of short TAFs.

at Schiphol airport for five different lead times (results for lead times +2, +4, +6, +7, and +8h are not available). The RPS values used in the SL formula are the 12-month averages for 2001. It is remarkable to note the largest losses are at the beginning and end of the TAF period. It is believed that this is a result of the suppression of change groups in the TAF encoder at the edges of the time window. As discussed in section 4a, this suppression has recently been removed from the encoder. The latest verification results, for 2003, have shown slight improvements in the code at the edges of the TAF time window.

Figure 14 shows the verification results for visibility and cloud ceiling forecasts at Schiphol airport as a function of the forecast lead times relative to the start time of the TAF valid period. The 1999 average RPS scores for the short and the long TAFs are presented. Note that the short TAF begins 1 h after issue time while the long TAF begins 8 h after issue time. In the figures, the RPS values are presented for the TAF guidance, the forecasters' TAF, and the persistence forecast. In this context the persistence forecast is defined as the persistence of the observation at issue time.

According to the scores, visibility and cloud ceiling forecasts in the TAF guidance clearly have better skill than visibility and cloud ceiling in the forecasters' TAF after lead times of +3 to +4 h. Note that with respect to this turnover point, the TAF guidance has a disadvantage of being issued 1 h earlier than the forecasters' TAF. On the other hand, we should note that the forecaster is restricted to its limited code while the guidance contains a complete forecast probability distribution over all categories. The skill of the persistence forecast is added for comparison. Note in the figure that the +0 h lead time is effectively +1 h for the persistence forecast and +2 h for the guidance. As a result the persistence forecast shows slightly better skill for the +0 h lead time.

d. Verification results for the TREND guidance and AUTOTREND code

The verification of the TREND products, guidance, and code has been carried out in the same way as the TAF verification. RPS values for the TREND guidance have been computed for a dataset containing 1 yr of data (July 2000–June 2001). In Fig. 15 the results for visibility and cloud ceiling forecasts at Schiphol airport are shown. The lead time is relative to the issue time of the METAR observation and the coupled TREND. The results of the guidance are compared to the persistence forecast and the forecasters' TREND.

Persistence has been calculated relative to the issue time, and therefore the RPS value at lead time = 0 h is zero by definition. In order to demonstrate the impact of the 30-min time lag of the guidance, the scores for the TREND guidance based on the actual METAR observation also have been included in Fig. 15. This additional guidance is labeled in the figure as TREND guidance +30. By definition, this guidance has no time lag, which results in a zero RPS value for lead time = 0. The verification results show that the RPS values of the persistence forecast and the forecasters' TREND grow rapidly with lead time. The RPS of the TREND guidance increases more slowly with lead time. As a result, at the end of the TREND period, the RPS of the TREND guidance is lower than the RPS of both the persistence forecast and the forecasters' TREND. The differences with the RPS of the forecasters' TREND, however, are small at the end of the TREND period. Figure 15 also shows that the RPS values for the addi-

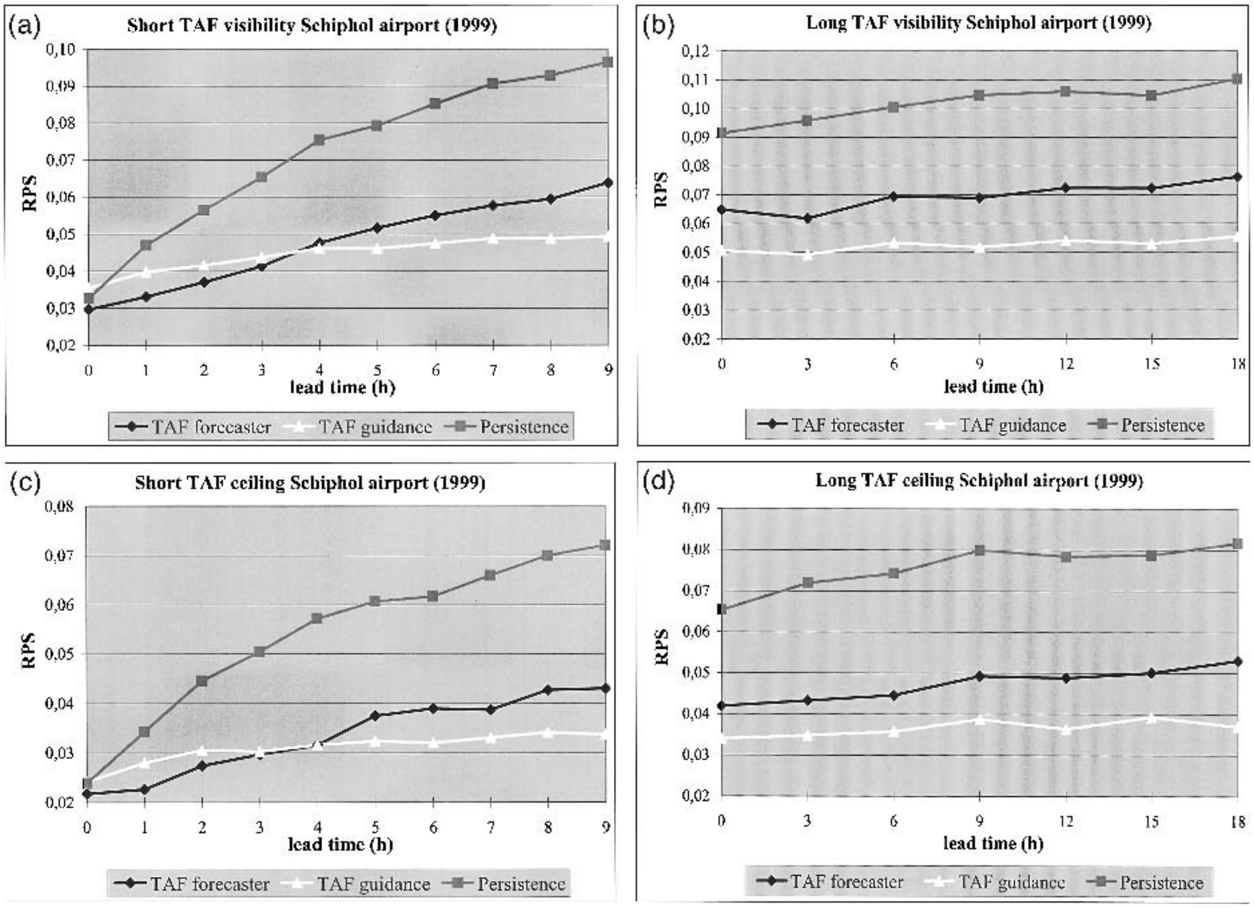

FIG. 14. Average RPS values for (a), (b) TAF visibility and (c), (d) cloud ceiling forecasts at Schiphol airport for 1999 as a function of forecast lead time. The lead times are relative to the start time of the TAF period, which is 1 h after the issue time of the short TAF and 8 h after issue time of the long TAF. The results are presented for (a), (c) the short TAFs and (b), (d) the long TAFs.

tional TREND guidance +30 forecasts are lower than the RPS values of the TREND guidance and persistence forecasts for all lead times. For visibility forecasts, the RPS values are the lowest of all. Obviously the 30-min time lag reduces the skill of the TREND guidance for all lead times. The differences are especially large during the first hour of the TREND period. However, whether these differences are significant for the aviation users depends on how the TREND guidance is used in a particular decision making process of the user.

In an operational environment, the TREND guidance is always used with a certain time lag. According to the scores, we can expect that a reduction in time lag leads to an improvement of forecast skill. In Fig. 15, we

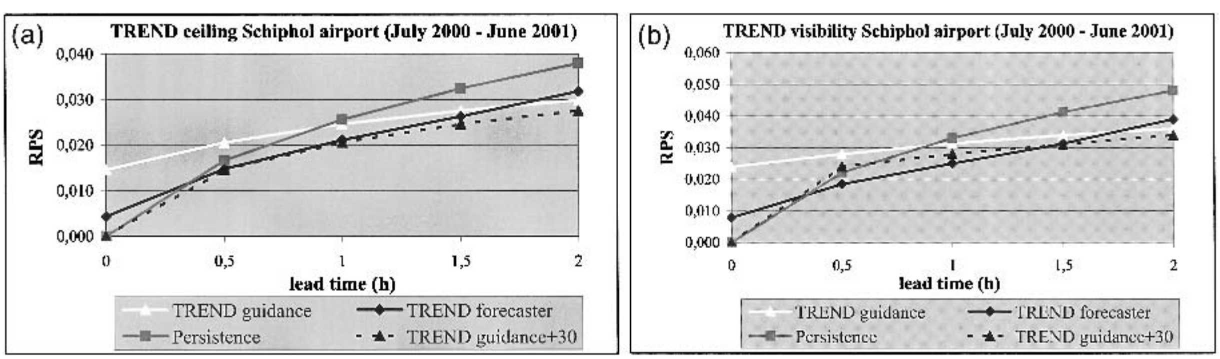

FIG. 15. Verification results for TREND (a) cloud ceiling and (b) visibility forecasts at Schiphol airport.

have compared the full forecast probability distribution of the TREND guidance with the TREND code of the forecaster. We should expect that the encoded AUTOTREND will have a larger RPS than the TREND guidance. In other words, as our experience with TAF encoding has shown, we lose quality by encoding. In this paper the verification results for the AUTOTREND code are not presented. The AUTOTREND code has been evaluated only for a 3-month period (April–June 2001), too short of a period to obtain objective verification results. Nevertheless, this first evaluation of the AUTOTREND code has shown that skill loss is less dramatic than for the encoding of the TAF guidance.

6. Discussion and recommendations

A numerical guidance system consisting of postprocessing NWP model data in combination with local and upstream observations, and high-resolution topographical data, is able to provide more detailed and accurate meteorological information on changing weather conditions at an airport. By providing this guidance to the aviation forecaster, TAF and TREND-type weather forecasts, and in particular forecasts of cloud ceiling, visibility, and wind, can be produced more efficiently and with higher accuracy. The guidance products are the result of recent developments in the semiautomation of the TAF and TREND production at civil airports in the Netherlands. Among the products are 1) a numerical guidance system for TAF and TREND production, supplemented with encoding software that translates the TAF and TREND parameters in the guidance into the required aeronautical codes; and 2) a numerical guidance for the local average wind, in particular crosswind and tailwind components in the takeoff and touchdown areas at the airport, provided by downscaling of NWP model winds. The TAF and TREND guidance, including the automatic encoding software, have been integrated into the operational meteorological production process. A graphical user interface and a code editor are available, and enable the forecaster to modify the suggested AUTOTAF and AUTOTREND code.

a. Results and limitations

The downscaling of winds is still in an evaluation phase. First verification results, however, show that downscaling reduces the representativeness error of the grid-box-averaged NWP model winds. For a +3 h lead time, downscaling yields a significant reduction in ME of the local wind forecasts for each wind direction, together with a small improvement in SD (Fig. 6). The results are valid for different atmospheric stability conditions (unstable, neutral, and stable).

The TAF and TREND guidance have been evaluated for their ability to forecast cloud ceiling and visibility at Amsterdam's Schiphol airport. The guidance forecasts are categorical forecasts, where a probability of occurrence is forecast for ranked categories. The forecasts are verified against METAR observations: 3.5 yr for the TAF guidance (January 1999–July 2002) and 1 yr for the TREND guidance (July 2000–June 2001). The corresponding AUTO codes have been verified for a shorter period, 1.5 yr for the AUTOTAF code (January 2001–July 2002), and 3 months for the AUTOTREND code (April–June 2001). The computed RPS skill score represents how close the forecast probability distribution ranks to the observed category. The quality of the guidance forecasts is compared to the code of the forecaster and the persistence forecast. The results show that the TAF guidance parameters have better skill than the forecasters' TAF after a lead time of +4 hs. In the present operational configuration of the TREND production system (Fig. 11), there is a time lag of 30 min between the observations used to produce the TREND guidance and the observations used for TREND encoding. This 30-min time lag results in a significant reduction of the forecast skill and is even worse than the skill of the persistence forecast for lead times up to +1 h. If the time lag is reduced to zero, the forecast skill improves for all lead times. In that case, the resulting cloud ceiling parameter in the guidance has better skill than the persistence forecast and the forecasters' TREND for all lead times.

The automatic encoding of the TAF guidance parameters for visibility and cloud ceiling leads to skill loss. This defect is due to limitations in the selection of relevant features and grouping of different phenomena in the forecast weather changes. The skill loss is greatest (about 25%–30%) at the beginning and end of the TAF valid time window. It is believed that this is due in part to the configuration settings of the encoding system. This configuration limits the number of words allowed in a TAF code. A limited number of words results in suppression of change groups, in particular at the edges of the valid time window. Recently, the configuration settings of the encoding system have been adjusted, allowing for more change groups in the code. The latest verification results of the AUTOTAF code, during 2003 have shown slight improvements at the beginning and end of the TAF valid period. The verification period for the AUTOTREND code is too short, 3 months, for an objective verification. Nevertheless, this short evaluation has shown that skill loss is less dramatic for automatic encoding of the TREND guidance.

b. Suggestions for improvements and future research

According to the RPS scores, at lead times from +4 h, the TAF guidance provides more accurate cloud ceiling and visibility forecasts than do those provided by the forecaster. For shorter lead times the differences in RPS seem small. The RPS numbers, however, do not tell whether the differences in skill are of any practical relevance. This depends on how the guidance products are actually used. For aviation, the economic value of reliable weather forecasts is high. Bad weather at an airport can reduce airport capacity, leading to enormous economic costs. For this user group, a small gain in forecast skill could lead to substantial economic benefits. The probabilistic information in the guidance allows these users to optimize their cost–benefit ratio in a particular decision making process during operations. At present, airport authorities and an airline at Schiphol airport, in cooperation with KNMI, have developed a prototype system for inbound capacity forecasts. The forecast system is based on technical runway availability and weather forecasts (wind, visibility, and cloud ceiling) from the numerical TAF guidance. The purpose of this research is to determine if and how the guidance forecasts can be used to reduce the impact of bad weather on the inbound airport capacity.

At the time of the development of the TREND guidance, the METAR observation had the highest temporal resolution (30 min) of all real-time observations routinely disseminated at civil airports in the Netherlands. In order to benefit optimally from the detailed information available in the TREND guidance, the update frequency of the guidance must be increased, and the time lag between the production time of the guidance and its use for encoding minimized. For this purpose, observations from fully automated stations will be used. These are available in a central, real-time, 10-min database.

In many areas or countries, high-density observing networks are not available. Moreover, various weather phenomena, such as convective showers, are likely to occur on spatial scales much smaller than the spatial density of the observing network. Other sources of detailed observation information, such as radar data and satellite images, could be beneficial as predictive information instead. In particular precipitation information from weather radar is very interesting in this respect, since it acts like a high-density and high-frequency observing network. The evaluation of the TREND guidance has revealed that the capability to forecast rainshowers at an airport is often limited by the inability to detect those showers in advance at observing stations in the airport's vicinity. The precipitation radar could serve to fill the gaps between the observing sites. Recently, a new method has been developed for short-term, up to 2 h, prediction of precipitation intensity and precipitation phase (liquid, solid, or freezing) at an airport. The method is based on extrapolation of the observed precipitation radar images, which are available every 5 min.

Currently, structural verification of all TAF and TREND-related intermediate and final products, such as the TAF and TREND guidance, the corresponding AUTO codes, and the forecasters' codes, has not yet been arranged. This does have the highest priority. The latest verification results of the TAF guidance have led to several suggestions for improvements in the computation of the guidance parameters and the automatic encoding. Originally, the usage of visibility and cloud observations for the computation of the guidance parameters had been optimized only for the low classes, or low values. Recently, an optimization has been carried out for all classes, which resulted in an increased average accuracy of the guidance parameters for lead times up to +9 h. This improvement is expected to further reduce the turnover point, where the TAF guidance and the forecasters' TAF have equal skill.

Acknowledgments. Hans van Bruggen is gratefully acknowledged for providing the TAF and TREND verification scores. In 1998 he implemented the TAF and TREND verification system that has been used for verification of all TAF- and TREND-related products presented in this paper. Recent discussions with him about the verification scores of the TAF guidance have led to some major improvements in short-term visibility and ceiling forecasts of the TAF and TREND guidance. Ruud Ivens, Sander Tijm, and Jeanette Onvlee are acknowledged for their comments on the first draft of this paper.

REFERENCES

Carr, M. B., 1988: Determining the optimum number of predictors for a linear prediction equation. *Mon. Wea. Rev.,* **116,** 1623–1640.

De Rooy, W. C., and K. Kok, 2004: A combined physical–statistical approach for the downscaling of model wind speed. *Wea. Forecasting,* **19,** 485–495.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.,* **11,** 1203–1211.

Gordon, N. D., 1989: Verification of aerodrome forecasts. Preprints, *Third Int. Conf. on the Aviation Weather System,* Anaheim, CA, Amer. Meteor. Soc., 264–269.

ICAO, 1998: Annex 3 to the Convention on International Civil Aviation (July 1998): Meteorological service for international air navigation. International Standards and Recommended Practices, 13th ed. ICAO, Montreal, PQ, Canada, 92 pp.

Kilpinen, J., 1994: Computer-aided weather forecasting system set to enter operation in Scandinavia. *ICAO J.,* **49** (8), 17–18.

Knüppfer, K., 1997: Automation of aviation forecasts. The projects AUTOTAF and AUTOGAFOR. Preprints, *Seventh Conf. on Aviation, Range, and Aerospace Meteorology,* Long Beach, CA, Amer. Meteor. Soc., 444–449.

Leyton, S. M., and J. M. Fritsch, 2003: Short-term probabilistic forecasts of ceiling and visibility utilizing high-density surface weather observations. *Wea. Forecasting,* **18,** 891–902.

Lynn, K. J., 1997: International survey of TAF automation systems. Forecasting systems 9, Met Office Internal Rep., 20 pp.

Reid, J. D., 1978: Verification of ceiling and visibility forecasts using the ranked probability score. *Atmos.–Ocean,* **16,** 177–186.

Undén, P., and Coauthors, 2002: HIRLAM-5 scientific documentation. Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden, 144 pp.

Verkaik, J. W., and A. Smits, 2001: Interpretation and estimation of the local wind climate. *Proc. Third European and African Conf. on Wind Engineering,* Eindhoven, Netherlands, Amer. Geophys. Union, 43–56.

Vislocky, R. L., and J. M. Fritsch, 1997: An automated, observations-based system for short-term prediction of ceiling and visibility. *Wea. Forecasting,* **12,** 31–43.

WMO, 1998: Code forms. Section A, Part A—Alphanumeric codes, *International Codes,* Vol. I.1, *Manual on Codes,* WMO-No.306, Geneva, Switzerland, 169 pp.