Springer Link

Meteorology and Atmospheric Physics
All Volumes & Issues
ISSN: 0177-7971 (Print) 1436-5065 (Online)

In this issue (10 articles)

Meteorol Atmos Phys (2018) 130:265–277
https://doi.org/10.1007/s00703-017-0513-1

ORIGINAL PAPER

Fuzzy logic-based analogue forecasting and hybrid modelling of horizontal visibility

Zoltán Tuba[1] · Zsolt Bottyán[2]

Received: 17 March 2016 / Accepted: 30 December 2016 / Published online: 25 February 2017

Abstract Forecasting visibility is one of the greatest challenges in aviation meteorology. At the same time, high accuracy visibility forecasts can significantly reduce or make avoidable weather-related risk in aviation as well. To improve forecasting visibility, this research links fuzzy logic-based analogue forecasting and post-processed numerical weather prediction model outputs in hybrid forecast. Performance of analogue forecasting model was improved by the application of Analytic Hierarchy Process. Then, linear combination of the mentioned outputs was applied to create ultra-short term hybrid visibility prediction which gradually shifts the focus from statistical to numerical products taking their advantages during the forecast period. It gives the opportunity to bring closer the numerical visibility forecast to the observations even it is wrong initially. Complete verification of categorical forecasts was carried out; results are available for persistence and terminal aerodrome forecasts (TAF) as well in order to compare. The average value of Heidke Skill Score (HSS) of examined airports of analogue and hybrid forecasts shows very similar results even at the end of forecast period where the rate of analogue prediction in the final hybrid output is 0.1–0.2 only. However, in case of poor visibility (1000–2500 m), hybrid (0.65) and analogue forecasts (0.64) have similar average of HSS in the first 6 h of forecast period, and have better performance than persistence (0.60) or TAF (0.56). Important achievement that hybrid model takes into consideration physics and dynamics of the atmosphere due to the increasing part of the numerical weather prediction. In spite of this, its performance is similar to the most effective visibility forecasting methods and does not follow the poor verification results of clearly numerical outputs.

Responsible Editor: M. Kaplan.

✉ Zoltán Tuba
Tuba.Zoltan@uni-nke.hu

1 Doctoral School of Military Engineering, National University of Public Service, Ludovika tér 2, 1083 Budapest, Hungary

2 Institute of Military Aviation, National University of Public Service, Kilián út 1, 5008 Szolnok, Hungary

1 Introduction

Most of the operational thresholds in aviation and air traffic controlling are related to cloud ceiling and horizontal visibility. Unfortunately, the complex physical processes of these parameters or the existing deficiencies of numerical weather prediction (NWP) models make them difficult to forecast as Herzegh (2006), Bankert and Hadjimichael (2007) and later Ghirardelli and Glahn (2010) clearly showed. Obviously, the accuracy of these forecasts, regardless of size of aircrafts in general aviation (Bateman and Herzegh 2007), has more significant impact on aviation costs (Keith and Leyton 2007), efficiency (Bergot et al. 2005) and safety (Groff and Price 2006; Gultepe et al. 2007; Souders and Showalter 2008) than prediction of other less important meteorological variables. During special military missions, often not the operational aviation thresholds but the special limits of reconnaissance or surveillance tasks restrict the mission execution (Darnell 2006). This once again highlights the importance of accurate visibility forecasts. Potential development in performance of visibility prediction will be under detailed examination in this paper due to its key role. To achieve this development, there are two fundamental methods which are often applied in combination. In the first, NWP model outputs are developed in different ways, and in the second statistical forecasting methods are applied.

As Gultepe et al. (2007) summarized in their extended overview on fog forecasting in connection with numerical prediction: "Owing to the complex interactions of various thermodynamic processes, in principle, an accurate 3D model is needed for reliable fog forecasting". 3D models can take into account all the physical processes and interactions involved in fog forecasting both in horizontal and vertical directions of NWP models. Unfortunately, on one hand fog formation, which has significant impact on visibility forecasting, usually happens on sub-grid scales and on not well-described topography (Müller et al. 2007); on the other hand, the physical processes and interactions involved (e.g. aerosol–droplet relationship, Gultepe and Isaac 1999 or aerosol–visibility relationship, Molnár et al. 2016) are not known in the extent necessary. Additionally, very high resolution can increase computational costs extremely; therefore, these types of models are not in use extensively.

1D models of planetary boundary layer (PBL) are widely used to model fog formation and dissipation numerically. The basic assumption of these models is horizontal homogeneity of thermodynamic variables and structure of turbulence (Gultepe et al. 2007), which allows the one-dimensional approach. It means that some of the dynamic processes (e.g. horizontal advection) which can have significant role in fog formation are neglected in simulations because of the initial assumption. Despite this deficiency, they were widely used (Musson-Genon 1987; Teixeira and Miranda 2001), because of their simplicity and their low computational cost.

In the past decade, 1D and 3D models are often used in combination with success (Gultepe et al. 2006a; Müller et al. 2007). With these solutions, computational cost can keep on an acceptable level and the processes are simulated in the detail necessary at the same time.

Development of statistical visibility forecasting methods often connected to NWP model outputs. Post-processing procedures (Wantuch 2001; Zhou et al. 2004) or improving the schemes of derived NWP model outputs (Gultepe et al. 2006b; Gyöngyösi et al. 2013) are commonly applied. These are often based on clearly statistical techniques (Marzban et al. 2006; Bang et al. 2008) or combine different methods (Vislocky and Fritsch 1997; Jacobs and Maat 2005). Methods independent from numerical weather prediction models are also known. Fabbian et al. (2007) showed the applicability of artificial neural networks in visibility forecasting, while Hansen (2007) and Bottyán et al. (2015) applied fuzzy logic-based analogue forecasting methods successfully.

In this paper, the improvement of statistical-based analogue forecasting is the starting point of a complex development of visibility forecasts, which combine statistical and numerical outputs to compose more accurate deterministic forecasts (Bottyán et al. 2013, 2015, 2016; Tuba 2014). The verification results show the efficiency of development of analogue and hybrid forecasts. The further results of different forecasting methods and their comparability help to determine the limits of application. Finally, the paper shows the performance of the different methods in the function of visibility and lead time, which also helps forecasters to find the most efficient forecasting method between given conditions.

2 Methods

In this section, analogue and hybrid forecasting is discussed in detail. These methods are closely connected to each other, but while analogue forecasting can provide deterministic visibility prediction individually, until then hybrid needs analogue output to combine its own forecast. The numerical weather prediction part of the process will only be shown superficially in hybrid forecasting section, because its development was not part of this research. The main steps of the entire forecasting process are shown in Fig. 1.

During the development of hybrid forecasting, there was essential to determine the different weights and parameter values applied, but for its operation it is enough to have the observations of the last 6 h, the historical database used for finding similar cases and the valid visibility output of the NWP model, as it can be seen in Fig. 1.

2.1 Observational data used

To set up a new forecasting system, it was essential to create an appropriate dataset which could be a good basis of analogue forecasting based on similarity (Bottyán et al. 2012, 2015). This aviation climate database contains only routine aviation weather report (METAR)-based information in optional format to help any professional or non-professional user processing data. Even meteorological information of METAR is not the most accurate in case of some variables (e.g. dry-bulb temperature, dew-point temperature), but at the same time it gives the opportunity to easily adopt the complete system to another forecast location.

The observational data of METAR reports are mainly from meteorological sensor measurements (wind speed, dry-bulb temperature, etc.), but in some cases they are from visual observations (e.g. prevailing visibility, significant

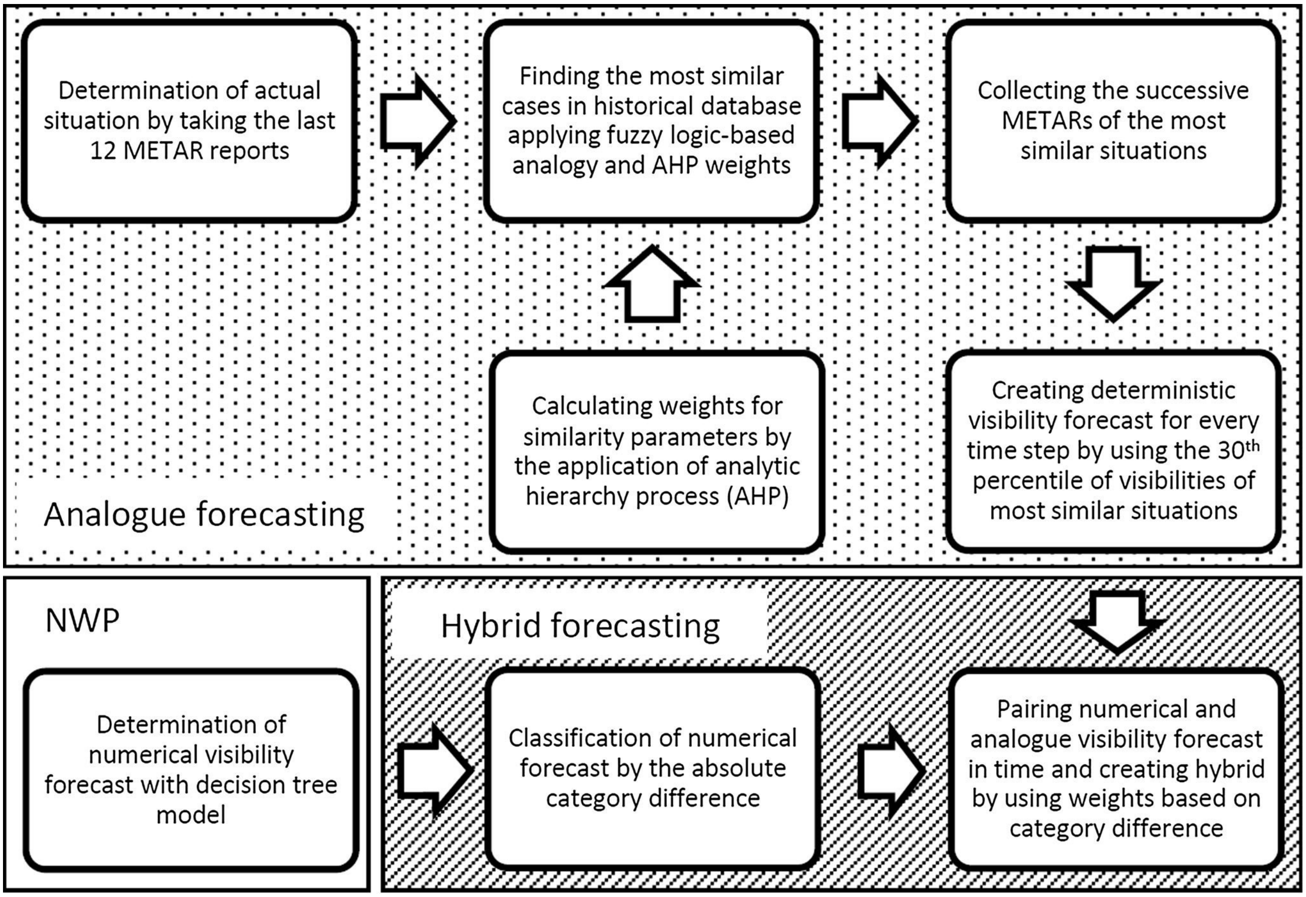


Fig. 1 Schematic flow chart of the forecasting process

weather). However, nowadays instrumental-based visibility data (MOR—meteorological optical range, RVR—runway visual range) are already commonly available, but their historical dataset is not long enough for such kind of applications in Hungary.

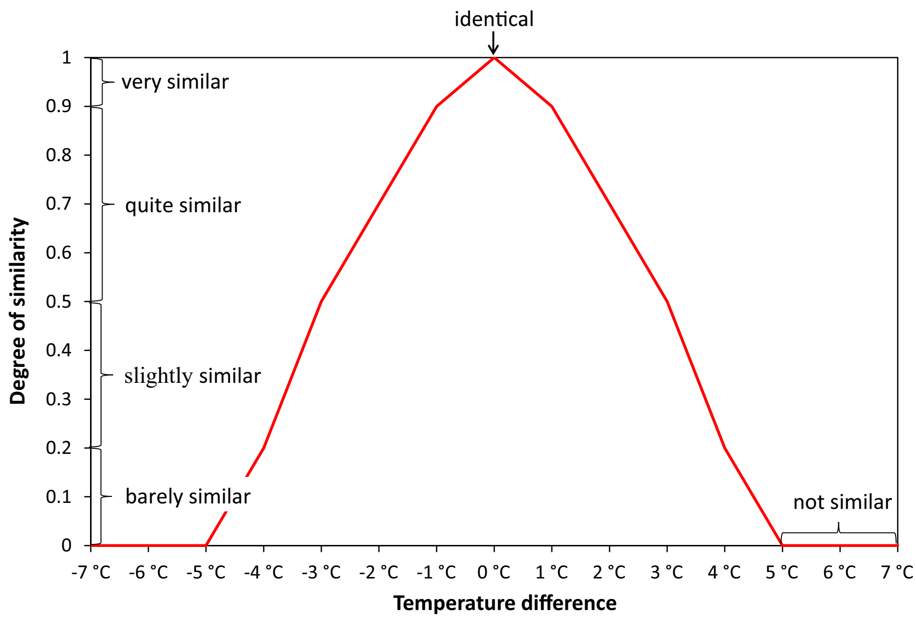
2.2 Analogue forecasting

The basic principle of analogue forecasting is to find similar weather situation to the current one in an appropriate database and give a prediction for the next period based on the degree of similarity. The phrase of weather situation means couple of hours' continuous observation in this case and in the followings. Usually the projection period in analogue forecasting is very short, often only up to 6 (Petty et al. 2000), 12 (Riordan and Hansen 2002) or 24 h (Hansen 2007), because it can work when the conditions are homogeneous (e.g. no expected dynamical changes) which is the substance of this type of forecasts. Additionally, the analogue method reflects variations in independent variables (e.g. location), which means that it is location dependent as Hansen (2007) described.

Like in Hansen's paper, fuzzy logic-based algorithm was applied to find similar situations. In this process, every single weather situation is represented by 12 METARs; therefore, the algorithm compares the meteorological and time variables of 6-h-long periods. Determination of the degree of similarity is based on fuzzy sets defined by forecast experts, which is a commonly used method in development of fuzzy systems (Meyer et al. 2002). It is based on conditional climatology statistics which were already applied in aviation meteorology in the 1960s (McCabe 1968). Individual similarities ranging from 0 to 1 can be determined using fuzzy sets for every single parameter at any given time. The smaller the difference between compared variables in a given time step, the higher the value of the degree of similarity, as it is shown in the example of Fig. 2. When the temperature difference equals zero, the similarity value is the highest (1). If the mentioned difference reaches 5 °C, then similarity will be zero. The intermediate values (0.9, 0.7, 0.5, and 0.2) connected to the expressions of similarity (very similar, between very and quiet similar, quiet similar, slightly similar), respectively.

In the following, phrase of time step represents the backward distance of a given METAR from the starting time of forecast (t), which is the time of the latest

Fig. 2 Example of a fuzzy set: degree of similarity (S) in the function of temperature difference



observation. In accordance with the definition above, the time steps go from present time $(t-0)$ till the eldest information compared $\{t-(k-1)\}$, where k is the number of the time steps.

The degree of similarity of different parameters can be summarized on several ways. For example, Hansen (2007) calculated overall similarity by taking the minimal value of individual similarities of examined variables. In a subsequent study, Tuba et al. (2013) showed the efficiency of application of Analytic Hierarchy Process (AHP) in determination of weights of weighted averaging in construction of total similarity. It is based on the assumption that forecast accuracy can be improved by highlighting the importance of individual elements in the forecasting process. Since AHP is a method of multi-criteria decision making, it is needed to make a connection between AHP and the original problem—to find the most similar cases—to be solved. Firstly, it is necessary to model this weighting problem as a hierarchy, where the criteria are the meteorological variables to be weighted and the possible alternatives are the individual weather situations. It is easy to see that the large number of alternatives does not let the complete application of AHP. Therefore, it was partially applied only for determining the applicable weights for the different parameters.

It can be supposed that there are given n objects (in our case different meteorological parameters) and the aim is to find weights w_i for $i = 1, \ldots, n$ such that the weight w_i refers to the importance of the ith object. Initially, there is only an estimation on the ratio of the importance of the ith and the jth object for each pair i,j. Let us denote these estimates by a_{ij} for $i,j = 1, \ldots, n$. Then, of course $a_{ij} = \frac{1}{a_{ji}}$ is required. Value of a_{ij} was calculated by expert judgements of pairwise comparison of the importance of ith and jth objects.

Note that these might be inconsistent, i.e. $a_{ij} \cdot a_{jk} = a_{ik}$ not necessarily holds. For example, let us assume that the first object is two times more important than the second one, the second is also two times more important than the third, but the first is only three times more important than the third. In such situation, it cannot be expected that the weights will perfectly reflect the ratios. Thus, the goal is to find w_is such that the difference between a_{ij} and $\frac{w_i}{w_j}$ is as small as possible.

Let us denote by $\mathbf{A}$ the $n \times n$ matrix of the ratios. It is easy to see that if the given ratios are consistent, then the desired vector of weights $\boldsymbol{w} = (w_1, \ldots, w_n)$ is an eigenvector of the mentioned matrix $\mathbf{A}$ with eigenvalue n, or in other words $\mathbf{A}\mathbf{w} = n\mathbf{w}$ and in such a case n is the maximal eigenvalue. Because $\sum_{i=1}^{n} \lambda_i = tr(\mathbf{A}) \equiv$ sum of the diagonal elements $= n$, where all the eigenvalues $\lambda_i, i = 1, \ldots, n$ of matrix $\mathbf{A}$ are zero except one. Therefore, only one of the λ_i, which we call $\lambda_{\max}$, the maximal eigenvalue of $\mathbf{A}$, equals n and $\lambda_i = 0$, and $\lambda_i \neq \lambda_{\max}$. Surprisingly, the converse also holds; Saaty (1977) proved that for any matrix of the type described above the matrix is consistent if and only if $\lambda_{\max} = n$. Moreover, he proved that even in the case of inconsistency the best choice for the weight vector is the eigenvector belonging to $\lambda_{\max}$ in the sense that it minimizes a certain function of the differences between a_{ij} and $\frac{w_i}{w_j}$ (for the detailed proof, see Saaty 1977). To determine the eigenvector, the standard power iteration method was used. Saaty advised to measure the consistency of the given matrix with $\frac{\lambda_{\max}-n}{n-1}$ which $\frac{\lambda_{\max}-n}{n-1}$ turned out to be a reasonable estimate.

Using the above statements, the AHP process goes as follows:

- assignation of the estimated ratios a_{ij} for $i,j = 1, \ldots, n$ yielding the matrix $\mathbf{A}$;
- calculate the eigenvector $\boldsymbol{w}_{\mathbf{max}}$ of $\mathbf{A}$ belonging to the maximal eigenvalue $\lambda_{\max}$ by iteration;
- norming $\boldsymbol{w}_{\mathbf{max}}$ gives the desired vector $\boldsymbol{w} = (w_1, \ldots, w_n)$ such that w_i is the weight of the ith object.

For more detailed description of this process, please refer to the cited study (Tuba et al. 2013) and Saaty (1977, 1991). The applied weights in this forecast system can be seen in Table 1. The level of the mentioned inconsistency is 2.5% which is less than the tolerable 10% (Saaty 1991), so the results are significantly reliable.

Due to the convincing verification results (detailed later), in this research this procedure was applied to determine w_i weight of the ith parameter. Following the earlier introduced notations, the total similarity ($S_{\mathrm{TOTAL}\ t-j}$) of the $(t-j)$th time step, which represents the summarized similarity of applied parameters of the examined METAR in the given time step, is

$$S_{\mathrm{TOTAL}\ t-j} = \sum_{i=1}^{n} w_i \cdot S_{ij} \tag{1}$$

where S_{ij} is the individual similarity value of the ith examined parameter determined by fuzzy sets in the $(t-j)$th time step. Knowing the calculated weights, the similarity of the individual time steps under investigation can be determined by weighted averaging of the single parameters' similarity. Finally, the overall similarity (S_{overall}) of the examined weather situation can be calculated from the weighted averaging of the similarity of time steps. General description of the applied weighted averaging is the following:

Table 1 Applied weights (w_i) determined by analytic hierarchy process and used for calculating similarity in analogue forecasting

Variable compared	Applied weight
Visibility	0.311
Ceiling	0.231
Hour of the day	0.096
Precipitation type	0.095
Date of the year	0.048
Pressure	0.046
Dry-bulb temperature	0.036
Dew-point temperature	0.036
Wind direction	0.034
Wind speed	0.034
Wind gust	0.033

$$S_{\mathrm{overall}} = \frac{\sum_{j=0}^{k-1}\left(2^{k-j-1} \cdot \sum_{i=1}^{n} w_i \cdot S_{ij}\right)}{2^k - 1} \tag{2}$$

where k is the number of the time steps applied ($k = 12$) and n is the number of variables compared in fuzzy logic-based algorithm.

The current observation ($(t-0)$th time step) gets the largest weight and this weight decreases rapidly as we go through time steps. It provides that the most similar cases probably contain the dynamic changes and guarantee the convergence in similarity during the examined time period.

Once the overall similarity is calculated for all the weather situations of the database, the most similar cases can be identified. After the selection of these situations, the visibility values of the next METARs are collected for every time step of the projection period. For deterministic forecast composition, 30th percentile of visibility values of the thirty most similar weather situations is obtained. In this case, it means that the tenth value of the ascending ordered visibility list gives the deterministic visibility prediction in a given time step. Forecast projection time is 9 h which is matched to the short terminal aerodrome forecast (TAF) time interval. Choosing 30th percentile as deterministic forecast value is based on the preliminary verification results which will be detailed later.

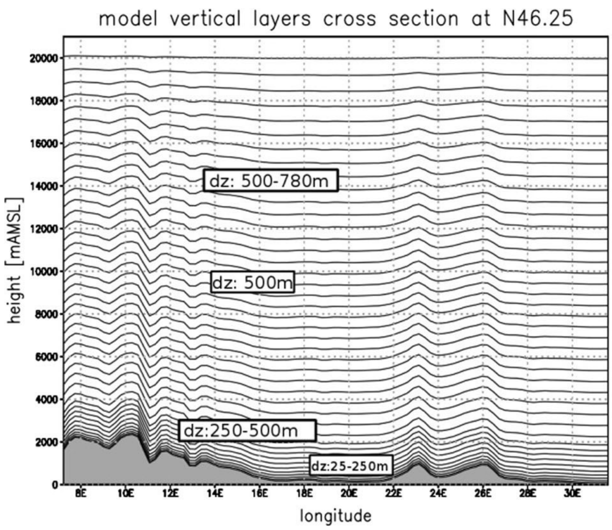
2.3 Hybrid forecasting

It can be stated that neither analogue nor numerical forecasting is perfect in prediction of visibility alone. Both of them have advantages and disadvantages which are collected in Table 2.

One of the biggest advantages of analogue forecasting is that when rapid amendment needed (e.g. unexpected changes in weather conditions), then a new adjusted forecast is available after next routine observation usually in half hour. In contrast, NWP models generally run in every 6–24 h, which means the same loss of time if amendment is necessary because the forecast based on NWP model outputs is incorrect. Forecast location is limited in analogue forecasting, because an appropriate database is needed to compose forecasts, which is only available where the observation programme is continuous. Fortunately, intended locations of analogue visibility forecasts (e.g. airports) usually have this kind of dataset. Performance of analogue forecast quickly decreases because of its basic concept; therefore the efficient length of projection period is only couple of hours. Efficiency of numerical products barely changes in projection period, but its performance is much weaker in the operationally important initial hours. Thanks to the fact that verification of visibility forecasts is only feasible where routine observation is available; both types of forecast can be verified in the same locations.

Table 2 Typical advantages and disadvantages of analogue and numerical forecasting

	Analogue forecasting	Numerical forecasting
Refresh rate	Fits observation frequency (0.5–1 h)	6–24 h
Forecast location	Limited: historical database needed	Optional in model domain
Performance in projection period	Quickly decreases	Quasi constant
Length of efficient projection period	Couple of hours	Couple of days
Verification	Complete: local observation needed for forecast	Limited: only where observation available



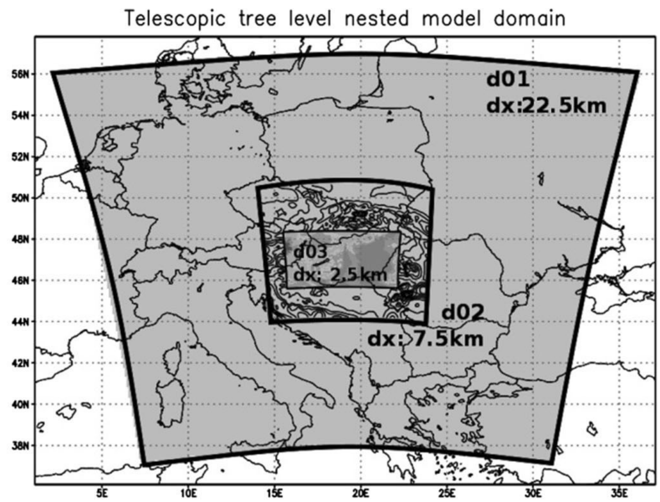


Fig. 3 Vertical levels with higher resolution in the lower levels and deeper layers in the upper portion (*right panel*). Three level telescopic nested domain setup for high-resolution modelling of the Carpathian Basin in the Integrated Forecast System. Horizontal resolution is 22.5, 7.5 and 2.5 km, grid size 97 × 97, 97 × 97 and 202 × 121, respectively (*left panel*) (Bottyán et al. 2015)

Usually NWP models are not fine tuned to special needs of aviation meteorology. Often there are no outputs for direct support during planning or executing special tasks of aviation or air traffic controlling. These parameters have the most influence on flight safety (Groff and Price 2006; Gultepe et al. 2007; Souders and Showalter 2008), but usually these are not important in general weather forecasting (e.g. icing, turbulence, ceiling, etc). Gyöngyösi et al. (2013) developed a complex dynamical modelling system focused on the special needs mentioned above. Its core is a Weather Research and Forecasting (WRF) limited area model (LAM) which was verified by targeted case studies and in situ measurements. It allowed choosing the best combination of parameterization schemes (e.g. Bretherton and Park (2009) planetary boundary layer scheme; Hong et al. (2004) micro-physics scheme) and post-processing methods. The model domain and the resolution (Fig. 3) were tailored to the formulated purposes above.

The hourly visibility output of WRF LAM is post-processed by Wantuch's perfect prognosis decision tree method (Wantuch 2001). This method inspects every single case concerning synoptic situation, surface and lower tropospheric temperature and humidity conditions, etc. and, based on these information, chooses the appropriate post-processing procedure of visibility in a given location.

To keep advantages and eliminate disadvantages described above and following the earlier considerations, Tuba (2014) and Bottyán et al. (2015) introduced a so-called hybrid model output which combines analogue and numerical visibility forecast:

$$\mathrm{Visibility}_{Hybrid} = a_j \cdot \mathrm{Visibility}_{Analogue} + b_j \cdot \mathrm{Visibility}_{Numerical} \quad (3)$$

where $a_j + b_j = 1$; $a_j, b_j \in [0; 1]$ and j is the number of time steps of forecast. To ensure the correction of initial inaccuracy of numerical prediction, value of a_j should increase with increasing category differences. To determine category difference at $t + 0000$ time, it is required to divide visibility values of observation and numerical forecast into categories.

(0–1000; 1000–1500; 1500–3000; 3000–5000; 5000 m). The gradual transition between analogue and numerical methods is ensured by the monotonically decreasing value of a_j over the projection period. Values of a_j and b_j are defined directly by operational forecasters' joint opinion considering

Table 3 The applied a_j weights in composition of hybrid visibility forecast

	$T + 0100$	$T + 0200$	$T + 0300$	$T + 0400$	$T + 0500$	$T + 0600$	$T + 0700$	$T + 0800$	$T + 0900$
Absolute category difference									
4	1.00	1.00	1.00	0.90	0.80	0.65	0.50	0.35	0.20
3	1.00	1.00	0.90	0.80	0.70	0.55	0.45	0.30	0.20
2	1.00	0.90	0.85	0.75	0.65	0.50	0.40	0.25	0.15
1	0.90	0.85	0.80	0.70	0.60	0.45	0.35	0.20	0.10
0	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10

the criteria above. Based on these values, a weight matrix (Table 3) was composed which allows to calculate the hybrid visibility forecast. The bigger the category difference, the worse the NWP forecast at initial time step and it is independent of the direction of difference. Due to this, the original weight matrix is symmetric and can be simplified with using absolute category difference.

The main advantage of hybrid forecast is that the incorrect visibility prediction can be amended with the refresh rate of analogue forecasts; furthermore, it keeps the potential dynamical changes from numerical model outputs in increasing proportion due to the applied method of hybrid composition.

3 Verification results

In this section, the complete verification procedure and its results will be shown and discussed. In the first part of the section, the time period of verification, the geographical locations and the method applied will be described in detail. In the second part, the main results of the entire research will be discussed.

3.1 Verification method

It should be noted in advance that the same verification procedure was applied for the different forecasts of visibility in order of their comparability. The performance of fuzzy logic-based analogue forecasting with and without AHP method, post-processed dynamical model prediction, hybrid forecast, terminal aerodrome forecast and naive forecast of persistence was investigated and compared in the first 9 h of the forecast period. Persistence forecast was chosen as a standard of reference or, in other words, a competitive benchmark which is commonly applied in the field of short-term forecasts (Murphy 1992). Analogue forecasting without AHP method means that the calculated weights are not applied; parameters involved into forecasting process have equal weight. One hour was chosen as a basic time step of verification on account of time step of dynamical model's output, and visibility values were not handled as representing time periods but characterizing instantaneous values of a given time. This procedure is maintained in case of each forecast to preserve comparability.

To complete the categorical verification, the whole statistical database (from 1 August 2005 to 31 July 2014) was divided into two independent parts by designating a verification period from 1 August 2013 to 31 July 2014 in accordance with the semi-operative period of the mentioned numerical model. In this time frame all of the forecasts mentioned above is completely available. TAFs are only issued in every third (LHSN and LHKE) or every sixth (LHPA and LHBP) hour (Fig. 4 shows the geographic locations of the examined airports), so verification was adjusted to this schedule. In the following, visibility categories will be represented with the upper limit of the given category (category limit), because the lower limit is equal to zero in each category. It is in accordance with the practice in aviation, which applies operational thresholds. It means that if the visibility is less than the concerning threshold (limit), then different flight rules or procedures need to be used (e.g. instrumental flight rules—IFR). These limits can depend on, for instance, the pilot's qualifications, the takeoff minimums of the airport or the aircraft. During the verification process, observed and forecast visibilities are under examination whether they are inside the given category or not. In the verification template, the upper limit of the examined category is parameterized, so the dependency of the results due to different visibility categories is easily controlled. It means that if there are special limits to be verified (e.g. in accordance with military missions) or sequence of categories for averaging needs to be verified, then the verification results are simply available.

The main steps of verification process are shown in Fig. 5. As it can be seen, the inner part of the process (from data categorization till calculating verification parameters) is repeatable as necessary. Thus, the average verification results are easily produced for the desired categories.

Doswell et al. (1990) showed that there was no omnipotent verification method. It is advised to use several skill scores and verification parameters [α, *BIAS*, *POD*,

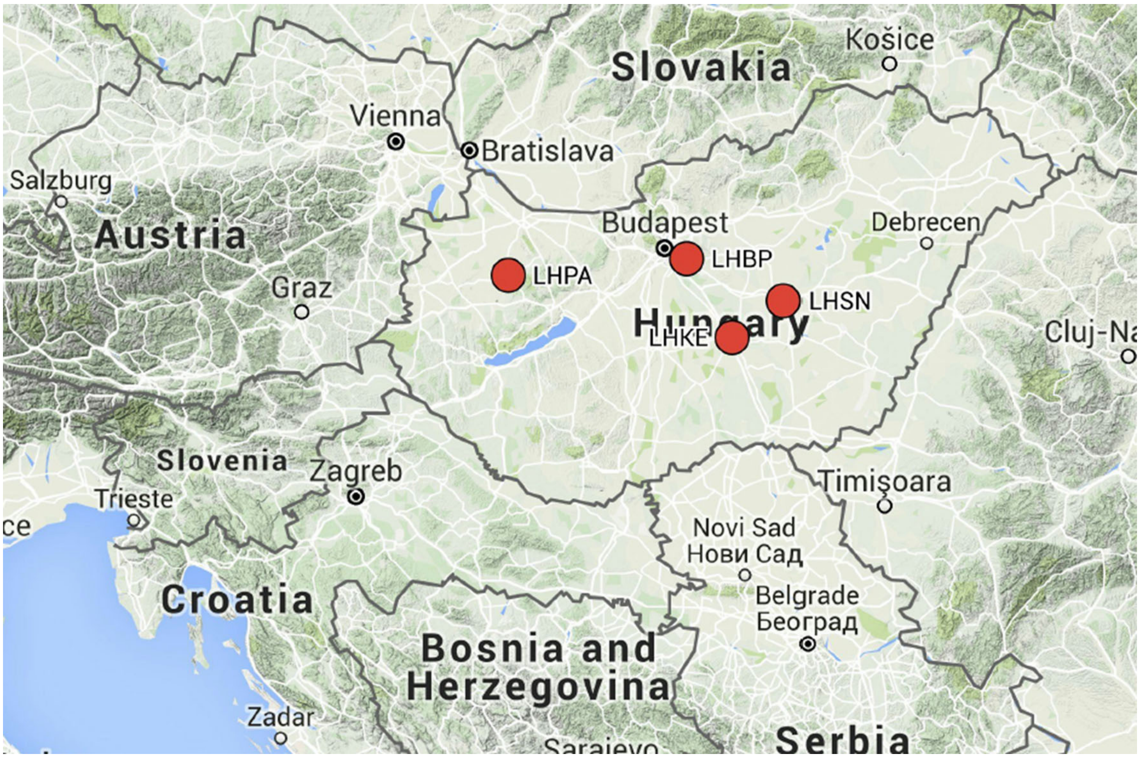
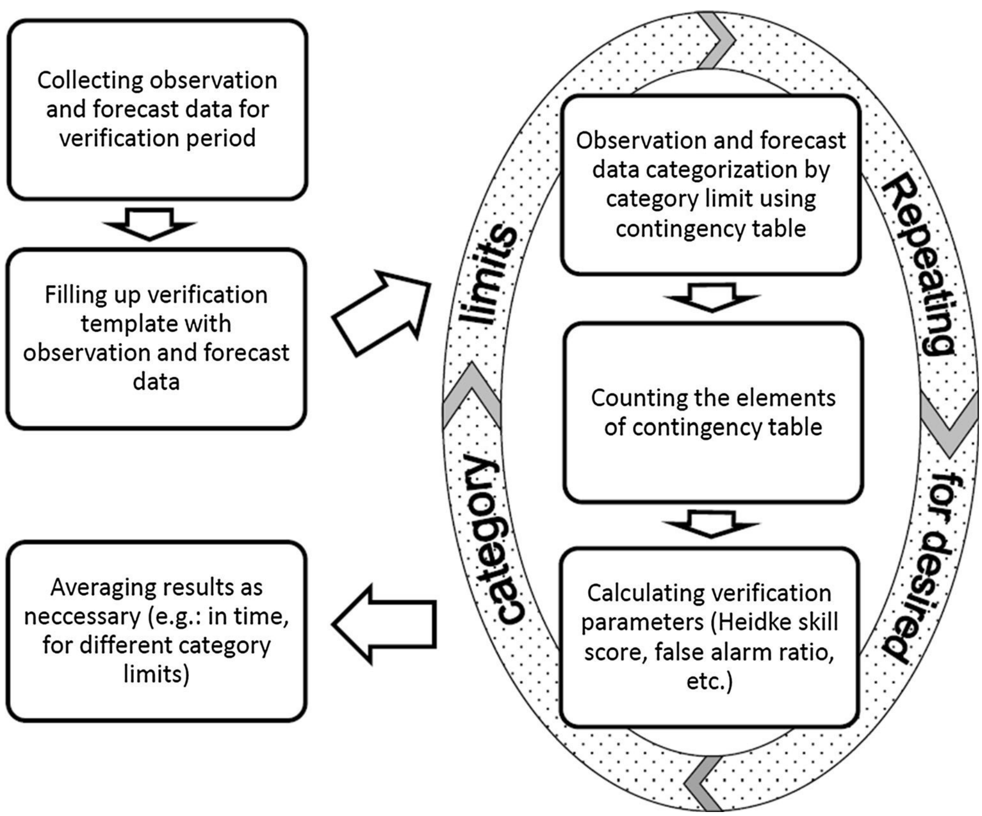


Fig. 4 The geographic locations of the examined airports (*LHPA* Pápa, *LHBP* Budapest, *LHKE* Kecskemét, *LHSN* Szolnok)

Fig. 5 Flow chart of the verification process



POFD, *FAR*, *HIT*, *CSI*, *TSS*, *HSS*, etc.—for detailed description, please see Nurmi (2003)] for comprehensive verification of forecasts. A 2 × 2 contingency table of different categories of visibility (Table 4) was used to calculate these verification parameters. As Bankert and Hadjimichael (2007) described: "Heidke skill score (HSS) is computed to measure the performance of each algorithm relative to random chance". Positive, zero or negative HSS value indicates better, no better or worse forecast performance than random chance, respectively. It is very important to note that HSS values remain correct with verification of rare events, which is typical in case of low visibilities. According to the reasons above, the HSS values of visibility forecasts of the different prediction methods are presented.

$$\mathrm{HSS} = \frac{2 \cdot (a \cdot b - b \cdot c)}{(a + c) \cdot (c + d) + (a + b) \cdot (b + d)} \tag{4}$$

Dividing the forecast period into short intervals is a commonly applied method in the international practice of TAF's verification (Kluepfel 2005; Mahringer 2008). According to the instantaneous visibility values of observations, this method can contain errors necessarily if category forecasts are verified with category limits different from ICAO amendment guidance (ICAO 2013). Additionally, takeoffs or beginning of missions usually have exact time, prior to that users want to know whether they delay or carry out their tasks. Due to the reasons above, a new approach was applied for TAF verification which also helped to preserve comparability of different forecasts. In this verification, the observed visibility of the previous hour's last METAR was chosen for comparison in the given hour of forecast period. In case of numerical and hybrid forecasts, it means that $t + 0100$ time step was verified with $t + 0045$ observation, $t + 0200$ with $t + 0145$ and so on. In case of analogue, persistence and terminal aerodrome forecasts predicted visibility at $t + 0045$, …, $t + 0845$ is clearly definable even when change groups are present.

In terminal aerodrome forecasts change groups are used when permanent transitions ("becoming"—BECMG) or temporary (TEMPO) changes are predicted in values or states of weather elements. These groups are followed by the time interval in which the forecast change is valid (Mahringer 2008). If change indicator BECMG is present in TAF, the predicted visibility value with better verification performance was used when changing group is valid. If both of them failed to predict correctly or both of them were successful, then the predicted visibility prior to changing group was used. If temporary fluctuations are forecast, namely TEMPO group is present firstly its length was examined. If this period is not longer than 2 h, then the predicted visibility value with better verification performance was used when any changing group is valid. If its length reached 3 h, then every hour of the changing period was assigned a 3-h-long period in the following way. The first and last hour was paired to the first and last 3 h of the period, respectively. Thus, the middle hours were paired to the period determined by the previous, the given and the following hours. Note that during the following verification steps, all the available METARs were examined. If all the predictions were correct in this 3-h-long period, then the application of BECMG would have been appropriate instead of TEMPO, so the predicted visibility (which is probably incorrect) prior changing group is taken into consideration as well. If none of the predictions were correct of this 3-h-long period, then the application of TEMPO was not provoked so this visibility (which is probably incorrect) was taken into consideration as well beside the predicted visibility prior to changing group. In other cases, predicted visibility value with better verification performance is used when any changing group is valid. If probability of occurrence is determined with temporary fluctuations, namely in case of the presence of PROB30 TEMPO or PROB40 TEMPO groups in TAF, which means a potential occurrence of the alternative value with 30 or 40% probability, then the same procedure as in case of TEMPO was applied, but their verification results were taken into consideration with smaller weights: 0.3 and 0.4, respectively. If probability of occurrence (PROB30 or PROB40) of an alternative value of visibility is predicted not as a temporary fluctuation, then the value with better verification performance was used.

The demonstrated method obviously overestimates the performance of terminal aerodrome forecasts a little due to the processing of changing groups and probability of occurrence. Considering that the verification was carried out for every hundred meter less than 5000 m, this also means that alternative forecaster tools can be certainly designated based on TAF's verification performance because of this overestimation.

Table 4 Two-category contingency table of categorical forecasts

Event forecast	Event observed	
	Yes	No
Yes	a (Hit)	b (False alarm)
No	c (Missed)	d (Correct rejection)

3.2 Results

First of all, it should be noted that on the figures of this section forecast time does not represent the lead time of forecasts. It simply means hours passed from the starting time of verification. In case of numerical forecast and TAF, lead time differs from forecast time. As it was mentioned earlier, numerical forecasts were made only 1–4 times a day, while TAFs were composed at certain times and these are not adjusted to verification schedule as forecasts usually are not adjusted to aviation tasks.

As it was described in the end of analogue forecasting section, the deterministic visibility forecast is based on the 30th percentile of visibility values of the most similar weather situations. This was determined by preliminary verification which was carried out before the described verification process using the data of 2012. In this procedure, the prediction was completed on the data of the period lasted from 1 August 2005 to 31 December 2011 and for all the four airports mentioned above. It meant 26,262 examined forecasts. All the results are averaged for the examined airports and for the 9 h long forecast period. As it can be seen in Fig. 6, the minimum of the number of incorrect forecasts ($b + c$) corresponds with the maximum values of HSS and the difference between the number of hits and number of incorrect forecasts ($a - (b + c)$). This situation shows the optimal application of 30th percentiles in case of quite rare events as poor visibility.

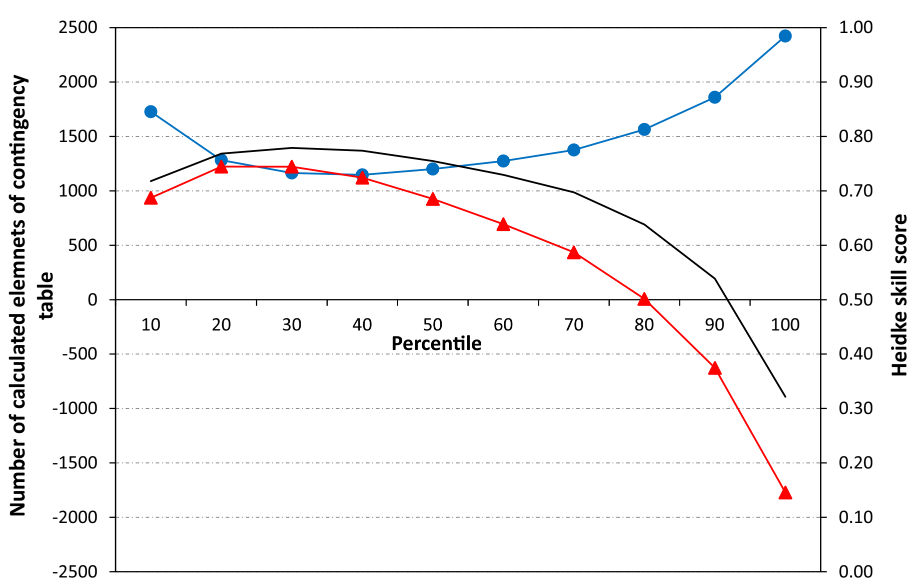
($a - (b + c)$) and the number of incorrect forecasts ($b + c$) in the function of applied percentiles in forecast composition are represented by bold line, triangles and circles, respectively.

The upper limits of the examined categories were defined for every 100 m up to visibility value of 5000 m. Then, HSS values were calculated for all the methods and all these limits. Category limits under 1000 m are undefined at numerical and hybrid forecasts due to the applied post-processing method. Its visibility output gives deterministic prediction only above 1000 m; therefore, appropriate category limits are used for forming averages to keep comparability of the different methods.

As it can be seen in Fig. 7, the application of AHP weights improved significantly the performance of analogue forecasting in the first 6 h. Additionally, the greatest increase of HSS happens in the case of very poor visibility ($\leq$1500 m), when accurate visibility forecast has more critical role. In addition, analogue and hybrid forecasting produces similar HSS values in the whole forecasting period despite the increasing weight of NWP model output in hybrid visibility. It proves that highlighting advantages and eliminating disadvantages determined in the previous Section were successful. Both graph beats naive forecast of persistence with an increasing difference. Performance of TAFs gains hybrid forecasting upon only the last hours of projection period, but as it was mentioned earlier these values are averages. In case of lower category limits, TAFs become competitive around $t + 0400$. The clear numerical visibility output has well-balanced performance, but it does not reach even the performance of the persistence during the whole forecasting period.

Values of different verification parameters are calculated from the elements of contingency table in the verification process of categorical forecasts. Obviously, if the changes of category limits do not generate changes in the elements of the contingency table, then the value of examined verification parameter remains the same. This can be easily discovered in Fig. 8 where the values of HSS often remain the same in the function of category limit. This is a consequence of the airports' observation and forecast practice, which is usually adjusted to ICAO SPECI and TAF amendment guidance, operational thresholds of

Fig. 6 *HSS*, the difference between number of hits and number of incorrect forecasts



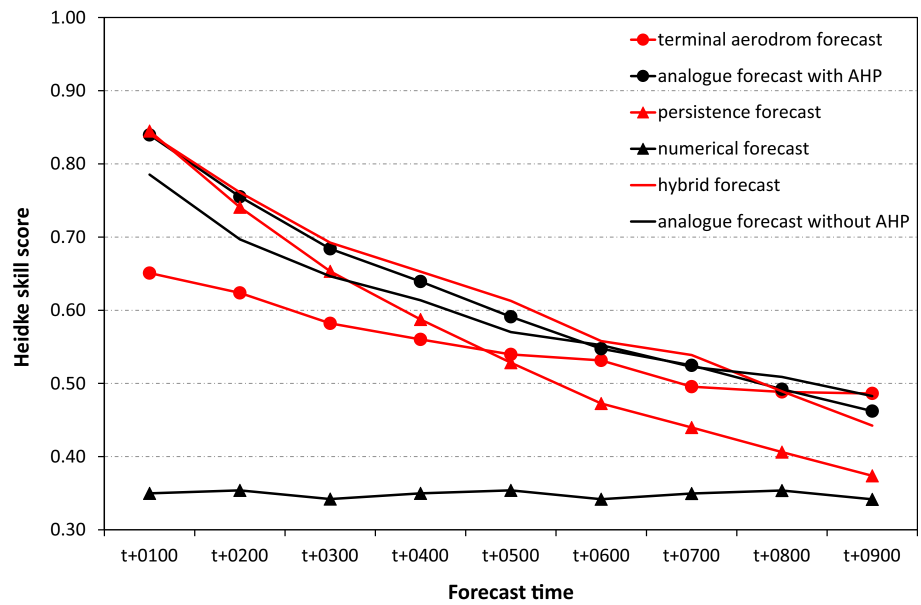


Fig. 7 Average HSS of different category limits (1000–5000 m in 100 m steps) and examined Hungarian airports for the applied forecast methods in the function of forecast time

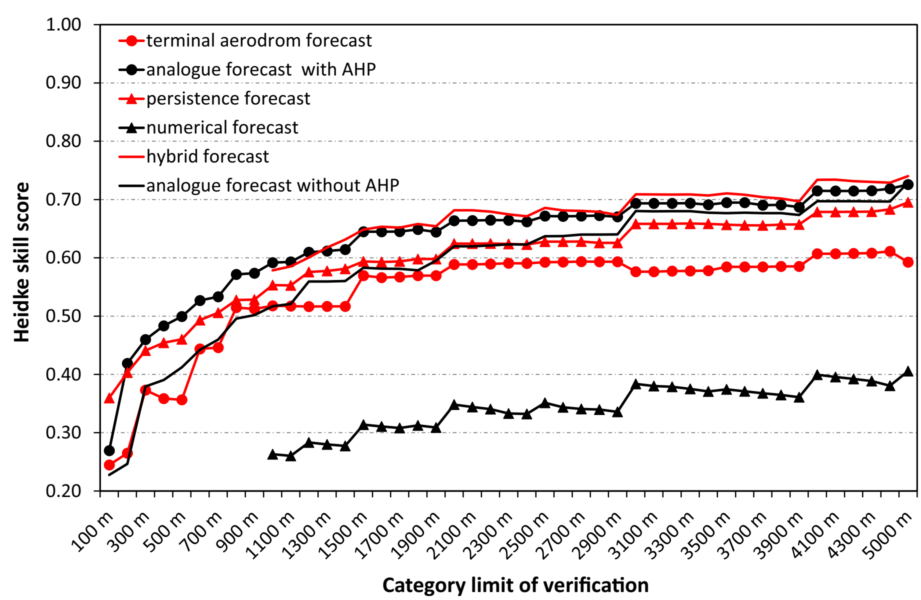


Fig. 8 HSS values of every examined category limits averaged for the first 6 h of forecast period, in case of examined Hungarian airports and for the applied forecast methods

airport and the distance of the reference points of visibility observation. Consequently, the relative frequency of visibility values does not show continuous distribution in case of observation or forecast data. This also explains the little jumps of HSS around the generally used values of visibility. It also means that in case of categorical verification, where distribution of observed or forecast data is not continuous, very detailed examination of category limits is necessary to filter fluctuations.

Except of 100 m visibility, HSS value of persistence is convincingly beaten by analogue and hybrid forecasting. They show quite similar, slowly increasing performance with increasing category limits. Despite the poor results of numerical forecast, hybrid has excellent performance according to the weighting procedure applied. In the first part of projection period in hybrid forecast, the numerical part is mainly from good quality forecasts, where the initial category difference is low. It means that numerical forecasts involved in hybrid creation probably exceed the performance of the analogue part.

HSS value of TAF is closest to leading performance of analogue or hybrid forecasting in case of lower visibility values. In the last third of the forecast period, performance of TAF gradually takes the lead in case of increasing visibility as well.

4 Summary and conclusions

In this work, a new hybrid forecasting method which produces deterministic forecasts of visibility and its complete verification has been described. This method uses fuzzy logic-based analogue forecasting and NWP limited area model prediction to compose hybrid visibility forecasts. Based on the outcomes and the detailed verification results of this research, the following conclusions can be drawn: (1) Application of weights of the applied variables in analogue forecasting significantly increases the forecasting performance of visibility. (2) In the first 6 h of projection period, TAF is beaten both analogue and hybrid forecast, so their applicability is well founded in this time horizon. (3) Hybrid and analogue forecasts have similar performance despite the increasing numerical part, because the applied NWP forecasts are filtered by the initial category difference between observation and numerical forecast. (4) Even the HSS value of the forecasts shows strong dependency on the category limits, the superiority of analogue and hybrid forecast is not a question. It means that their applicability does not depend on category limits.

Based on the results above, analogue and hybrid visibility products are provided for operational military aviation forecasters semi-operationally. Forecasters of military airports and aviation meteorology centre gave positive feedbacks; as they found it could be a useful tool in direct meteorological support of general and special (for example UAS) flight missions as well.

Possible future extension of this work can involve further parameters. In preliminary experiments, the similar hybrid forecasting of cloud ceiling also seems to be a promising approach. In addition, the most similar cases selected during analogue forecasting could be the starting point of a future probabilistic forecast.

Acknowledgements This research was supported by the European Social Fund (TÁMOP-4.2.1.B-11/2/KMR-2011-0001, Research of Critical Infrastructure Defense). The project was realized through the assistance of the European Union, with the co-financing of the European Social Fund.

References

Bang CH, Lee JW, Hong SY (2008) Predictability experiments of fog and visibility in local airports over Korea using the WRF model. J Korean Soc Atmos Environ 24:92–101

Bankert RL, Hadjimichael M (2007) Data mining numerical model output for single-station cloud-ceiling forecast algorithms. Weather Forecast 22:1123–1131

Bateman RE, Herzegh P (2007) An exploration of an automated analysis product for ceiling and visibility: potential impacts on general aviation weather decision making. 23rd Conference on IIPS, San Antonio, 15–18 Jan 2007

Bergot T, Carrer D, Noilhan J, Bougeault P (2005) Improved site-specific numerical prediction of fog and low clouds: a feasibility study. Weather Forecast 20:627–646

Bottyán Z, Wantuch F, Tuba Z, Hadobács K, Jámbor K (2012) Creation of a new climatic database for aviation meteorological support system of unmanned aerial vehicles. (in Hungarian). Repüléstudományi Közlemények 24(3):11–18

Bottyán Z, Wantuch F, Gyöngyösi AZ, Tuba Z, Hadobács K, Kardos P, Kurunczi R (2013) Development of a complex meteorological support system for UAVs. World Acad Sci Eng Tech 76:1124–1129

Bottyán Z, Gyöngyösi AZ, Wantuch F, Tuba Z, Kurunczi R, Kardos P, Istenes Z, Weidinger T, Hadobács K, Szabó Z, Balczó M, Varga A, Bíróné Kircsi A, Horváth G (2015) Measuring and modeling of hazardous weather phenomena to aviation using the hungarian unmanned meteorological aircraft system (HUMAS). Idojaras 119(3):307–335

Bottyán Z, Tuba Z, Gyöngyösi AZ (2016) Weather forecasting system for the unmanned aircraft systems (UAS) missions with the special regard to visibility prediction in Hungary. In: Critical infrastructure protection research: results of the first critical infrastructure protection research project in Hungary: topics in intelligent engineering and informatics, vol 12. Springer International Publishing, pp 23–34. doi:10.1007/978-3-319-28091-2_2

Bretherton CS, Park S (2009) A new moist turbulence parameterization in the community atmosphere model. J Clim 22:3422–3448

Darnell KM (2006) Analysis of weather forecast impacts on United States Air Force combat operations. MS Thesis, Department of Meteorology, Naval Postgraduate School, Monterey

Doswell CA III, Davies-Jones R, Keller DL (1990) On summary measures of skill in rare event forecasting based on contingency tables. Weather Forecast 5(4):576–585

Fabbian D, de Dear R, Lellyett S (2007) Application of artificial neural network forecasts to predict fog at Canberra International Airport. Weather Forecast 22(2):372–381

Ghirardelli JEG, Glahn B (2010) The meteorological development laboratory's aviation weather prediction system. Weather Forecast 25:1027–1051

Groff LS, Price JM (2006) General aviation accidents in degraded visibility: a case control study of 72 accidents. Aviat Space Environ Med 77(10):1062–1067

Gultepe I, Isaac GA (1999) Scale effects on averaging of cloud droplet and aerosol number concentrations: observations and models. J Clim 12(5):1268–1279

Gultepe I, Cober SG, King P, Isaac G, Taylor P, Hansen B (2006a) The fog remote sensing and modeling (FRAM) field project and preliminary results. In: AMS 12th Cloud Physics Conference, Madison, 9–14 July 2006

Gultepe I, Müller MD, Boybeyi Z (2006b) A new visibility parameterization for warm-fog applications in numerical weather prediction models. J Appl Meteor Climatol 45(11):1469–1480

Gultepe I, Tardif R, Michaelides SC, Cermak J, Bott A, Bendix J, Müller MD, Pagowski M, Hansen B, Ellrod G, Jacobs W, Toth G, Cober SG (2007) Fog research: a review of past achievements and future perspectives. J Pure Appl Geophys 164:1121–1159

Gyöngyösi AZ, Kardos P, Kurunczi R, Bottyán Z (2013) Development of a complex dynamical modeling system for the meteorological support of unmanned aerial operation in Hungary. Proceedings of International Conference, Atlanta, 28–31 May 2013, doi: 10.1109/ICUAS.2013.6564668

Hansen BK (2007) A fuzzy logic-based analogue forecasting system for ceiling and visibility. Weather Forecast 22:1319–1330

Herzegh P (2006) Development of FAA national ceiling and visibility products: challenges, strategies and progress. 12th Conference on Aviation Range and Aerospace Meteorology on 30 January–02 February 2006, Atlanta

Hong SY, Dudhia J, Chen SH (2004) A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. Mon Weather Rev 132:103–120

International Civil Aviation Organization (ICAO) (2013) Meteorological service for international air navigation. Annex 3 to the Convention on International Civil Aviation, 18th edition, Montreal

Jacobs AJM, Maat N (2005) Numerical guidance methods for decision support in aviation meteorological forecasting. Weather Forecast 20:82–100

Keith R, Leyton SM (2007) An experiment to measure the value of statistical probability forecasts for airports. Weather Forecast 22(4):928–935

Kluepfel CK. (2005) TAF verification in the US National Weather Service. NWS Instruction 10–1601, http://www.nws.noaa.gov/directives/010/archive/pd01016001c.pdf. Accessed 10 Aug 2016

Mahringer G (2008) Terminal aerodrome forecast verification in Austro Control using time windows and ranges of forecast conditions. Meteorol Appl 15(1):113–123

Marzban C, Sandgathe S, Kalnay E (2006) MOS, perfect prog, and reanalysis. Mon Weather Rev 134(2):657–663

McCabe JT (1968) Estimating conditional probability and persistence. No. AWS–TR–208. Air Weather Service Scott AFB IL, USA. http://www.dtic.mil/get-tr-doc/pdf?AD=AD0671506. Accessed 22 Feb 2017

Meyer MA, Butterfield KB, Murray WS, Smith RE, Booker JM (2002) Guidelines for eliciting expert judgment as probabilities or fuzzy logic. In: Ross TJ, Booker JM, Parkinson WJ (eds) Fuzzy logic and probability applications: bridging the gap. Society for Industrial and Applied Mathematics, Philadelphia, pp 105–123

Molnár Á, Párkányi D, Imre K, Gácser V, Czágler E (2016) A closure study on aerosol extinction in urban air in Hungary. Idojaras 120(2):163–181

Müller MD, Schmutz C, Parlow E (2007) A one-dimensional ensemble forecast and assimilation system for fog prediction. J Pure Appl Geophys 164:1241–1264

Murphy AH (1992) Climatology, persistence, and their linear combination as standards of reference in skill scores. Weather Forecast 7(4):692–698

Musson-Genon L (1987) Numerical simulation of a fog event with a one-dimensional boundary layer model. Mon Weather Rev 115(2):592–607

Nurmi P (2003) Recommendations on the verification of local weather forecasts. ECMWF Tech. Memorandum 430. http://www.ecmwf.int/sites/default/files/elibrary/2003/11401-recommendations-verification-local-weather-forecasts.pdf. Accessed 10 Aug 2016

Petty K, Carmichael B, Wiener G, Petty M, Limber M (2000) A fuzzy logic system for the analysis and prediction of cloud ceiling and visibility. 9th Conference on Aviation, Range, and Aerospace Meteorology, Orlando, Am Meteor Soc pp 331–333

Riordan D, Hansen BK (2002) A fuzzy case-based system for weather prediction. Eng Intell Syst Elec 10(3):139–146

Saaty TL (1977) A scaling method for priorities in hierarchical structures. J Math Psychol 15:234–281

Saaty TL (1991) Some mathematical concepts of analytic hierarchy process. Behaviormetrika 18(29):1–9

Souders CG, Showalter RC (2008) Transformation of NAS to NEXTGEN and FAA's weather architecture impacts: an update. Aerospace Meteorology, AMS 88th Annual Meeting on 20–24 January 2008, New Orleans

Teixeira J, Miranda PM (2001) Fog prediction at Lisbon airport using a one-dimensional boundary layer model. Meteorol Appl 8(4):497–505

Tuba Z (2014) Selected questions of unmanned aerial vehicles (UASs) and visibility (in Hungarian). Repüléstudományi Közlemények 26(2):94–105

Tuba Z, Vidnyánszky Z, Bottyán Z, Wantuch F, Hadobács K (2013) Application of analytic hierarchy process (AHP) in fuzzy logic-based meteorological support system of unmanned aerial vehicles. Acad Appl Res Mil Sci 12(2):221–228

Vislocky RL, Fritsch JM (1997) An automated, observations-based system for short-term prediction of ceiling and visibility. Weather Forecast 12(1):31–43

Wantuch F (2001) Visibility and fog forecasting based on decision tree method. Idojaras 105(1):29–38

Zhou B, Du J, McQueen J, Dimego G, Manikin G, Ferrier B, Toth Z, Juang H, Hart M, Han J (2004) An introduction to NCEP SREF Aviation Project. In Preprints, 11th Conference on aviation, range and aerospace meteorology, Hyannis, 4–8 October 2004. http://s3.amazonaws.com/academia.edu.documents/45442921/81314.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1487752227&Signature=T6EN1q%2FiKsjJRalhb7sKV4UcVZc%3D&response-content-disposition=inline%3B%20filename%3DAn_Introduction_to_NCEP_SREF_Aviation_Pr.pdf. Accessed 22 Feb 2017