9.14

AN EXPERIMENT TO MEASURE THE VALUE OF STATISTICAL PROBABILITY FORECASTS FOR AERODROMES

Ross Keith[1]
Bureau of Meteorology and James Cook University, Townsville, Australia

and

Stephen M. Leyton[2]
Center for Analysis and Prediction of Storms
University of Oklahoma, Norman, Oklahoma

1. INTRODUCTION

Over the past couple of decades, considerable research has been performed to investigate alternatives to traditional numerical weather model prediction of surface weather conditions. In particular, observations-based statistical techniques have shown great promise for improving short-term forecasts of surface weather conditions. Vislocky and Fritsch (1997) demonstrated that such a statistical forecasting system has superior skill to numerical models for the short-term prediction of ceiling and visibility. This system considered a network of surface observations surrounding an observing site (often an airport) to produce probabilistic forecasts of ceiling and visibility for that observing site. The results indicated that such a system has greater skill than model output statistics (MOS) derived forecasts as well as persistence climatology out to a lead time of 6h. Leyton and Fritsch (2003, 2004) extended this work by demonstrating a further increase in skill when utilizing higher density and higher frequency surface observations for these short-term forecasts.

While observations-based statistical systems have shown greater forecast skill over traditional techniques, little has been done to investigate whether the *economic value* of the forecast has been improved as well. Keith (2003) demonstrated how human subjective forecasts of the probability of the weather being below alternate minimum should increase the value to airlines compared to the traditional categorical aerodrome forecasts. The alternate minimum is that level of visibility and cloud base which requires an aircraft to carry extra fuel, which, in turn, incurs a cost.

[1] *Corresponding author address:* Ross Keith, 14 Lawson St., Townsville 4812, Australia;
Email address: keiths14@bigpond.net.au

[2] *Corresponding author address:* Stephen M. Leyton, Center for Analysis and Prediction of Storms, Sarkeys Energy Center Room 1110, 100 East Boyd St., Norman, OK 73019;
Email address: sleyton@ou.edu

Considering the promising results of these recent studies, it was of interest to combine the two concepts and investigate whether observations-based probabilistic forecasts can further increase the value of the forecast to aviation interests. The concepts behind these two studies are synergistic, as the lead times utilized in the observations-based forecasts coincide with the lead times utilized for the planning and flight period of the majority of domestic flights around the world (typically 2 to 6 hours). It is important to note that human forecasters produce aerodrome forecasts that generally cannot match even raw persistence out to between 3 to 6 hours ahead (depending on how this forecast skill is measured).

This study focused on probabilistic forecasts of low ceiling and/or reduced visibility and the corresponding impact on forecast value for flights arriving at three major airports in the United States[3]. As a basis of comparison, National Weather Service Terminal Aerodrome Forecasts (TAFs) that correspond to the developed statistical forecasts have been obtained for each location. In addition, costs for several flights arriving at each airport have been provided by a domestic, commercial carrier. These have been utilized to determine the optimal forecast probability above which extra fuel should be carried. The combination of improved short-term forecasts and identification of optimal forecast probabilities will lead to greater forecast value, potentially saving the aviation industry millions of dollars per year.

2. METHODOLOGY

2.1 *Data*

Two data sets were utilized to develop the statistical forecast equations. The first contained hourly surface weather observations for the period of January 1982 through December 1996. The second data set contained hourly observations spanning from January 1997 to July 2003. Each

[3] The identification of the three airports has been specifically omitted to protect the privacy of the commercial airline that assisted in this study. Each airport serves as a "hub" for that airline.

dataset included hourly observations of temperature, dew point, wind speed and direction, cloud cover (from which ceiling can be derived), visibility and present weather at over 1500 automated weather-observing sites around the United States. The earlier dataset is used solely for the development of climatological values, while the latter is utilized for both climatological purposes as well as the statistical forecast development. Leyton and Fritsch (2003, 2004) provides a detailed explanation of the data processing procedures followed.

In order to assess the skill of the statistical forecast equations, a database of Terminal Aerodrome Forecasts (TAFs) was obtained for a basis of comparison. These TAFs were produced by the National Weather Service office that is responsible for each of the three airports being examined in this study. Of interest to this experiment is whether the TAF predicts adverse weather such that flight dispatchers must specify that a given flight should add extra fuel. This archive of TAFs spans from April 2002 to May 2003.

A database of operating costs for flights arriving into each of the three major airports was obtained from a commercial airline. This data provided cost per flight information for a total of 18 daily flights, in 2003 dollars. These operating costs represent average costs for each flight during the April 2002 to May 2003 period; approximately 7500 flights in all.

The flights used in this study were determined in consultation with the airline. They represent varying distances and departure/arrival times during the day. Thus, this study is not focused simply on a specific type of flight but rather a broad representation of an airline's typical daily flights.

2.2 Methodology

There are two primary stages to this experiment. First, probabilistic forecasts of low ceiling and/or reduced visibility must be created. Second, an optimal probability threshold must be calculated for each flight using the operating costs data. Both the probabilistic forecast and the TAF for each flight are then compared to the calculated optimal probability for each flight, determining whether either forecast would require additional fuel carriage. Each request for additional fuel incurs a cost, since extra fuel is needed simply to carry the requested additional fuel. Therefore, the cumulative cost of each forecast leads to the assessment of the economic value of the statistical forecasts.

2.3 Forecast Equation Development

Statistical forecast equations were developed in the same manner as described by Leyton and Fritsch (2003, 2004). For this experiment, the forecast of interest was the probability of ceiling ≤ 2000 ft. and/or visibility < 3 SM. The reason for these thresholds is that they are current FAA criteria for determining whether a given flight should add additional jet fuel in case the flight has to divert to an alternate landing site.

In order to duplicate the impact of the weather forecast on the decision processes for assessing fuel requirements, careful consideration must be given to the initialization and valid times for the observations-based forecasts. This is mainly because decisions on the loading of jet fuel typically are made during the dispatcher's flight planning, which usually takes place 1-2h prior to the flight departure. Another complicating factor is due to the nature of the observations-based forecasts as well. As described by Leyton and Fritsch (2003, 2004), these forecasts can only be initialized and verified at the top of the hour. However, scheduled flights rarely arrive/depart exactly at the top of the hour. Therefore, taking all of this information into consideration, a unique forecast initialization time and valid time was determined for each flight.

It is important to note that while thunderstorms have a significant impact on aviation operating costs, this study focused only on low ceiling and/or reduced visibility. Since only 18 daily flight departure/arrival times were considered, analysis of the observational data showed that very few thunderstorms occurred during these times. Moreover, of those convective events that did occur at a time of interest, nearly all resulted in a reduction in ceiling/visibility. However, we believe that the absence of storm forecasts does not significantly affect the degree of savings extracted by the probability forecasts.

2.4 Optimal Threshold Development

The determination of optimal threshold for each flight follows the techniques described by Keith (2003). When allocating costs, alternate fuel was assigned whenever the probability forecast was greater than or equal to the optimal decision probability for the particular flight. The traditional method of planning fuel was used for the TAF, as carried out by airline dispatchers under FAA regulations.

3. RESULTS

Table 1 displays a breakdown of the cumulative costs per flight using the traditional TAFs and the observations-based probabilistic forecasts, in 2003 dollars. These costs represent total costs from April 2002 to May 2003. It is important to note that for each of the 18 flights, the use of the observations-based probabilistic forecasts resulted in lower cumulative costs than the traditional TAFs. In fact, the average savings per flight is $23K over 14 months. Projecting this

Approx. Flight Distance to Major Airport	Costs Using TAFs	Costs Using Obs-Based Forecasts	Savings Using Obs-Based Forecasts
1100 miles to airport #1	**$54406**	**$31387**	**$23019**
1400 miles to airport #1	**$71388**	**$59914**	**$11474**
1600 miles to airport #1	**$113996**	**$97688**	**$16308**
2500 miles to airport #1	**$88743**	**$78272**	**$10501**
200 miles to airport #2	**$28848**	**$12444**	**$16404**
1650 miles to airport #2	**$30738**	**$14721**	**$16017**
250 miles to airport #2	**$21706**	**$8859**	**$12847**
550 miles to airport #2	**$50136**	**$8370**	**$41766**
1450 miles to airport #2	**$54636**	**$14076**	**$40560**
550 miles to airport #2	**$77768**	**$35688**	**$42080**
1400 miles to airport #3	**$37106**	**$10704**	**$26402**
1600 miles to airport #3	**$50983**	**$26776**	**$24207**
500 miles to airport #3 (1)	**$65430**	**$38730**	**$26700**
500 miles to airport #3 (2)	**$36943**	**$28795**	**$8148**
450 miles to airport #3 (1)	**$82723**	**$42067**	**$40656**
450 miles to airport #3 (2)	**$57849**	**$34048**	**$23801**
250 miles to airport #3	**$43049**	**$19543**	**$23506**
1050 miles to airport #3	**$44441**	**$21846**	**$21846**
TOTAL COSTS	***$1010889***	***$583928***	***$426961***

Table 1: Comparison of cumulative costs per flight, for the period of April 2002 to May 2003, using TAFs and observations-based probabilistic forecasts of low ceiling and/or reduced visibility. The average savings per flight is approximately $23,000 during the 14-month period. The dollar amounts are in 2003 U.S. dollars. The (1) and (2) represent the same flight route but different departure/arrival times.
Note: To protect the privacy of the airline that assisted in this study, each flight is identified simply by the distance traveled and to which airport each arrived.

savings across 2000 to 3000 daily flights typical of a major carrier, the annual savings approaches $50M per year. This significant savings is simply due to a transition from a deterministic TAF forecast to a probabilistic forecast.

4. SUMMARY

Low ceilings and reduced visibilities impact departures and arrivals at airports across the nation. Airline dispatchers must account for the possibility of delays due to such impeding weather phenomena and decide whether extra fuel should be loaded onto an aircraft. Of course, this decision on future weather conditions 2h or more after the flight departure must be made 1-2h prior to the plane's departure. Traditionally, these decisions have been made using National Weather Service (NWS) issued TAFs, which supply categorical forecasts of weather for aviation interests.

Dallavalle and Dagostaro (1995) showed that simple persistence forecasts are highly competitive with subjective NWS categorical forecasts of ceiling and visibility at short-range lead times. Moreover, Vislocky and Fritsch (1997) demonstrated that an observations-based statistical forecast technique not only proved superior to persistence but persistence climatology and Model Output Statistics (MOS) as well. However, while the statistical forecast clearly demonstrated improved forecast skill, it remained to be seen whether the forecast value had been improved.

The goal behind this experiment was to determine whether statistical, probabilistic weather forecasts possess more economic value than traditional, categorical forecasts. To do this, 18 daily domestic flights that arrive at three major airports were examined. Using operating costs data provided by the airline, an optimal threshold for determining whether additional fuel should be added was calculated for each flight. Then, using probabilistic forecasts and the TAF forecasts, cumulative operating costs were calculated for decisions based on each forecast method.

The analysis indicates that by using statistical, probabilistic forecasts rather than categorical forecasts over 14 months, significantly less money would be spent on fuel by utilizing probabilistic forecasts. In fact, an average of $23K is saved per flight during this period. Projecting these figures over one year and all flights for a major airline, a potential savings of approximately $50M in operating costs has been demonstrated. This is a significant annual savings for simply transitioning from traditional categorical forecasts to a probabilistic forecasting system; a system that could easily be automated and updated continuously.

Acknowledgments: The anonymous commercial airline is graciously thanked for providing the operating costs of the flights used in this study. The TAFs were provided by several National Weather Service Forecast Offices. We also thank Weathernews Americas, Inc. for its involvement in this effort.

5. REFERENCES

Dallavalle, J. P., and V. J. Dagostaro, 1995: The accuracy of ceiling and visibility forecasts produced by the National Weather Service. Preprints, *Sixth Conf. on Probability and Statistics in the Atmospheric Sciences*, Toronto, ON, Canada, Amer. Meteor. Soc., 201-206.

Keith, R., 2003: Optimization of value of aerodrome forecasts. *Wea. Forecasting*. **18**, 808–824.

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.

________, 2004: The impact of high-frequency surface weather observations on short-term probabilistic forecasts of ceilings and visibility. *J. Appl. Meteor.*, **43**, 145-156.

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.