8.11 DEMONSTRATION OF THE MARINE STRATUS FORECAST (MSF) PRODUCT

Cynthia B. Fidalgo*, Danny L. Sims, Starr L. McGettigan, and Jeffrey Weinrich
Federal Aviation Administration William J. Hughes Technical Center,
Atlantic City, New Jersey

1. INTRODUCTION

Marine stratus clouds effect the San Francisco Bay area approximately 100 days each year from May through October. Low clouds in the vicinity of the approach zone into San Francisco International Airport (SFO) prohibit dual approaches to the closely spaced parallel runways during peak arrival times and may effectively reduce airport capacity by half. Air traffic managers at the Traffic Management Unit at the Oakland Air Route Traffic Control Center (ZOA ARTCC) need to anticipate future changes in capacity (at least one hour in advance) so that they can plan accordingly. Forecasts of marine stratus dissipation, or “burn off” are provided to them via meteorologists in the ZOA Center Weather Service Unit (CWSU). Inaccurate forecasts lead to either wasted capacity when cloudiness dissipates earlier than expected, or unacceptable levels of airborne holding delay and aircraft diversions when marine stratus clouds burn off later than anticipated.

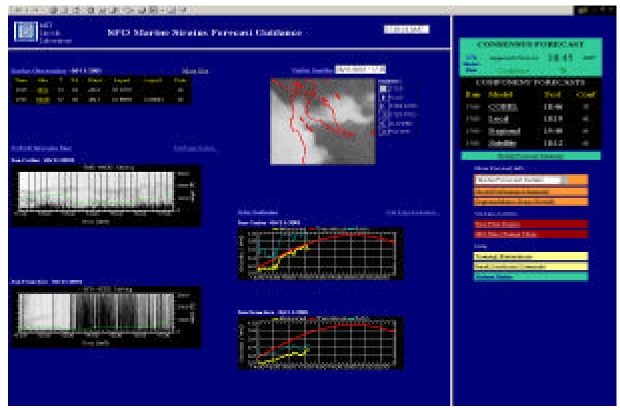
In response to the marine stratus problem at SFO, the Federal Aviation Administration (FAA) Aviation Weather Research Program (AWRP), managed by AUA-430, has sponsored the Massachusetts Institute of Technology Lincoln Laboratories (MIT/LL) in the development of the Marine Stratus Forecast (MSF) product. Development goals of the MSF were to help improve forecasts of marine stratus burn off in the approach zone of SFO, and provide demonstrated merit as a forecast guidance product. The FAA William J. Hughes Technical Center, Weather Branch (ACT-320) conducted a demonstration of the MSF product during the 2001 summer stratus season with participation from meteorologists from the ZOA CWSU and United Airlines (UAL). Overall, demonstration objectives were to assess the utility, task benefit, interface design and performance of the MSF product and components.

2. SYSTEM DESCRIPTION

The MSF product relies on weather observations from a network of sensors surrounding the Bay region. Data are collected from these sensors and transferred to a database computer located at San Jose State University. There the data are processed for display, and for input into a suite of algorithms designed to forecast the time of approach zone clearing. The display of observations and the automated forecast guidance products were made available to users via the Internet.

The MSF display was divided into two major sections: 1) Observation Display Window, and 2) Forecast Display Window, located on the left and right hand sides of the display, respectively. Figure 1 shows an overall view of the main MSF display.

Figure 1. Overall MSF Display.



The Observation Window displayed current observations of data relative to stratus conditions. Observation products included the following: Visible Satellite Depiction; SODAR (SOnic Detection And Ranging) Inversion Base graphs for SFO and the San Mateo Bridge (SMB); Solar Radiation Plots from radiometers; and Surface Weather Observations for SFO, SMB, and San Carlos (SQL). Due to latencies in collecting and processing remote data observations, the latency of the most recent data viewed on the Observation Window would typically range from 10 to 25 minutes.

The Forecast Window featured automated forecast guidance for predicting the time of stratus clearing in the SFO approach zone. There were four "Component" forecast models and one Consensus forecast. Forecast models included: 1) COBEL Model; 2) Local Statistical Forecast Model (SFM); 3) Regional SFM; 4) Satellite SFM; and 5) Consensus forecast, which weighed the other component forecasts to provide a single forecast (Clark, 2002).

3. PARTICIPANTS

Principal participants in the MSF 2001 Demonstration were the ZOA CWSU meteorologists. CWSU meteorologists provide ZOA traffic management with marine stratus burn off forecasts for the SFO approach zone. Traffic management decisions for SFO are based on these forecasts, which are issued every morning (if required) at 1300 and 1700 UTC.

Corresponding author address: Cynthia B. Fidalgo, ACT-320/Raytheon, FAA Wm. J. Hughes Technical Center, Atlantic City Intl. Airport, NJ 08405; e-mail: Cynthia.B-CTR.Fidalgo@tc.faa.gov

To a lesser extent, meteorologists staffing the West Desk at UAL's Center Weather Unit also participated. The West Desk provides forecasts for UAL's Western tier hubs, including SFO. UAL dispatchers controlling flights into SFO base flight operation decisions on the West Desk forecasts. Accurate stratus forecasts allow for more effective, strategic flight planning into SFO.

4. DEMONSTRATION APPROACH AND METRICS

The MSF Demonstration consisted of two distinctly different methodologies to determine Demonstration objectives. One method, the MSF Usability Study, evaluated the MSF product from the user perspective. Feedback from CWSU and UAL users included perceptions on the accuracy, utility, interface design, and benefit of the MSF.

In preparation for the usability portion of the MSF Demonstration, baseline data was collected at the ZOA CWSU to gather information on current practices involved in creating marine stratus burn off forecasts. Data from this visit formed the basis for the objectives of the usability study portion. Baseline data was also collected from three other sites where the MSF would be operational: 1) the Weather Unit at the FAA Air Traffic Control System Command Center (ATCSCC), Herndon, VA; 2) UAL Center Weather Unit , Chicago, IL; and 3) the National Weather Service (NWS) Forecast Office, Monterey, CA.

Prior to demonstration conduct, personnel from MIT/LL provided MSF training to forecasters at both the ZOA CWSU and UAL's Weather Center Unit. Due to unavoidable circumstances, training was not provided to ATCSCC or NWS users. (Note: Neither the NWS or the ATCSCC sites participated in the demonstration due to questionable or limited MSF usage).

The Demonstration period was divided into two data collection phases. Phase 1 occurred from August 6 – 9, 2001 (CWSU only). Phase 2 was conducted from November 12 – 17, 2001 (CWSU and UAL).

Usability data collection metrics for both phases of the demonstration included: questionnaires; structured interviews; and post-stratus telephone interviews

A Forecast Accuracy Comparison Study, measuring forecast performance, was the second demonstration method used to determine MSF effect on CWSU forecast skill. This was achieved by comparing the accuracy of the CWSU stratus burn off forecasts during the 2001 stratus season (while utilizing the MSF), to forecast accuracy data collected during 1998 and 1999's stratus seasons when the MSF was unavailable (McGettigan, 2001). Another objective of using forecast performance measures was to derive 2001 performance statistics for the CWSU and the forecast models.

All performance data was derived from CWSU forecast logs containing information on weather conditions, 1300 and 1700 UTC forecasts, and the time SFO went to side-by operations (i.e., the time that stratus has cleared over the SFO approach zone allowing dual approaches to the airport).

In order to quantify forecast skill, data from these sources were analyzed. Forecast performance measures were defined as the amount of error, in time, between forecast model and/or CWSU forecasted times to the time that stratus actually dissipated (delta time). Mean error scores were derived for the following categories: mean values for burn off and forecasted times; mean absolute error; underforecasting error; and overforecasting error.

Currently, forecast performance results are still the subject of analysis. They will be presented at the conference.

5. RESULTS

5.1 Questionnaire Results

Questionnaire results were summarized, using the median as the measure of central tendency for rating information. In general, CWSU results were mostly consistent with UAL results.

All MSF interface items displayed on both the Observation and Forecast Windows were acceptable. Interface characteristics included: meaningful item arrangement; ease of access and navigation; and distinguishability and interpretability of graphics, colors and text.

There was a slight effect on shared situational awareness (SA) between the CWSU and UAL, the NWS, and the ATCSCC. Some components of shared SA included: similar perceptions of stratus behavior; discussion of MSF information when forecast collaboration took place; and the perception of sharing a common picture from MSF observation and model forecast data.

Operational suitability results for MSF products and components displayed on both the Observation and Forecast Windows were also derived. All components were rated on the dimensions of utility, ease of use, and readability. Forecast models and SODAR were also rated on accuracy.

Results from Phase 2 (see Table 1) indicated that all MSF observation products and components were acceptable on the dimensions of utility, readability, and ease of use, with the exception of readability of the SFO SODAR and the utility of the satellite depictions were neither acceptable nor unacceptable.

For Forecast Window products, results indicated that only COBEL and the Consensus forecast were considered operationally acceptable for accuracy (ratings of 4 or greater). The Regional SFM and the Local SFM were considered less accurate; whereas the accuracy of the Satellite SFM was operationally unacceptable. Perceptions of the accuracy of the Regional SFM, Local SFM, and Satellite SFM significantly decreased from Phase 1, wherein perceived accuracy was operationally acceptable. Other dimensions were mostly consistent with Phase 1 results (not shown).

Table 1. MSF Operational Acceptability CWSU Phase 2 Ratings.

Median Ratings for MSF Operational Acceptability				
Product	**Utility**	**Read-ability**	**Ease of Use**	**Accuracy**
Observation Window	4	4	4	
SODAR SFO	4.5	3.5	4	4
SODAR SQL	5	5	5	4.5
Satellite Depiciton	3.5	4	4	
Radiometer	5	5	5	
Surface Observations	4	4	4	
Forecast Window	4.5	4.5	4.5	
Consensus Forecast	4	4.5	4	4
COBEL	3.5	4.5	4	4
Local SFM	4	4.5	4	3.5
Regional SFM	3	4	4	3
Satellite SFM	3	4.5	3	2.5

Scale: 5 = Highly Acceptable, 4 = Acceptable,
3 = Neither Acceptable nor Unacceptable,
2 = Unacceptable, 1 = Highly Unacceptable

Frequency of use results from Phase 2 are shown in Table 2. SODAR and Radiometer data were almost always used. The Consensus forecast was used most frequently of all the forecast models. Satellite SFM and COBEL were used now and then. Frequency of use of COBEL diminished significantly from Phase 1 to Phase 2.

Table 2. Frequency of Use CWSU Ratings from Phase 2.

Median Ratings of CWSU Frequency of Use of MSF Products and Components	
Product	**Frequency**
Observation Window	5
SODAR SFO	5
SODAR SQL	4
Satellite Depiciton	4
Radiometer	5
Surface Observations	4
Forecast Window	4
Consensus Forecast	4.5
COBEL	3.5
Local SFM	4
Regional SFM	4
Satellite SFM	3.5

Scale: 5 = Almost Always, 4 = Frequently,
3 = Now and Then, 2 = Seldom, 1 = Hardly Ever

UAL questionnaire results were consistent with CWSU results with the following exceptions: satellite observations were highly rated; all forecast models were considered operationally acceptable on all dimensions; and frequency of use was less than that for the CWSU.

5.2 Interview Results

Results from the structured interviews indicated that, in general, all CWSU users were positive about the MSF product. However, the most positive perceptions pertained to observation data, particularly SODAR data. The desire to maintain the SODAR depictions in the future was oftentimes reported. Although the SFO SODAR depiction was reportedly unclear and the inversion height was difficult to distinguish, it was still considered very useful. The satellite image was thought to be unnecessary, since the CWSU has satellite depictions through its primary weather information workstation.

Overall, model accuracy was questionable, and confidence was not high. In particular, the Satellite SFM seemed to have difficulty when mid or high level clouds existed above the marine stratus. The Consensus forecast was viewed most favorably among all forecast models. Views on COBEL were mixed in that initially it seemed to perform well, but over time seemed inconsistent. Perceptions of the Regional and Local SFM performance were somewhat acceptable, although judgments on their overall performance or when they performed poorly or well, were unknown. Consistency of all model forecast output (when all or most models are predicting the same forecasted burn off time), and confirmation of the CWSU derived forecasts, promoted higher model confidence. Most forecasters indicated that when the models can out-perform the CWSU, then model confidence would increase.

Model use, initially, was generally infrequent, although model output was looked at and sometimes used to confirm CWSU forecasts. All users reported that the initial model output at 1300 UTC was too late to be factored into their initial forecast and that earlier output would be desired. Many forecasters noted that they had not had enough experience with the models to understand their strengths and weaknesses.

Forecasters indicated that model forecasts tended to be accurate under typical conditions, i.e., high pressure, weak onshore winds, and solar radiation primarily responsible for burn off. These conditions were stated as easy for forecasters and models alike. Conditions involving instability, advection, wind changes, and upper level features were considered difficult for both forecasters and models.

Suggested MSF improvements included: improved model forecast accuracy (considered a priority for future development); improved SFO SODAR depiction; larger or better monitor placement for easier viewing; and earlier and later model runs.

Interview data from UAL was limited to only two meteorologists. Their remarks were mostly consistent with the CWSU's. However, contrary to the CWSU, the satellite depiction was considered very useful. This may be a factor of UAL having a different weather information workstation.

Results from the CWSU post-stratus telephone interviews were mixed. On days identified by the CWSU as "difficult" stratus forecast days, perceptions of model performance were negative overall. For more typical or easier stratus days, model performance perception was more positive.

6. CONCLUSIONS AND RECOMMENDATIONS

CWSU users indicated that the most useful attributes of the MSF were the observations, especially SODAR and radiometer data. Maintenance of these products for future operational use was highly advocated.

Overall perceptions of forecast model performance and accuracy were questionable. Performance perceptions of the Satellite SFM were mostly negative, while performance of the Consensus model was viewed most favorably. According to users, future MSF development should focus on improved model forecast accuracy.

Some of the more salient suggestions for future MSF development included: earlier forecast model output for assistance with the CWSU 1300 UTC forecast; later model runs (beyond 1800 UTC) for when stratus lingers or lasts all day; and a clearer depiction of the SFO SODAR output.

The forecast models were mostly used to confirm or check against meteorologists' derived burn off forecasts. This in part is a function of time, since the model forecasts are not available until 1300 UTC when the initial CWSU forecast has already been issued. Model use beyond forecast confirmation was rarely reported.

When the models reflected forecasted times closer to CWSU derived forecasts, confidence in the model output increased. Model confidence also seemed to increase when all model output was consistent.

Overall, forecasters have indicated that greater familiarity with the MSF forecasts is needed to make informed judgments on their performance. This would also increase sensitivity to conditions where models perform well or poorly and may lead to greater use.

CWSU users tended to assume that, overall, model forecasts were less accurate than the CWSU forecasts. Model usefulness and confidence was often predicated on the following criteria: if it outperforms the human forecaster, then confidence in the forecasts will increase. According to forecasters, there was no indication of this occurring. It is anticipated that performance statistics will indicate whether or not this perception is correct.

7. ACKNOWLEDGEMENTS

The authors wish to thankfully acknowledge the participation of the ZOA CWSU Meteorologists and UAL West Desk Meteorologists for giving their time and effort during the MSF demonstration. Special thanks are offered to Walt Strach, ZOA CWSU Meteorologist–in-Charge for his efforts, not only by offering valuable feedback, but by coordinating his staff and helping to ensure smooth demonstration progress. The efforts of Garry Hinds, UAL Weather Center Operations Manager in providing information and coordinating data collection are also appreciated. Thanks are also extended to MIT/LL MSF developers, Wes Wilson (now at the National Center for Atmospheric Research) and Dave Clark for their responsiveness and willingness to provide data and other information regarding the MSF product.

8. REFERENCES

Clark, D.A., 2002: The 2001 Demonstration of Automated Cloud Forecast Guidance Products for San Francisco International Airport. Preprints, Tenth Conference on Aviation, Range, and Aerospace Meteorology, Portland, OR, American Meteorological Society, Boston.

McGettigan, S. and D. Sims, 2001: Ceiling & Visibility Marine Stratus Forecast Baseline Effort for San Francisco Airport. Unpublished report, ACT-320, FAA William J. Hughes Technical Center, Atlantic City, NJ.