# EVALUATION OF ENROUTE CONVECTIVE WEATHER AVOIDANCE MODELS BASED ON PLANNED AND OBSERVED FLIGHT*

Michael P Matthews†
Rich DeLaura
*Massachusetts Institute of Technology, Lincoln Laboratory*
*244 Wood Street, Lexington, MA 02420-9185*

## 1. INTRODUCTION

The effective management of convective weather in congested airspace requires decision support tools that can translate weather information available to air traffic managers into anticipated impact on air traffic operations. The Convective Weather Avoidance Model (CWAM) has been under development at Lincoln Lab under sponsorship of NASA to develop a correlation between pilot behavior and observable weather parameters (Chan et al, 2007). To date, the observable weather parameters have been derived primarily from the Corridor Integrated Weather System (CIWS) (Klingle-Wilson and Evans, 2005 and Evans and Ducot, 2006) high resolution Vertically Integrated Liquid (VIL) precipitation map and the CIWS Echo Top product. Deviations have been identified using a database of planned and actual en route flight trajectories based on data from the Enhanced Traffic management System (ETMS). CWAM has used a simple model that is based upon finding weather encounters and then comparing the distance between the planned and actual flight trajectories to define pilot deviations. Due to a large number of false deviations from this model, a significant amount of hand editing was required to use the database.

This paper will present two areas of work to improve the performance and usefulness of the enroute convective weather avoidance models. First, an improved automated algorithm to detect weather-related deviations that significantly reduces the percentage of false deviation detections will be presented. This new model includes additional information on each deviation, including the location where the decision was made to deviate. The additional information extracted from this algorithm can be used to evaluate the conditions at the decision time which may impact the severity of weather pilots are willing to penetrate. The new deviation detection algorithm has also reduced the amount of hand editing required by limiting the deviation detections to only include weather encounters that occur within 15 minutes of the time the aircraft deviated from the planned flight path.

The second focus of this paper will be the evaluation of the Convective Weather Avoidance Model. Six weather impact days from 2007 and 2008 have been added to the existing case set from 2006, tripling the number of flight trajectories used in validating the model. The deviation prediction skill for several variations of the CWAM will be compared, and results will be presented as a function of air traffic control region, weather type and forecast lead time.

## 2. DEVIATION MODEL

The Convective Weather Avoidance Model described in the initial version (DeLaura and Evans, 2006 and DeLaura et al., 2008) defines weather encounters by looking for minimum weather thresholds along a planned flight trajectory. Once an encounter was established along the planned trajectory, the actual trajectory points nearest to the encounter were collected and the mean and maximum distance between the two paths were calculated and compared to a deviation distance threshold to determine if the aircraft was deviating. These thresholds were set based upon the typical maneuverability of flights observed along a route in clear air conditions. The initial version required review by an analyst and a significant amount of hand editing of the deviation flags.

The original CWAM deviation database developed in 2006 consisted of approximately 1500 weather encounters, with almost 500 of those deviating due to the weather. In order to perform a more comprehensive evaluation of the CWAM and look towards additional variables or situations which could improve the model, an expansion of the deviation database was required.

*This work was sponsored by the National Aeronautics and Space Administration (NASA) under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

†*Corresponding author address*: Michael P Matthews, MIT Lincoln Laboratory, 244 Wood Street, Lexington, MA 02420-9185; e-mail: mpm@ll.mit.edu

However, to expand the database, improvements would be required to the deviation detection model to reduce the level of effort required to review and edit the data. An improved deviation detection model has been developed which reduces the deviation classification error rate from roughly one-third to just under ten percent. Several examples of the deviation classification done by the algorithm are shown in figure 1.

The major improvement to the deviation detection algorithm is to perform processing of the planned and actual trajectories independently from the weather encounters to identify deviations. A deviation is defined as having a decision point and an end point. The decision point is the point at which the actual trajectory separates from the planned trajectory. The end point is the point at which the actual trajectory rejoins the planned trajectory or 15 minutes after the decision point, whichever comes first. The 15 minute limit is chosen with the assumption that the pilot is likely unaware of the specific characteristics of the weather in the more distant future. This limit also eliminates a significant number of cases that would have required hand editing to eliminate short cuts, airborne holding, or other unplanned reroutes that heavily impacted the results of the initial algorithm.

The deviation detection algorithm begins by computing the distance between each point on the planned trajectory and the nearest actual trajectory point. The mean deviation threshold (nominally 20 kilometers) representing the limits of normal operational variations in flight trajectories is then applied to the distance computation to flag any trajectory point in which the aircraft is likely deviating. If at least six contiguous points are larger than the mean deviation threshold, representing 60 seconds of flight time, a deviation is declared. Next, deviations with less than 30 seconds (3 data points) between them are joined together. Finally, the algorithm will identify the beginning of each deviation (first point greater than mean deviation threshold) and search backwards along the planned trajectory to the point where the planned and actual trajectories are less than 2 kilometers apart. This point is defined as the decision point.

The deviation detection algorithm also performs several functions to identify scenarios in which the algorithm is unable to adequately characterize or detect deviations. These scenarios include aircraft that have entered airborne holding patterns and deviations that occur prior to the aircraft entering the region of interest. Deviations that may be detected in these scenarios will be flagged as invalid in the database.

The improved deviation detection algorithm does not modify the method used to identify weather encounters. A weather encounter is defined as a portion of a planned trajectory that passes through *either* VIL level 2 or greater *or* echo tops of 25 kft or greater for at least 2 minutes. Multiple weather encounters may be defined along a flight path if the weather along the planned trajectory drops below the encounter thresholds (VIL level 2 and echo tops of 25 kft) for at least one minute.

Once all of the weather encounters and deviations are determined, the algorithm will merge the two, looking for overlapping deviations with weather encounters. For weather encounters that coincide with deviations the weather encounter will be flagged as a deviation and three critical points will be stored in the database: the decision point along with the beginning and ending points of the weather encounter. For weather encounters that do not coincide with deviations the encounter will be flagged as a non-deviation and the beginning and ending points of the weather encounter will be stored. For non-deviations an estimate of the decision point is computed that is nominally four minutes of flight time prior to the beginning of the encounter. Deviations that do not coincide with weather encounters are assumed to be unrelated to the weather.

In addition to reducing the amount of hand-editing required to use the database for the CWAM, the identification of a specific decision point supports future analysis of pilot decision making; for instance, comparing IFR vs. visual conditions when making the decision to deviate, estimating the distance from the deviation decision point to weather encounter, etc. The classification algorithm also identifies the reason for rejection of each trajectory that is edited out of the dataset, supporting further review of algorithm performance.

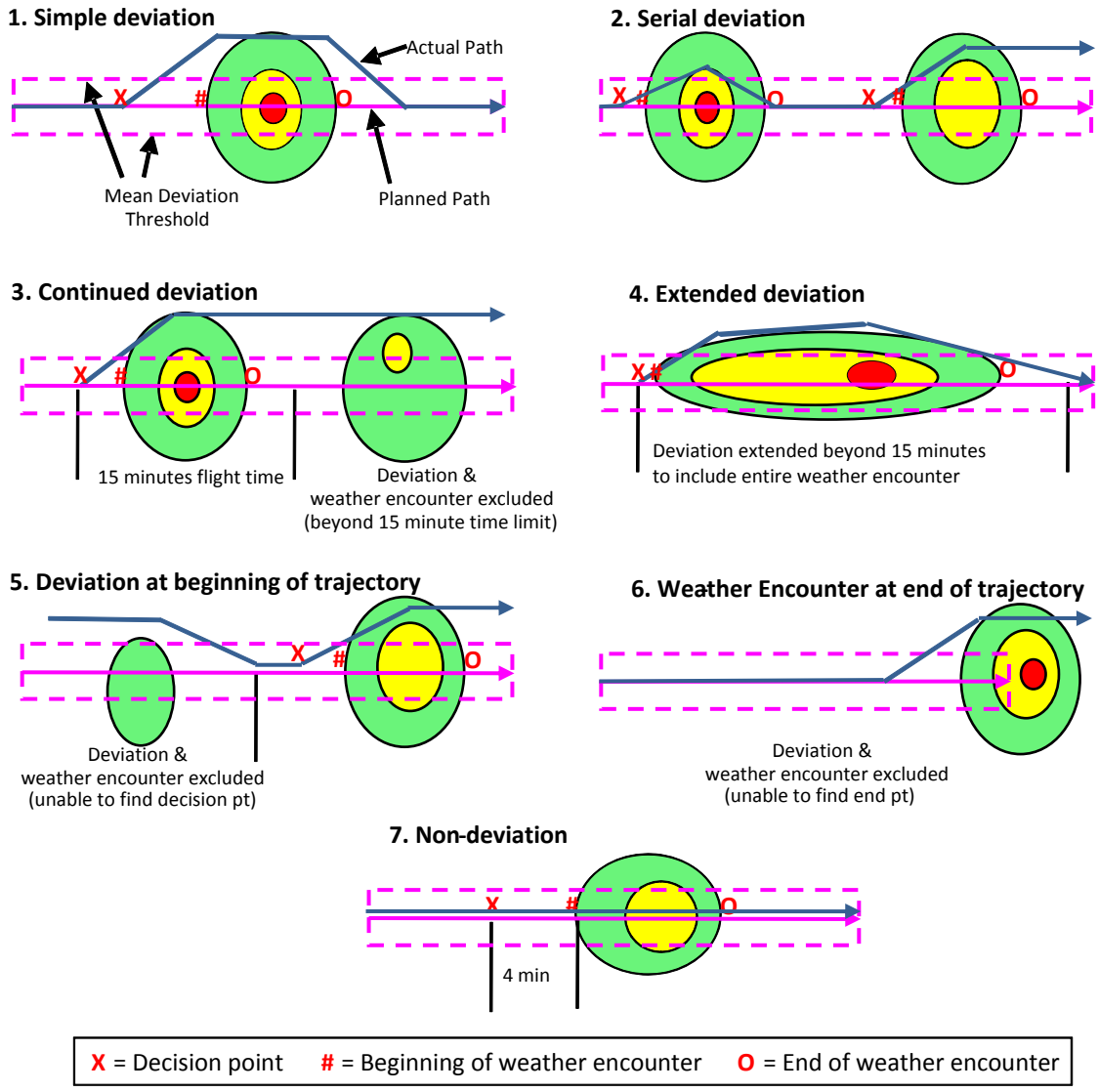



*Figure 1.Example illustrations of deviations, non-deviations and deviation filters from improved deviation detection model.*

## 3. DESCRIPTION OF THE CONVECTIVE WEATHER AVOIDANCE MODEL AND EVALUATION DATASET

The Convective Weather Avoidance Model translates deterministic weather information grids from systems such as the Corridor Integrated Weather System (CIWS) into estimates of the likelihood of pilot deviation. The CWAM output is a three-dimensional Weather Avoidance Field (WAF), which estimates the probability (0 to 100%) that a pilot will deviate around the convective weather at each point in the WAF grid. From the WAF, automated decision support tools could estimate the capacity reduction due to weather and provide alternative route suggestions to avoid the convective weather. The CWAM evaluated here applies to flights above 25kft in the en-route airspace.

Figure 2 shows the steps involved in generating a WAF from the CWAM. Spatial filters are run on each weather input to generate deviation predictors. The two best predictors of deviation were identified, using a Gaussian classification algorithm. Two-dimensional histograms, based on the observed deviation statistics, give the observed probability of deviation as a function of the two best predictors. These histograms are smoothed and extrapolated to create a deviation probability lookup table that is used to generate the WAF.

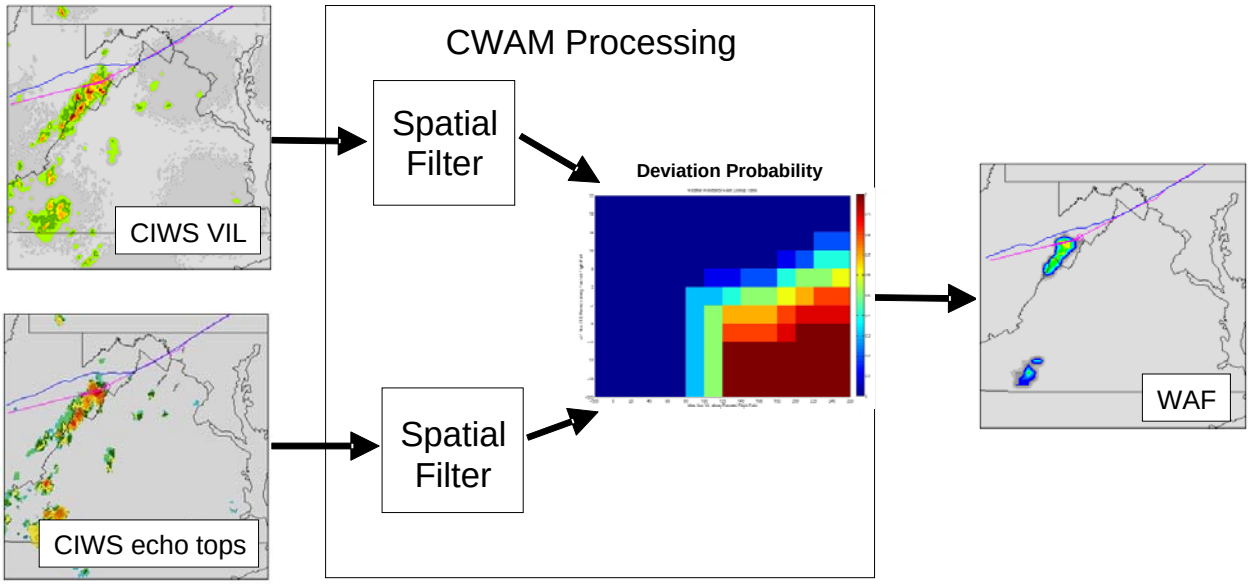



*Figure 2. Generation of Weather Avoidance Field (WAF) using the Convective Weather Avoidance Model (CWAM).*

For this evaluation, the CWAM database was expanded from the original 1,500 encounters from six case days to just over 5,200 encounters, with the addition of data from five case days during the summer of 2007 and one case day during the winter of 2008. Flight trajectories were collected from three Air Route Traffic Control Centers (ARTCC); ZID, ZOB, and ZDC. Figure 3 illustrates the ARTCCs, showing the major jet routes and fair weather traffic in each. The majority of ZOB traffic flow is carried along several parallel and closely spaced East-West oriented jet routes. The ZID traffic is evenly distributed among several jet routes with different orientations. The ZDC traffic is primarily orientated Northeast to Southwest and carries a heavy demand of aircraft leaving the Northeast corridor. Table 1 presents a summary of the case dates and times, the ARTCCs analyzed on each day, and the number of deviations and non-deviations.

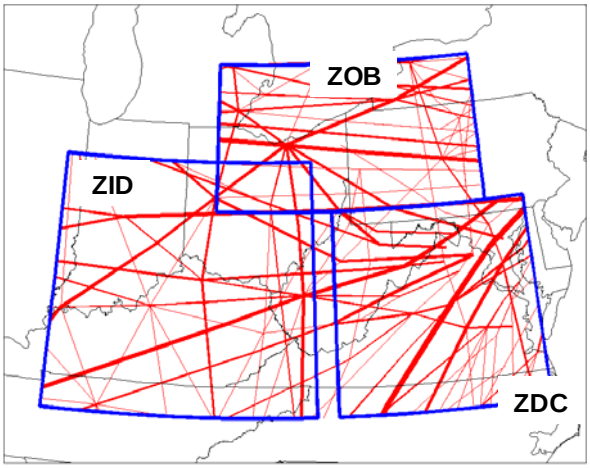



*Figure 3. ARTCCs studied in the CWAM. Red lines are jet routes; thicknesses are proportional to peak traffic load.*

**Table 1. Summary of case days in Convective Weather Avoidance Model.**

| Date | Start (GMT) | End (GMT) | ARTCCs | Deviations | Non-Deviations |
|---|---|---|---|---|---|
| June 1, 2006 | 1700 | 2400 | ZDC,ZID,ZOB | 141 | 206 |
| June 19 , 2006 | 0000 | 2400 | ZDC,ZID,ZOB | 125 | 565 |
| June 23, 2006 | 0000 | 0800 | ZDC,ZOB | 30 | 29 |
| July 12, 2006 | 0800 | 2100 | ZID | 7 | 93 |
| July 14, 2006 | 1900 | 2400 | ZDC,ZID,ZOB | 153 | 176 |
| September 22, 2006 | 1100 | 2400 | ZID | 7 | 81 |
| June 4, 2007 | 1600 | 2400 | ZDC,ZID,ZOB | 164 | 279 |
| June 27, 2007 | 1700 | 2100 | ZDC,ZID,ZOB | 132 | 149 |
| July 10, 2007 | 1300 | 2400 | ZDC,ZID,ZOB | 171 | 166 |
| July 19, 2007 | 1500 | (20)0200 | ZDC,ZID,ZOB | 227 | 626 |
| August 8, 2007 | 0000 | (9)0300 | ZDC,ZID,ZOB | 373 | 635 |
| February 5, 2008 | 0900 | (6)0500 | ZDC,ZID,ZOB | 34 | 666 |

The additional days represent a wider variety of weather regimens than were in the original CWAM data set, which were heavily weighted toward well organized convection. For instance, June 4, 2007 saw a cold pool of air settling into the region in the upper atmosphere that triggered a large region of convectively induced cells that did not grow very tall or produce a significant amount of precipitation. Figure 4 depicts a flight deviating around one of these many cells.

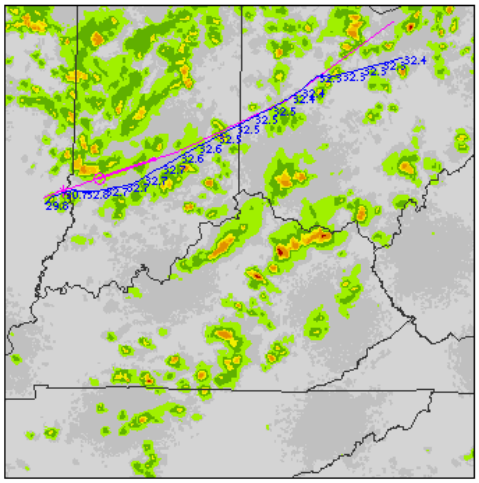


*Figure 4. CIWS precipitation image and aircraft trajectory on June 4, 2007 at 20:18:40Z.*

The February 5, 2008 case represented a winter event with a low pressure system tracking through the Midwest with the associated warm and cold fronts. Although the VIL levels were fairly intense with the cold front, the echo tops were at or below 40kft. Figure 5 depicts a flight penetrating the cold front as was common on this day.

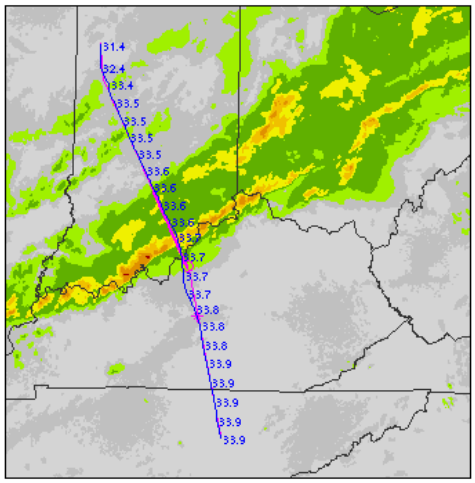


*Figure 5. CIWS precipitation image and aircraft trajectory on February 5, 2008 at 12:40:29Z.*

Figure 6 depicts a deviating flight from June 27, 2007. On this day the cells appear similar in size and scale to the 4th of June, however these cells grew to over 50kft as is typical on a warm summer day. Finally, figure 7 depicts a flight deviating around a mesoscale convective complex on July 10, 2007. On this day a significant number of aircraft were deviating around the weather.

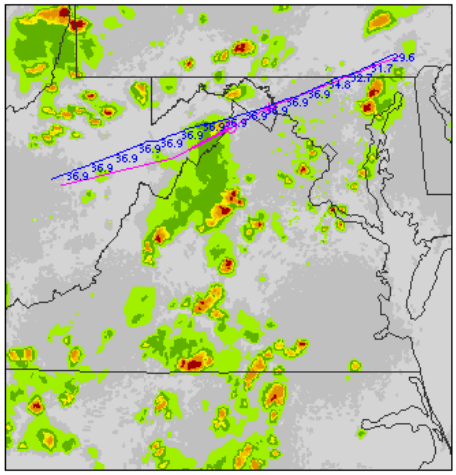


*Figure 6. CIWS precipitation image and aircraft trajectory on June 27, 2007 at 20:03:40Z.*

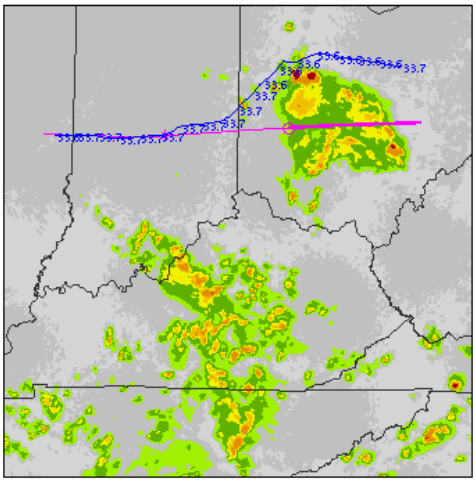


*Figure 7. CIWS precipitation image and aircraft trajectory on July 10, 2007 at 17:16:20Z.*

## 4. CWAM EVALUATION

The original CWAM (CWAM-ORIG) identified the difference between flight altitude and echo top height as the primary predictor of deviation to avoid convective weather, and the percent of area covered by VIL>=level 3 as a secondary predictor. The echo top height used in the difference calculation was the 90$^{th}$ percentile calculated over a 16 x 16 km kernel. The VIL coverage kernel was 60 x 60 km. Logically, these two filters represent the regional coverage of heavy weather and whether the flight was above or below the tallest nearby echo tops.

In order to validate CWAM-ORIG against the new test dataset from 2007 and 2008 case days, the deviation and non-deviation probability distributions as a function of maximum WAF deviation probability encountered along the planned trajectory were calculated and compared (figure 8). Probability distributions from the training and the test datasets are similar and well-calibrated, suggesting that CWAM-ORIG may be applied to a wide range of weather regimes with consistent results. All evaluations described in the remainder of this paper were performed on the complete trajectory dataset (2006, 2007, 2008).

CWAM deviation prediction errors are greatest where the uncertainty is greatest (probability of deviation is in the 30 – 70% range, illustrated on the two-dimensional deviation probability histogram in figure 9). The best deviation predictors, resulting in the lowest CWAM prediction error, are characterized by the smallest range of uncertainty. In an effort to further improve CWAM performance, three variations were defined, and their deviation prediction errors were compared to CWAM-ORIG.

CWAM-ORIG-LITE is based on the same two deviation predictors as CWAM-ORIG, but uses smaller spatial filter kernels on the echo top (4 x 4 km) and VIL (16 x 16 km) fields. CWAM-1KM does not apply spatial filters to either weather input; WAF deviation probabilities at each grid point are based on echo top and VIL values at that grid point only. Finally, CWAM-16KM-MAX uses the 90$^{th}$ percentile value of both echo top and VIL in a 16 x 16 km kernel as deviation predictors. Figures 10-13 present the two-dimensional histograms of observed deviation probability, and the smoothed / extrapolated CWAM deviation probability lookup tables for CWAM-ORIG, CWAM-ORIG-LITE, CWAM-1KM, and CWAM-16KM-MAX, respectively. Figure 14 shows the WAF outputs from each of the four CWAM variations for the illustrated VIL and echo top fields.

The skill of CWAM deviation predictions was evaluated by turning CWAM deviation probabilities into deterministic deviation predictions by applying a deviation probability threshold ($P_{dev}$). First, each trajectory in the database was analyzed to find the maximum WAF ($P_{max}$) along the planned path. A deviation was predicted for flights whose $P_{max}>=P_{dev}$. Finally each deviation prediction was classified as a hit (correct prediction of deviation), a miss (no deviation predicted for trajectory classified as a deviation), or a false (deviation predicted for a trajectory that did not deviate).

For each CWAM, the probability of correct deviation prediction (PoD), probability of incorrect deviation detection (false alarm rate, or FAR), and critical skill index (CSI) were calculated for different values of $P_{dev}$. These quantities are defined as:

PoD = hits / ( hits + misses )
FAR = false / ( hits + false )
CSI = hits / ( hits + misses + false )

POD vs. FAR plots can be created, with the optimum performance being the closest point to the top left corner. The highest WAF threshold of 100 will be in the lower left hand corner (low false rate and low detection rate) and the lowest WAF threshold of 10 will be in the upper right hand corner (high false alarm rate and high detection rate). The CSI, which provides one number to identify the optimum performance, allows easy identification of the best WAF threshold ($P_{dev}$) for deviation predictions. The POD vs. FAR and CSI results for the four CWAM versions at 10 WAF thresholds are shown in figure 15.

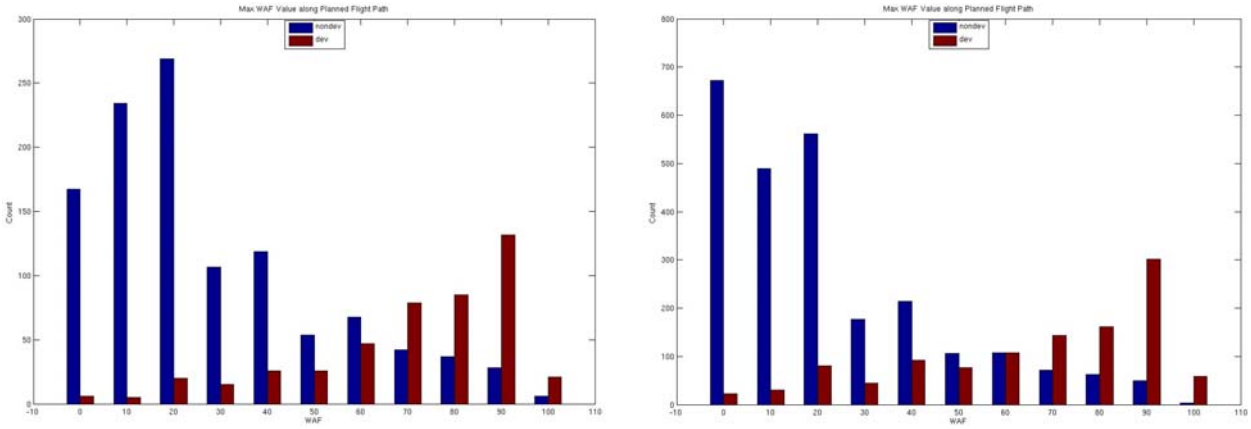



*Figure 8. Deviation and non-deviation probability distributions for CWAM-ORIG from the initial dataset (a) and the new test dataset (b).*

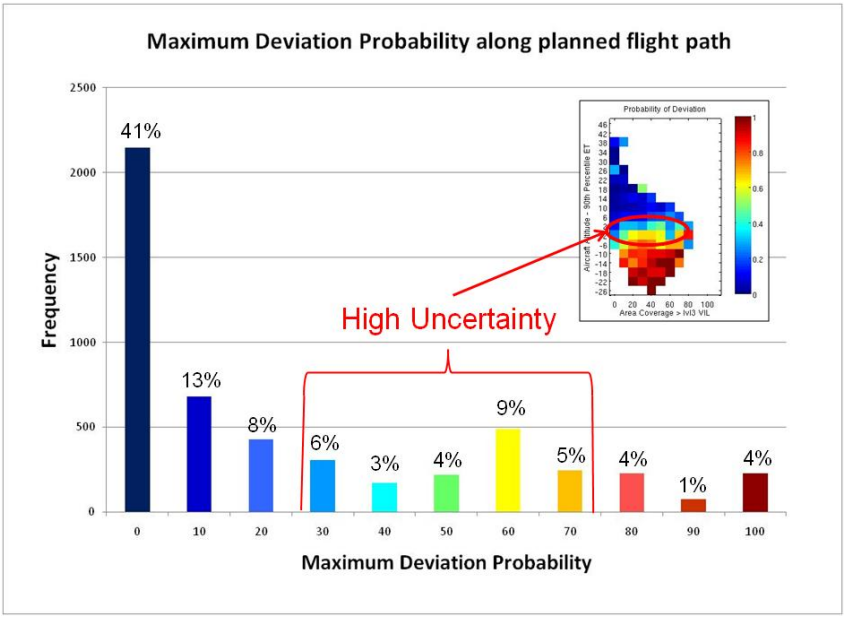



*Figure 9. Maximum Deviation Probability distributions for CWAM-ORIG from the new test dataset denoting the region of uncertainty for the deviation prediction model between 30 and 70 percent. Percentages give the percent of total encounters in the dataset whose planned trajectories encountered the specified maximum deviation probability.*

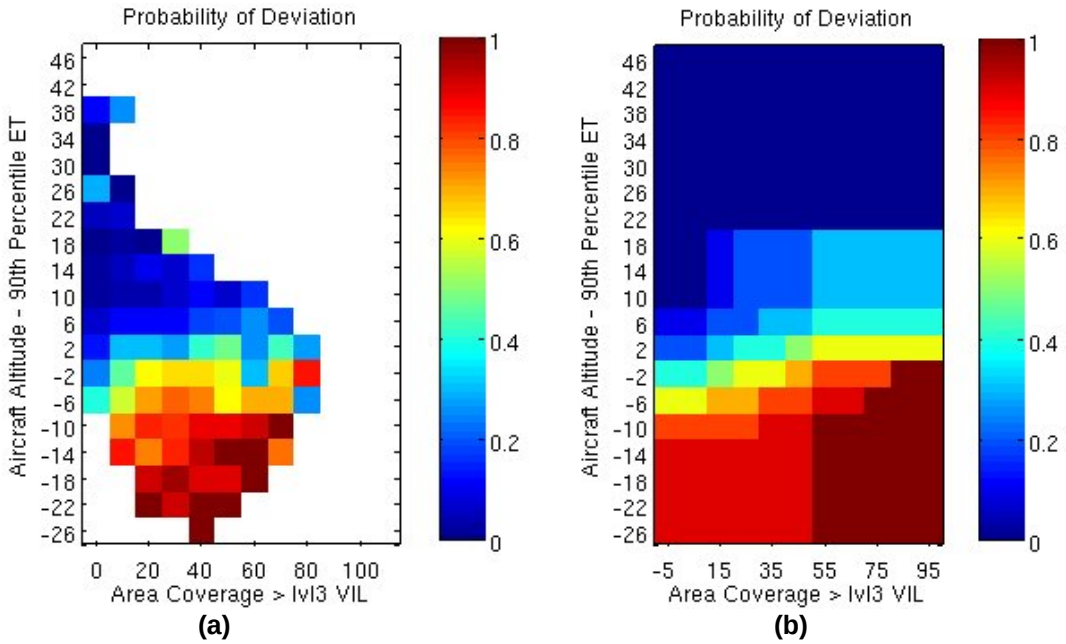



*Figure 10. Observed deviation probability for CWAM-ORIG (a) and smoothed / extrapolated CWAM deviation probability lookup table (b).*

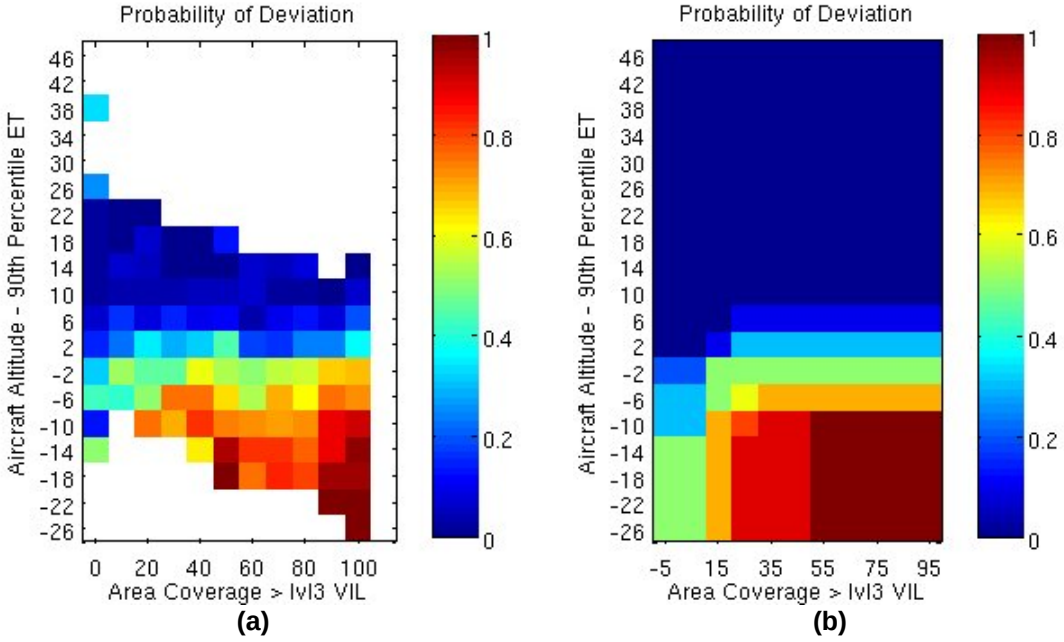



*Figure 11. Observed deviation probability for CWAM-ORIG-LITE (a) and smoothed / extrapolated CWAM deviation probability lookup table (b).*

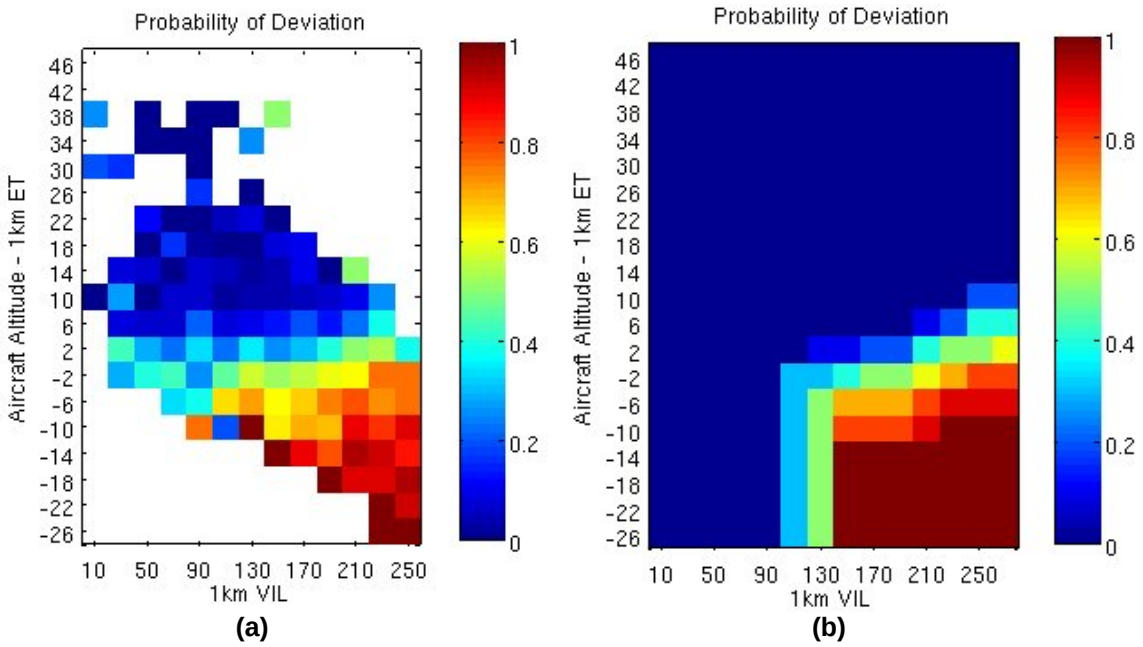



*Figure 12. Observed deviation probability for CWAM-1KM (a) and smoothed / extrapolated CWAM deviation probability lookup table (b).*

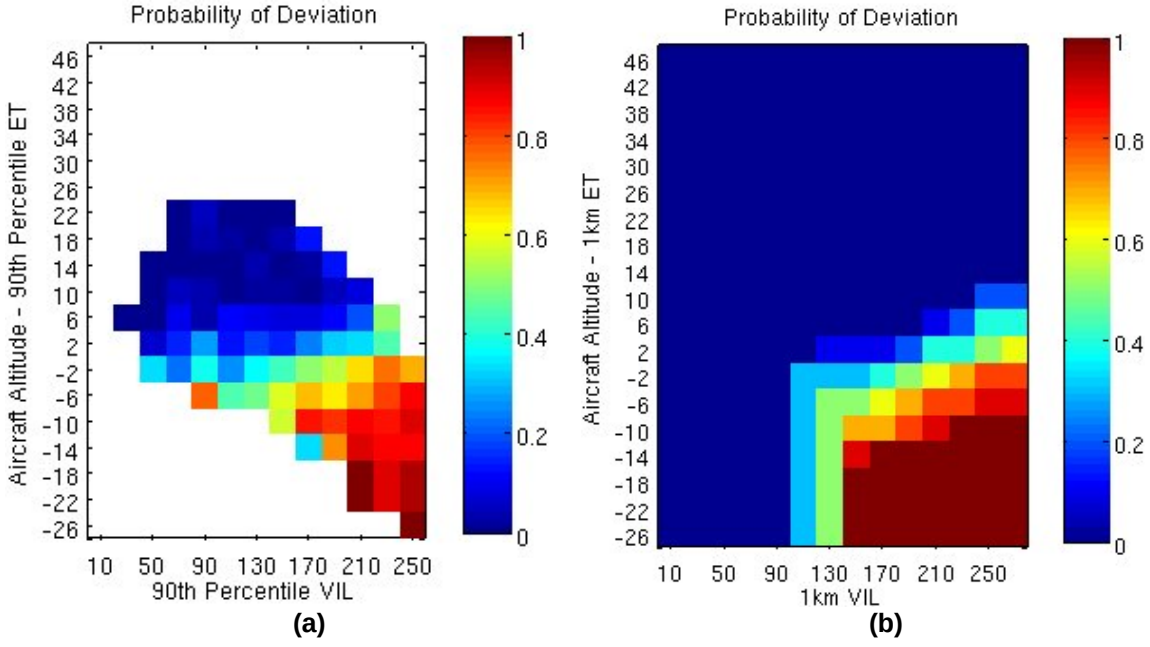



*Figure 13. Observed deviation probability for CWAM-16KM-MAX (a) and smoothed / extrapolated CWAM deviation probability lookup table (b).*

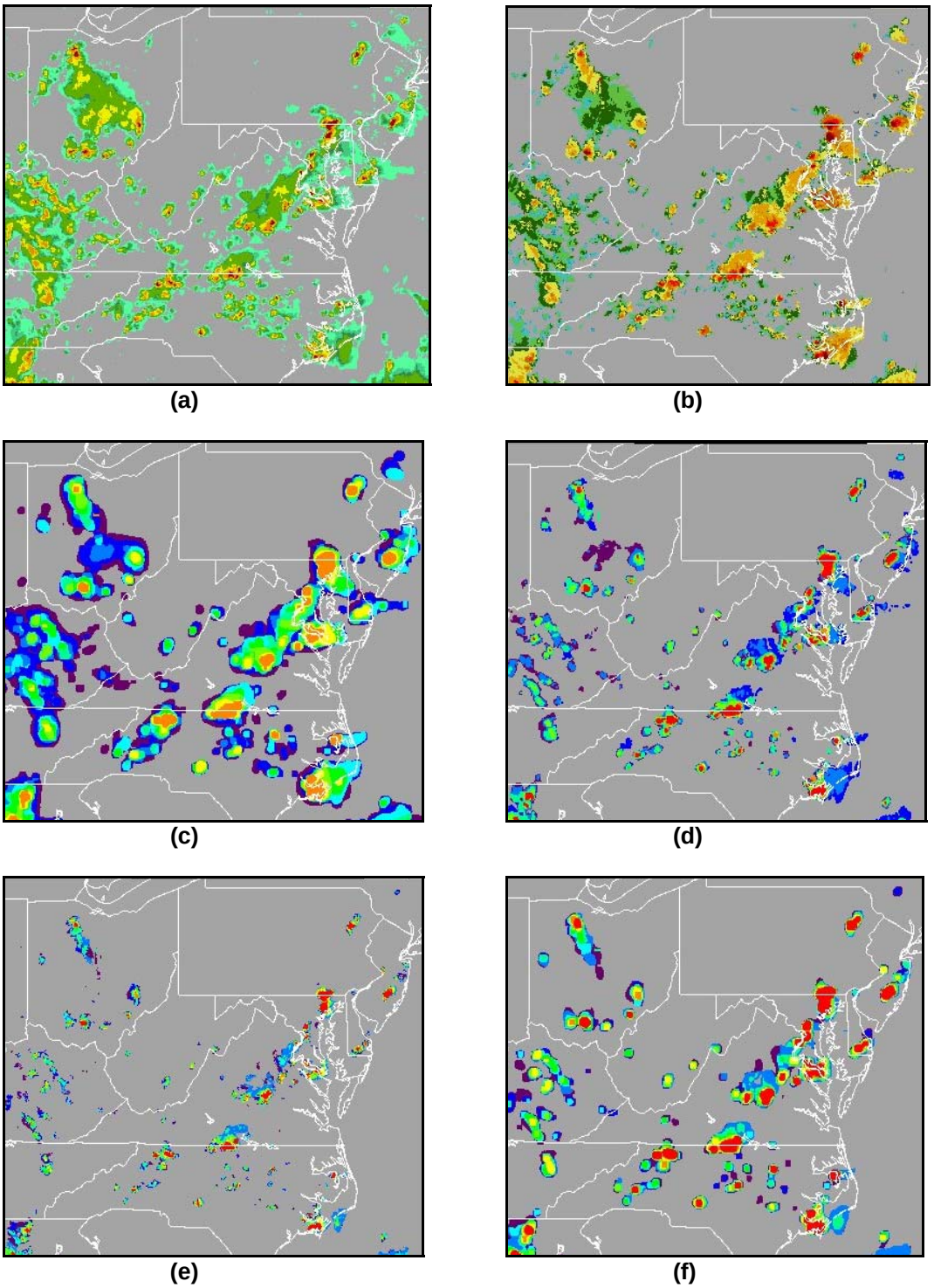
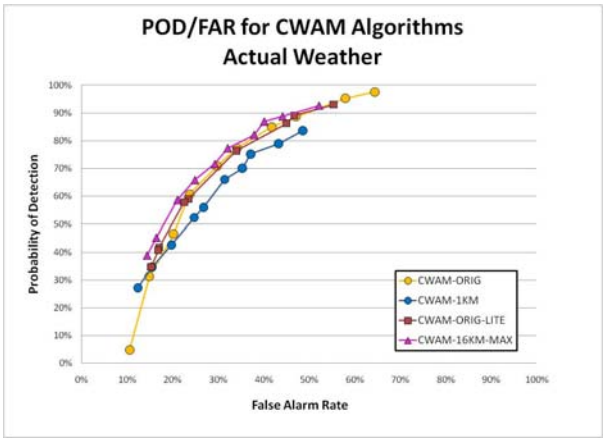
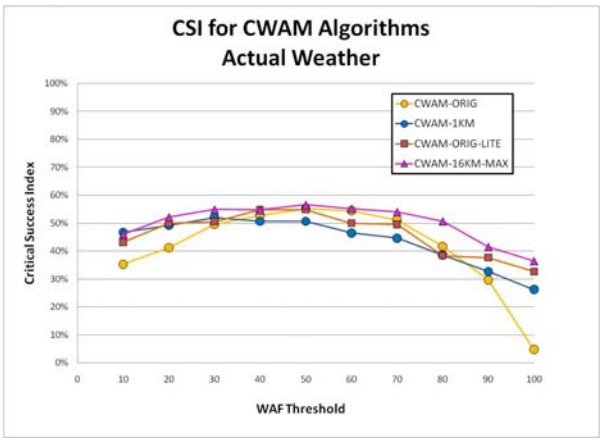



*Figure 14. Input VIL (a) and echo top (b) fields, and output WAF from CWAM variations CWAM-ORIG(c), CWAM-ORIG-LITE(d), CWAM-1KM(e) and CWAM-16KM-MAX(f) for July 10, 2007 18:40Z.*




**(a)**




**(b)**

*Figure 15. Probability of Detection vs. False Alarm Rate **(a)** and Critical Success Index **(b)** for original CWAM, CWAM Lite, 1km CWAM, and 16km Max CWAM.*

A comparison of the four CWAM versions in figure 15 shows the performance of the 1km CWAM is not as good as the other three versions. This indicates that performing a spatial filter on the weather is important for optimum results. However, any differences between the three versions using a spatial filter appear to be negligible. CWAM-16KM-MAX does show slightly better results suggesting that a 90$^{th}$ percentile filter on VIL may be better than using a level 3 area coverage filter. The results also indicate that the optimal WAF deviation probability threshold of 70 produces a POD of ~65% and a FAR of ~25% for CWAM-16KM-MAX.

The preceding analysis evaluates the deviation prediction skill of the CWAM variants based on actual weather. However, for planning purposes, air traffic control managers need to know which regions will be impacted by weather well in advance. To do this, ATC makes use of weather forecasts provided by CIWS and other sources. Ultimately, the CWAM will be combined with weather forecasts from CIWS to predict future regions that will be closed or impacted due to the weather.

The same methods used to measure the performance of CWAM using actual weather can be expanded to use forecasted weather. Figure 16 shows the POD vs. FAR from CWAM-ORIG WAF calculated from the actual weather (FCST000), the 60 minute forecast (FCST060) and the 120 minute (FCST120) forecast. As expected the performance decreases with lead time due to errors in the location or intensity of the forecasted weather.

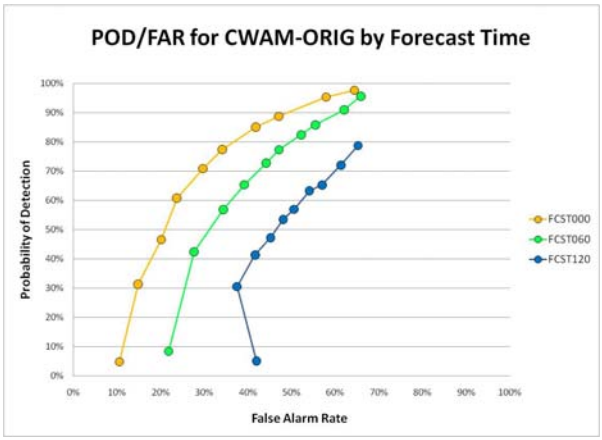
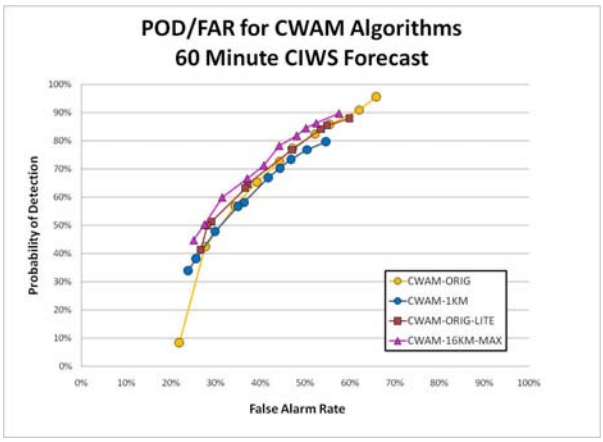
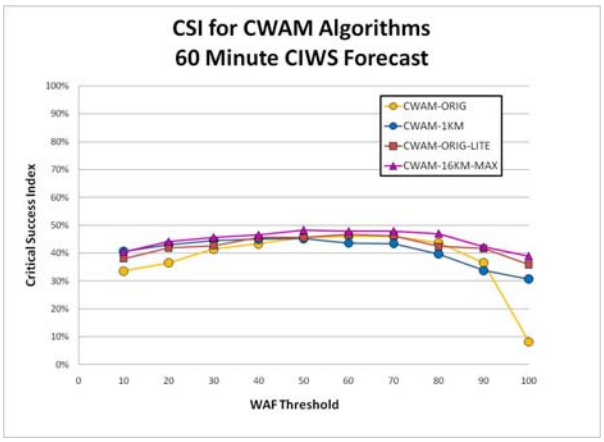



*Figure 16. Probability of Detection vs. False Alarm Rate for the original CWAM at encounter time (T=0) and 60 and 120 minute forecasts.*

Figures 17 and 18 compares the POD vs. FAR and CSI curves calculated for all four CWAM using 60 minute (figure 17) and 120 minute (figure 18) forecast WAF. As the forecast time horizon increases, the differences in deviation prediction errors associated with the different CWAM variants decreases. The deviation prediction performance of the three spatially filtered CWAM variants is virtually identical for the two hour forecasts. This finding suggests that the the characteristics of the forecast – spatial smoothing and forecast error – have a greater impact on CWAM deviation prediction accuracy then the choice of spatial filter applied in the CWAM itself.

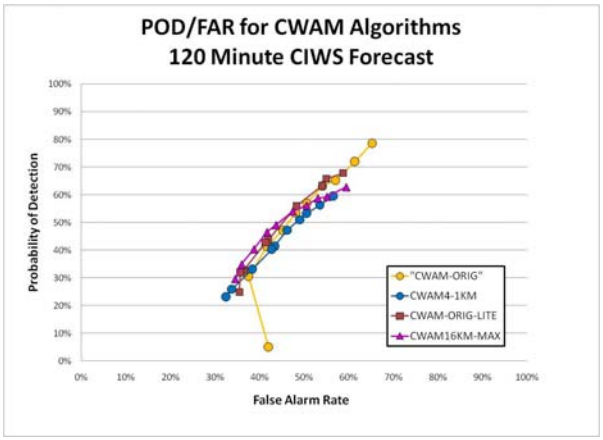



**(a)**

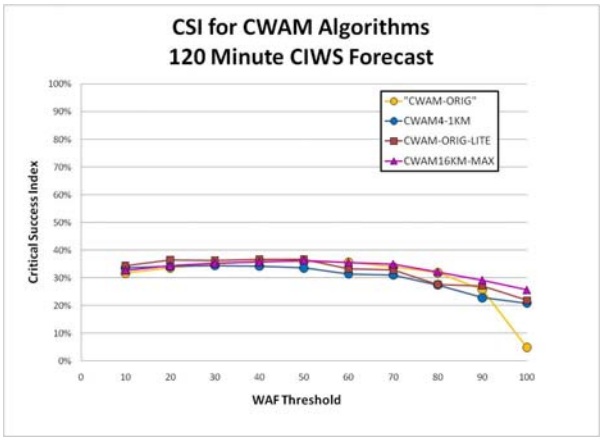



**(b)**

*Figure 17. Probability of Detection vs. False Alarm Rate**(a)** and Critical Success Index (**(b)** for four versions of CWAM using the 60 minute CIWS VIL and echo tops forecasts.*




**(a)**




**(b)**

*Figure 18. Probability of Detection vs. False Alarm Rate**(a)** and Critical Success Index (**(b)** for four versions of CWAM using the 120 minute CIWS VIL and echo tops forecasts.*

This result, if it is confirmed by additional studies, has significant consequences for real time applications of CWAM. The choice of spatial filter will determine the computational resources needed to run the CWAM. In particular, using no spatial filter on the weather inputs greatly reduces the computational load and complexity of running the CWAM to create real time WAF predictions in operational systems.

CWAM evaluation was also partitioned by region and case day, in an effort to gain insight into the set of factors that influence deviation prediction accuracy. Figure 19 shows that there are clear differences in predictive skill for weather impacts in different ARTCCs, but the reasons for these differences are not readily evident. Factors that may be related to these differences include differences in the predominant weather type in each ARTCC (e.g. severe thunderstorm cores vs. weak or moderate high topped convection), differences in the geometric relationship of the route structure and weather orientation (e.g., do routes cross or parallel major weather features), or the perceived willingness of air traffic control to accommodate deviations (possibly affecting pilot behavior).

Figure 20 shows that clear variations in performance are also evident on different case days. Three days show particularly high false alarm rates, which may be related to the predominant weather type on those days. The strong distinction between these high false alarm days and more 'typical' days suggests that weather may be classified according to CWAM performance. Weather characteristics that differentiate between typical and false alarm days could be incorporated directly into CWAM to improve its performance, or to identify days when CWAM will perform poorly.

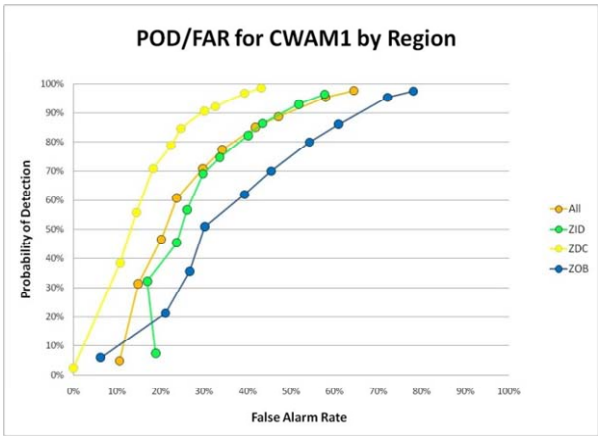



*Figure 19. Probability of Detection vs False Alarm Rate for the original CWAM partitioned by ARTCC.*

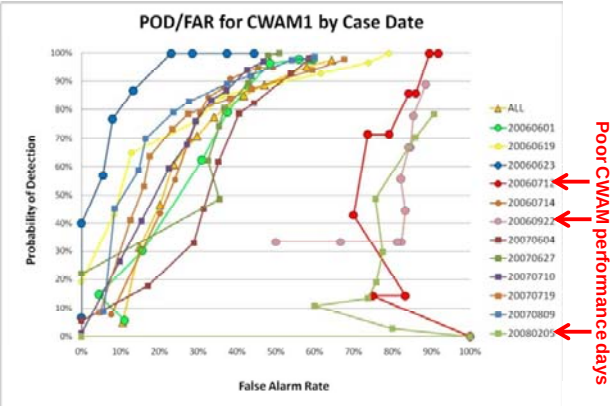



*Figure 20. Probability of Detection vs False Alarm Rate for the original CWAM partitioned by case day.*

Figure 21a illustrates a false alarm (predicted deviation, actual penetration) from a day with a high false alarm rate, September 22, 2006. The high false alarm rate in this case is associated primarily with flights below the echo tops (between 25 and 35kft). The weather demonstrates several characteristics which may influence the pilot decision making process. First, the weather (greater than level 1) has a large spatial coverage that may make it difficult for the pilot to see alternate routes. Second, the weather along the trajectory does not appear to be 'convective'. Although it is intense (> level 3), it does not appear to be cellular in nature and the echo tops are low (<35kft) compared to intensely convective weather. For these lower altitude flights, the actual echo top height and organization of the weather, not the height of the echo top relative to flight altitude, may suggest to the pilot that the weather is benign and can be penetrated.

Figure 21b illustrates a trajectory from a day with a low probability of detecting deviations, June 4, 2007. Looking at each storm cell independently, CWAM classifies this weather as having a low probability of deviation. However, pilots appear to be choosing to maneuver around these small storms. One might speculate that the appearance of several of these storms from the cockpit window suggests that the weather is very convective in nature. The current CWAM does not capture such details of storm organization. Note that the characteristics used by CWAM to estimate the severity of weather impacts – echo top and precipitation intensity – are very similar in both illustrations. What is different is the spatial organization and strength of convection of the storm – factors that are poorly captured by the spatial filters currently used in CWAM variations.

The difference in POD vs. FAR results by case day of CWAM suggests a possible correlation between the type of weather and the algorithm performance. With the expanded data set used in this study, CWAM cases can be partitioned into one of three weather types; Strong organized convection (strong cold fronts, MCCs), synoptic scale events (low pressure systems, warm fronts, weak cold fronts), and cellular convection (small scale cells, short life cycle).

The 2D deviation probability histograms from the 1km CWAM were regenerated using the entire data partitioned by weather type, and are shown in figure 22. The cellular convection consists of all valid weather encounters on June 4, 2007. The synoptic scale events are all encounters from July 12, 2006, September 22, 2006 and February 5, 2008. The eight remaining case days were typical days with strong organized convection. The results from days with strong organized convection are similar to the overall results presented in figure 14. However, very different results are observed in the cellular convection and synoptic scale event days. The region highlighted with a red box shows the most significant difference. This region covers the cases with VIL greater than level 3 and flight altitudes at or slightly above the echo tops. For synoptic scale events, the pilots are unlikely to deviate at or below the echo top heights. For the cellular convection weather type pilots are more likely to deviate even when well above the echo tops. In fact, on days with cellular convection the likelihood of pilot deviation is roughly 50% over a large segment of the 2D probability of deviation histogram, suggesting that the current CWAM predictors are a poor choice in cellular convection.

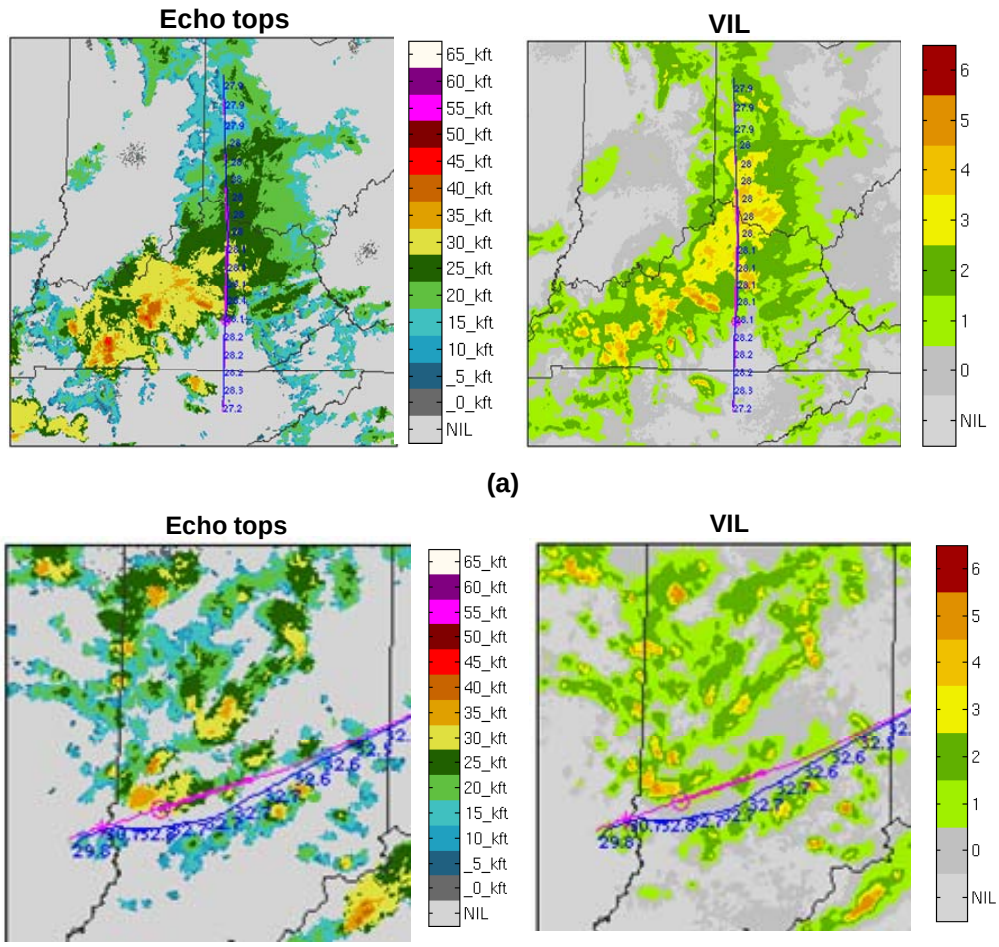



*Figure 21. Illustration of a **(a)** deviation prediction false alarm from September 22, 2006 and a **(b)** missed deviation from June 4, 2007. Magenta track is planned trajectory, blue is actual trajectory. Numbers give the flight altitude in kft., every minute.*

The comparison of CWAM accuracy, based on both true and forecast weather, also provides some insight into what may be operationally meaningful measures of uncertainty. The CWAM forecast error is a convolution of two terms: deviation prediction errors in the CWAM itself (based on deviation probabilities calculated using true weather as the CWAM inputs), and weather forecast errors. The comparison of CSI scores for CWAM based on true and forecast weather provides a basis for the assessment of the relative contributions of CWAM prediction error and weather forecast error to the total deviation prediction error that is observed in operational use. In essence, the comparison of CWAM performance based on forecast and truth may be used as a weather forecast uncertainty metric.

Figure 23 shows the deviation prediction CSI scores for true weather, one hour and two hour weather forecasts, for each of the 12 case days. Figure 23a shows 'typical' forecast behavior: deviation prediction skill decreases as the forecast time horizon increases. Figure 23b shows examples of excellent weather forecasts: deviation predictions based on one and two hour forecasts are as accurate as those based on the actual weather. It may be desirable to develop a forecast scoring model based on weather characteristics that correlate well to deviation prediction forecast performance.

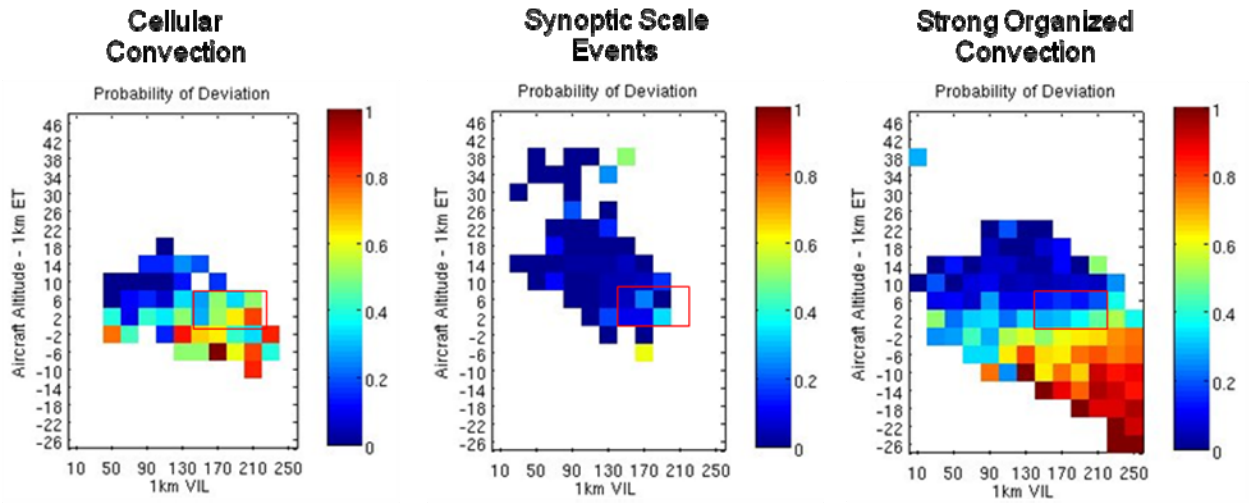
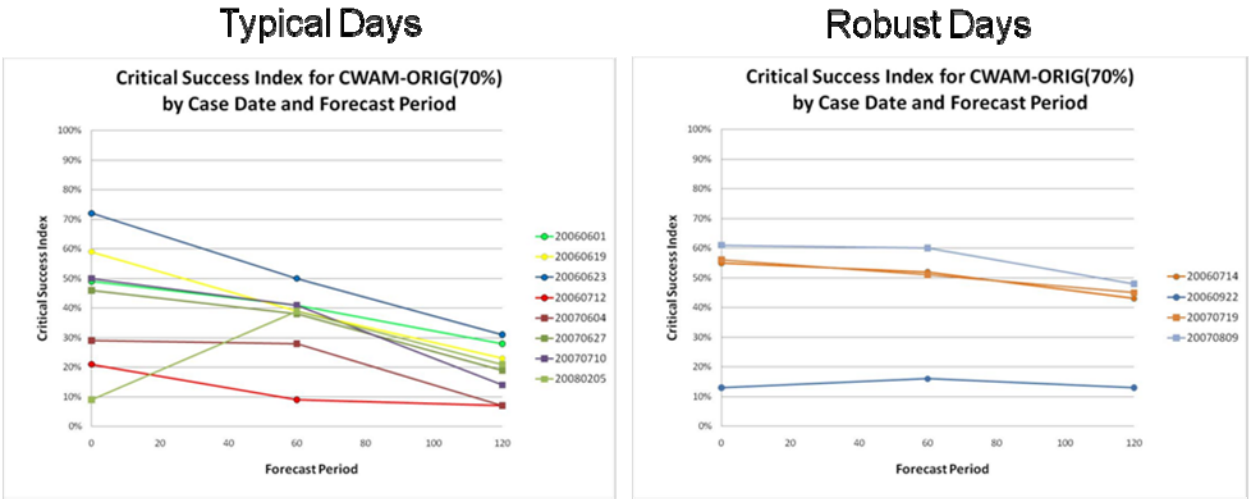



*Figure 22. 2D histograms of observed probability of deviation (percentage of flights in each histogram bin that deviated) for* **(a)** *cellular convection,* ***(b)*** *synoptic scale events, and* ***(c)*** *strong organized convection. The region of the 2D histogram with the most significant differences is highlighted in red.*




**(a)** **(b)**

*Figure 23. CSI for deviation prediction using original CWAM, as a function of forecast time for* ***(a)*** *typical forecast performance days and* ***(b)*** *robust forecast performance days.*

The ability to predict the contributions of CWAM deviation prediction errors and weather forecast errors to overall uncertainty during operations can be useful in providing decision support (figure 24). Clearly, it is important to be able to identify the type of weather in which both CWAM and the weather forecasts are likely to be correct (decision makers can follow decision support guidance with high confidence). It is also useful to know when CWAM accuracy is high and weather forecast uncertainty increases over time. In these cases, there will be a limited time horizon during which CWAM predictions can be used with high confidence. Conversely, when weather forecasts are good but CWAM predictions are poor, decision makers know that the weather forecast on its own provides valuable information, but CWAM guidance is likely to be unreliable.

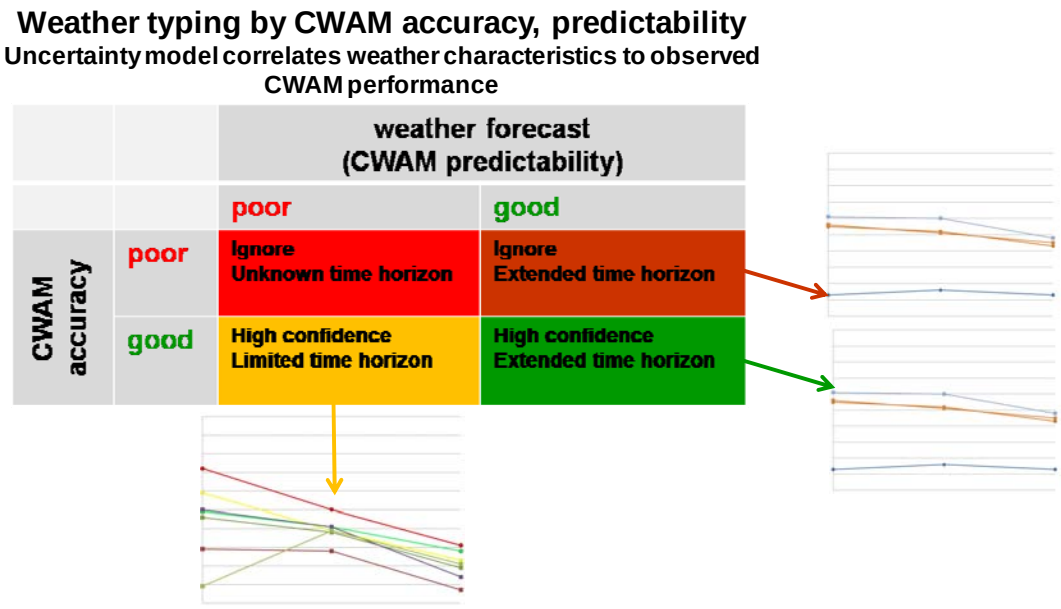



*Figure 24. Application of CWAM and weather forecast uncertainty models for decision support.*

## 5. CONCLUSIONS AND FUTURE WORK

This paper presented the results of ongoing work to improve the usability of the Convective Weather Avoidance Model (CWAM) that has been under development at Lincoln Lab under the sponsorship of NASA Ames Laboratory. CWAM is being developed to identify the correlation between pilot behavior and observable and predictable weather parameters. The CWAM deviation database of en-route weather encounters was expanded from the initial data set of 1,500 encounters to over 5,200. The automated deviation detection model used to classify flight trajectories was redesigned to improve the ability of the algorithm to automatically detect deviations that are associated with weather encounters. The redesigned model has reduced the hand editing required from over 30% to less than 10%.

The performance of four different variations of the CWAM was evaluated. The original CWAM developed in 2006 used two spatial filters to process the VIL and echo tops data into parameters found to provide the best correlation with pilot behavior. Two of the CWAM variations used different spatial filters to derive deviation predictors from the weather inputs; a third variation used unfiltered weather data. The accuracy of deviation predictions, based on the different CWAM, were compared for both true (observed) and forecast weather inputs.

Deviation prediction accuracy based on true weather, was similar for the three versions of CWAM that used filtered weather inputs, although the CWAM version based on filters that extracted the 90$^{th}$ percentile of echo top and VIL intensity over a 16 x 16 km kernel was slightly better than the other filtered versions. The CWAM based on unfiltered weather was less accurate. However, when using forecasted weather data in the CWAM, the difference in performance between all four versions was almost insignificant.

Finally, the data were partitioned by case day and weather type. Significant differences in CWAM performance by case day suggested that weather types that correlate to CWAM performance could be identified. This information could be incorporated directly into CWAM or provide information on the usefulness of the CWAM under different circumstances. Differences in CWAM skill based on forecast weather were also evident for different case days. This suggests that weather type may also be correlated to weather forecast performance, with CWAM-based deviation prediction as the forecast metric. Further research into the definition of weather types based on the predictability of pilot behavior for use in weather and weather impact uncertainty models is suggested as future work.

## 6. REFERENCES

Chan, William, Mohamad Refai, Rich DeLaura, **Validation of a Model to Predict Pilot Penetrations of Convective Weather**, 7$^{th}$ AIAA Conference on Aviation, Technology, Integration and Operations, Belfast, Ireland, 2007.

DeLaura, R. A., Evans, J. E., **An Exploratory Study of Modeling Enroute Pilot Convective Storm Flight Deviation Behavior**, 12th Conference on Aviation, Range, and Aerospace Meteorology (ARAM), Atlanta, GA, *Amer. Meteor. Soc.*,2006.
[Paper]

DeLaura, R. A., Robinson, M., Pawlak, M. L., Evans, J. E., **Modeling Convective Weather Avoidance in Enroute Airspace**, 13th Conference on Aviation, Range, and Aerospace Meteorology (ARAM), New Orleans, LA, *Amer. Meteor.Soc.*,2008.
[Paper]

Evans, J. E., Ducot, E. R., **Corridor Integrated Weather System**, MIT Lincoln Laboratory Journal, Volume 16, Number 1, 2006. [Journal Article]

Klingle-Wilson, D., Evans, J. E., **Description of the Corridor Integrated Weather System (CIWS) Weather Products**, Project Report ATC-317, MIT Lincoln Laboratory, Lexington, MA, 2005.
[Report]