Predicting Airspace Capacity Impacts Using the Consolidated Storm Prediction for Aviation

Carl Russell[*]

NASA Ames Research Center, Moffett Field, CA, 94035

Convective weather is currently the largest contributor to air traffic delays in the United States. In order to make effective traffic flow management decisions to mitigate these delays, weather forecasts must be made as early and as accurately as possible. A forecast product that could be used to mitigate convective weather impacts is the Consolidated Storm Prediction for Aviation. This product provides forecasts of cloud water content and convective top heights at 0- to 8-hour look-ahead times. The objective of this study was to examine a method of predicting the impact of convective weather on air traffic sector capacities using these forecasts. Polygons representing forecast convective weather were overlaid at multiple flight levels on a sector map to calculate the fraction of each sector covered by weather. The fractional volume coverage was used as the primary metric to determine convection's impact on sectors. Results reveal that the forecasts can be used to predict the probability and magnitude of weather impacts on sector capacity up to eight hours in advance.

I. Introduction

ONE of the primary goals of Traffic Flow Management (TFM) is to mitigate delays caused by weather. Aircraft delays in the national airspace system in 2009 totaled approximately 21.5 million minutes.[1] This translates to nearly 41 years of aircraft delay! Of those delays, 65 percent were caused by weather.[1] En route convection is the source of approximately one quarter of all weather related delays.[2] Simple arithmetic reveals that en route convection is therefore responsible for nearly 3.5 million minutes of aircraft delay per year, at a cost of millions of dollars for both airlines and passengers.[3] Early knowledge of convection location and intensity can improve TFM mitigation strategies, so accurate weather forecasts with long look-ahead times to determine TFM impacts are desirable for use in the Next Generation Air Transportation System (NextGen).[4]

Previous studies have examined ways to estimate airspace capacity impacts based on weather forecasts with look-ahead times of 0 to 2 hours;[5,6] however, little research has been done to generate capacity impacts based on forecasts with look-ahead times greater than 2 hours, even though such forecasts do exist. One newer tool that shows promise for TFM applications is the Consolidated Storm Prediction for Aviation (CoSPA). CoSPA generates grid-based predictions of cloud water content and echo top heights at look-ahead times of 0 to 8 hours.[7] This data can then be used to generate a model predicting long-range TFM impacts based on weather coverage of sectors.

This study focused on translating CoSPA weather forecasts into sector impact estimates up to 8 hours in advance using a simple model based on weather coverage. Predictions of sector coverage were compared against observed sector coverage to see how well the estimates performed as the look-ahead time was increased. The analysis first used a qualitative approach comparing the locations of predicted and observed weather impacts. Quantitative analysis then showed how well the predicted weather coverage correlated to the observed weather coverage. The quantitative analysis was split into three parts. The first examined how forecast weather coverage for a particular sector varied as a function of time. The data showed large and rapid fluctuations, which would make the forecasts difficult to use, so a time-based rolling average was used to smooth the weather coverage estimates. The second step in the quantitative analysis calculated the Pearson correlation coefficient to see how well the forecast weather coverage agreed with the observed weather coverage as the look-ahead time was increased. For the third step, a probabilistic model of weather impact was developed to show the likelihood that a sector will be affected by weather given a predicted level of impact at a particular look-ahead time.

---
[*] Research Aerospace Engineer, Systems Modeling & Optimization Branch, Mail Stop 210-15, Member AIAA.

The remainder of this paper is organized as follows. Section II describes the data used to perform the analysis. Section III describes the process used to generate sector impacts from the raw data. The results of the data analysis are presented in Section IV, followed by concluding remarks in Section V and future work in Section VI.

II.   Data Description

The CoSPA data consists of predicted precipitation water content, or vertically integrated liquid (VIL), and storm echo top altitudes. For look-ahead times of 0 to 2 hours, the forecast currently uses the Corridor Integrated Weather System (CIWS).[8] Beyond 2 hours, three techniques are used to generate forecasts out to 8 hours. The first is an extrapolation technique similar to that used to generate the CIWS forecasts. For CoSPA, three spatial scales are used for the extrapolation to capture both the rotational and translational motion of storms and their embedded cells. The second technique is based on the experimental High Resolution Rapid Refresh (HRRR) model.[9] The HRRR forecast provides numerically generated high-resolution VIL and echo top information with lead times out to 12 hours. The third technique seamlessly blends the extrapolation with the numerically predicted VIL and echo top forecasts into 0- to 8-hour forecasts with 15 min updates at 15 min time steps and 3 km spatial resolution. For a thorough description of CoSPA forecasts, see Ref. 7. Data for this study was provided by the National Center for Atmospheric Research.

CoSPA data from three consecutive bad weather days, August 19-21, 2009, was used to estimate sector impacts for this study. The evening of August 19, in particular, had numerous storms decreasing sector capacities. As shown in Fig. 1, several bands of thunderstorms were scattered across the Midwest. In addition, there were localized storms scattered throughout the Southeast. The image on the right shows the digital VIL, and the image on the left shows the echo top heights. The echo top heights indicated by the color bar in the left image are in thousands of feet. The VIL indicated by the color bar in the right image is an 8-bit digital representation of the cloud water content. This 8-bit number translates to both a VIL level and a physical measurement of water content. A digital VIL above 133, for example, corresponds to a VIL level of 3 or greater and a cloud water content of 3.47 kg/m^2. A full lookup table of digital VIL and their corresponding VIL levels is contained in Ref. 7. The area enclosed by the curved yellow rectangle shows the domain of the CoSPA forecasts. Starting in April 2010 after this study was initiated, CoSPA was improved, and its domain was expanded to include the entire continental US; however, for this study, the area was limited to that shown in Fig. 1.
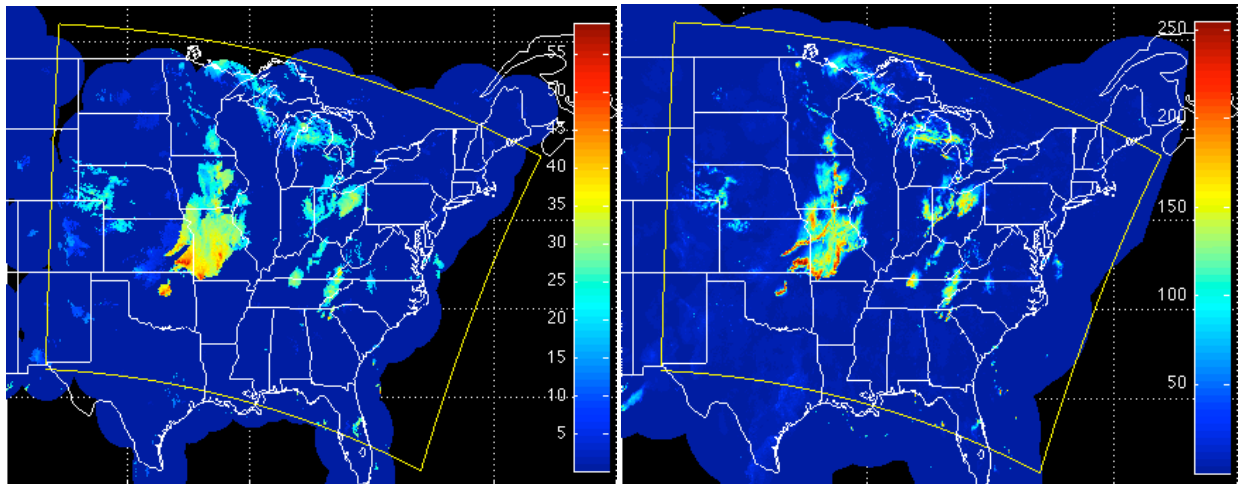
Figure 1. Observed echo tops (left) and digital VIL (right) at 00:00 EST, August 20, 2009

As expected with any weather forecast, the accuracy of the predictions generally improves as the look-ahead time decreases. This can be seen qualitatively in Fig. 2, which shows the evolution of the VIL forecast from an initial look-ahead time of 8 hours to a look-ahead time of 1 hour. The forecast look-ahead time is indicated in the top right corner of each forecast in Fig. 2. All eight forecasts represented in Fig. 2 are valid at 0:00 EST on August 20, 2009, meaning that the 8-hour look-ahead forecast (upper left) was issued at 16:00 on August 19, 2009. Similarly, the 1-hour look-ahead forecast (bottom right) was issued at 23:00.
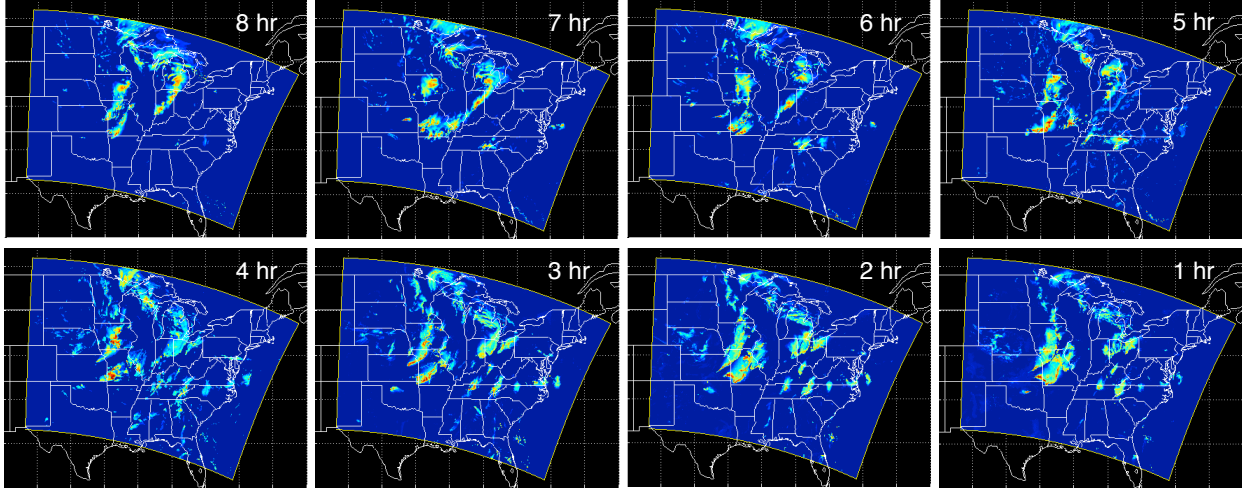
Figure 2. VIL predictions from 8 hour to 1 hour look-ahead

To get an idea of the accuracy of these forecasts, the images can be compared to the observed weather, shown in the right panel of Fig. 1. As the look-ahead time increases, the resemblance degrades, but the structure, location, and intensity of the predicted storms remains similar. This indicates that while the forecasts with longer look-ahead times may not be completely accurate, they give an idea of where storms are going to be located.

III.   Converting Weather to Sector Impact

In order to generate estimates of the impact of convective weather on capacity, the CoSPA data was used to identify regions that aircraft will presumably avoid. There is no model that definitively shows whether or not a pilot will fly through an area of convection based solely on echo top heights and VIL. Too many variables affect a pilot's decision of whether to fly through or around a weather cell, but previous studies suggest that in general, they will not fly below the echo tops in a region of VIL greater than 3.[10] This model may be viewed as a simplification of the Convective Weather Avoidance Model (CWAM), which provides probabilities of pilot deviation, given echo tops and VIL.[11] The simplified model can be easily implemented in Matlab with less computational complexity than CWAM, so it was used for the current study.

Sector impact can be measured by the three-dimensional Weather Avoidance Altitude Field coverage, or *3D WAAF*, described by Song, et al. and shown in Eq. 1.[5]

$$3DWAAF = \sum_{k=1}^{N_{FL}} w_k \times BandCoverage_k \tag{1}$$

N_{FL} is the number of flight levels within a given sector. *BandCoverage$_k$* is the fraction of the sector area covered by weather at the kth flight level, and is determined using CWAM. Each flight level is assigned a non-dimensional weight w_k based on the importance of that flight level. Results in Ref. 5 suggest that in general, uniform weighting performs as well as flight-level based weighting, so the current study used a uniform weight of $1/N_{FL}$ as well as the simplified weather coverage, S_k. This paper will refer to this implementation of *3D WAAF* as the Weather Coverage Index, or WCI. WCI is the fraction of a sector's volume affected by weather, and its formulation is given in Eq. 2.

$$WCI = \frac{1}{N_{FL}} \times \sum_{k=1}^{N_{FL}} S_k \tag{2}$$

This is a simplistic model, but as Fig. 2 shows, the specific location, size, and shape of each storm is not necessarily known at longer look-ahead times, so more sophisticated flow-based metrics that rely on accurate knowledge of storm location may not work. Additionally, Song, et al. have shown that *3D WAAF* generally correlates as well as flow-based metrics to sector capacity, so the simple WCI model should be sufficient.[5]

The conversion of the CoSPA forecasts to areas of capacity impact was performed in two steps. First, two sets of polygons were generated: one showing areas of VIL greater than level 3, and the other showing areas where the echo tops were above a given altitude. The second step in generating the avoidance polygons was calculating the overlap of the VIL and echo top contours. Figure 3 shows an example of how the polygons were generated at 30,000 ft for the observed weather shown in Fig. 1. Red areas in the leftmost frame of Fig. 3 are areas where the echo tops are above 30,000 ft, and red areas in the middle frame are located where the VIL is above level 3. The frame on the right is a combination of the two, with red indicating that both the VIL level is greater than 3 and the echo tops are above 30,000 ft. Since this study focused on en route airspace, similar polygons were computed for altitudes of 25,000 ft through 45,000 ft in increments of 1,000 ft. To predict the weather impact on an entire sector, the fraction of the sector's total area that was impacted by weather at each altitude band was combined into the 3D WCI given in Eq. 2.
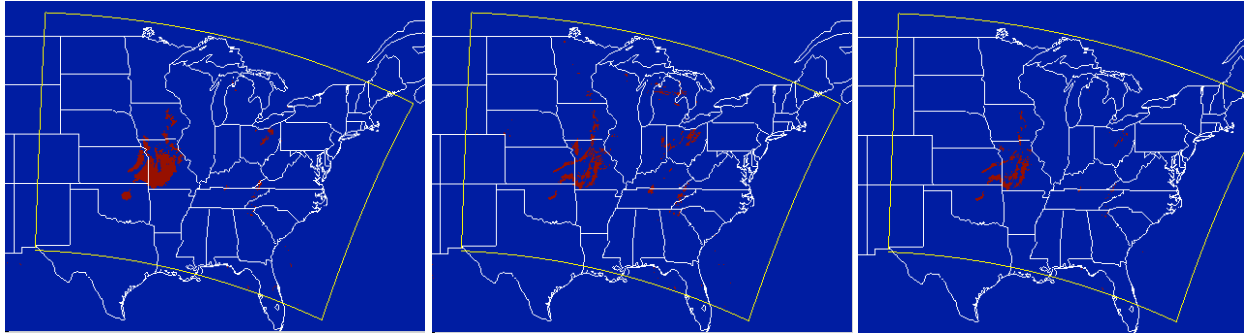
Figure 3. Echo tops > 30,000 ft (left), VIL > 3 (middle), and overlapping VIL and echo top contours (right)

The analysis included sectors from the 10 centers enclosed by the CoSPA domain. These centers are shown in Fig. 4. Some sectors were filtered out; to avoid sectors that are mostly below en route airspace, only sectors with at least five flight levels between 25,000 ft and 45,000 ft were included. Also, the Oceanic sectors in New York Center were not included, because they mostly fall outside of the CoSPA domain.
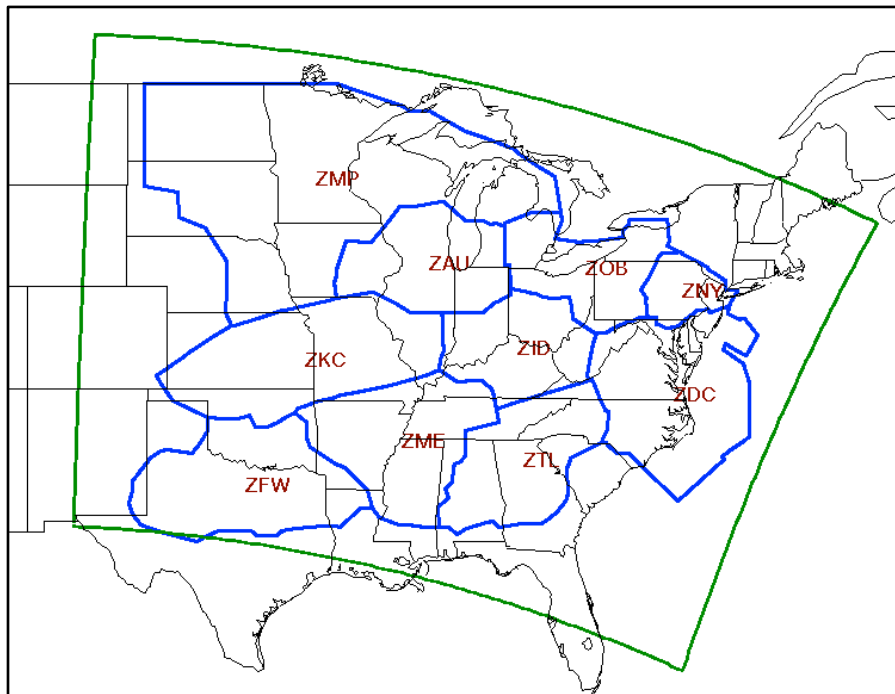

Figure 4: Centers included

To determine the usability of CoSPA-based capacity impact forecasts for air traffic flow management, the predicted WCI was compared against the observed WCI using several methods. Predictions for weather-impacted sectors were compared against observations. Plots showing the correlation between the predicted and observed WCI were then generated to quantify the accuracy of the forecasts. Finally, a probabilistic model was used to determine the likelihood of weather impact given a prediction.

IV. Results

The results section is split into three parts. The first discusses the qualitative performance of the forecasts. The second discusses the mathematical correlation between the forecast and observed weather. The probabilistic model is covered in the third portion of the results section.

A. Qualitative Results

The plots in Fig. 5 show an example of how the combined VIL and echo top contours evolve from the 8-hour to the 1-hour look-ahead time. These plots show the impacted areas at 30,000 ft. The forecast issue times are the same as those in Fig. 2; the forecasts are valid at 0:00 EST, August 20, 2009. The arrangement of plots is the same as that in Fig. 2, with the 8-hour on the top left and the 1-hour on the bottom right.
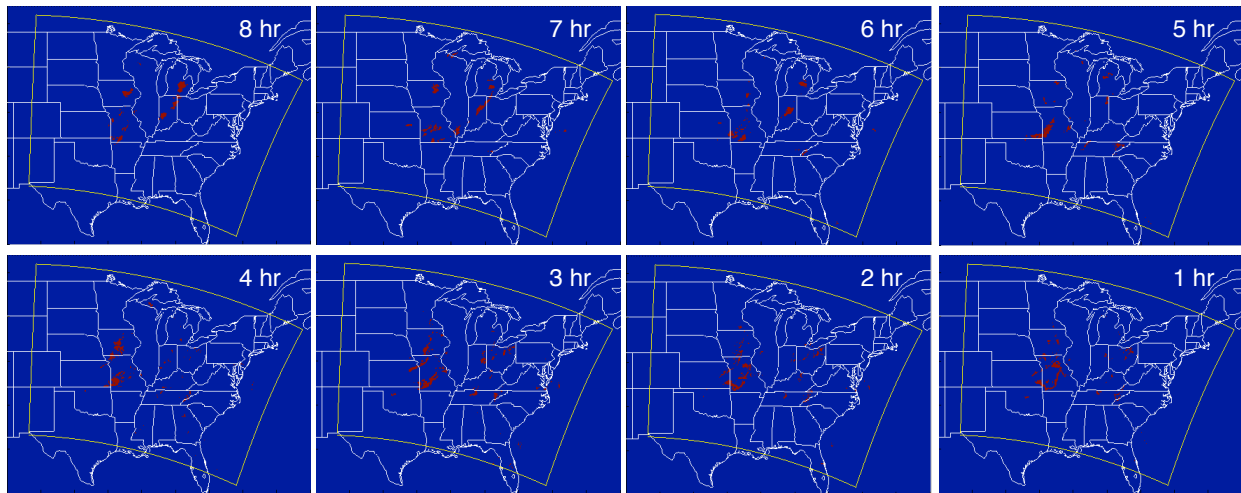

Figure 5. Combined VIL and echo tops from 8 to 1 hour look-ahead

The plots in Fig. 5 show there is variation in the forecasts as a function of look-ahead time, but the locations of the predicted storms generally stay the same. As the look-ahead time nears zero, the predicted weather impacted areas move closer to the observations, shown in the rightmost panel of Fig. 3. Based on Fig. 5, it is clear that the CoSPA forecasts do not pinpoint the location of a storm eight hours in advance. Instead, they provide an idea of where the storms are going to occur. As the look-ahead time decreases, the predicted storm location and size gets closer to the observed weather.

Figure 5 only shows the weather impacts for a single altitude, but the analysis included 21 discrete altitudes from 25,000 ft to 45,000 ft. In general, the weather impacted areas at lower altitudes are dominated by the location of VIL greater than 3, but at higher altitudes, the echo tops become the determining factor. This feature led to smaller weather-impacted areas at higher altitudes.

The second qualitative analysis compared the predicted WCI against the observed WCI. WCI can be calculated for an arbitrary simple polyhedron, but this analysis used the existing sector geometry. Figure 6 shows an example map of the predicted WCI for two look-ahead times along with the observed WCI for sectors that exist at 30,000 ft. Once again, the valid time for the forecasts is 0:00 EST, August 20, 2009. The image on the left shows the forecast WCI that was issued eight hours in advance, while the middle image shows the forecast issued just two hours in advance. The image on the right shows the WCI calculated from the observed weather. WCI has a range of 0 to 1— or 0% to 100% sector volume coverage for this study. The color scale of Fig. 6 only goes up to 0.5, because a sector coverage greater than 50% was never observed in the three days of weather data used for this analysis.
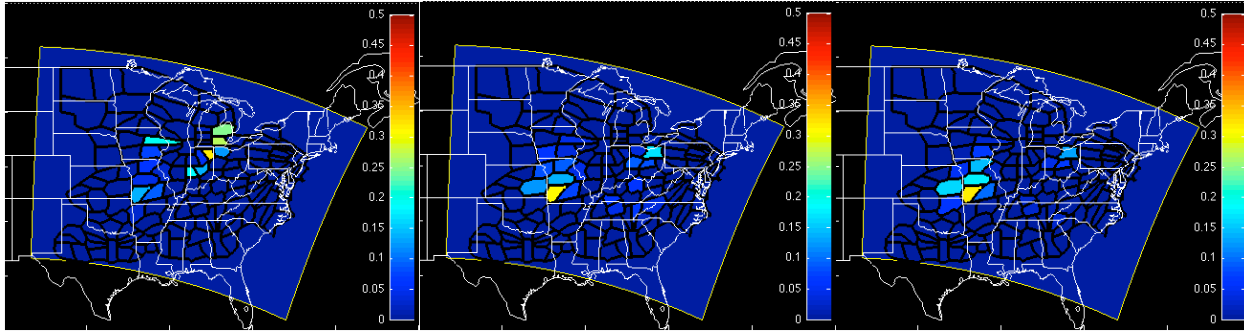
Figure 6. Predicted WCI with 8-hr (left) and 2-hr (middle) look-ahead and observed WCI (right) at 30,000 ft

Like the maps of impacted airspace shown in Fig. 5, the maps of WCI in Fig. 6 show uncertainty in the location of weather, especially as the look-ahead time is increased. The general location of weather impact is predicted up to eight hours in advance, but it does not appear possible at longer look-ahead times to confidently determine whether a given sector will be impacted. Once the look ahead time decreases to two hours, the predictions are much better. In fact, the map of WCI predicted two hours in advance and the map of observed WCI are almost identical.

The qualitative analysis suggests that the WCI forecasts are not very accurate at their longest look-ahead times, but the accuracy increases as the look-ahead time decreases. Also, the analysis suggests that although the forecasts with long look-ahead times do not pinpoint the location of storms, they make reasonable predictions of where convective weather is likely to appear. The next step in the analysis was quantifying how "reasonable" these predictions really are.

B. Correlation of Results

One of the purposes of this research is to determine how an individual sector's predicted WCI correlates to its observed WCI. One of the more heavily impacted sectors on August 20, 2009 was Memphis Center's ZME22, which extends from 23,900 ft to 34,900 ft. This sector experienced increasing weather coverage starting at midnight, Eastern Standard Time. Just before noon, the sector's WCI peaked just under 0.4, followed by a rapid clearing of weather around 13:00. A time history of the observed WCI in sector ZME22 is plotted in Fig. 7 along with the forecasts for the same time period. Solid lines represent the shorter look-ahead times (0- to 4-hr) while dashed lines show the WCI for the longer look-ahead times. Figure 7 shows large, rapid fluctuations in the predicted WCI due to variations in the forecast, which would make it difficult to use for decision making purposes.
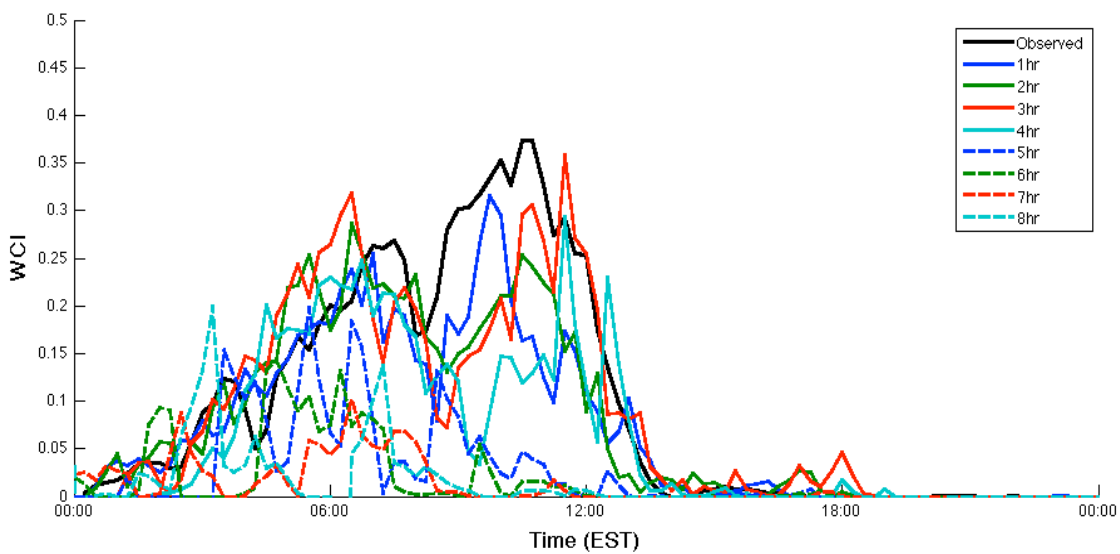
Figure 7. WCI time history for sector ZME22, August 20, 2009

American Institute of Aeronautics and Astronautics

The predicted WCI tracks the observations well for short look-ahead times, but this agreement degrades as the look-ahead time increases. Also, the large fluctuations in the predicted storm intensity do not appear in the observations. Still, even 8 hours in advance, the CoSPA forecast for this day predicts that there will be weather in the sector, though the intensity and timing of the storms is uncertain.

Smoothing the forecast data using a rolling average makes comparison with the observed weather easier. The smoothing operation uses data from 30 minutes before and 30 minutes after the issue time. This means that each forecast is based on five discreet predictions separated by fifteen-minute intervals. Since the rolling average requires 30 minutes of data after the issue time, the forecast issue is delayed by 30 minutes (for example, the 8-hr look-ahead forecast cannot be issued until 7 hours and 30 minutes before the valid time). The smoothed data is plotted in Fig. 8.
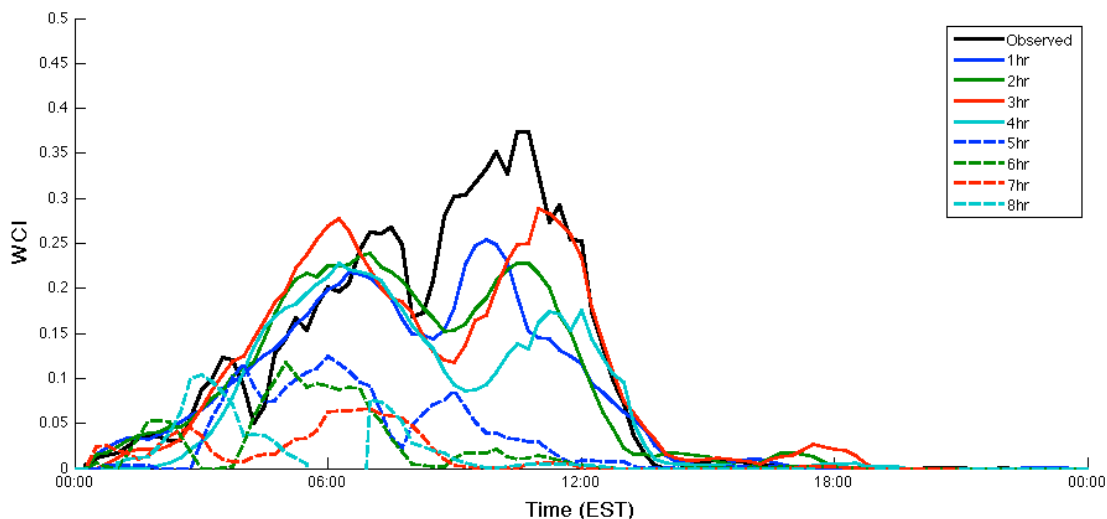
Figure 8. WCI time history for sector ZME22, August 20, 2009 with ±30 min rolling average

The smoothed data track the observed WCI more closely. Again, the performance drops off as the look-ahead time increases, especially above 4 hours. As with the raw data, while the smoothed predicted WCI data does not always accurately predict the intensity or timing of storms, it indicates that weather will be present, which is useful information to traffic flow managers.

While Figs. 7 and 8 provide a lot of information about a single sector, it is important to see how well WCI predicts sector impacts as a whole. Since this study included 269 sectors, it would have been impractical to generate similar plots for every sector. Figure 9 contains a series of scatter plots showing the correlation of the observed to the predicted WCI for multiple look-ahead times for all sectors. The top left plot is observed WCI plotted against itself, so it falls in a straight line with a slope of 1. The rest of the plots are labeled with their respective look-ahead times. The predicted WCI is plotted on the x-axis, and the observed WCI is plotted on the y-axis. Both axes range from 0 to 1. Although a WCI above 0.5 was never observed, the predictions with longer look-ahead times occasionally had WCI values as high as 0.8.

With perfectly accurate forecasts, all nine scatter plots in Fig. 9 would be identical to the plot of the observed weather (top left); ideally, the data would fall in a straight line with a slope of 1. The Pearson correlation coefficient, ρ, gives an indication of how well the CoSPA forecasts predict the observed WCI, and is included in the scatter plots of Fig. 9. ρ quantifies how linear the relationship is between predicted and observed WCI, where $\rho = 1$ indicates a perfectly linear relationship, and $\rho = 0$ indicates no linear relationship.

Figure 10 shows how the correlation varies as a function of look-ahead time. There is strong correlation at short look-ahead times, but the value of ρ decreases nearly monotonically as the look-ahead time increases. The decrease is also nearly linear from the 0-hr to the 4-hr look-ahead. After a knee at 4-hr look-ahead, the decrease is still nearly linear, but has a shallower slope. This data indicates that the forecast WCI 4 hours in advance is not much more accurate than the forecast WCI 8 hours in advance. At 8 hours in advance, ρ shows a weak correlation between the forecast and the observed WCI, and it is questionable whether the WCI predictions are still useful this far in advance. The behavior of the data shown in Fig. 10 resembles a plot of Critical Success Index (forecast skill) for CoSPA for August 2009, indicating that WCI does a good job capturing the accuracy of the CoSPA forecasts.[12]
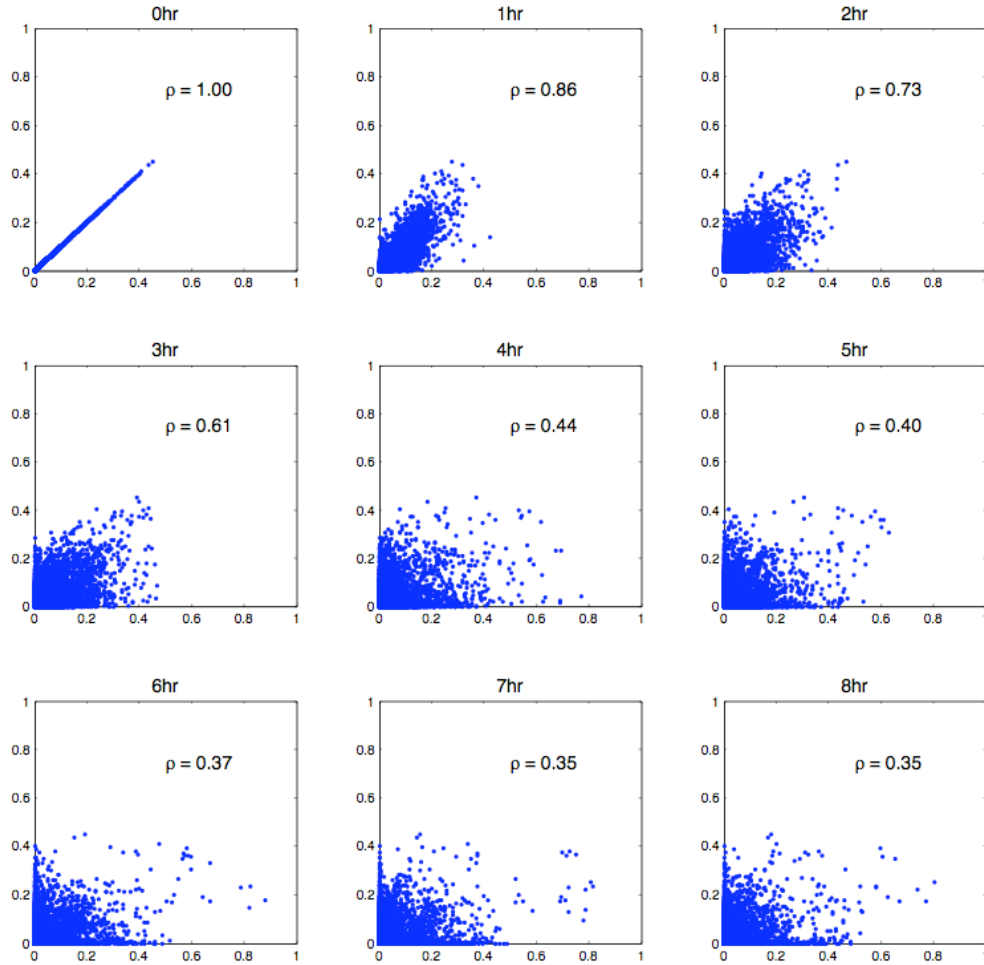
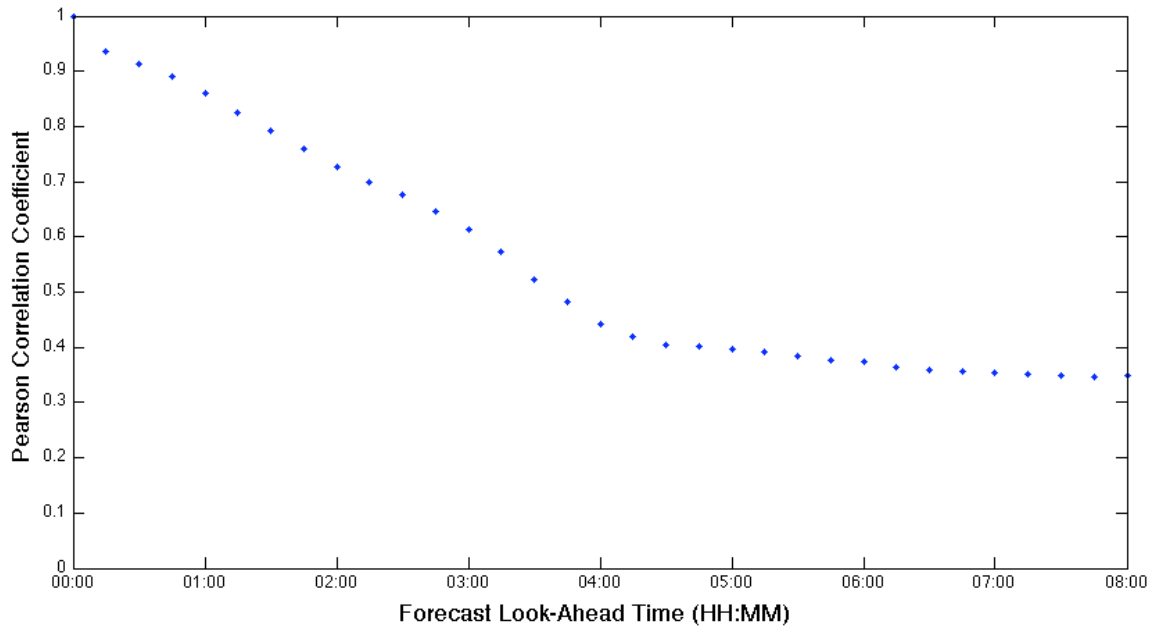**Figure 9. Scatter plots of observed WCI vs. predicted WCI**

Figure 10. Pearson correlation coefficient as a function of look-ahead time

American Institute of Aeronautics and Astronautics

C. Probabilistic Model

Although the forecasts at longer look-ahead times show only weak correlation to the observed weather, they still provide useful estimates of when and where sectors will be impacted by weather. The aim of the probabilistic model is to use CoSPA forecast data to determine the likelihood that a sector will be impacted by weather. The model is based on all three days of CoSPA data from August 19-21, 2009. Given a look-ahead time and a forecast WCI, the model shows the probability that the observed WCI will be above a specified threshold. Figure 11 shows a heat map that visualizes the probabilistic model for a look-ahead time of 2 hours.
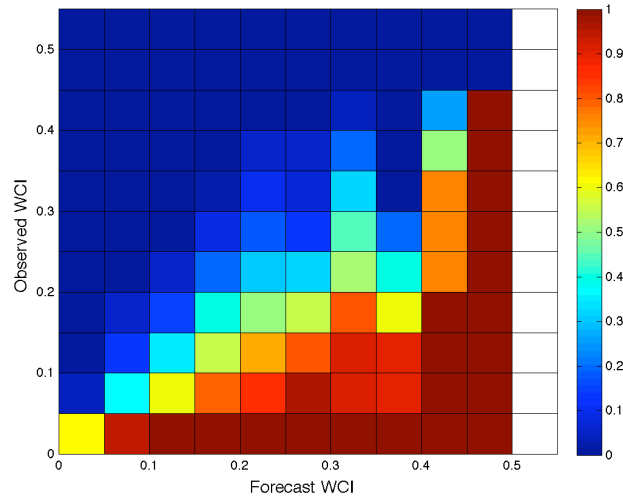

Figure 11. WCI probability with a 2-hr look-ahead time

In order to read a probability of impact from Fig. 11, find the bin containing the forecast WCI on the x-axis. To find the probability that the observed WCI will be above a certain threshold, read up along the y-axis. The color above the intersection represents the probability that the observed WCI will be greater than what is specified on the y-axis. For example, if the 2-hour look-ahead forecast predicts a WCI between 0.2 and 0.25, there is approximately a 70% chance that the observed WCI will be greater than 0.1, a 30% chance that the observed WCI will be greater than 0.2, and so on. Figure 12 is a similar heat map that shows WCI probabilities using the 8-hr look-ahead forecast.
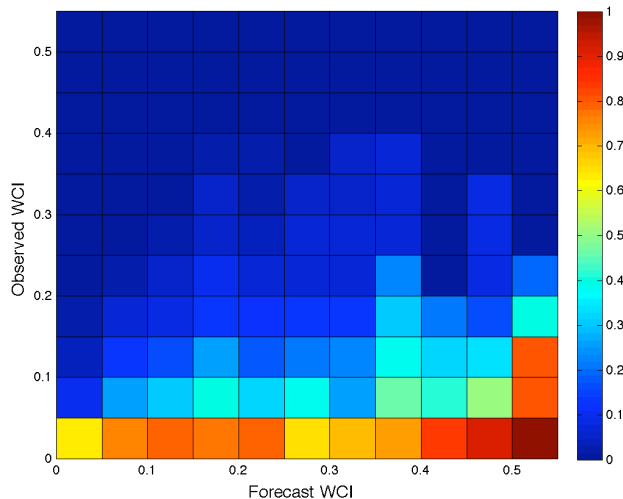

Figure 12. WCI probability with an 8-hr look-ahead time

Figures 11 and 12 show that although CoSPA may not accurately predict the magnitude of weather's impact on a sector up to 8 hours in advance, the forecasts can be used to determine the probability that the impact will be above a certain threshold.

One of the surprising features of the heat maps shown in Figs. 11 and 12 is the lack of smoothness in the color transitions, especially at higher WCI thresholds. Given the fact that over 75,000 data points are used for each look ahead time to generate the heat maps, smooth color transitions might be expected, barring a systemic problem with the data. The reason for this behavior is shown in Fig. 13, which displays the number of times each WCI was observed. Note that the color scale for this plot is logarithmic. White boxes indicate that a WCI at the given threshold was never observed. Figure 13 shows that the vast majority of the WCI values observed were between 0 and 0.05. In fact approximately 75 percent of both the predicted and observed WCI values were zero. In contrast, values of WCI above 0.4 were only observed or predicted a handful of times. The white column in Fig. 11 exists because the 2-hr forecast never produced a WCI above 0.50.
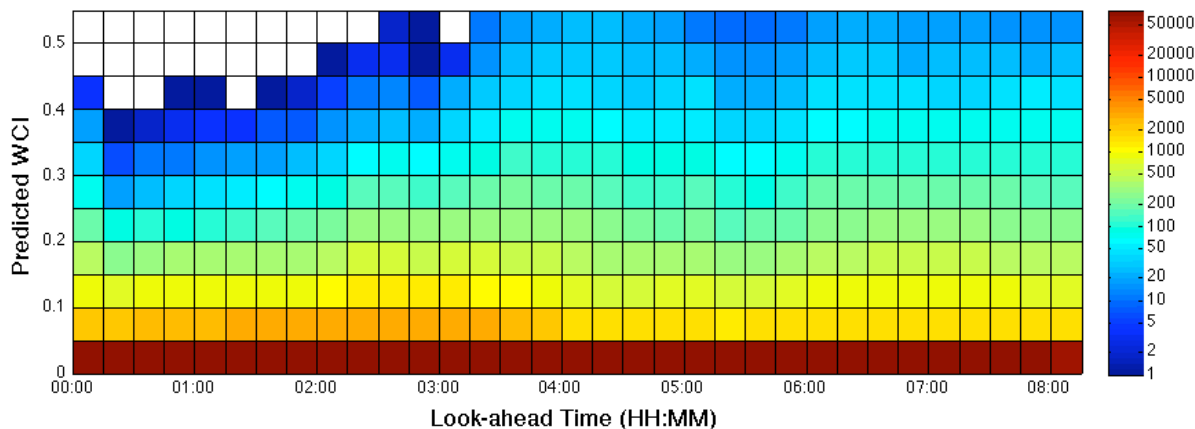
Figure 13. Number of WCI samples observed

This shows that even over the course of three bad-weather days, the number of sectors with very high WCI values is quite low. The small sample size causes uncertainty for higher WCI values leading to large jumps in the indicated probabilities. In order to smooth the probabilities shown in Figs. 11 and 12, several more days of data need to be analyzed. Still, the results in Figs. 11 and 12 show that it is possible to use CoSPA to determine the probability and magnitude of weather impact on a sector.

V. Conclusions

This study analyzed three days worth of data generated by the Consolidated Storm Prediction for Aviation to see how CoSPA can be used to predict convective weather impacts on en route airspace up to 8 hours in advance. CoSPA can accurately predict storm location and intensity at short look-ahead times, but decreases in accuracy as the time horizon increases. Rather than attempt to use the forecasts to specifically identify the location and intensity of storms, this analysis focused on sector-level impacts. The forecasts provide predictions of cloud water content, or VIL, and echo top heights, which were used to generate polygons that aircraft will likely avoid. The forecast sector impacts were modeled by a simplified weather coverage index using the weather polygons to calculate the affected volume of each sector. The results show that the observed sector impacts are strongly correlated to the predictions up to 2 hours in advance, but the correlation decreases nearly monotonically as the look-ahead time increases. Finally, a probabilistic model was developed to show how the predicted sector impacts can be used to estimate the likelihood and magnitude of convective weather impact on a sector. This study suggests that CoSPA will be a valuable tool for traffic flow managers by providing early identification of sectors that will have their capacities reduced by convective weather in the national airspace system.

VI. Future Work

This study assumed that a sector is only affected by the weather contained within its boundaries, which is not an entirely valid assumption. For example, if a major route feeding a sector is blocked by weather, but the sector itself is not blocked, it will see a large decrease in traffic. Future work will examine how blockages of major air routes can be predicted by WCI, based on CoSPA forecasts. Future studies may also use CWAM and differently weighted flight levels to calculate a more accurate WCI. Follow-on work to this study will use data from the newest version of CoSPA, which contains improvements to the forecast models as well as an enlarged domain that covers the entire continental US.

American Institute of Aeronautics and Astronautics

Acknowledgements

References

[1] Federal Aviation Administration, "The Operations Network (OPSNET)", http://aspm.faa.gov/opsnet/sys/main.asp, January 2010.

[2] Cook, L., Wood, B., Klein, A., Lee, R., and Memarzadeh, B., "Analyzing the Share of Individual Weather Factors Affecting NAS Performance Using the Weather Impacted Traffic Index", *9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*, Hilton Head, SC, Sept. 21-23, 2009.

[3] Schumer, C.E. and Maloney, C.B., "Your Flight Has Been Delayed Again," US Senate Joint Economic Committee, May 2008.

[4] Robinson, M., Moser, W., and Evans, J., "Measuring the Utilization of Available Aviation System Capacity in Convective Weather," *13th Conference on Aviation, Range, and Aerospace Meteorology*, New Orleans, LA, Jan. 2008.

[5] Song, L., Greenbaum, D., and Wanke, C., "The Impact of Severe Weather on Sector Capacity", *8th USA/Europe Air Traffic Management Research and Development Seminar*, Napa, CA, June 29-July 2, 2009.

[6] Wolfe, S. and Kulkarni, D., "Predicting Peak Sector Occupancy with Two-Hour Convective Weather Forecasts," *AIAA Guidance Navigation and Control Conference*, Toronto, Ont., Aug. 2-4, 2010.

[7] Dupree, W., et al., "The 2008 CoSPA Forecast Demonstration", *89th AMS Annual Meeting ARAM Special Symposium on Weather – Air Traffic*, Phoenix, AZ, Jan. 11-15, 2009.

[8] Evans, J.E. and Ducot, E.R., "Corridor Integrated Weather System", *Lincoln Laboratory Journal*, Vol. 16, pp. 59-80.

[9] Benjamin, S.G., et al., "The HRRR 3-km Storm-Resolving, Ratar-Initialized, Hourly Updated Forecasts for Air Traffic Management", *89th AMS Annual Meeting ARAM Special Symposium on Weather – Air Traffic*, Phoenix, AZ, Jan. 11-15, 2009.

[10] Kuhn, K. "Analysis of Thunderstorm Effects on Aggregated Aircraft Trajectories", *Journal of Aerospace Computing, Information, and Communication*, Vol. 5, April 2008, pp. 108-19.

[11] DeLaura, R., Robinson, M., Pawlak, M., Evans, J., "Modeling Convective Weather Avoidance in Enroute Airspace", *AMS 13th Conference on Aviation, Range and Aerospace Meterology*, New Orleans, LA, Jan. 2008.

[12] Pinto, J., et al., "Advances in the Consolidated Storm Prediction for Aviation," *14th Conference on Aviation, Range, and Aerospace Meteorology*, Atlanta, GA, Jan. 18-21, 2010.