THE ADVANCED STORM PREDICTION FOR AVIATION FORECAST DEMONSTRATION*

W. Dupree[1], J. Pinto[2], M. Wolfson[1], S. Benjamin[3], S. Weygandt[3], M. Steiner[2], J. K. Williams[2], D. Morse[1], X. Tao[1], D. Ahijevych[2], H. Iskenderian[1], C. Reiche[1], J. Pelagatti[1], and M. Matthews[1]

[1]MIT Lincoln Laboratory, Lexington, MA
[2]NCAR Research Applications Laboratory, Boulder, CO
[3]NOAA ESRL Global Systems Division, Boulder, CO

1. INTRODUCTION

Air traffic congestion caused by convective weather in the US is a serious national problem. Several studies have shown that there is a critical need for timely, reliable and high quality forecasts of precipitation and echo tops with forecast time horizons of up to 12 hours in order to predict airspace capacity (Robinson et al. 2008, Evans et al. 2006 and FAA REDAC Report 2007). Yet, there are currently several forecast systems available to strategic planners across the National Airspace System (NAS) that are not fully meeting Air Traffic Management (ATM) needs. Additionally, the use of many forecasting systems increases the potential for conflicting information in the planning process, which can cause situational awareness problems between operational facilities, ultimately leading to more potential delays and perhaps safety problems.

One of the goals of the Next Generation Air Transportation System (NextGen) is to eliminate these redundant and sometimes conflicting forecast systems and replace them with a common weather picture for ATM. The FAA initiated an effort to begin consolidating these systems in 2006, which led to the establishment of a collaboration between MIT Lincoln Laboratory (MIT LL), the National Center for Atmospheric Research (NCAR) Research Applications Laboratory (RAL), the NOAA Earth Systems Research Laboratory (ESRL) Global Systems Division (GSD) and NASA, called the Advanced Storm Prediction for Aviation[1] (ASPA; Dupree et al. 2009, and Wolfson et al. 2008). The on-going collaboration is structured to leverage the expertise and technologies of each laboratory to build a forecast capability that not only exceeds all current operational forecast capabilities and skill, but provides enough resolution and skill to meet the demands of the envisioned NextGen decision support technology. The Advanced Storm Prediction for Aviation system for 0-8 hour forecasts is planned for operation as part of the NextGen Initial Operational Capability (IOC) in 2013. The Advanced Storm Prediction for Aviation is funded under the FAA's Aviation Weather Research Program (AWRP) and the Reduced Weather Impact (RWI) program within the Federal Aviation Administration (FAA) Air Traffic Organization (ATO) planning division.

The Advanced Storm Prediction for Aviation prototype demonstrations were conducted starting in the summer of 2008 and have been running in real-time continuously. Technologies from the Corridor Integrated Weather System (CIWS) (Evans and Ducot 2006), National Convective Weather Forecast (NCWF) (Megenhardt et al. 2004), and NOAA's Rapid Update Cycle (RUC) (Benjamin et al. 2004) and High Resolution Rapid Refresh (HRRR) (Benjamin et al. 2009) models were consolidated into a single high-resolution forecast. A website[2] was provided to the research community and evaluation and monitoring of the system is ongoing.

*This work was sponsored by the Federal Aviation Administration under Air Force Contract No. FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.
This research is in response to requirements and funding by the Federal Aviation Administration (FAA). The views expressed are those of the authors and do not necessarily represent the official policy or position of the FAA.

[1] ASAP was formally called the Collaborative Storm Prediction for Aviation (CoSPA)
[2] ASPA Website: http://cospa.wx.ll.mit.edu

The ASPA forecast has shown tremendous promise for greatly improving strategic storm forecasts for the NAS. This paper discusses the system infrastructure, the forecast display, the forecast technology and performance of the 0-8 hr VIL and Echo Tops forecast.

2. FORECAST PRODUCTS

2.1 Overview

There are two mechanisms for displaying the ASPA forecast: a dedicated Situation Display (SD) and a password protected website. The SD is provided in the air traffic facilities and is used in real-time operations, the website is provided to any other users that need access. The dedicated SD and website have similar interfaces but the SD provides more agile manipulation and is not subject to latencies that are common on the internet. The website also has two additional modes not available on the SD, a playback data archive; and a forecast analysis tool called the ASPA Analysis Tool. A link to a forecast analysis tool is provided for comparison and analysis of various forecast products. It can be used to compare forecasts such as the ASPA Blended VIL, MIT LL Extrapolated VIL, Localized Aviation MOS Product (LAMP), and the Collaborative Convective Forecast Product (CCFP).

The ASPA situation display was added onto the CIWS situation display; this has the advantage of a high degree of code reuse, cost reduction, and benefits from years of user feedback from demonstrations and operational use in CIWS. Additionally, once the forecast is ready for prototype testing, this platform can be used for real-time application with little training since it is already a familiar interface to ATM.

2.2 Situation Display and Products

The situation display can be used to show the current weather situation and forecasts of VIL and Echo Tops via static forecast images, animation loops and contour overlays.

The SD functionality is very robust, agile panning and zooming are available via mouse control. An operator can zoom the SD to the terminal level with pixel resolution of 1 km (0.5nm) and continuously zoom out to full CONUS with 7 km resolution (4 nm) pixels.

Users also have the option of viewing the forecast in both tactical and strategic modes (Figure 1a). The tactical mode shows CIWS 0-2 hour forecasts on the full CONUS and is primarily intended for shorter term operational decisions. Here 2 hours of past weather is available in an animation loop and the forecast frame rate is adjustable from 5 min to 1 hour.

The strategic mode was designed for long-term planning as well as short-term situational awareness. For the strategic mode the display can run in an animation loop mode with 8 hours of past weather that transitions into a forecast with 8 hours of future weather. The user can also adjust the frame rate frequency continuously from 15 min, 30 min, to 60 min if desired.

For the strategic mode the available domain size depends on the forecast time horizon being displayed. Weather information is available over the CONUS from the current time to 8 hours prior. For the first 0-2 hours, the CONUS CIWS forecast is displayed; as the display transitions to the 2-8 hour forecast the domain becomes restricted to the Midwest and Eastern US (Figure 1b), which is bounded by the current computational limit of the HRRR domain. Full 2-8 hour CONUS forecasts are planned for 2010.

The growth and decay trends, “G&D Trends” product, shows storm growth and decay areas that have been detected over the past 15 to 18 minutes. Recent storm growth is shown in orange areas with a black cross hatch, and recent storm decay is shown in navy blue areas (Figure 1c). Echo Tops Tags can be enabled to show cloud top height in units of hundreds of feet. This product is useful for determining whether or not aircraft can fly over storms that may appear to be untraversable based on VIL intensity alone. The user can also display the forecast using contours on the current weather. The “Lightning” product displays recent cloud-to-

ground lightning strikes using real-time data from the National Lightning Detection Network (NLDN) (Figure 1d).

The "Storm Motion" product displays black vectors estimating storm speed and direction in knots and Storm Extrapolated Positions (SEP) are displayed in aqua-colored markings predicting where the edge of level 3+ VIL is projected to be in 20 minutes (Figure 1e). Finally previous forecasts can be evaluated by activating the "Verification" contours product in which forecasted contours of VIL level 3 or Echo Tops 30 Kft valid at the current time are overlaid on the current VIL and Echo Tops images respectively.

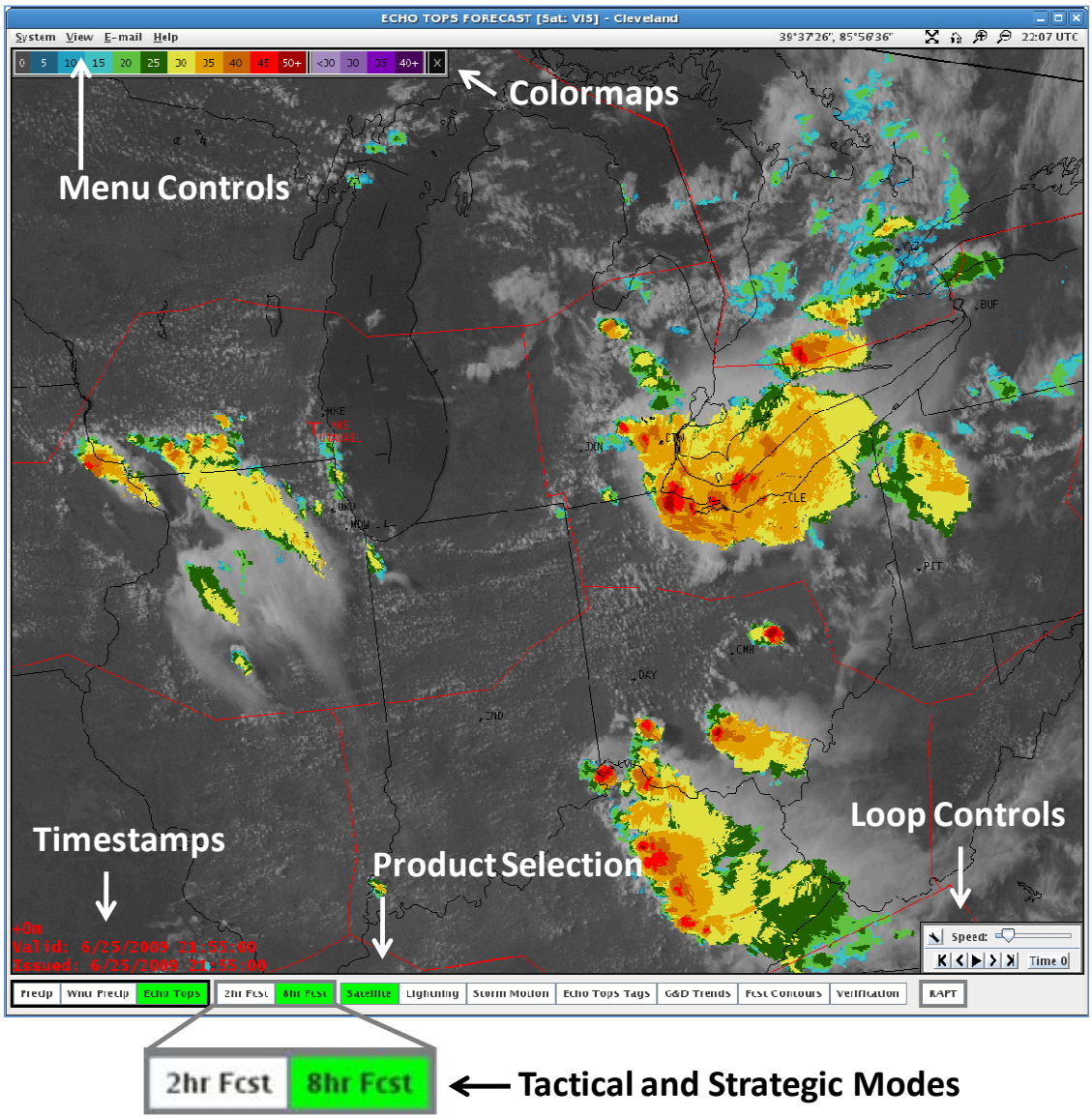

Figure 1a. The situation display for ASPA is shown for 2155 UTC on 25 June 2009 with Echo Tops showing significant convection occurring in the Midwest. The display can run as an animation loop with 8 hours of past weather that transitions into a forecast out to 8 hours. The tabs at the bottom show the various products that can be displayed. The ASPA display has been integrated with the CIWS display. The user can switch to a strategic mode that shows the 0-8 hour forecast products that have been tailored for strategic planning. The 0-2 hour portion of the 0-8 hour ASPA forecast comes directly from CIWS. These products were developed specifically for ATM use through years of user interaction.

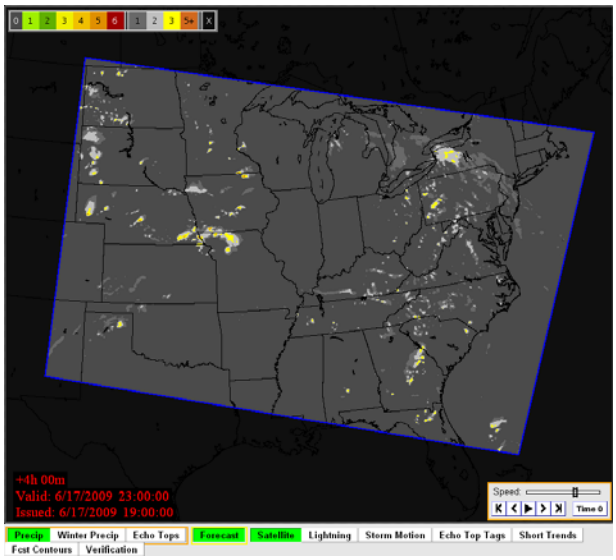

Figure 1b. The situation display for ASPA is shown with the 4 hour VIL forecast at 19 UTC on 17 June 2009. 2-8 hour forecasts are currently limited to the Central and Eastern US as depicted by the domain boundary.

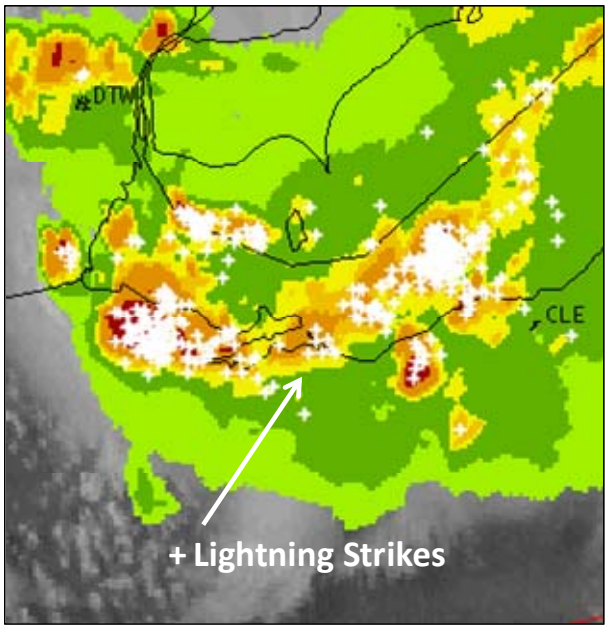

Figure 1d. This zoomed in section shows lightning strikes depicted with white crosses. A dense cluster of dense strikes clearly shows weather that is to be avoided.

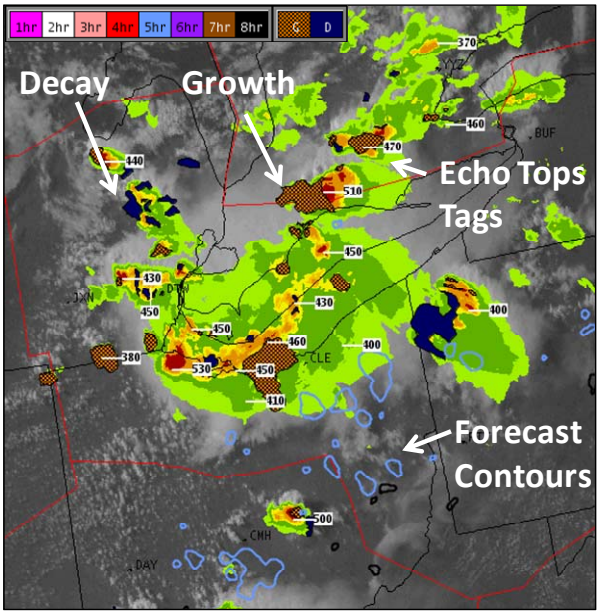

Figure 1c. This figure shows short term growth and decay trends (G&D Trends), forecast contours for the 5 hour VIL forecast and tags depicting the Echo Top with numerical values in units of hundreds of feet superimposed on current time VIL.

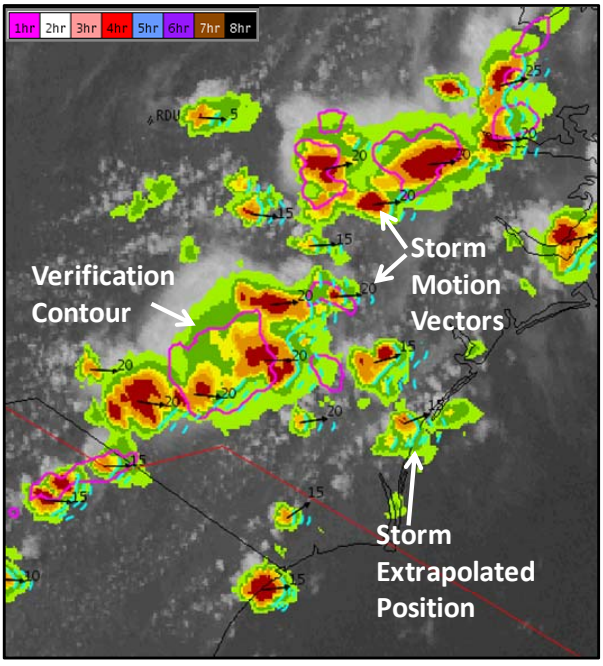

Figure 1e: This figure shows storm motion vectors that depict the direction of cell in knots and Storm Extrapolated positions (SEP). The SEPs are shown in the aqua-colored lines and indicate where the edge of level 3+ VIL is moving.

3. SYSTEM ARCHITECTURE

In line with the concepts of a virtual distributed system as envisioned by NextGen (NextGen ConOps, 2007), the ASPA system was designed as a distributed set of processing nodes that are linked together by a network. In order to seamlessly exchange data between contributing organizations, the NextGen Network Enabled Weather (NNEW) working group has been exploring a number of data formats and web services that can be used to exchange data across this system. We are using the common gridded data format NetCDF4 for ASPA as this format will likely be adopted by FAA systems. The data flow diagram is provided in Figure 2. Many sources of sensor and meteorological data are ingested by MIT LL, NCAR and NOAA/GSD for the heuristic and NWP models. For the 0-2 hr forecasts we are using the CIWS forecasts (Evans and Ducot, 2006). Primitives from the CIWS system in addition to new large scale tracking and advection modules are used to produce a 0-6 hour heuristic extrapolation forecast at MIT LL.

Once the MIT LL extrapolation and HRRR model data are available, they are ingested into the NCAR blending algorithm described in Section 4.3, which combines the heuristic extrapolation forecasts with the numerical model forecasts. Upon completion, the blended forecast data are post-processed for display on the ASPA situational display (SD) and website.

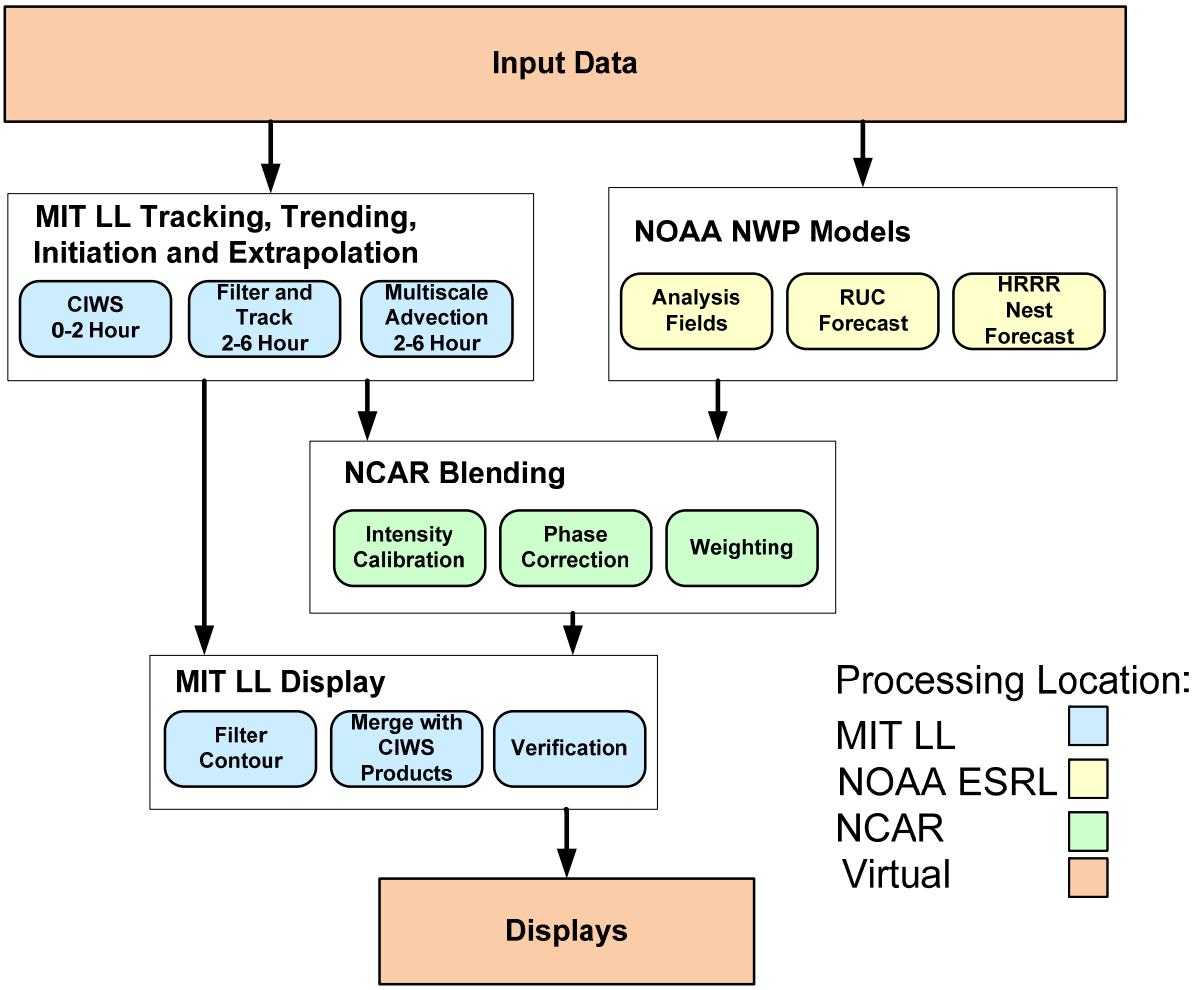

Figure 2: Shown are the combined functions and dataflow for the ASPA system. The arrows represent dataflow between processes that are either local or remote.

4. FORECAST TECHNOLOGY

This section discusses the three main components of the ASPA forecast: the heuristic long range extrapolation forecast, the HRRR numerical model and the blending algorithm. For the 0-2 hour forecast, CIWS technology is used; for a review of CIWS see Wolfson et al. 2004, Dupree et al. (2005, 2006) and Wolfson and Clark (2006).

4.1 Extrapolation Forecast

Storm forecasting on multiple scales remains an area of active research (Bellon and Zawadzki, 1994, Wolfson et al., 1999, Seed and Keenan, 2001, Lakshmanan et al., 2003 and Dupree et al., 2002, 2005). These studies show large-scale features are more predictable than small-scale features, and large-scale features can be extrapolated to longer time horizons with greater accuracy than small-scale features. Additionally, the longer the forecast time horizon, the larger the minimum scale at which meaningful motion data may be extracted.

The motion prediction consists of three fundamental steps: 1) filtering and tracking, 2) interpolation of motion fields, and 3) advection of the weather. First, for the filter and track step, the motion of a storm system must be determined and distilled into motion vector fields at several scales. To create the raw motion vectors from the observed data, the input precipitation (VIL) images are filtered with a set of mean filters followed by cross correlation on a time series of the images. Three scales are used for the extrapolation; these are the cell, envelope and synoptic scales shown in Figure 3a. Two of the three motion scales are created in the CIWS system: the cell scale, a 13 km diameter circular mean filter with a 6 minute correlation time, and the envelope scale, a 13x69 km rotated elliptical filter with an 18 minute correlation time. A new scale, created for ASPA, applies for longer time horizons: the synoptic scale, a 101x201 km filter with a 45 minute correlation time. For the interpolation step, each set of raw motion vectors is interpolated to create a smooth vector map for each scale.

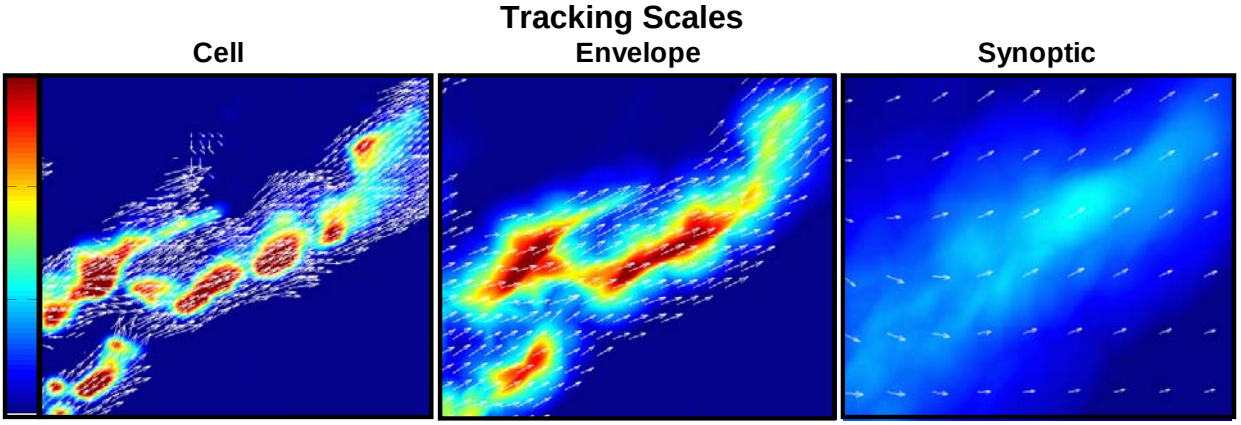

Figure 3a: Three scales used to create raw motion vectors from smallest to largest: cell, envelope and synoptic. Motion vectors are in white, background is spatially filtered interest image. VIL is passed through the filters, then correlated with previous images to calculate the motion vector.

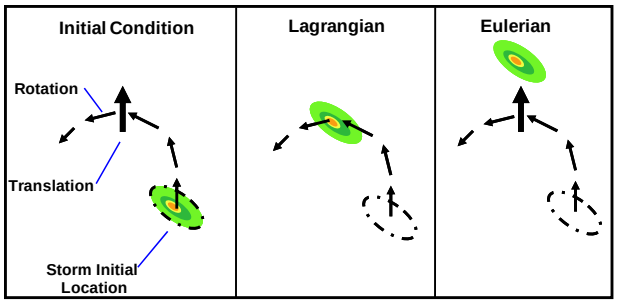

Figure 3b: Schematic depicting multiscale advection technique. Initial image is advected in many short steps using successive small scale cell and envelope vectors, then in one large step using large scale (synoptic scale) vectors. Small arrows represent small scale vectors used in the small scale step, the large arrow represents vectors used in large scale step.

The advection process uses two steps to move the separate scales (Figure 3b). First, pseudo-Lagrangian[3] advection is applied to the cell and envelope motions, and second, an Eulerian advection step or translation is applied to the synoptic scale. For the first step, the synoptic motion is subtracted from the cell and envelope scales, and the resulting field is applied in a pseudo-Lagrangian sense to the forecast image. The method works as follows: a pixel is advected with a small time step, and then placed at a new location. The pixel is then advected again for the next time step with the motion field in the area of its new location. The pixel therefore should approximately follow the streamline of the small scale (rotational) motion field. The cell vectors are used out to the 10-minute time horizon, this transitions to the envelope vectors which are used out to a 90-minute time horizon at which point their influence is progressively diminished. After the rotation step is complete an Eulerian step is applied using the synoptic-scale motion vectors to accomplish the final translation step.

[3]Here we refer to the method as pseudo-Lagrangian because no mathematical Lagrangian operators are utilized in the algorithm.

4.2 *High Resolution Rapid Refresh (HRRR)*

An experimental version of the Weather Research and Forecasting (WRF) model called the High Resolution Rapid Refresh (HRRR) model (Benjamin et al. 2009a and 2009b, and Weygandt et al. 2009) is being run at NOAA's ESRL GSD laboratory. The HRRR model is a 3-km resolution model that is nested inside an experimental version of the 13-km Rapid Refresh (RR) model. Running in parallel is a version of the 13-km Rapid Update Cycle (RUC) that assimilates the three-dimensional radar reflectivity data using a diabatic Digital Filter Initialization (DFI) technique. The HRRR model benefits from the RUC radar data assimilation through the lateral boundaries throughout the forecast as well as in improved initial conditions. In addition, the high resolution of the HRRR explicitly resolves convection, allowing the model to produce realistic convective structures vital for improved forecast fidelity without the need for sub-grid parameterizations.

The HRRR model is restarted once an hour and generates forecasts out to 12 hours. VIL and Echo Tops forecasts are made available at a special 15 minute frequency for the ASPA forecast system so that multiple valid times can be used in the blending technology.

For the 2008 summer demonstration, HRRR was run in the Northeast corridor domain; in 2009 the domain was expanded to cover a large portion of the Midwestern and Eastern US, shown in Figure 4. The HRRR has shown remarkable skill at depicting storm organization and evolution. In particular, the HRRR typically provides clear guidance in distinguishing between scattered and organized convection, which is critical information for aviation planning.

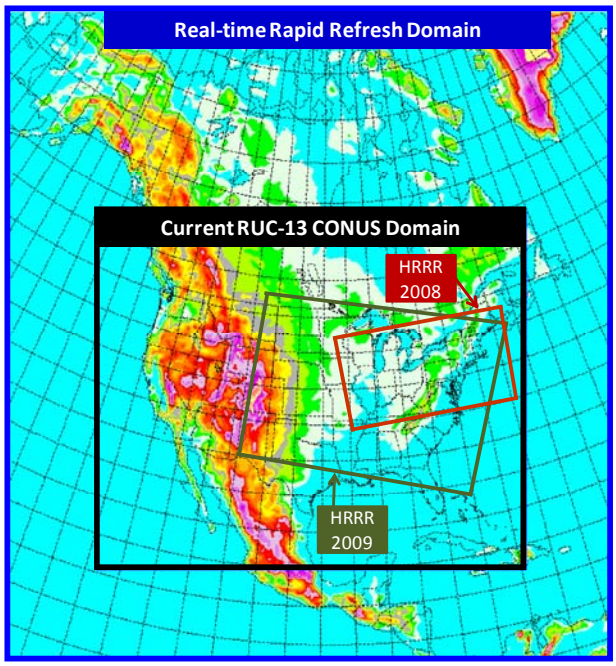

Figure 4: HRRR model nested in the WRF-Rapid Refresh and Rapid Update Cycle (RUC) models. Depicted are the experimental Northeast domain over which the HRRR was run during 2008 and the expanded Midwestern and Eastern domain model for 2009. Beginning in 2010 we plan to expand the HRRR to cover the CONUS.

4.3 *Blending*

A blending algorithm has been developed to combine extrapolation-based and NWP forecasts of VIL and Echo Tops to produce a seamless, rapidly-updating 0-8 hour forecast of weather intensity and storm top heights. This is done through a calibration of model data, a phase correction to remove location errors in the modeled precipitation field and a statistically-based weighted average. In ASPA, heuristic extrapolation forecasts of VIL and echo tops from MIT LL are blended with VIL and echo tops forecasts from the HRRR model.

The storm intensity and echo tops are "retrieved" from the modeled profiles of snow, rain and graupel water mixing ratio (note that the model currently does not predict hail). The radar reflectivity is calculated at each model level using the equations given by Thompson et al. (2004). This quantity is then converted to liquid water and vertically-integrated to give a total VIL for each atmospheric column. It is found that using the radar equations results in more realistic looking storm structures. Echo tops are determined by finding the highest vertical level at which the reflectivity exceeds 18 dBZ.

This radar-retrieved model VIL field is then converted to digital VIL (0-254 using an 8 bit integer scale) and calibrated by performing a frequency matching procedure that reduces intensity biases in the modeled VIL values. This technique is described in more detail by Pinto et al. (2009).

Spatial offsets between modeled digital VIL (digVILm) and the observed digital VIL (digVILo) are then reduced using a phase correction technique based on the minimization of the squared errors following Brewster (2004). The phase error vectors are then applied to both the modeled digVILm and modeled echo tops. An example of the corrections impact on a 3 hour forecast is shown in Figure 5a.

Timing of the real-time data feeds is a critical aspect of the system. A new forecast is generated every 15 min with forecast output frequency of 15 min out to 8 hours. The forecast blending hinges on the latency of the model forecast, which is typically 2-3 hours old by the time it is available. The phase correction procedure then compares the current radar mosaic data with the closest valid forecast lead time and determines the amount and direction of shift at each grid point. These "shifts" are then applied to all the forecast lead times as a constant offset.

Results using this technique are similar to those obtained using other image morphing techniques (Figure 5b). Details on how the Brewster technique has been implemented are described in Pinto et al. (2009).

Time-varying weights are then used to blend the phase-shifted model with the heuristic extrapolation forecasts. The weights are determined by the relative performance of the two forecasts. The performance is determined by looking at a combination of Bias and CSI scores. Generally, the model is given more weight at the longer lead times, with equal weighting at about 4 hours. The

weights are allowed to vary as a function of the time of day, with the model receiving more weight during the period of most rapid storm initiation and growth over the CONUS (i.e., 10-15 UTC) as this period of rapid change is difficult to handle through observation-based approaches.

The final resulting forecast aims to optimally combine extrapolation and heuristics with high resolution NWP output. Forecasts that accurately depict storm evolution and morphology are critical for making well-informed decisions related to routing air traffic across the NAS.

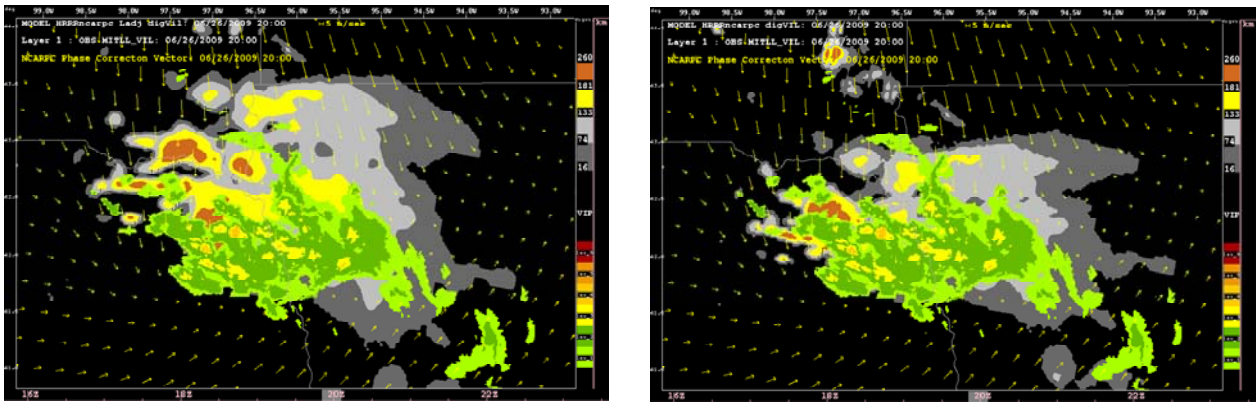

Figure 5a: The model-predicted (left) and phase-corrected (right) forecast VIL fields shown in the gray-yellow shading while the observed VIL is given in the green-yellow-red scale. The arrows indicate the direction and magnitude of the correction.

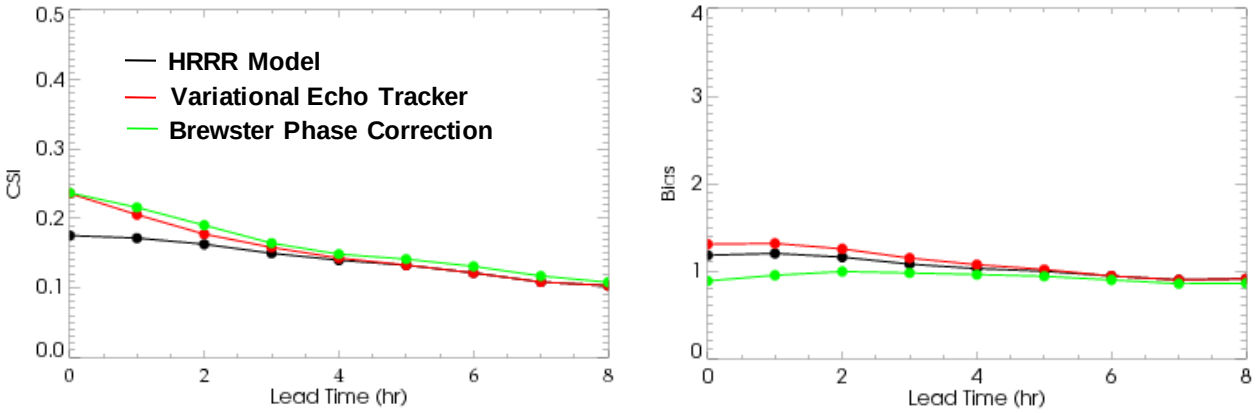

Figure 5b: Forecast skill scores for 6 June 2009 for VIL thresholded at level 2 (left - CSI and right – Bias) and plotted as a function of lead time for HRRR model (black), Phase corrected model using a proprietary variational echo tracker technique (red) and using the phase correction technique after Brewster, 2003 (green), described above. The phase shifts were calculated assuming a model latency of 2 hours.

5. VERIFICATION AND PERFORMANCE

The performance of the ASPA forecast was monitored throughout summers of 2008 and 2009 at both MIT LL and NCAR using a variety of methods. The methods include: 1) Real-time verification code that produces forecast statistics and verification contours, 2) onsite assessment by meteorologists, and 3) post performance analysis using playback and analysis tools.

It has become clear that ASPA has forecast skill and performs well in most situations. Examples of routine performance are shown in Figure 6. Four and 6 hour forecasts issued at 7 UTC on 20 June 2009 of a large-scale line are shown in Figure 6a. Note that the line begins to decay at about 11 UTC and is fully collapsed by 13 UTC. ASPA captured this large scale decay event; albeit late by about an hour, this forecast is still useful to aviation planning. Figure 6b shows an example from 23 June 2009 in which convection intensified during an 8 hour period. The 8-hour forecast issued at 15 UTC shows an example of large-scale, widespread initiation. The location of the large-scale instability region is captured by the model, as is the scattered airmass mode of the convection. Even though the precise locations of the individual cells are not forecast correctly, the forecast clearly shows that air traffic could move through this weather, since there are many open flow paths.

Figure 6c shows an ASPA VIL forecast issued at 15:30 UTC on 26 June 2009. Note at the issue time there is very little preexisting precipitation in the domain. A line begins to develop in New York and Pennsylvania at 16:30 UTC and becomes pronounced at 18:30 UTC. ASPA begins to resolve this evolution but does not immediately catch the linear organization until around 20:30 UTC, approximately 1-2 hours late. However the 8 hour forecast predicts a strong line impacting New York airspace at about 23:30 UTC. The prediction was late with a small time lag of approximately 30 minutes. The NY airspace did experience a high impact weather event at 23:00 UTC which resulted in significant delays. Also note that scattered airmass storms are well predicted; extending from Washington DC in a Southeastern direction out to Kentucky.

A route blockage algorithm has been developed for scoring in ASPA (Matthews et al., 2009). This method, known as the Blockage algorithm, is a directional approach to spatial filtering of the weather. A blockage image or map is created by finding a preferred path for route segments given the precipitation field; a distance weighted mean of the weather within the segment centered on this path is calculated. This is done for multiple orientations about a pixel and the average value is saved. The result is a filtered image that isolates precipitation that typically blocks air traffic routes. When comparing the blockage image to actual air traffic flight paths, one observes that values approximately >= 110 (on a 0-255 scale) blockage units serves as a good indicator of pilot deviations. Forecast verification is run in real time and scientific assessment and performance is monitored daily. Figure 6d shows scoring statistics for 13-15 June 2009 of blockage thresholded at 110. One can see that ASPA shows skill through the trends in the number of pixels between truth and forecasts for 2-8 hour time horizons. Statistical scores show good performance when approaching about a 20% CSI.

Storm location, structure, and scale have emerged as having an important role in ASPA performance. The forecast is more robust for larger scale convection, such as a line or bow echo, but the growth mode of smaller scale weather can also be well-forecasted out to 8 hours. The ASPA forecasts continue to be monitored, analyzed, and enhanced to improve overall performance.

20 June 2009

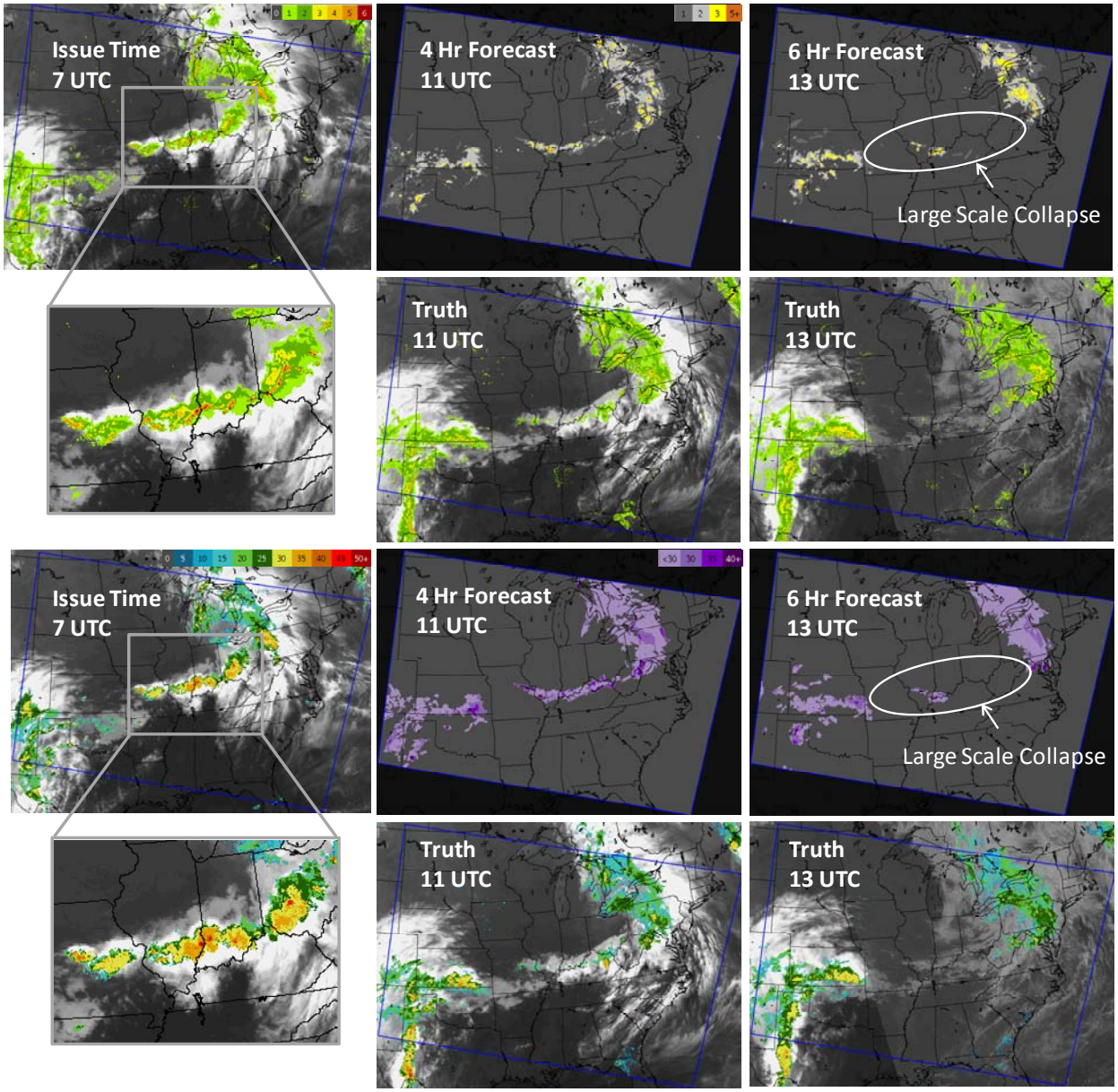

Figure 6a: This figure shows 4 and 6 hour forecasts and corresponding truth for both VIL (upper panels) and Echo Tops (lower panels) issued at 7 UTC on 20 June 2009. Initial conditions at 7 UTC show a large line that is mostly impenetrable. Note how ASPA captures the large scale decay of the line although perhaps late by about an hour. However this forecast clearly indicates that routes will be available later in the day during the heaviest air traffic.

23 June 2009

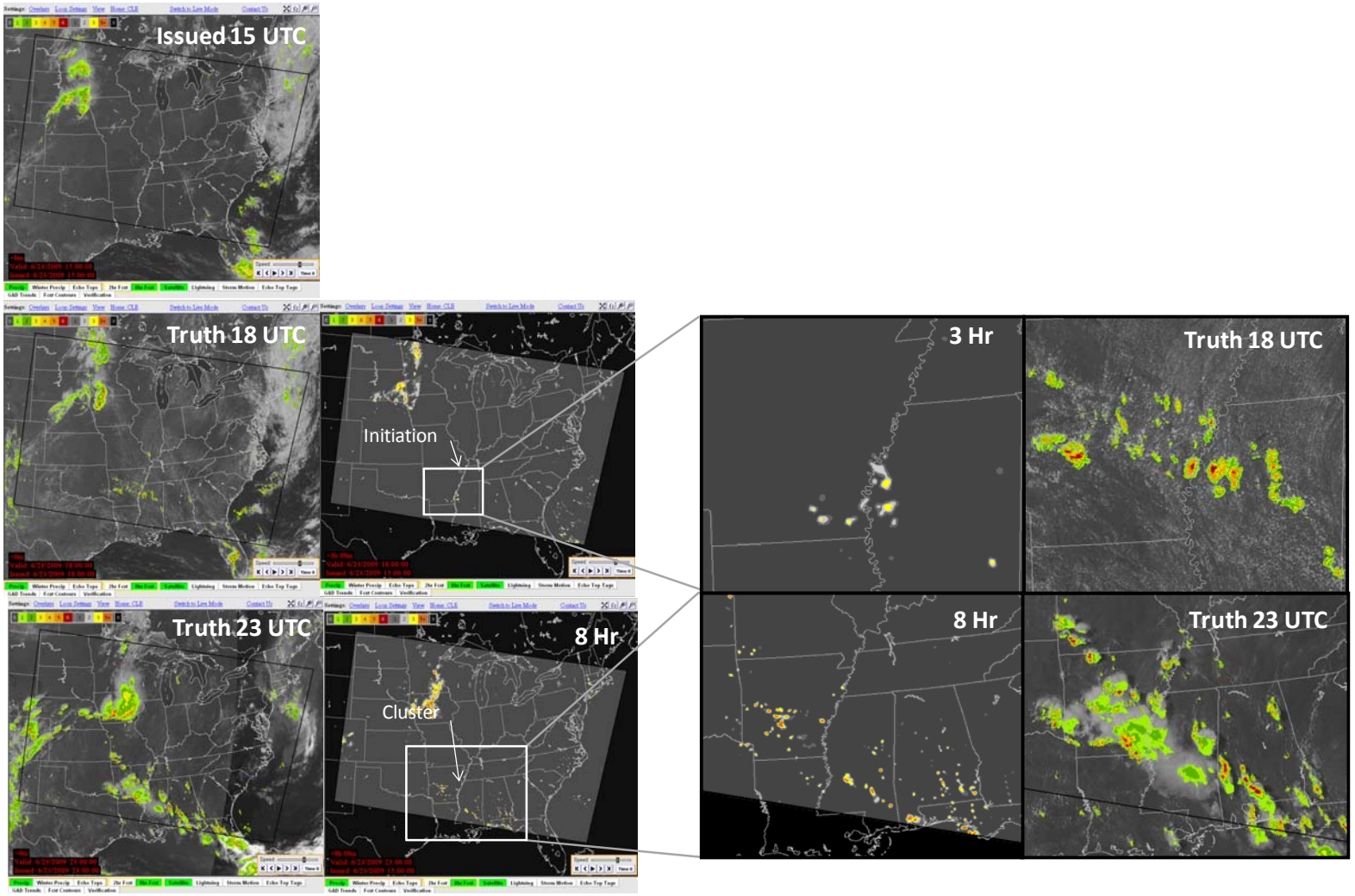

Figure 6b: Comparison of initial and observed VIL with the 3 and 8 hour blended forecast issued at 15 UTC on 23 June 2009. Note that the forecast predicted the initiation, growth, and aerial extent of the scattered airmass storms in this case. The zoomed in panels on the right hand side show the forecast and truth and show how the ASPA forecast resolves storm structure giving a depiction of traversable weather.

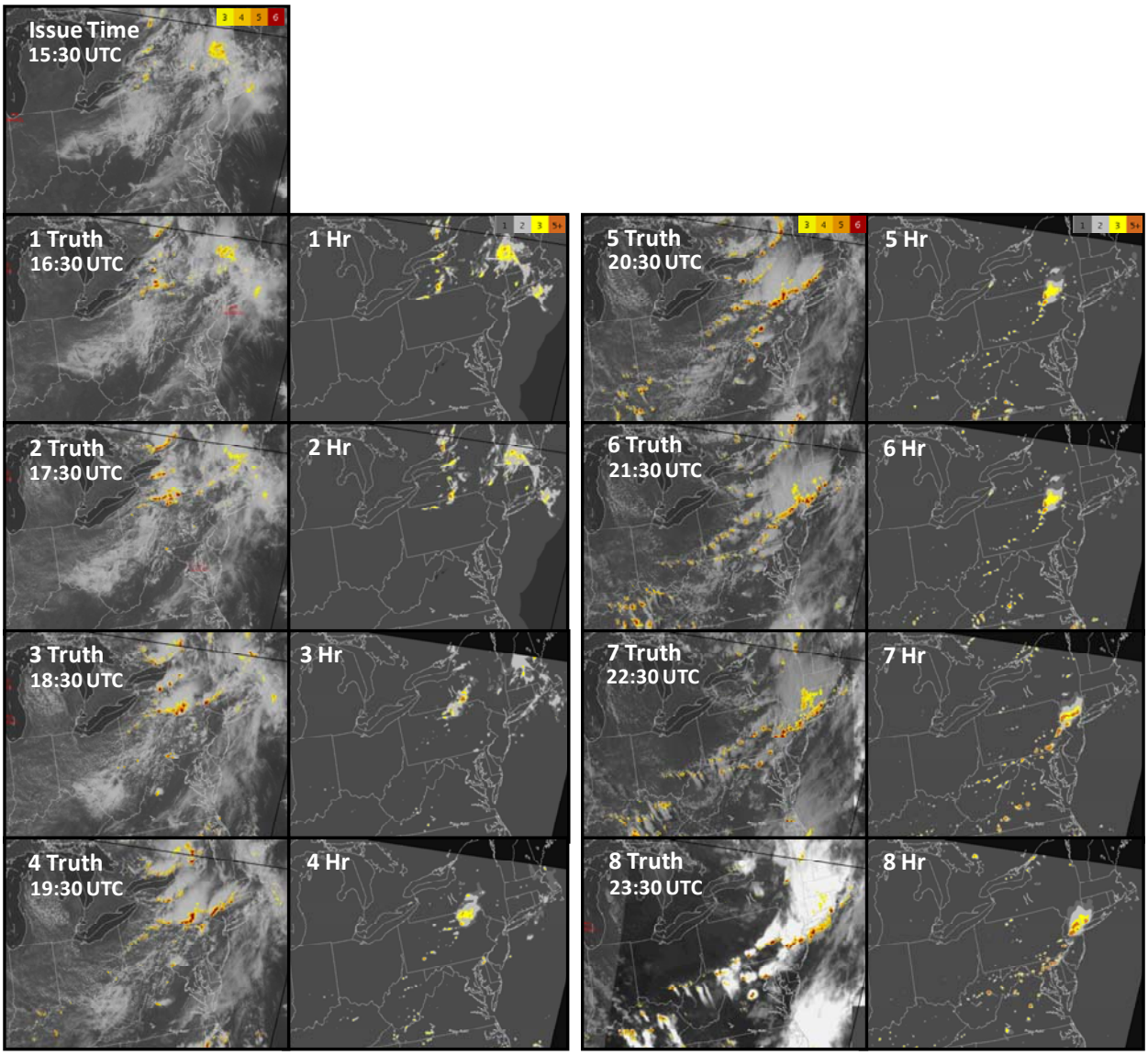

Figure 6c: This figure shows an ASPA VIL forecast issued at 1530 UTC on 26 June 2009. Note at the issue time there is very little preexisting precipitation in the domain. A line begins to develop in New York and Pennsylvania at 1630 UTC and becomes pronounced at 1830 UTC, ASPA resolves this evolution but does not produce the linear organization until around 2030 UTC, approximately 1-2 hour late. However the 8 hour forecast predicts a strong line impacting New York airspace at about 2330 UTC. NY did experience a high impact weather event at 2300 UTC which resulted in significant delays. Also note that scattered airmass storms are predicted extending from Washington DC in a Southeastern direction out to Kentucky; although the precise locations are not perfectly matched, this forecast correctly shows that the airspace is flyable.

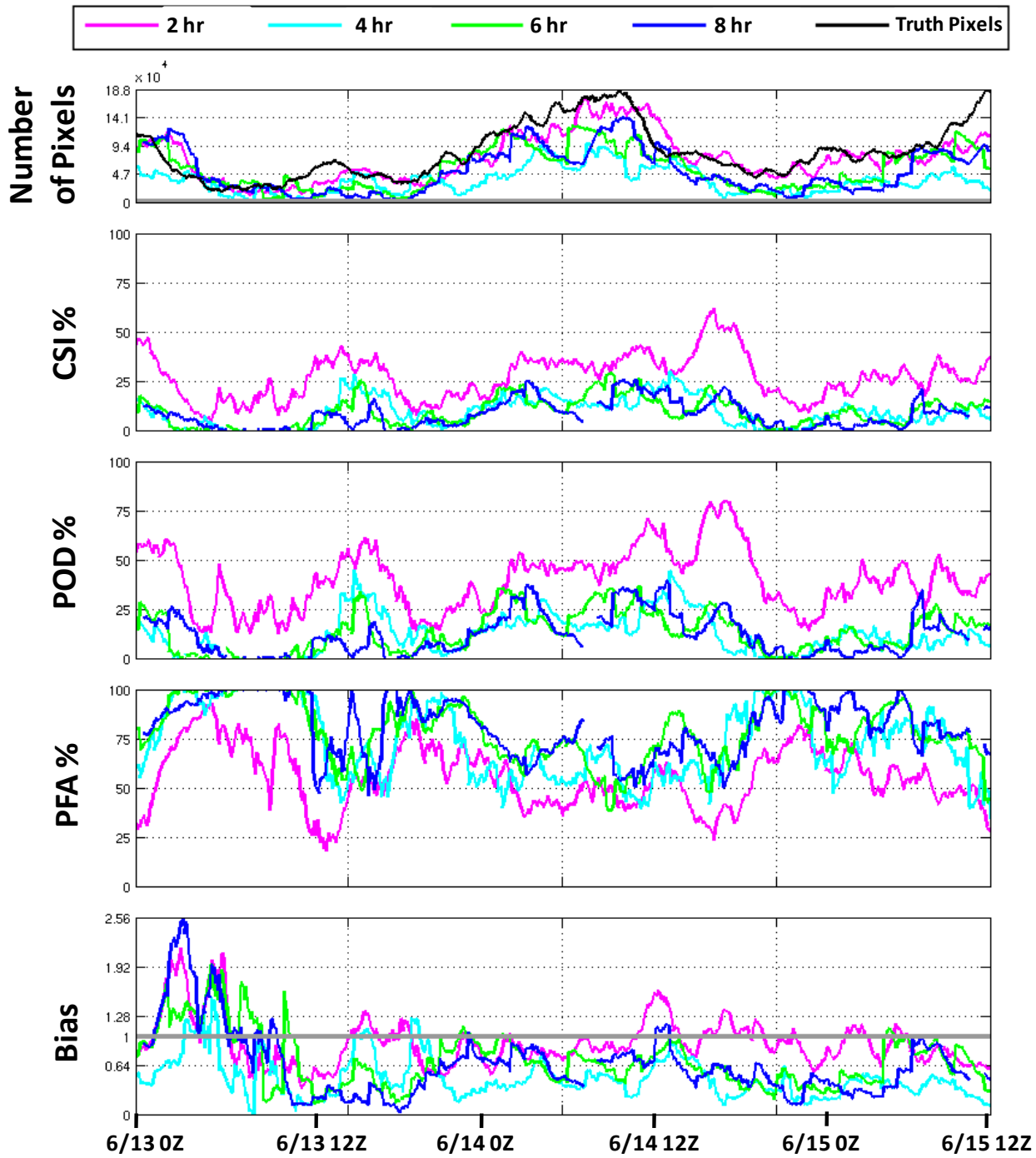

Figure 6d: Statistical scores for 13-15 June 2009 of Blockage thresholded with a value (110) which corresponds to the lowest values that are avoided by aircraft. The trends in the number of pixels between truth and forecasts are indicative of development and decay of weather events and the ASPA forecasts match this development very well. Statistical scores show good performance approaching the 20% CSI level. Blockage scoring is running in real time and scientific and performance assessment is performed daily by a monitoring staff.

6. FUTURE WORK

6.1 Convective Initiation

Convective initiation (CI) remains a nowcasting challenge at all forecast time horizons. At short forecast time horizons (under 1 hour), infrared imagery from the Geostationary Operational Environmental Satellites (GOES) can be used to identify precursors to radar echoes, such as cloud top cooling rate (Roberts and Rutledge, 2003) and infrared channel differences. Using infrared indicators to forecast CI is the basis of the Satellite Convection Analysis and Tracking System (SATCAST) (Mecikalski and Bedka 2006). The SATCAST system uses a cloud mask component, an atmospheric motion vector component, and a nowcasting component to create eight satellite-derived CI indicators based on tracking and trending of cloud properties in multiple infrared channels. The eight indicators are combined into a single CI nowcast field with values from zero to eight, where pixels with higher values indicate a higher confidence in CI. SATCAST development was funded by the NASA Advanced Satellite Aviation-weather Products (ASAP) program which provides satellite-derived meteorological products and expertise to the FAA weather research community.

The SATCAST system is currently being transitioned to and tested in the ASPA environment to evaluate its benefit to convective forecasts for aviation. As part of this transition and evaluation, algorithms have been developed that utilize the SATCAST CI indicators in ASPA. These algorithms use a combination of atmospheric variables (including stability and lower-tropospheric winds) and image processing to create enhanced CI interest in regions identified by SATCAST to be favorable for CI. Figure 7 shows an example of a forecast of airmass initiation using the SATCAST CI indicators. There is limited VIL over North Carolina at the time the forecast is made (Figure 7a). The visible cloud field shows little obvious structure. One hour later, isolated, cellular airmass cells have formed over the region (Figure 7b). The 1 hour forecast without the SATCAST CI fields (Figure 7c) does not capture this development. When the SATCAST indicators are included (Figure 7d), the forecast captures the initiation of the airmass cells. While on a cell by cell basis there are differences in the location and strength of the forecasted and observed cells, the characteristics of the convection that are important to aviation applications, namely the cellular mode and approximate location and area coverage of the cells, are captured. Future work will continue to improve the use of the SATCAST CI indicators in ASPA with the goal of improving short-term CI in convective forecasts.

6.2 Probabilistic Forecasts

To convey uncertainty and risk in using the forecast for strategic planning we are actively developing a probabilistic forecast for integration into aviation decision support tools. To do so we are exploring several techniques including: a statistical approach that predicts the probability of intense weather based on measured probability distribution functions, a novel decision tree technique called Random Forecast (RF), and time lagged ensembles (Pinto et al., 2009).

6.2.1 Statistical Approach

A probabilistic forecast field can be generated that expresses the probability of exceeding a Blockage threshold (see section 5) based on historical and recent forecast performance. A series of probability distribution functions are generated for a range of spatial coverage scales. The spatial coverage scale is accomplished by sampling the weather within several different sized boxes centered on a pixel, and then a fraction of the number of weather pixels is calculated within each box, the box with the largest fraction value determines the scale. The probability of exceeding a threshold Blockage value (which is calibrated against pilot behavior; Figure 8) can then be computed from a database of probabilities, one for each scale classification and forecast time horizon,

based on the forecast at that pixel. The database of probabilities will be generated and modified using archived and real-time data, and categorized by the spatial scale classification and forecast time horizon.

6.2.2 Random Forest

Methods for enhancing the fusion of observation and NWP model data have begun to be investigated as an approach to improving ASPA's short-range forecasts. These methods may help developers in providing probabilistic forecasts and estimates of forecast uncertainty, as well as evaluating the potential usefulness of new information sources.

Results suggest that a non-linear statistical analysis technique called *random forests* (Breiman 2001) holds promise for addressing these challenges (Williams et al. 2008a, 2008b, and 2009). A RF is a set of decision trees, created via an automated "training" process, that collectively relate a vector of predictor values (e.g., model fields and observation features at a map pixel) to a targeted output quantity (e.g., whether there will be a storm there one hour later) by "voting" on the correct classification. They also provide an assessment of the relative value, or "importance", of each predictor.

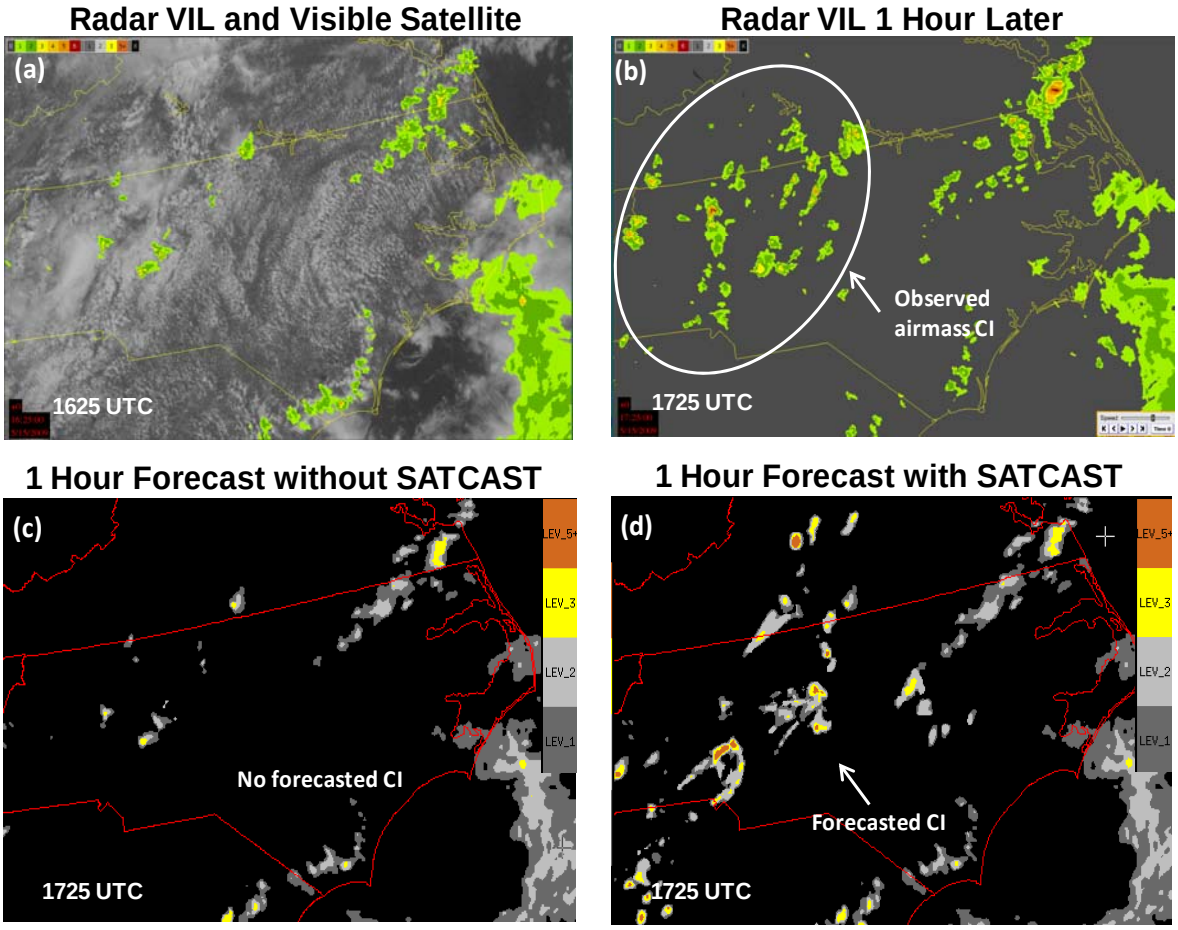

Figure 7: Example of airmass initiation using SATCAST in ASPA. a) Visible satellite and radar VIL over NC at 1625 UTC 15 May 2009 shows limited VIL in this region and little coherent structure in the cloud field. (b) The observed VIL 1 hour later shows that scattered, cellular airmass storms have initiated over NC. (c) The VIL forecast without SATCAST CI does not depict the newly-developed storms, whereas the forecast with the SATCAST CI (d) captures the convective initiation and the scattered nature of the convection.

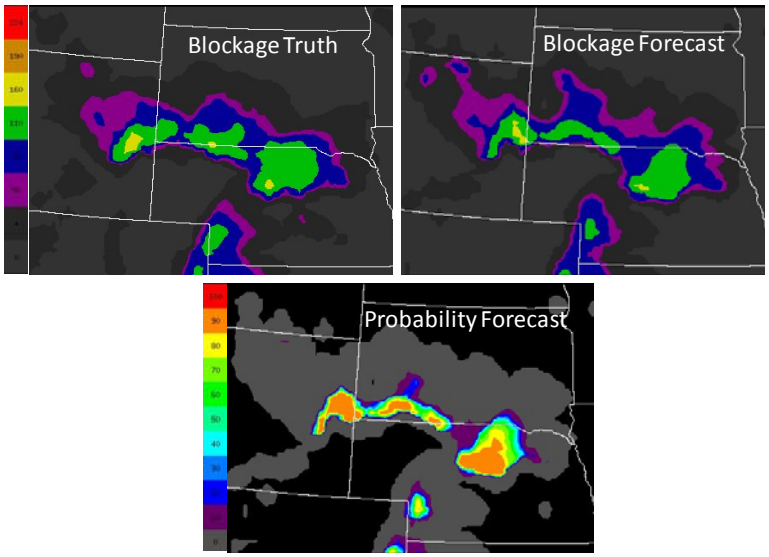

Figure 8: Probability of exceeding the Blockage threshold of 110 units (lower panel) based on the 60 minute Blockage forecast (upper panel) issued at 01:30:00 UTC on 14 June 2009. "Truth" Blockage is also shown (upper left panel) at the forecast verification time of 02:30:00 UTC.

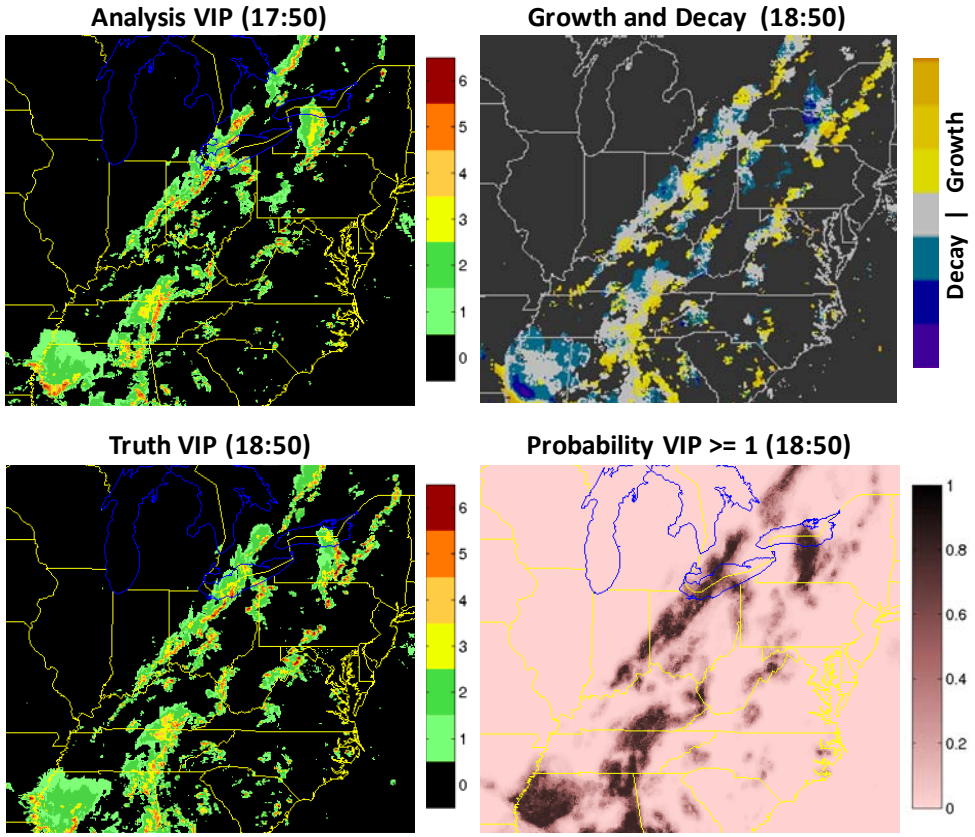

Figure 9: Case study results for a 1-hr RF probabilistic forecast from 19 June 2007 at 1805 UTC. (Upper left) Issue Time VIP; (lower left) actual truth VIP; (upper right) differences between Truth and Issue Time VIP at 1805 UTC (yellows without adjacent blues represent pure growth regions), and (lower right) calibrated probability at 1805 UTC. There is some weak probability in the pure growth regions, suggesting predictive skill.

Predictor fields obtained from observations and NWP models have been evaluated along with various derived features including distance computations and spatially-filtered versions of selected fields. The RF methodology has been used to examine the relevance of different fields, features and scales for different aspects of the storm nowcast problem (e.g., isolated storm initiation or prediction at different temporal lags or intensity thresholds). For example, RF importance rankings of candidate predictors were performed as a first step to understanding what weather regimes (e.g., "weather types" or distances from existing convection) were significant and would require separate forecast logic. A methodology for calibrating RF votes as probabilities was developed; based on the frequency of observed VIP level exceedance for each possible number of RF votes on an independent test set. It is based on running the RF logic on a large number of independent test cases and observing what happened one hour later for each possible RF prediction vote count. The calibration for that vote count is then taken to be the frequency of observed VIP level exceedance; the end result is a mapping from vote counts to probabilities. Figure 9 shows results from a case, 1805 UTC on 19 June 2007, including the RF model's probabilistic 1-hour nowcast of VIP level 1 exceedance and verification of the deterministic nowcast obtained by thresholding at the maximum CSI probability threshold determined from a test set. Experimental RF-based data fusion models are being run in real-time in the summer of 2009 to test their skill in predicting VIP 1 and VIP 3 exceedance probabilities.

The RF technique shows promise as a technique for creating probability and forecast uncertainly fields. It also appears useful as an approach to objectively identifying the potential contribution of new candidate predictor fields and their scale of contribution at different lead times. Thus, insights gained from the application of this technique may feed back into the development of ASPA data fusion logic, and the output of an RF empirical model may itself be used as a component of that logic or as a complementary uncertainty assessment.

7. SUMMARY

A research demonstration of 0-8 hour forecasts of VIL precipitation has been running since July 2008. The forecasts are a blend between tracking-trending-initiation and NWP forecasts and show promising skill at predicting aviation-specific content including storm mode and permeability structure. For the summer 2009 we added blended forecasts of Echo Tops, extended both the VIL and Echo Tops forecasts out to 8 hours, and extended the domain to cover most of the Midwestern and Eastern US. In 2010 we plan to provide CONUS coverage with companion forecast error estimates for probabilistic use of the forecast information.

8. REFERENCES

Bellon, A. and I. Zawadzki, 1994: Forecasting of hourly accumulations of precipitation by optimal extrapolation of radar maps, *J. of Hydrol.*, **157**, 211-233.

Benjamin S., M. Hu, S. Weygandt, D. Devenyi, 2009a: Integrated assimilation of radar, satellite, and METAR cloud data for initial hydrometeor/divergence fields to improve hourly updated short-range forecasts from the RUC, Rapid Refresh, and HRRR, WMO WSN Whistler, B.C., Canada

Benjamin, S. G., T. G. Smirnova, S. S. Weygandt, M. Hu, S. R. Sahm, B. D. Jamison, M. M. Wolfson, and J. O. Pinto, 2009b: The HRRR 3-km storm-resolving, radar-initialized, hourly updated forecasts for air traffic management. *AMS Aviation, Range and Aerospace Meteorology Special Symposium on Weather-Air Traffic Management Integration*, Phoenix AZ

Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Sminova, and R. Bleck, 2004: Mesoscale weather prediction with the RUC hybrid isentropic-terrain-following coordinate model. *Mon. Wea. Rev.*, **132**, 473-494.

Breiman, L., 2001: Random forests. *Machine Learning*, **45**, 5-32.

Brewster, K.A., 2004: Phase-correcting data assimilation and application to storm-scale numerical weather prediction. Part I: Method description and simulation testing. *Mon. Wea. Rev.*, **131**, 480-492.

Dupree, W., D. Morse, M. Chan, X. Tao, C. Reiche, H. Iskenderian, M. Wolfson, J. Pinto, J. K. Williams, D. Albo, S. Dettling, and M. Steiner, S. Benjamin and S. Weygandt, 2009: The 2008 CoSPA* Forecast Demonstration (Collaborative Storm Prediction For Aviation, *89th AMS Annual Meeting ARAM Special Symposium on Weather - Air Traffic* Phoenix, AZ.

Dupree, W.J., M. Robinson, R. DeLaura, R. A. P. Bieringer, 2006: Echo Tops Forecast Generation and Evaluation of Air Traffic Flow Management Needs in the National Airspace System, *AMS 12th Conference on Aviation, Range, and Aerospace Meteorology*, Atlanta, GA.

Dupree, W.J., M.M. Wolfson, R.J. Johnson Jr., R.A. Boldi, E.B. Mann, K. Theriault Calden, C.A. Wilson, P.E. Bieringer, B.D. Martin, and H. Iskenderian, 2005: FAA Tactical Weather Forecasting in the United States National Airspace, *Proceedings from the World Weather Research Symposium on Nowcasting and Very Short Term Forecasts*. Toulouse, France.

Dupree, W.J., R.J. Johnson, M.M. Wolfson, K.E. Theriault, B.E. Forman, R.A. Boldi, and C. A. Wilson, 2002: Forecasting Convective Weather Using Multiscale Detectors and Weather Classification – Enhancements to the MIT Lincoln Laboratory Terminal Convective Weather Forecast. *AMS 10th Conference on Aviation, Range, and Aerospace Meteorology*, Portland, Oregon, 132-135.

Evans, J. E. and E. R. Ducot, 2006: Corridor Integrated Weather System. *Lincoln Laboratory Journal*, **16**, 59-80.

Evans, J., M. Weber and W. Moser, 2006: Integrating Advanced Weather Forecast Technologies into Air Traffic Management Decision Support, *MIT Lincoln Laboratory Journal*, v. 16, n. 1, pp. 81-96 (available for download at http://www.ll.mit.edu/mission/aviation/publications/publications.html)

FAA REDAC, 2007: "Weather-Air Traffic Management Integration Final Report," Weather – ATM Integration Working Group (WAIWG) of the National Airspace System Operations Subcommittee, Federal Aviation Administration (FAA) Research, Engineering and Development Advisory Committee (REDAC). 3 October 2007 (to be available at http://research.faa.gov/redac/)

Lakshmanan, V., R. Rabin, and V. Debrunner, 2003: Multiscale Storm Identification and Forecast, *J. Atm. Res.*, 367-380.

Matthews, M, M. Wolfson, R. DeLaura, J. Evans and C. Reiche, 2009: Measuring the uncertainty of weather forecast specific to air traffic management operations. *AMS Aviation, Range and Aerospace Meteorology Special Symposium on Weather-Air Traffic Management Integration*, Phoenix AZ

Mecikalski, J.R. and K.M. Bedka, 2006: Forecasting convective initiation by monitoring evolution of moving cumulus in daytime GOES imagery. *Mon. Wea. Rev.*, **134**, 49-78.

Megenhardt, D. L., C. Mueller, S. Trier, D. Ahijevych, and N. Rehak, 2004: NCWF-2 Probabilistic Forecasts. *AMS Eleventh Conf. Aviat. Range Aerospace Meteorol.*, paper 5.2.

NextGen ConOps, Joint Planning and Development Office, 2007: Concept of Operations of the Next Generation Air Transportation System, http://www.faa.gov/about/office_org/headquarters_offices/ato/publications/nextgenplan/resources/view/NextGen_v2.0.pdf

Pinto J., M. Xu, D. Dowell, M. Steiner, and J. K. Williams, 2009: Assessment of convective forecast uncertainty using high-resolution model ensemble data, WMO WSN Whistler, B.C., Canada

Roberts, R. D. and S. Rutledge, 2003: Nowcasting storm initiation and growth using GOES-8 and WSR-88D data. *Wea. Forecasting*, **18**, 562-584.

Robinson, M., W. Moser, and J. Evans, 2008: Measuring the Utilization of Available Aviation System Capacity in Convective Weather. *AMS 13th Conference on Aviation, Range, and Aerospace Meteorology (ARAM)*, New Orleans, LA.

Seed, A. W. and T. Keenan, 2001: A Dynamic Spatial Scaling Approach to Advection Forecasting. *30th International Conference on Radar Meteorology*, 19-24 July 2001, 492-494.

Thompson, G., R. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part 1: Description and sensitivity analysis. *Mon. Wea. Rev.*, 132, 519-542.

Weygandt, S, Benjamin, Smirnova T., Kevin Brundage, Alexander, C., Hu M., Brian Jamison, B., Sahm S: 2009: Evaluation of the High Resolution Rapid Refresh (HRRR): an hourly updated convection resolving model utilizing radar reflectivity assimilation from the RUC / RR WMO WSN Whistler, B.C., Canada

Williams, J. K., D. Ahijevych, M. Steiner, and S. Dettling, 2009: Data mining for thunderstorm nowcast system development WMO WSN Whistler, B.C., Canada

Williams, J. K., D. A. Ahijevych, C. J. Kessinger, T. R. Saxen, M. Steiner and S. Dettling, 2008a: A machine-learning approach to finding weather regimes and skillful predictor combinations for short-term storm forecasting. *AMS 6th Conference on Artificial Intelligence Applications to Environmental Science and 13th Conference on Aviation, Range and Aerospace Meteorology*, paper J1.4.

Williams, J. K., D. Ahijevych, S. Dettling and M. Steiner, 2008b: Combining observations and model data for short-term storm forecasting. In W. Feltz and J. Murray, Eds., Remote Sensing Applications for Aviation Weather Hazard Detection and Decision Support. *Proceedings of SPIE*, 7088, paper 708805.

Wolfson, M.M., W. J. Dupree, R. Rasmussen, M. Steiner, S. Benjamin, S. Weygandt, 2008: Consolidated Storm Prediction for Aviation (CoSPA), *AMS 13th Conference on Aviation, Range, and Aerospace Meteorology*, New Orleans, LA, 2008.

Wolfson, M. M. and D. Clark, 2006: Advanced Aviation Weather Forecasts, *Lincoln Laboratory Journal*, Vol. 16, Number 1. 31-58.

Wolfson, M. M. Forman, B. E., Calden, K. T., Dupree, W. J., Johnson Jr., R. J., Boldi, R., Wilson, C. A., Bieringer, P., Mann, E. B., Morgan, J., 2004: Tactical 0-2 Hour Convective Weather Forecasts for FAA, *11th Conference on Aviation, Range, and Aerospace Meteorology (ARAM), Hyannis, MA, Amer. Meteor. Soc.*

Wolfson, M.M., B.E. Forman, R.G. Hallowell, and M.P. Moore, 1999: The Growth and Decay Storm Tracker, *AMS 8th Conference on Aviation, Range, and Aerospace Meteorology*, Dallas, TX, 58-62.