***J11.4***

# ***Probabilistic Forecasts of Convection: How do we do it?***

*Cynthia K Mueller, B.G. Brown, D. Megenhardt*
*National Center for Atmospheric Research*

# *Introduction*

When the abstract for this paper was prepared, our plans were to present a comprehensive study of potential methods to produce probability forecasts and their evaluation. Since then our focus has narrowed to a description of our current efforts which involve very-short term (0-2hr) probabilistic forecast of convection on the national scale. As part of the FAA Aviation Weather Research Program (AWRP) Convective Weather Product development team, the National Center for Atmospheric Research is in the process of developing and implementing (at NOAA's Aviation Weather Center) a 0-2 hr national-scale nowcast product called the National Convective Weather Forecast (NCWF-2). The NCWF development is sponsored by the Federal Aviation Administration's (FAA) Aviation Weather Research (AWR) program as part of the Convective Weather Product Development Team. The Convective Weather Product Development Team consists of MIT Lincoln Laboratories, National Severe Storms Laboratory (NSSL), National Weather Service's Aviation Weather Center (AWC), and NCAR.

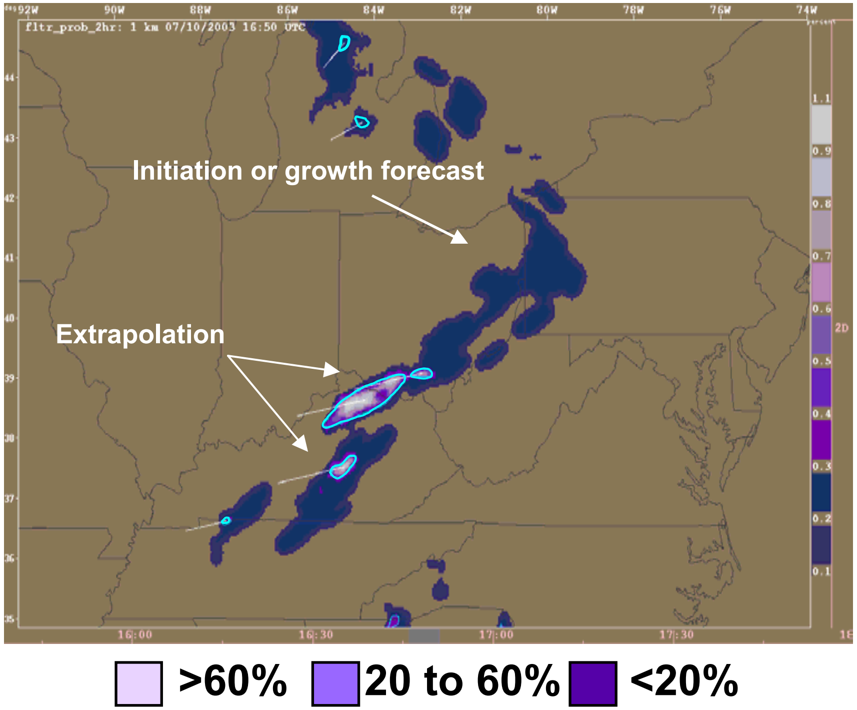
The current operational version of NCWF (Megenhardt et. al., 2000) provides a binary forecast of storm location with a one hour lead time, the new NCWF-2 provides probabilistic forecast out to 2 hrs, and the plan for the future is NCWF-6 which forecasts out to 6 hr. At these longer lead times, the forecasts become more uncertain because of the chaotic nature of convection. Forecasts need to express expected outcome and uncertainty. One methodology to express uncertainty is to use probabilities. In order for the NCWF to evolve to longer period (0-6 hr) forecasts, probabilities of some type are required. In this paper, we describe the methodology and verification methods we incorporate to produce the NCWF-2 probabilistic forecasts based primarily on observational analysis techniques.

*Megenhardt, Dan, C.K. Mueller, N. Rehak, and G. Cunning, 2000: Preprints 9th Conference on Aviation, Range, and Aerospace Meteorology, Amer. Met. Society, Orlando, Fl.*

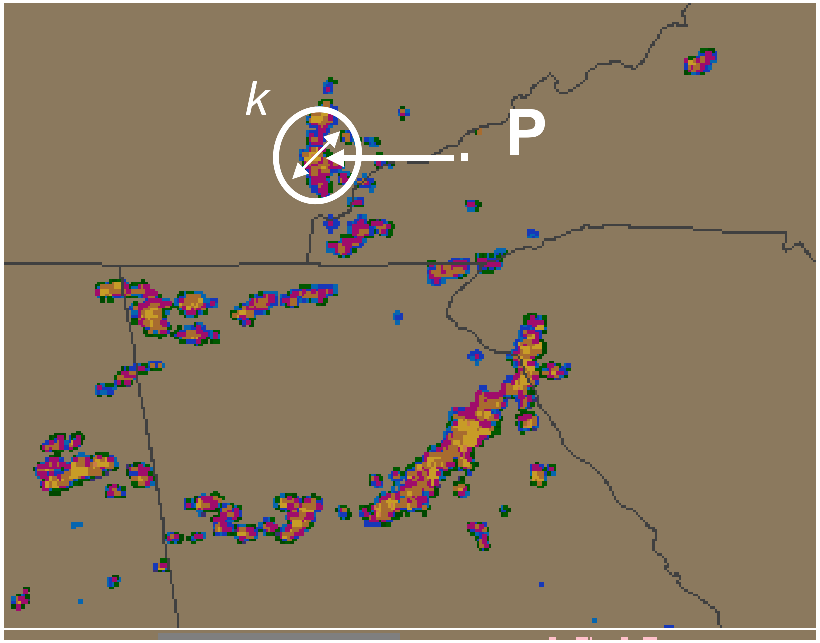
# *NCWF-2*

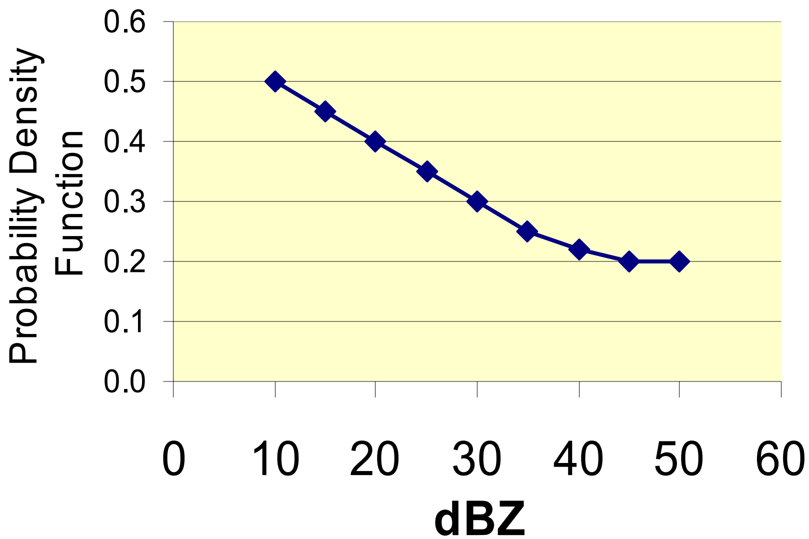
The image to the right shows an example of a two hour NCWF-2 probability forecast. The probabilities are indicated by shades of purple. The darker purples represent lower probability regions and lighter colors higher probabilities. The cyan contour represents the extrapolation-only forecast. The NCWF-2 was run in real time at NCAR using the WSI 2-km reflectivity data and national lightning detection network (cloud to ground) data. The NCWF-2 provides 1 and 2 hr probability forecasts for convection that are updated every 5 min. It ran in a test mode during the 2003 summer season at NCAR.

NCWF-2 Probability Nowcasts



# *Methodology – Local Lagrangian*




The methodology used to calculate a preliminary probability field is described in Germann and Zawadzki (2004 referred to as GZ) and illustrated in the images. GZ suggests several ways to calculate probability distributions of precipitation based on radar reflectivities, including synoptic eularian, local eularian, and local lagrangian. GZ shows that the local lagrangian technique provides the most skill and is the one that is incorporated in this paper.

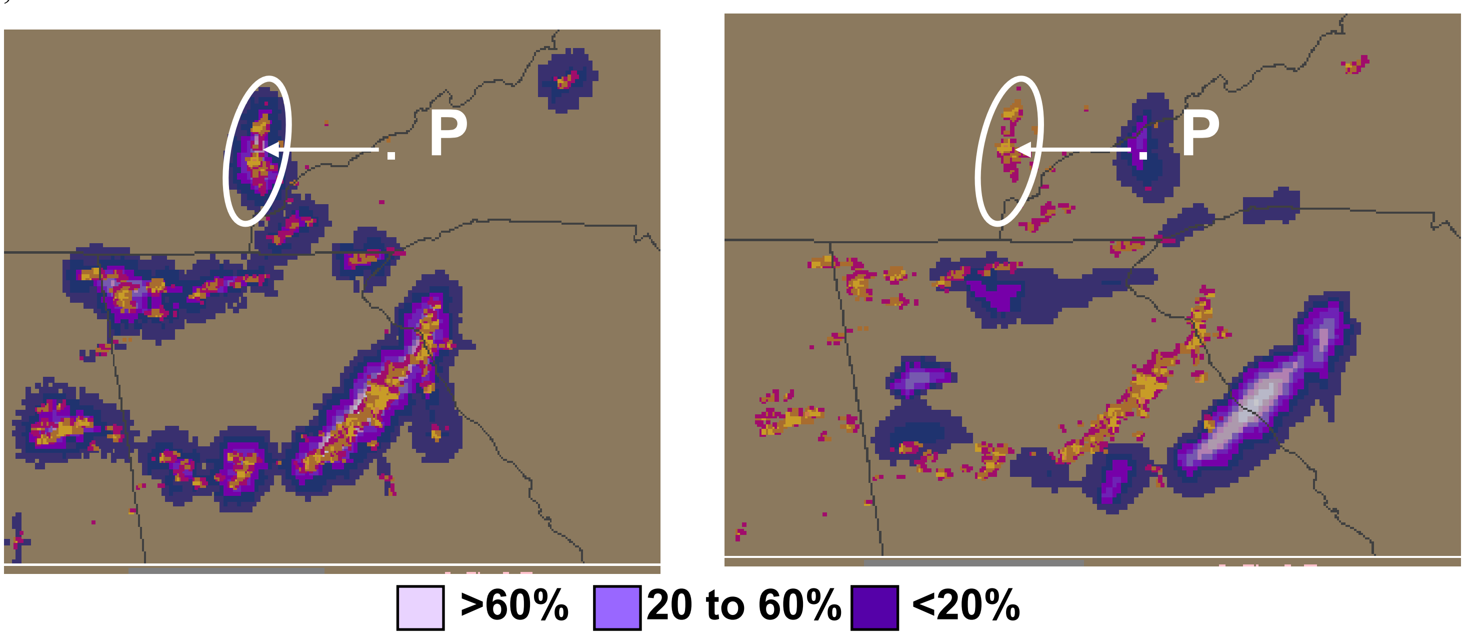
An example of the local lagrangian method is shown. The probability distribution function shown below for point P is calculated by following a trajectory based on the motion vector from neighboring convection to the trajectory origin. The probability function is then calculated by the spatial coverage of various rain rate thresholds within a distance (k – scale) of the origin. GZ calculates that the optimal scale as about 1 km per min (k=60 km for a 60 min forecast) based on evaluations using the conditional square-root of the ranked probability score.

*Germann and Zawadzki, 2004; Part II: Probability forecasts . To be published in Journal of Applied Meteorology*

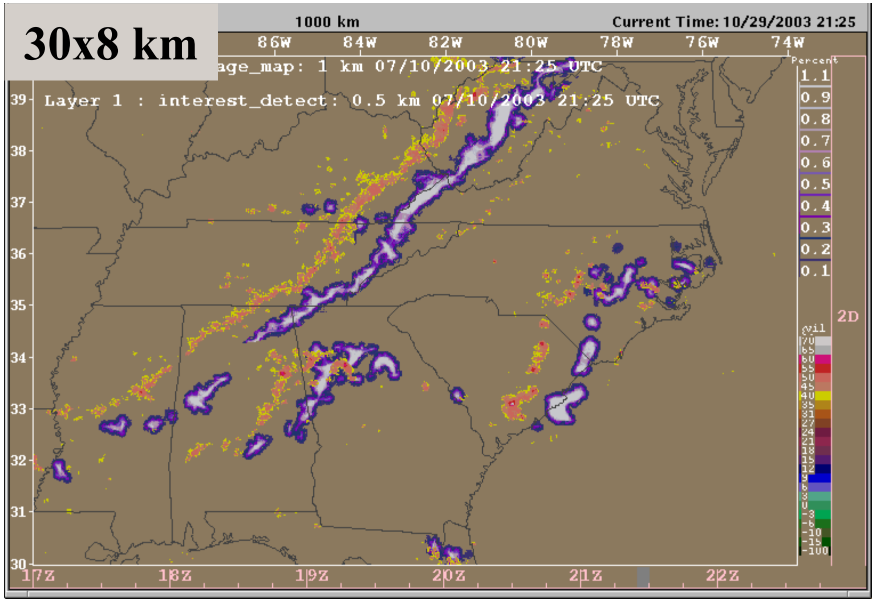
# *Methodology – NCWF-2 probabilities*

The images illustrate the methodology used in the NCWF-2 to calculate probabilties. For NCWF-2 a full probability density function at each point is not desirable. Instead a binary convection - no convection field is used for calculating the probabilities. The binary field combines lightning and radar vertically integrated liquid (vil) water and is overlaid on the images in shades of yellow. A convective – stratiform filter is used to remove stratiform echo from the vil. The binary field is thresholded such that either; (1) a convective-vil value of 3.5 units or greater or 3 or more lightning strokes within 8 km and 5 min are considered convection.

Probabilities are calculated by determining the area coverage of convection within an elliptical filter as illustrated in figure on the left. The elliptical filter is rotated at 10 degree intervals to determine the orientation with the maximum area coverage. The maximum area coverage is mapped as the probability level (purple shades). Probabilities are then advected based on storm motion vectors. In addition to extrapolation, NCWF-2 captures regions of growth by use of the RUC data along with radar trending and diurnal considerations. These regions are always indicated at low-probabilities. Dissipation is based on area trending. However as in all short-term forecast products, NCWF-2 probabilities are primarily the result of extrapolation. The goal of this study is to evaluate methodology and scales used for calculating the probabilities. Therefore, only the extrapolation component of the NCWF-2 is tested. Probability forecasts were calculated for 15, 30, 45, 60, 75, 90, 120 and 180 minutes based on 30x8, 60x16, and 120x32 km filters.

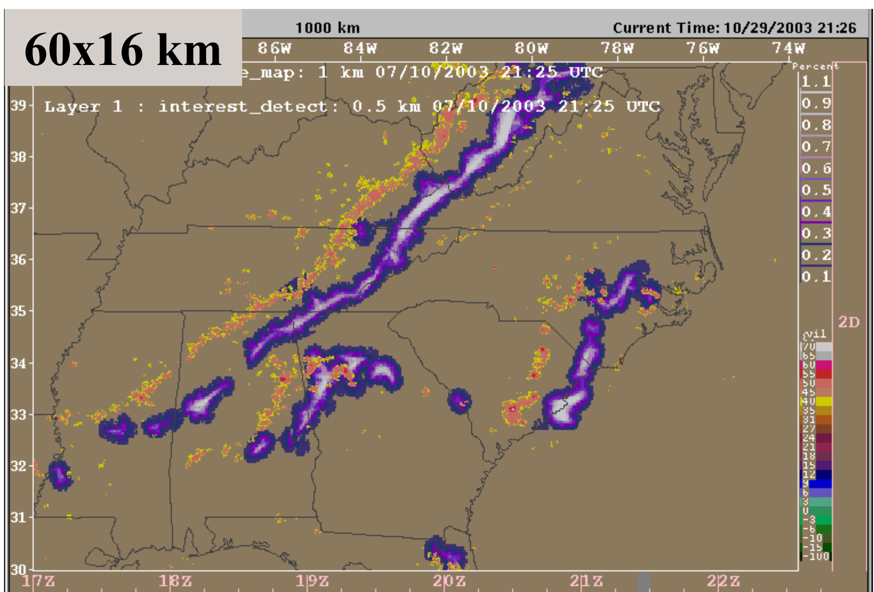


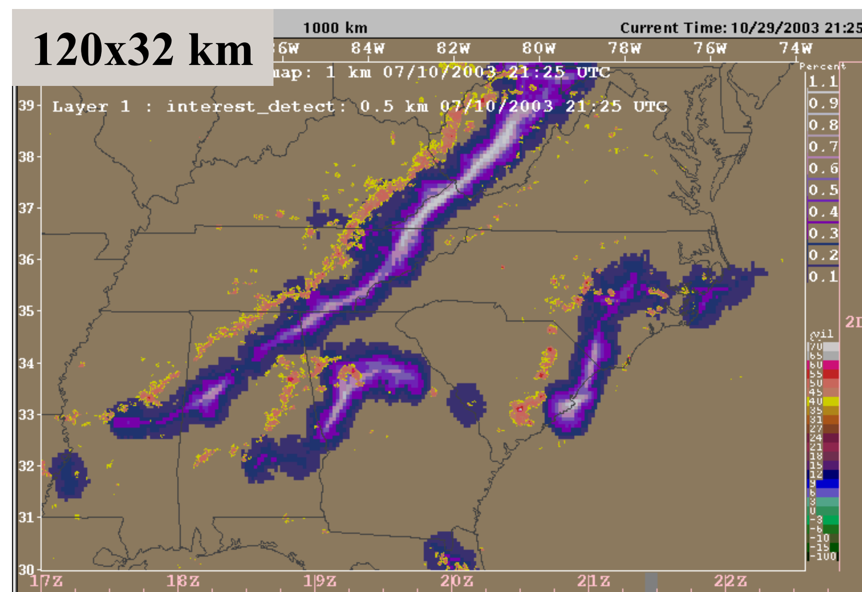
# *Example nowcasts*



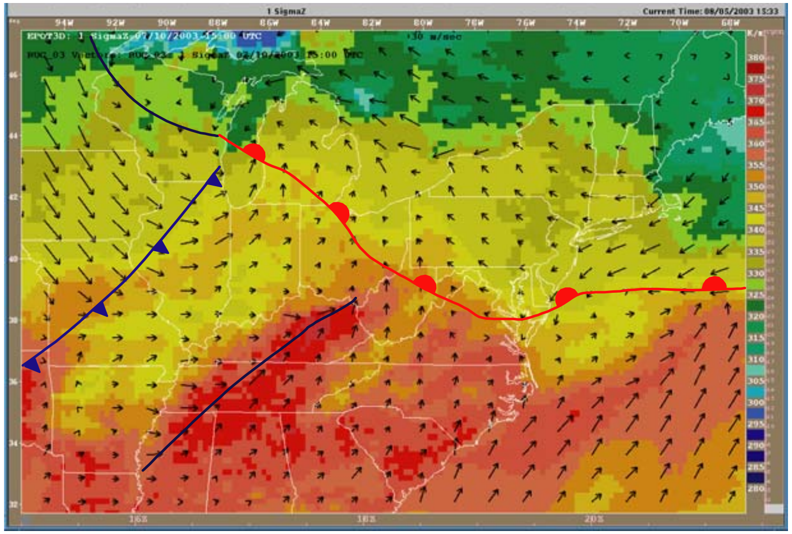

Images show probability nowcasts (purple shades) based on 30x8, 60x16, and 120x32 km filters. The reflectivity at forecast time is super-positioned on the nowcasts. The effect of the filter is evident in the analysis, the smaller filter retains the structure of the convection and has more high level (light color) probabilities. The larger filters, smooth out much of the convective structure and have fewer high probability regions.

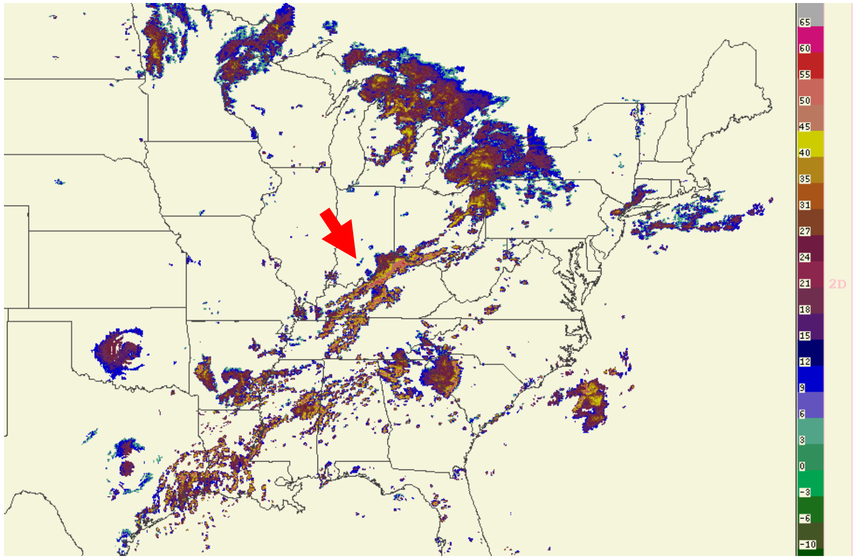
>60% 20 to 60% <20%



# *Case Day – July 10, 2003*





The statistics shown in this paper are all calculated based on a 24 hr period that started on July 10, 2003 at 12Z (0900 CDT). The RUC data from 15Z analysis is shown in the upper panel. The image shows equivalent potential temperatures (reds are >355$^{o}$K, yellows range from 330 to 355$^{o}$K, and greens are < 330$^{o}$K), the low-level wind vectors, and NWS hydrological prediction center (HPC) boundaries (the cold front is blue, warm front is red, and stationary boundaries are black lines).

Early in the time period there are a few isolated cells that were remnants of evening convection. During the afternoon (17Z to 23Z), a line of thunderstorms formed along the stationary boundary located in the warm sector. The resultant line is indicated by the red arrow super positioned on a composite WSI reflectivity data taken at 18Z (15Z CDT). This line of convection continued to grow, propagate to the east, and maintained itself for ~6 hr period. In addition there was considerable isolated storm activity during the afternoon as evident in the WSI radar data in the southern part of the country. By evening the storms systems began to dissipate. During the night there were large isolated areas of nocturnal convection.

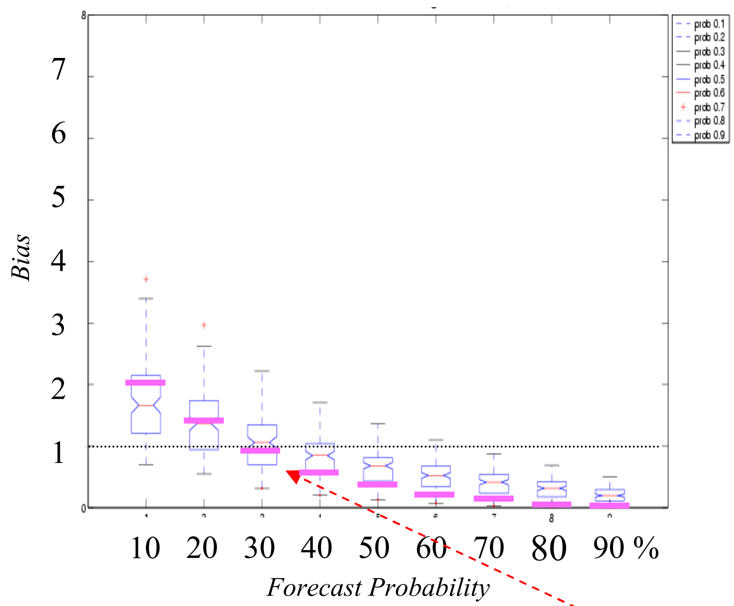
# *Evaluation techniques and results*

Standard forecast statistics are calculated using contingency tables that are based on a grid-to-grid comparison of the forecast and observed fields (Doswell, 1986). The observations and forecasts are extended five grids (10 km) in order to relax the stringent requirements of the grid-to-grid comparisons. Statistics include the probability of detection (POD), false alarm rate (FAR), critical success index (CSI) and bias. The bias is the ratio of the forecast area to the observed area. Generally a bias >1 indicates more area of thunderstorm is being forecast than observed (over-forecasting). A perfect score for POD is 100%, FAR is 0% and CSI is 100%. Statistics were calculated for probabilities at intervals of 10%.

Doswell,C.A., 1986: Short-range forecasting. Mesoscale Meteorology and Forecasting. P.Ray, Ed. Amer. Meteor. Soc., Boston, 793 pp..

# *Forecast Bias*

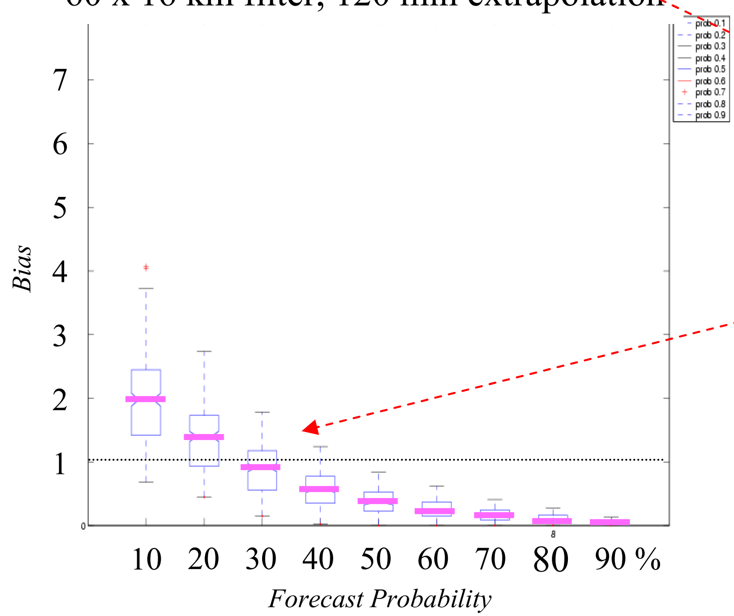
30 x 8 km filter, 120 min extrapolation



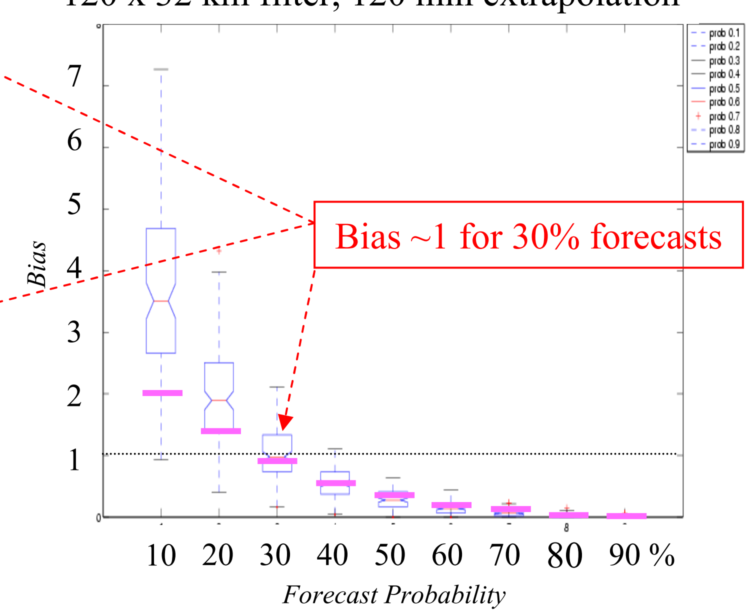

The bias statistics are shown for each probability level for the 120 min forecast period. Box plots indicate the mean (red lines for 30x8 and 120x32 km filters and magenta lines for the 60x16 km filter which are shown on all the plots for a reference), one standard deviation is indicated by the top and bottom lines of the box, and two standard deviations are shown by the short lines.

The effects of the different filter sizes are most evident for the low 10% probability levels where the bias is much larger with the 120x32 km filter (3.5) than with the smaller 30x8 km filter (1.8). The bias at high probability levels tends to decrease with the larger filter. The 30% probability level tends to be ~1 for all filter sizes. As seen in the next slide, contours of 30% probability looks very similar to a deterministic extrapolation forecast especially with the smallest filter.

60 x 16 km filter, 120 min extrapolation

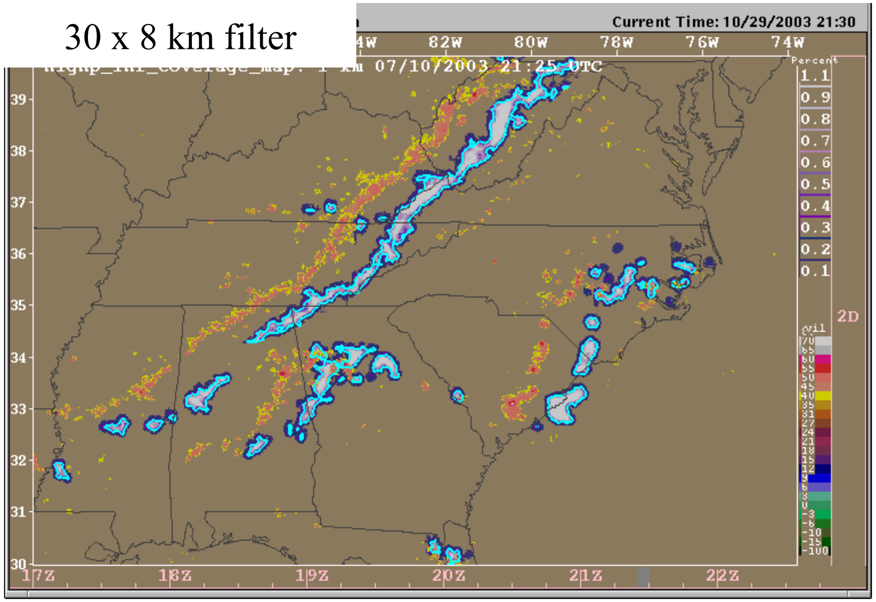



120 x 32 km filter, 120 min extrapolation



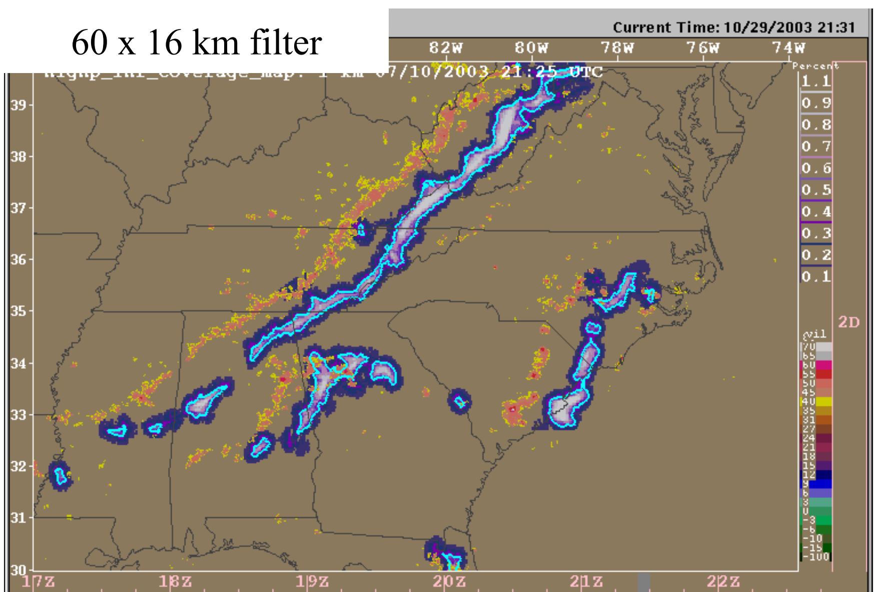
# *Example nowcasts using analyzed scales*

30 x 8 km filter

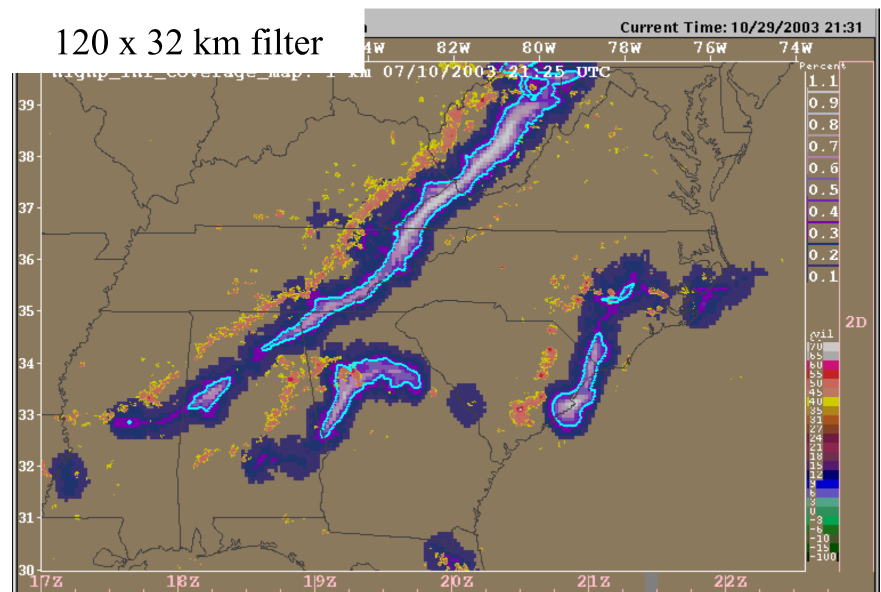



This slide highlights the 30% probability forecasts (shown as the cyan contour). The forecast period is the same as shown previously.

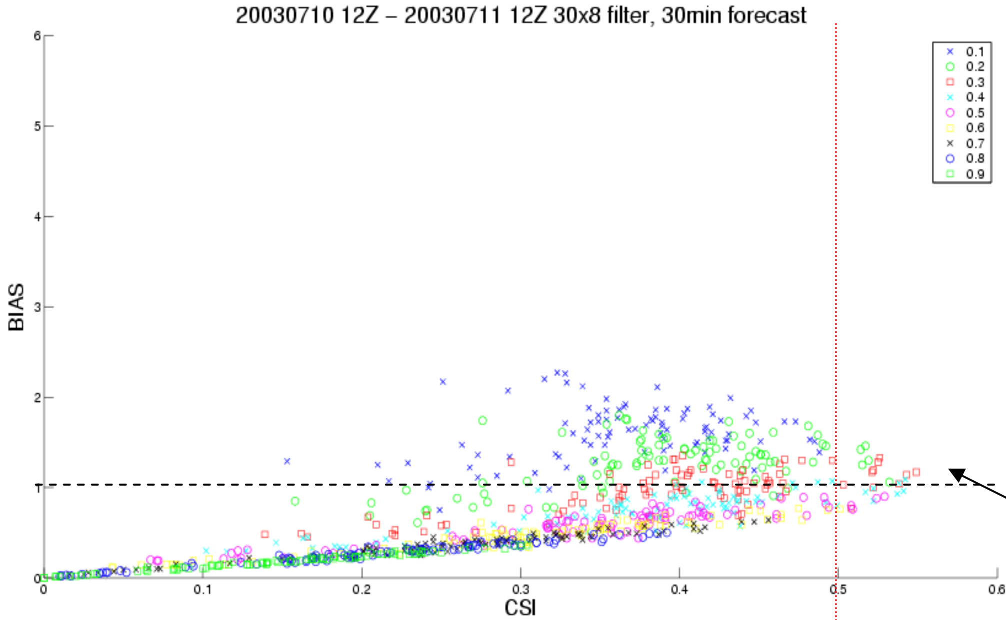
60 x 16 km filter




120 x 32 km filter

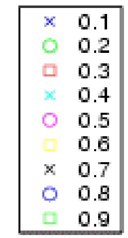


# *Bias vs. CSI*
# *30 min forecast comparisons*



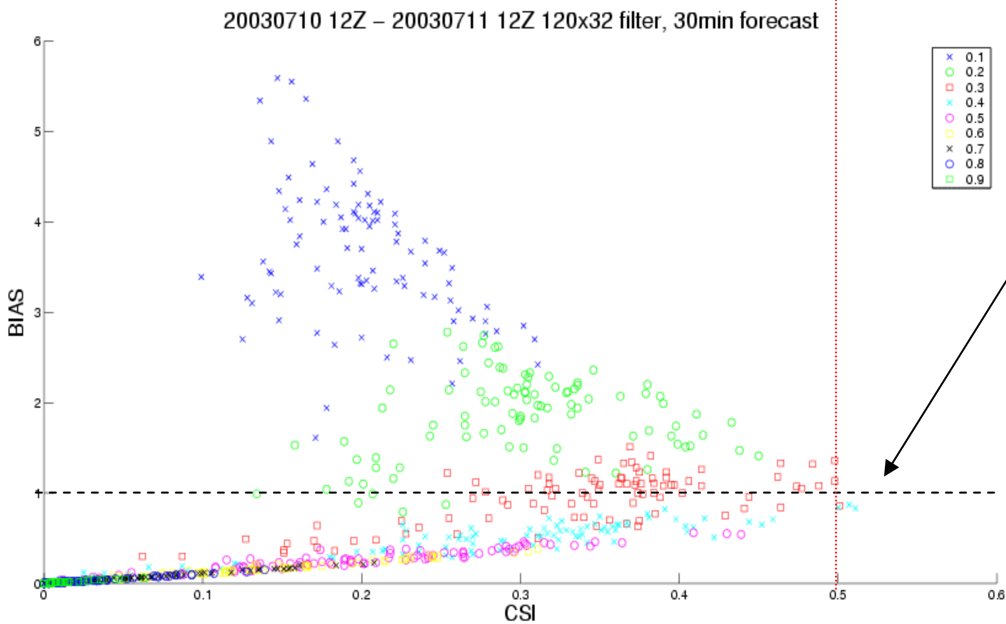

The graphs are scatter plots of BIAS vs. CSI for the 30 min nowcasts. The nowcast probabilities are color-coded by probability level. The colors for each level are shown on the key to the right.

Forecast Probability Levels



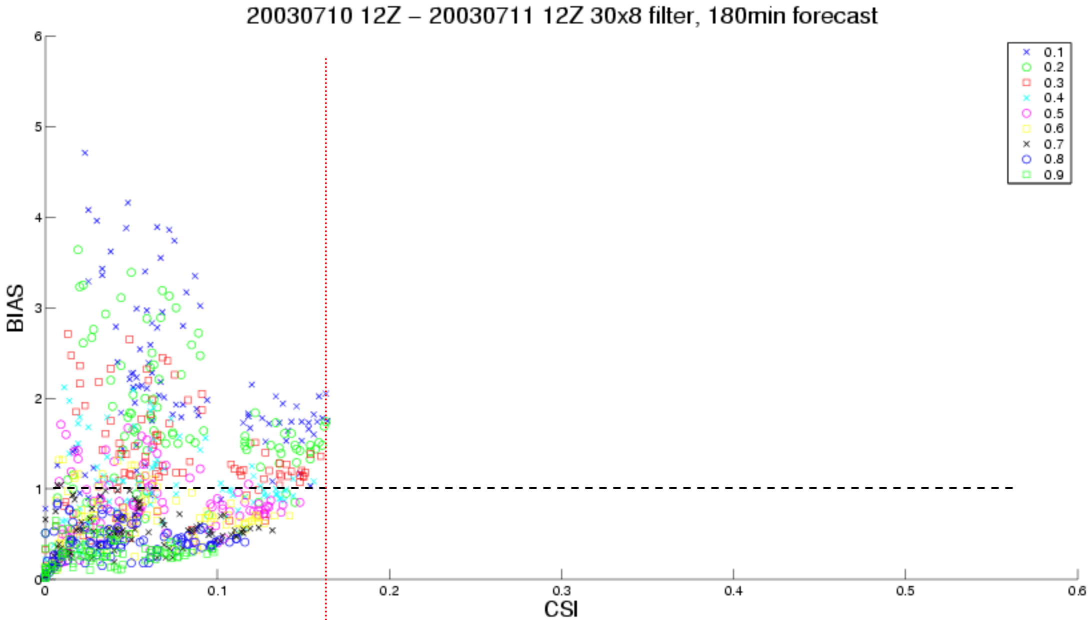

The 30% probability level (red boxes) tend toward the highest CSI values. They also are closest to bias of 1.

Maximum CSI values are slightly higher for the 30x8 km filter. But review of these statistics provides little insight into the best filter size to use.

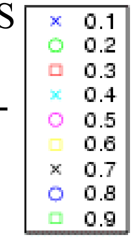



There is a close relationship between CSI and bias. Forecasts with maximum CSI scores have a bias slightly above 1. CSI sharply decreases with decreasing biases less than 1. The CSI also decreases for biases greater than 1 but the slope is not as steep. Or the CSI penalizes a forecast with a bias of 0.5 more than a bias of 1.5.

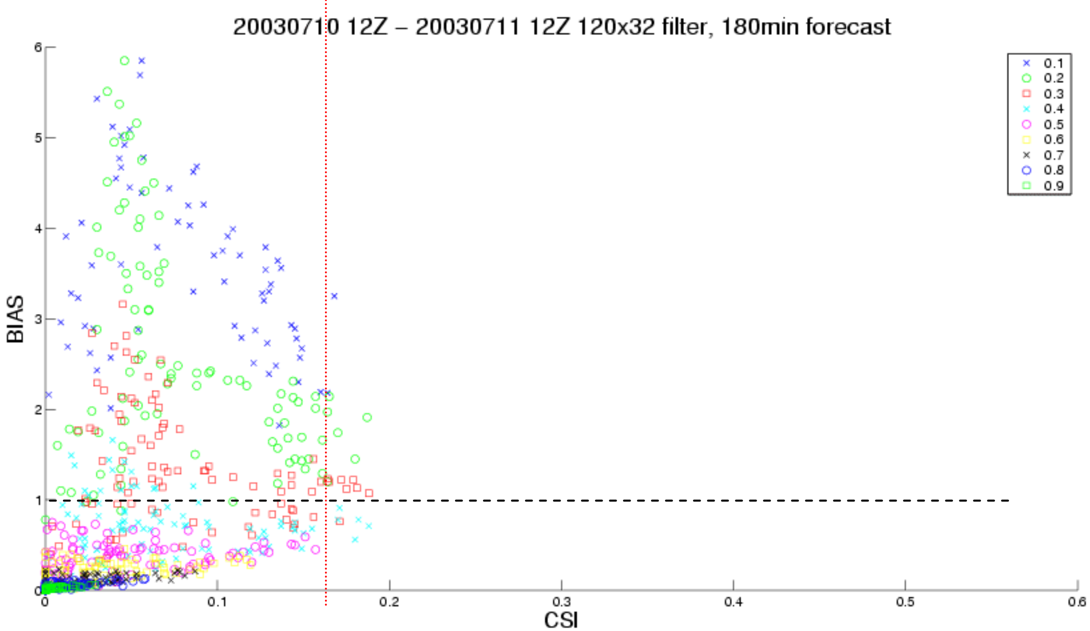
# *Bias vs. CSI*
# *180 min forecast comparisons*




The graphs are scatter plots of BIAS vs. CSI for the 180 min nowcasts. The nowcast probabilities are color-coded by probability level. The colors for each level are shown on the key to the right.

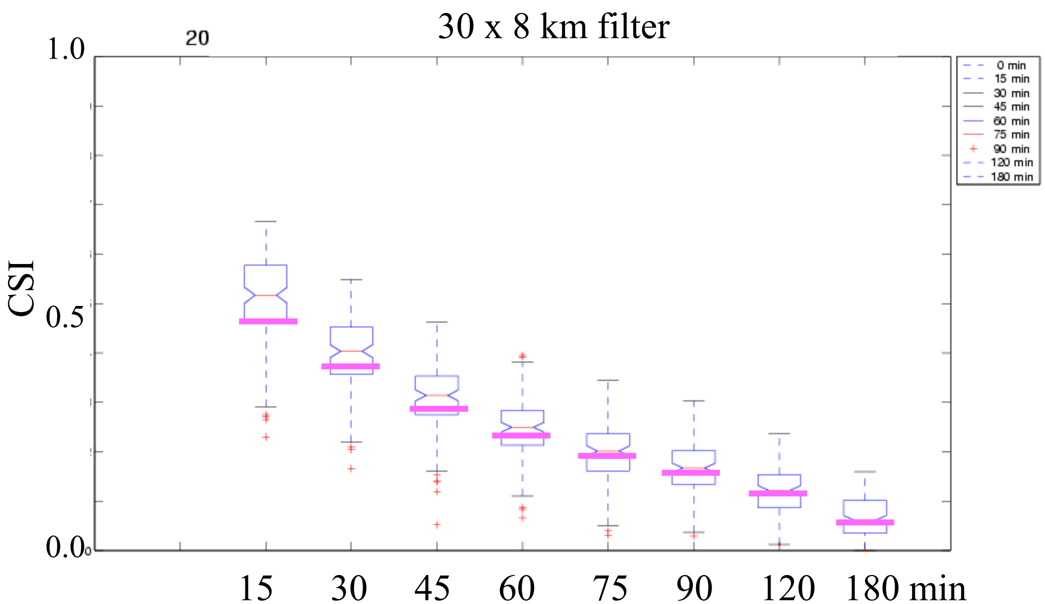



At 180 min, the 30% probability level (red boxes) also tends toward the highest CSI values.



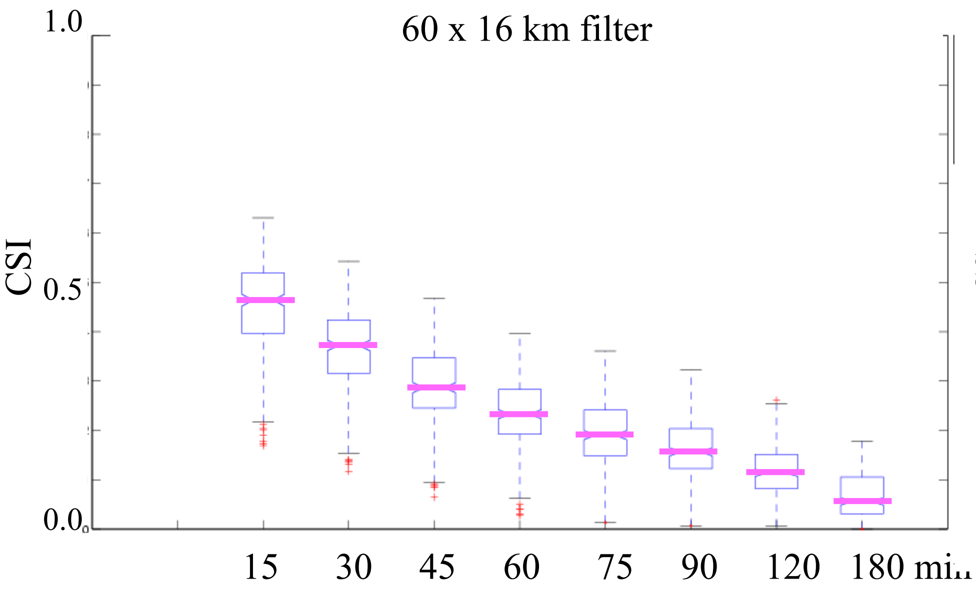

In this case maximum CSI values are slightly higher for the 120 x 32 km filter. But it is still difficult to determine which filter size should be used.

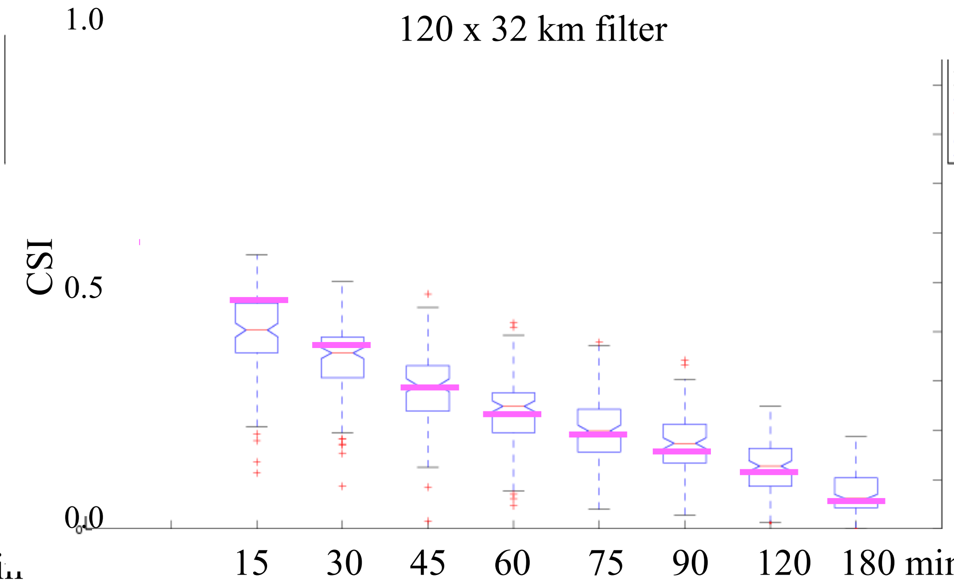
# *CSI for 30% probability level vs. Forecast Period*




Box plot statistics for 30% forecasts calculated using the different filter sizes verses lead time. The magenta lines are the mean values for the forecasts that are calculated at 60x16 km filter.

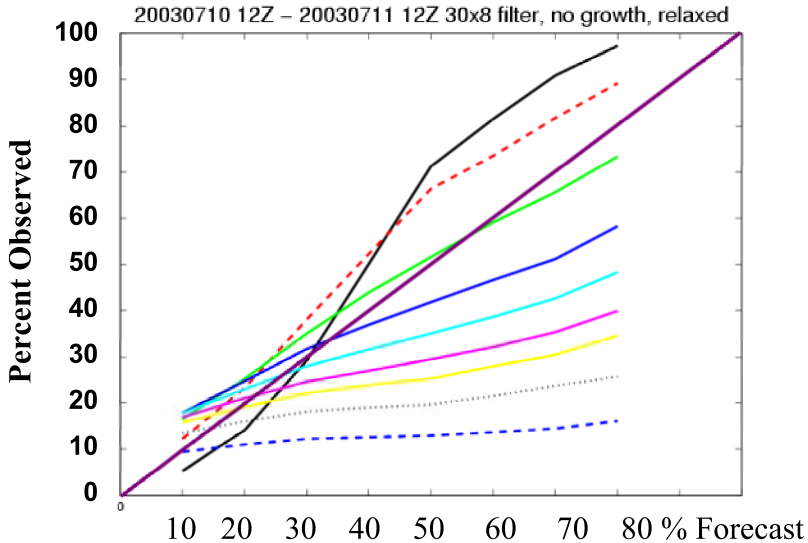
The CSI scores at short-periods 15, 30, 45, and 60 min indicate that the forecast calculated with the 30x8 km filter shows slight skill compared to the larger scale filters. At longer periods (over 75 mins) the CSI values are similar and do not differentiate between the forecasts.



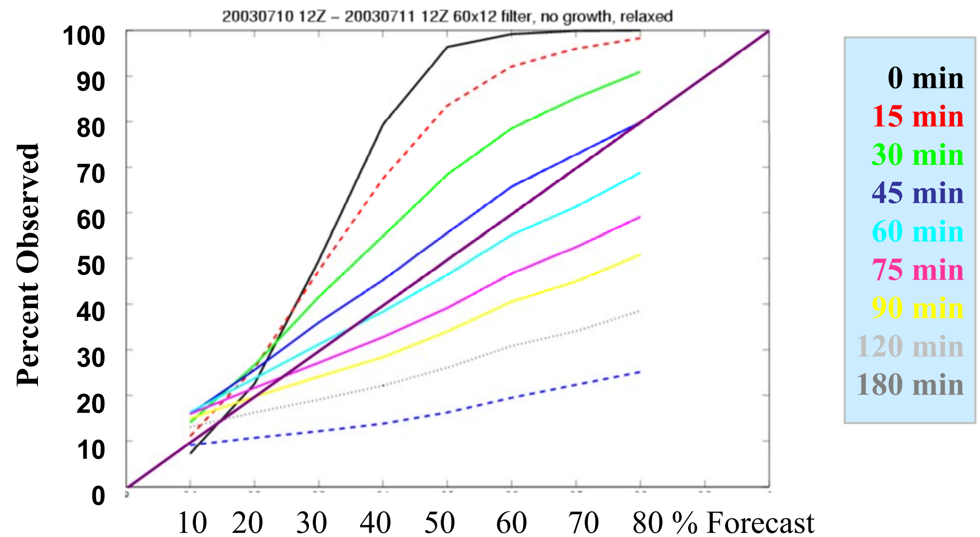



# Reliability (July 10-11, 2003)

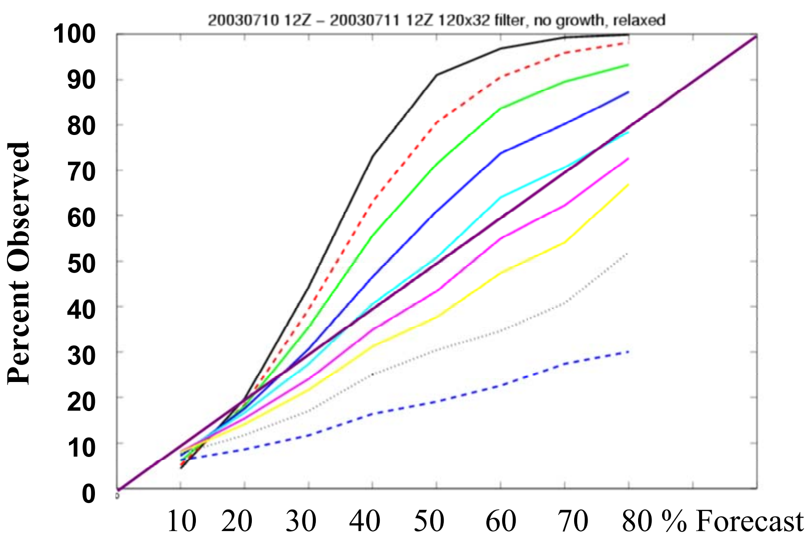
## Extrapolation, k=30km




## Extrapolation, k=60km




## Extrapolation, k=120km




These reliability charts are based on forecast derived with different filter sizes. A perfectly reliable forecast would fall along the black diagonal line. The plots are encouraging because they show that using a spatial filter to derive probabilities can be used to get good reliability in the forecast. Further work is needed to better understand and calibrate the probabilities. Other data sets will be used as well as methods to stratify convective organization.

## *Discussion - Probability Forecasts of Convection*

Forecaster perspective (algorithm and NWP developers too)

- Quantify uncertainty (e.g. we know that forecast accuracy decreases with time, probabilities allow mechanisms to express the degree of uncertainty in the forecast.)
- Provide a methodology (or common unit) to combine observational-based forecasts with NWP.
- Must be well defined (what is the probability forecast for storm location, severity, organization, etc.).
- Requires additional research to better quantify predictability of convective events and associated environmental processes.

User perspective

- ATM - The impact on airspace of weather and the effects of forecast at various probability levels and lead times needs to be understood.
- FAA user - Forecast need to provide more than a physical map of probabilities of convection instead they need to also include convective organization (expect small or large isolated storms, rapidly developing line, a slowly dissipating line, line with gaps).

Verification

- Observations (reflectivity and lightning data) are at a higher resolution than can be achieved by current forecast techniques – how should they be incorporated into the verification?
- Are there ways to quantify loss of accuracy due to offset errors in extrapolation verses initiation, growth and decay?

Challenges

To define a useful and realistic probability forecast.

This paper has dealt with forecasting probabilities of the location of convection at different time periods. This is only a first step, it is recognized that this information alone especially at longer forecast periods is not sufficient, other parameters such as echo height, convective organization, potential gaps in linear systems need to be explored.

## *Summary*

This paper;

- Presents the methodology used to calculate NCWF-2 probabilities.

- Explores verification techniques
  - Shows the inability of standard skill scores such as CSI to discern between various probability forecasts calculated at different scales.
  - Shows that the reliability plots do discern between different filter sizes and probability levels.

- Next steps
  - Analyze other cases.
  - Calibrate forecast reliability and use fewer probability levels
  - Explore forecast product, one possibility is to use probabilities to indicate probability of convection at a certain location and super-position a deterministic forecast (show example).