Machine Learning for Real-time Prediction of Damaging Straight-line Wind

Ryan Lagerquist[1], Amy McGovern[2], Travis Smith[1], Michael Richman[2]

[1]University of Oklahoma & Cooperative Institute for Mesoscale Meteorological Studies; [2]Univ. of Oklahoma; contact: ryan.lagerquist@ou.edu

1. Introduction

- Thunderstorms in the U.S. average > 100 deaths and $10 billion damage per year (Insurance Information Institute 2016).
- Many of these losses caused by straight-line (non-tornadic) wind.

- Machine learning (ML) has been successfully operationalized to predict other convective hazards: hail, tornadoes, aircraft turbulence.
- However, very few studies have applied ML specifically to straight-line wind.

- We created an ML system to forecast probability of damaging straight-line wind ($\geq$ 50 kt) for each storm cell in the CONUS at lead times up to 90 minutes.
- Output was shown in the Spring 2016 Hazardous Weather Testbed.

2. Input Data

Data Type	Sources	Resolution	Time Period
Radar images	Multi-year Reanalysis of Remotely Sensed Storms (MYRORSS)	0.01° (~1 km), 5 minutes	2000-11 (excluding 2009)
Model soundings	Rapid Update Cycle (RUC)	13-20 km, 1 hour	Apr 1994 – Apr 2012
	North American Regional Reanalysis (NARR)	32 km, 3 hours	1979-present
Near-surface wind observations	Meteorological Assimilation Data Ingest System (MADIS)	variable	July 2001 – present
	Oklahoma Mesonet	variable, 5 minutes	1994-present
	One-minute METARs	variable, 1 minute	2000-present
	NWS storm reports	variable	1955-present

- Radar imgs and soundings used to create predictors for the "event" (wind gust $\geq$ 50 kt).
- Wind obs are used to determine when and where event occurred.
- All datasets except Oklahoma Mesonet are CONUS-wide.
- We use 804 days for model development (training, validation, testing).

- We assume that all wind obs are straight-line (non-tornadic):
  1. NWS reports distinguish between tornadoes and straight-line wind.
  2. Tornadoes are much less common (less likely to hit weather station) and more intense (tend to destroy the anemometer when they do hit stations).

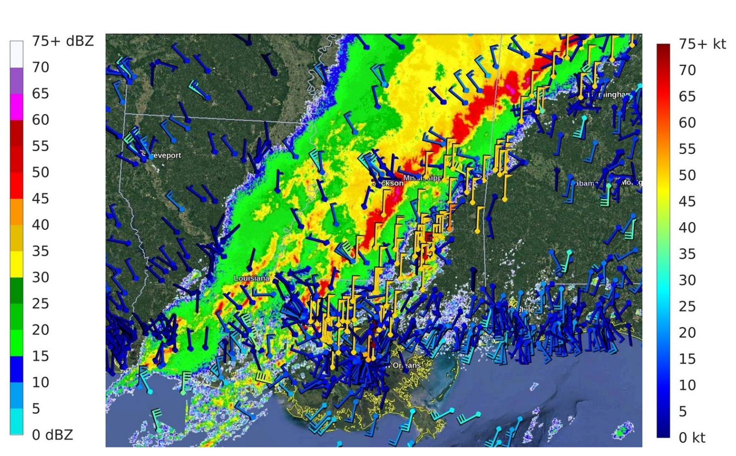

Figure 1: Wind observations and radar scan. Colour fill = composite reflectivity at t_0. Wind barbs = max gust at each location in 90 minutes after t_0. NWS reports have no direction, so direction is assumed north when plotting.

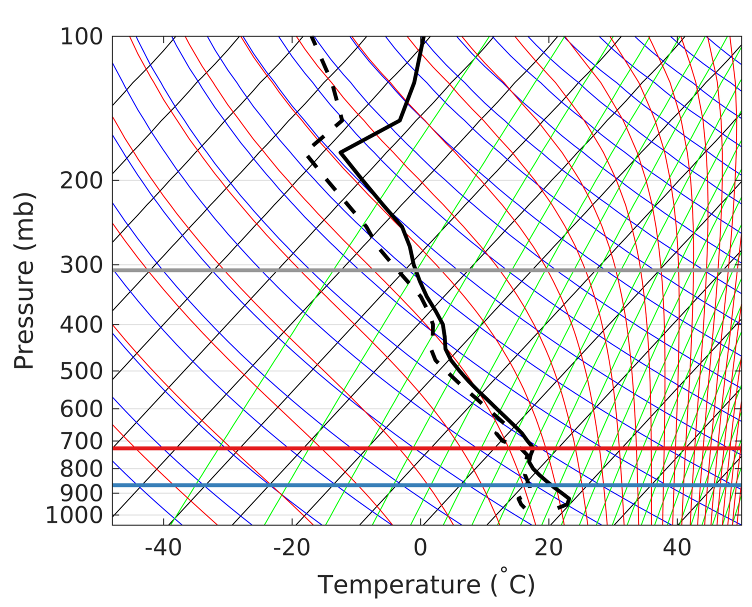

Figure 2: RUC sounding interpolated to center of storm object. Lifting condensation level (LCL) in blue; level of free convection (LFC) in red; equilibrium level (EL) in grey.

3. Data-processing

1. **Storm detection and tracking**
   - Storm detection (outlining of storm objects* in radar image) is done by `w2segmotionll` (Lakshmanan and Smith 2010).
   - Threshold of 30 dBZ for -10 °C reflectivity, 50 km² for storm area.
   - ❖ Storm object = one storm cell at one time step (Figure 3).

   - Prelim storm-tracking is done by `w2segmotionll`.
   - Final storm-tracking is done by `w2besttrack` (Lakshmanan *et al.* 2015), which improves results from `w2segmotionll` (Figure 4).

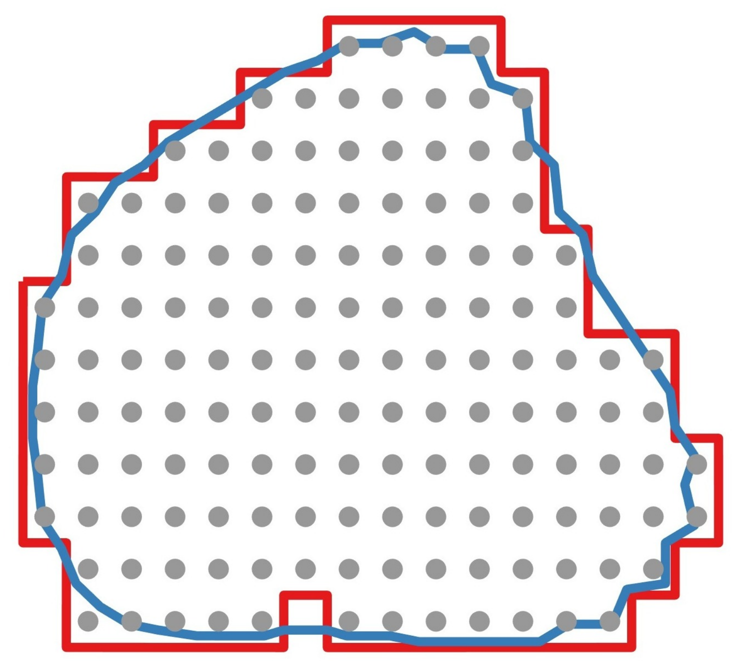

Figure 3: Storm object. Raw version (from `w2segmotionll`) is in red; smoothed version is in blue. Grey dots are radar pixels inside the raw storm object, used to calculate spatial statistics.

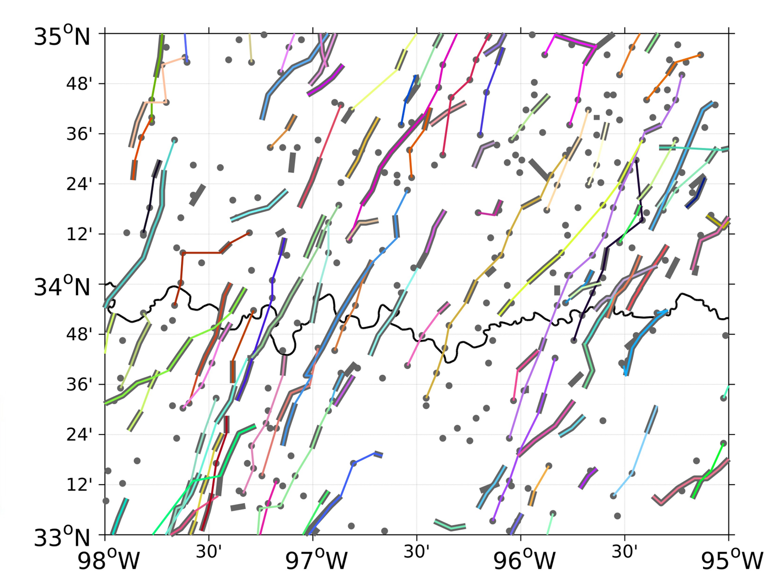

Figure 4: Storm tracks for a 24-hour period. Thick grey lines are from `w2segmotionll`; thin multi-coloured lines are from `w2besttrack`.

2. **Linking wind observations to storms**
   - Each wind observation is linked to the nearest storm cell (Figure 5).
   - Edge of storm cell must pass within 10 km.

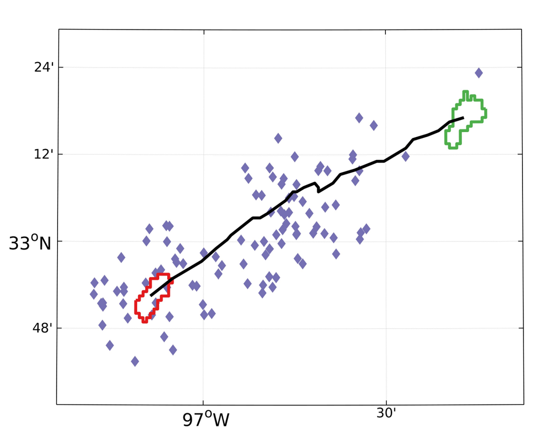

Figure 5: Linking wind observations to a storm cell. The green (red) polygon is the first (last) storm object in the track. The purple diamonds are wind observations linked to the storm (passing within 10 km of edge).

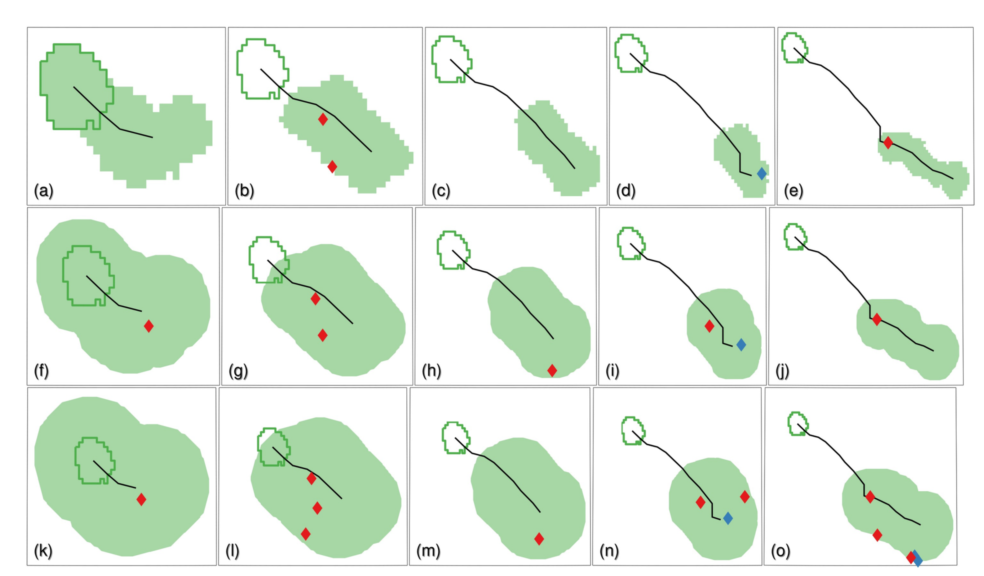

Figure 6: Labeling a storm object. Buffer distance d increases from 0 km (top) to 10 km (bottom). Lead time $[\Delta t_{min}, \Delta t_{max}]$ increases from 0-15 minutes (left) to 60-90 minutes (right). In each panel, green polygon is storm object to label (time t_0). Light green fill is area covered by distance d around storm objects in same track from times $[\Delta t_{min}, \Delta t_{max}]$ after t_0. Red and blue diamonds are wind obs in buffered area.

3. **Calculation of predictors**
   - Four types of predictors for each storm object.

   a) **Radar statistics**
      - Compute 11 spatial stats for each of 12 radar variables.
      - Based only on values inside storm object (Figure 3).
      - Spatial stats = mean, stdev, skewness, kurtosis, 7 percentiles.
      - Radar variables = composite reflectivity, VIL, MESH, etc.

   b) **Storm motion (speed and direction)**

   c) **Shape parameters**
      - Orientation, eccentricity, area, etc. of storm object.
      - Based on storm outline (Figure 3).

   d) **Sounding indices**
      - RUC sounding interpolated to center of storm object.
      - NARR sounding used if RUC data are unavailable.
      - 97 indices computed with SHARPpy software (Halbert *et al.* 2015).

3. Data-processing (continued)

4. **Creation of labels**
   - Each storm object is labeled for 3 buffer distances (0, 5, 10 km) and 5 lead times (0-15, 15-30, 30-45, 45-60, 60-90 minutes).
   - Label = 1 if there is a wind gust in buffered area, 0 otherwise (Figure 6).

4. Machine Learning

- Base model is trained to predict probability of severe wind (prob of label = 1).
- We found best of 5 base models for each buffer distance and lead time:
  - Logistic regression
  - Logistic regression with elastic net
  - Neural network
  - Random forest
  - Ensemble of gradient-boosted trees

- Training/validation/testing sets independent (separated by 24 hours).
- Best model was usually (12 of 15 times) a random forest or gradient-boosted trees.

5. Results

- Figures 7-8 show best model for each buffer distance at 15—30-minute lead time.
- Figures 9-10 are analogues for 60—90-minute lead time.

- Conclusions from Figures 7-10:
  - Performance (area under ROC curve [AUC], maximum critical success index [CSI], and Brier skill score [BSS]) drops with buffer distance and lead time, as expected.
  - AUC > 0.9 for all buffer distances and lead times except the longest (10 km and 60-90 minutes).
  - AUC > 0.9 is "excellent" (Luna-Herrera *et al.* 2003; Muller *et al.* 2005; Mehdi *et al.* 2011).
  - Vast majority of each reliability curve in positive-skill area (beats climatology).
  - Max CSI occurs for unbiased model (frequency bias = 1.0).
  - Forecast histogram has secondary peak at right for 15—30-minute lead time (can still forecast high probs for rare event).

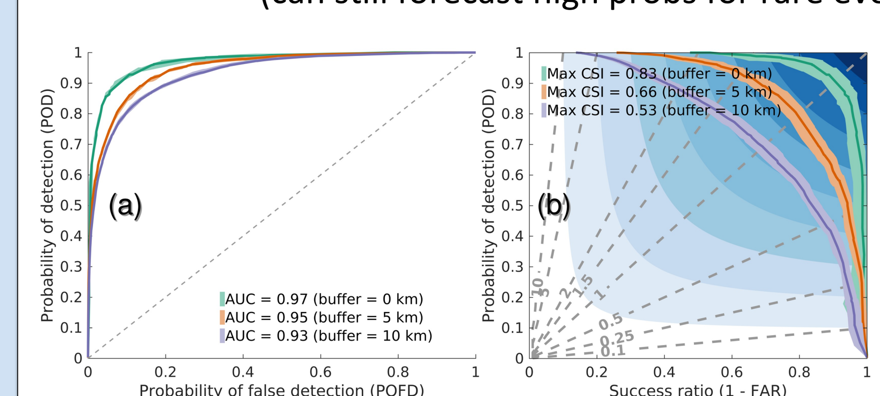

Figure 7: (a) ROC curves and (b) performance diagrams for best models at 15—30-minute lead time. Solid line (shading) is bootstrap mean (95% CI). Dashed grey line in (a) is ROC curve for a random model. Dashed grey lines in (b) are frequency bias; blue fill is critical success index (CSI). See Metz 1978 on ROC curves, Roebber 2009 on performance diagrams.

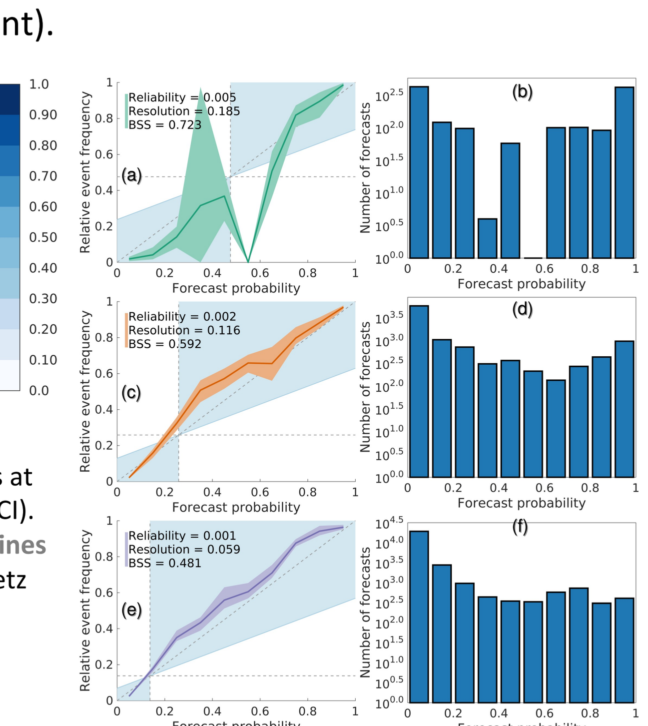

Figure 8: (a) attributes diagram and (b) forecast histogram for 0-km buffer; (c) and (d) for 5-km buffer; (e) and (f) for 10-km buffer. Lead time for all is 15-30 minutes. Blue shading is positive-BSS area (better than climatology). See Hsu and Murphy 1986 for more on attributes diagram.

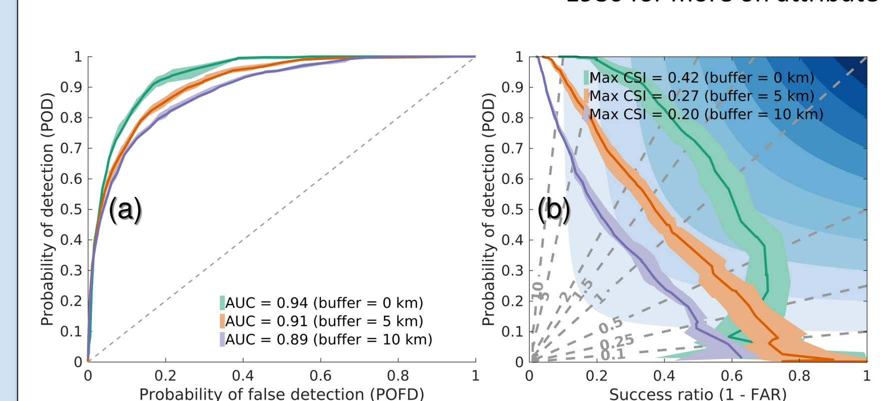

Figure 9: Analogous to Figure 7, but for 60—90-minute lead time.

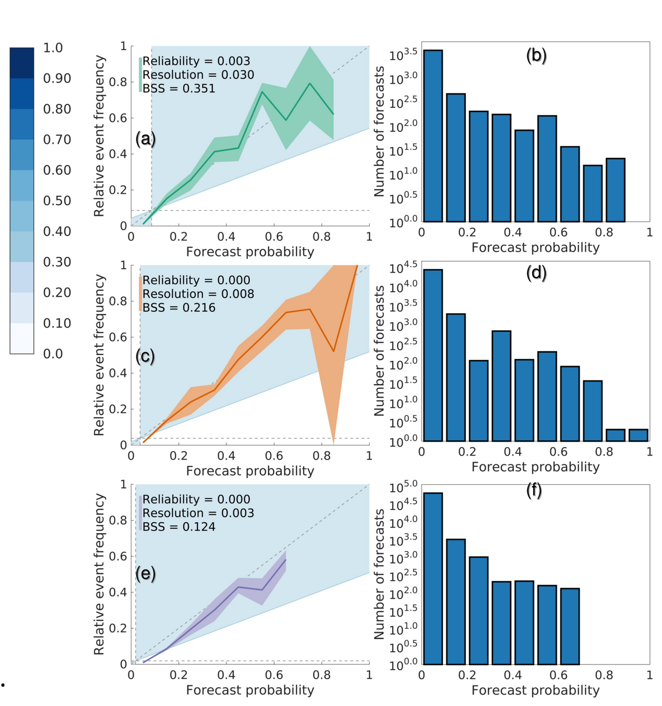

Figure 10: Analogous to Figure 8, but for 60—90-minute lead time.

6. Future Work

- Publish paper (submitting to *Weather and Forecasting* in the next few weeks).
- Interpolate storm-cell-wise probabilities to a grid (easier interpretation for forecasters).
- Use variable-ranking methods to gain insight into phys relations being exploited by ML.
- More detailed predictions (*e.g.,* real value of max wind; prob of $\geq$ 30 kt for aviation).

- **Funded by NOAA/Office of Oceanic and Atmospheric Research under NOAA – OU Cooperative Agreement #NA11OAR4320072, U.S. Department of Commerce.**