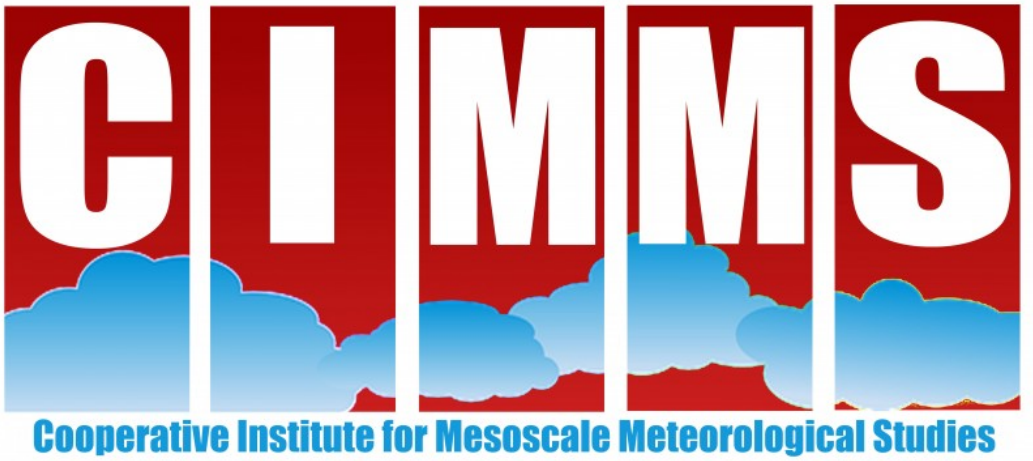
Variable Importance for Prediction of Damaging Straight-line Thunderstorm Winds

Ryan Lagerquist

Master's work at the University of Oklahoma

Committee members: Amy McGovern, Travis Smith, Michael Richman



Variable Importance for Damaging Straight-line Storm Winds

Section 1: Introduction

- In an average year in the U.S., thunderstorms cause over 100 deaths and $10 billion of damage[1].
- Most of this damage is caused by straight-line winds[2].
- Near-surface winds are non-linearly related to many environmental factors.
- Difficult for human forecasters and physical models to predict.
- Recent studies have successfully used machine learning to predict other thunderstorm-related hazards:
  - Tornadoes[3, 4, 5]
  - Hail[4, 6]
  - Lightning[7]
  - Aviation turbulence[5, 8]

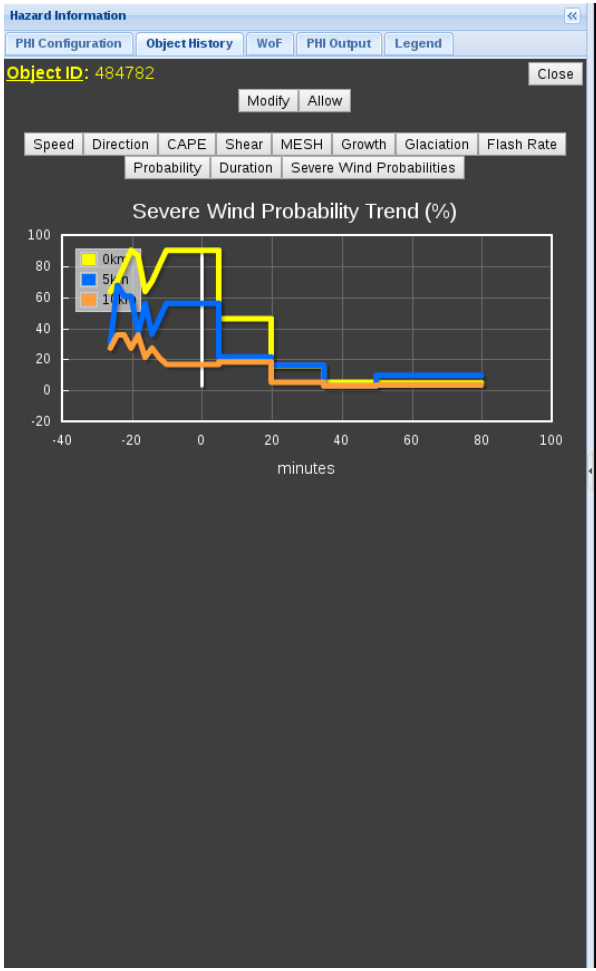
Variable Importance for Damaging Straight-line Storm Winds

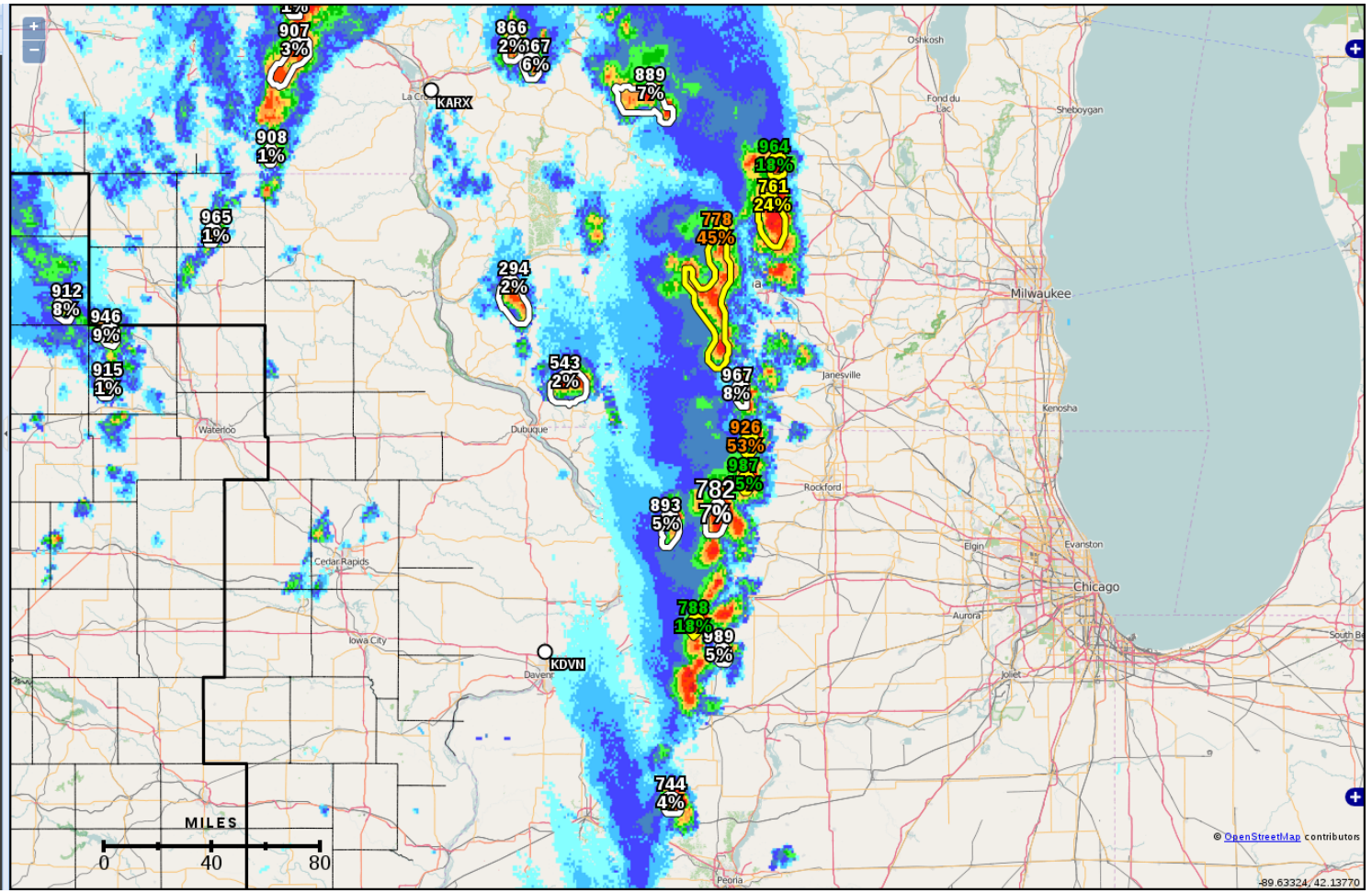
Section 1: Introduction

- Goal: develop a machine-learning system to predict damaging straight-line winds in real time.
- Research to operations.
- For each storm cell, we predict damaging straight-line winds at 3 buffer distances and 5 lead times:
  - 0, 5, 10 km
  - 0-15, 15-30, 30-45, 45-60, 60-90 minutes
- Our system ran during the Spring 2016 Hazardous Weather Testbed (May 2 – June 3) at the Storm Prediction Center in Norman OK.

Variable Importance for Damaging Straight-line Storm Winds

Section 1: Introduction





Variable Importance for Damaging Straight-line Storm Winds

Section 2: Input Data

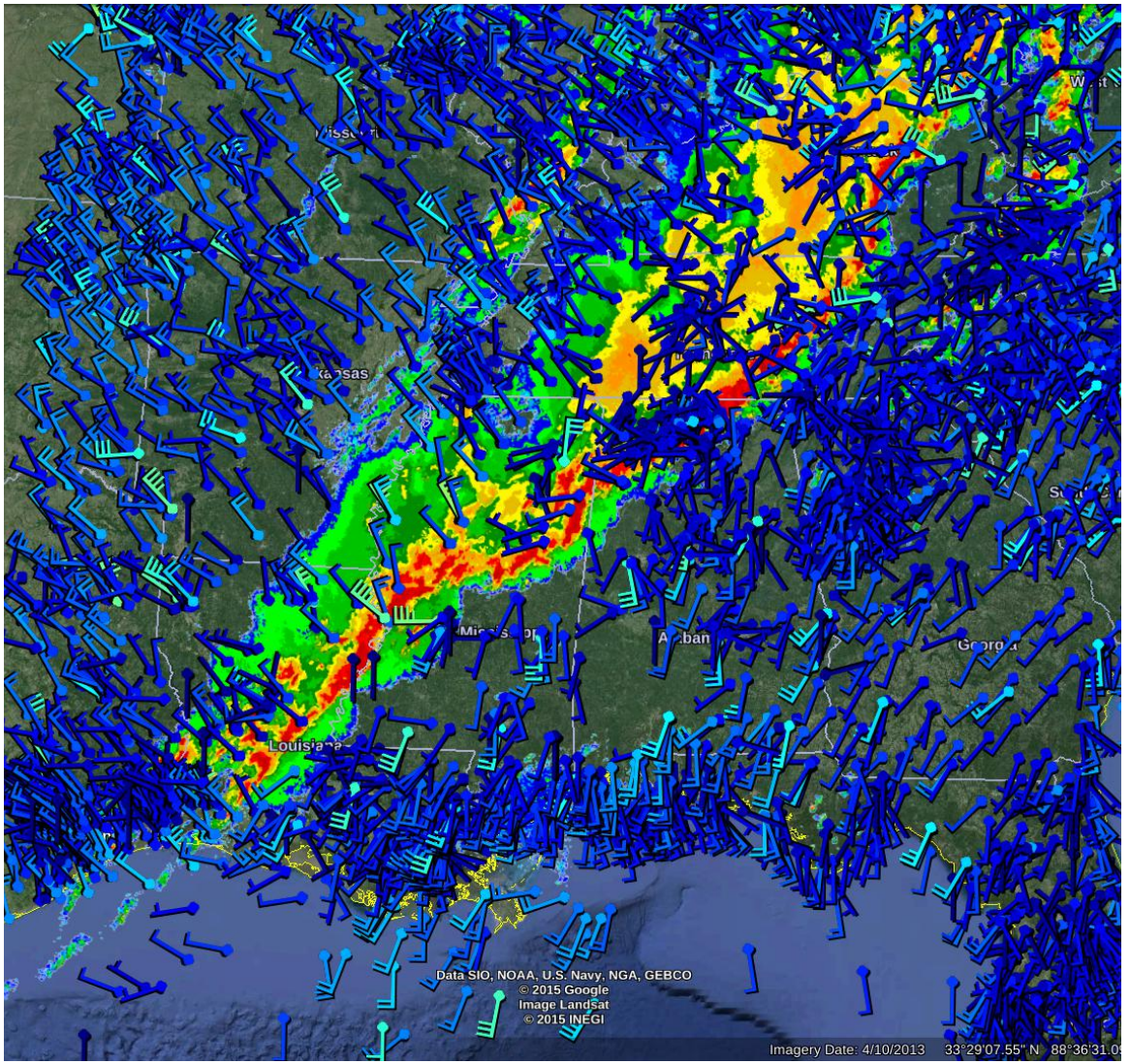
Data Type	Sources	Resolution (if applicable)	Time Period
Radar grids	Multi-year Reanalysis of Remotely Sensed Storms (MYRORSS[9])	1 km / 5 minutes	2000-11 (no 2009)
Model soundings	Rapid Update Cycle (RUC[10])	13 or 20 km / 1 hour	Apr 1994 – Apr 2012
Surface wind observations	Meteorological Assimilation Data Ingest System (MADIS[11]) Oklahoma Mesonet[12] One-minute METARs[13] National Weather Service local storm reports[14]		Jul 2001 – present 1994-present 2000-present 1955-present

- Radar grids and model soundings were used to create predictors for the “event” (storm cell produces winds ≥ 50 kt).
- Surface wind obs were used as “verification data,” to determine when and where the event occurred.

Variable Importance for Damaging Straight-line Storm Winds

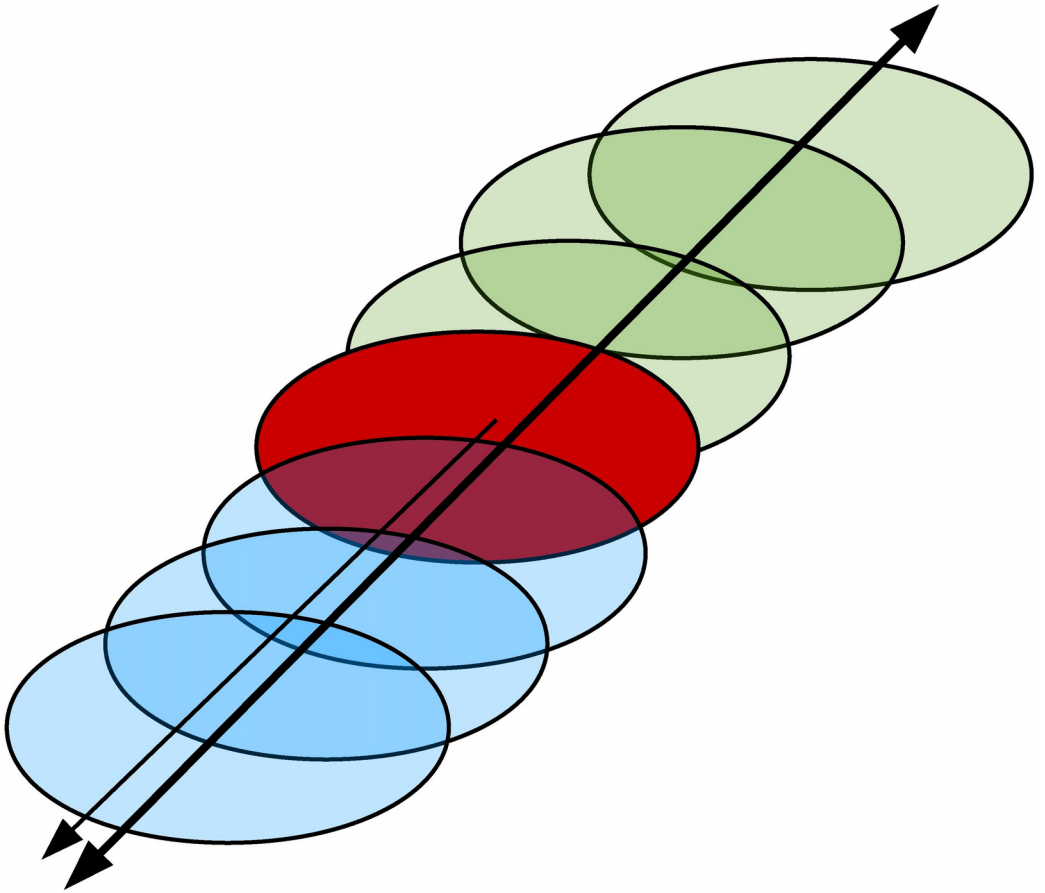
Section 2: Input Data

- Radar grids and model soundings were used to create predictors.
- Surface wind obs were used as verification data, to determine when and where the “event” occurred.
- “Event” = storm cell produces wind ≥ 50 kt.
- We used 804 training days (those with complete data and ≥ 30 severe-wind reports).
- For these 804 days we used all storm cells in the CONUS.



Variable Importance for Damaging Straight-line Storm Winds
Section 3: Storm Identification and Tracking

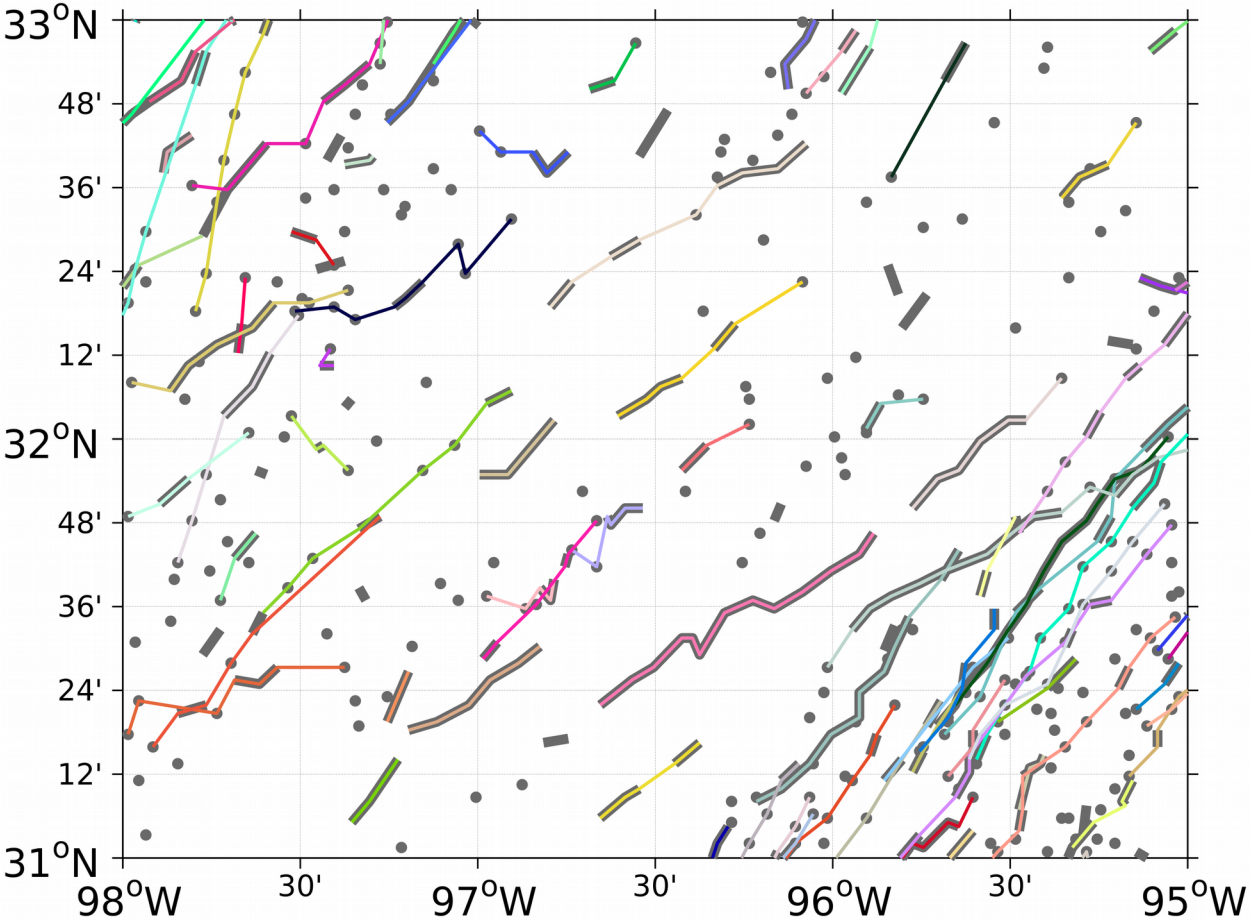
- Identification and real-time tracking is done by w2segmotionll[15] and w2besttrack[16].
- Identification is based on -10 °C reflectivity field, with a minimum storm-cell size of 50 km^2.
- To connect storm objects, real-time tracking algorithms (like w2segmotionll) may look only into the past.
- However, post-event tracking algorithms (like w2besttrack) may look into both the past and future.
- w2besttrack "corrects" the output from w2segmotionll, resulting in longer storm tracks.



Thin line = w2segmotionll
Thick line = w2besttrack
Red = storm object at time t_0
Blue = storm objects before t_0
Green = storm objects after t_0

Technical Note

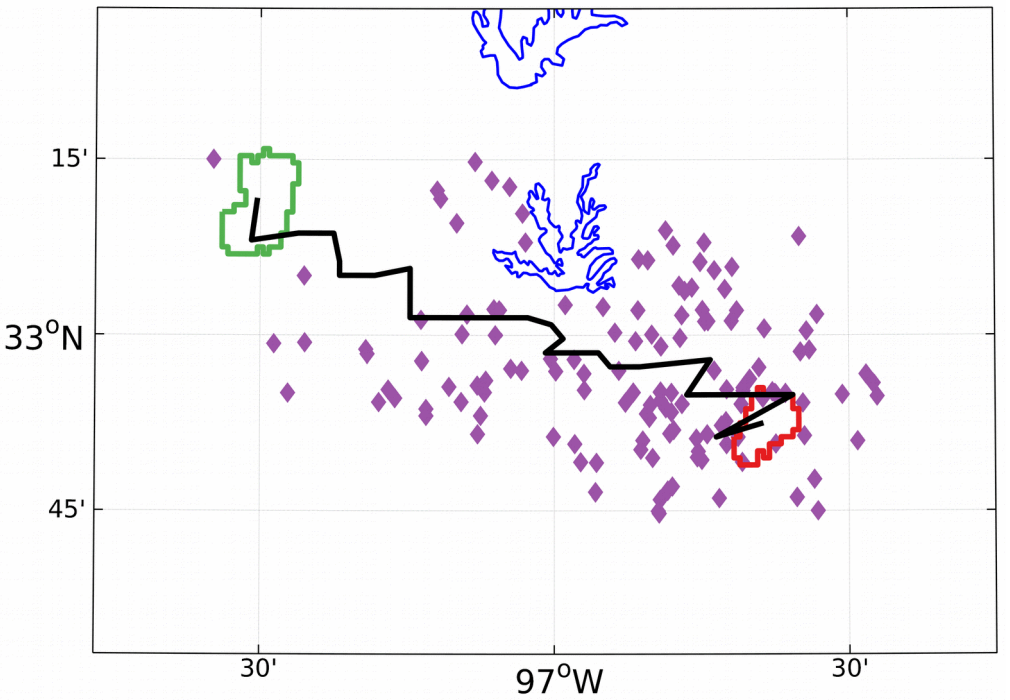
- “Storm cell” = a single thunderstorm, or a single updraft-downdraft pair, consistent with common usage.
- “Storm object” = one storm cell at one time step.
- “Storm track” = succession of storm objects linked in time (presumably all snapshots of the same storm cell).



Thick grey lines = w2segmotionll
Thin multi-colour lines = w2besttrack
For a 24-hour period.

Variable Importance for Damaging Straight-line Storm Winds
Section 4: Linkage Wind Observations to Storm Cells

- Causal attribution (which storm cells are responsible for which wind observations?).
- Procedure is described below, for a single wind observation *W*.
    1. If *W* does not occur at a radar-scan time, interpolate storm objects (along their respective tracks) to same time as *W*.
    2. Find storm object *(S)* with nearest polygon edge. Let the parent track be S^*.
    3. If *S* is within 10 km, link *W* to all objects in track S^*. Otherwise, do not link *W* at all.
- This is where tracking becomes important.
- Allows us to say that, *e.g.*, storm cell at 1830 UTC eventually produced severe winds at 2000 UTC.



Black line = storm track
Green polygon = first storm object
Red polygon = last storm object
Purple diamonds = wind obs linked to track

Variable Importance for Damaging Straight-line Storm Winds
Section 5: Calculation of Predictors

- 4 types of predictors for each storm object:
  - Radar statistics
  - Storm motion (speed and direction)
  - Shape parameters (describe outline of storm object)
  - Sounding parameters (from interpolated RUC soundings)

Variable Importance for Damaging Straight-line Storm Winds
Section 5: Calculation of Predictors

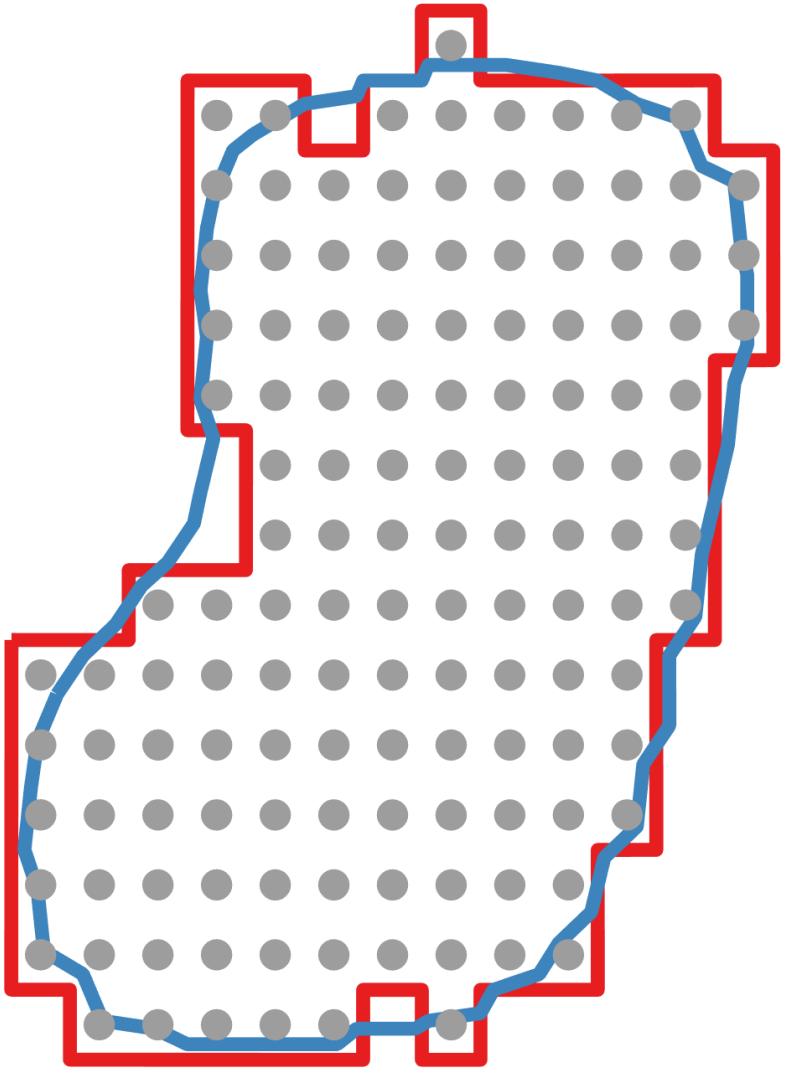
- 11 stats are calculated for 12 radar variables, using only values inside the storm object.
- This is done for both raw values and gradients.
- This leads to 264 (11 x 12 x 2) predictors.

Radar Variables	Stats Calculated for Each
Low-level (0—2-km) azimuthal shear	0th percentile (minimum)
Mid-level (3—6-km) azimuthal shear	5th percentile
18-dBZ echo top	25th percentile
50-dBZ echo top	50th percentile (median)
Max estimated hail size	75th percentile
-20 °C reflectivity	95th percentile
-10 °C reflectivity	100th percentile (max)
0 °C reflectivity	Mean (1st moment)
Composite reflectivity	Standard deviation (~ 2nd moment)
Lowest-altitude reflectivity	Skewness (~ 3rd moment)
Severe-hail index (SHI)	Kurtosis (~ 4th moment)
Vertically integrated liquid (VIL)	

Variable Importance for Damaging Straight-line Storm Winds

Section 5: Calculation of Predictors

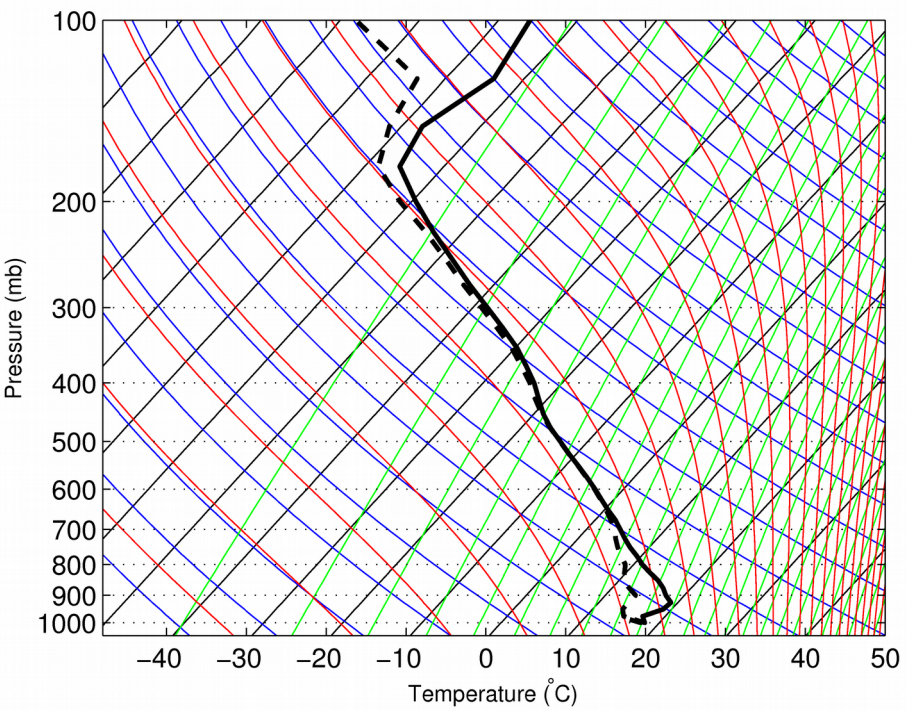
- Shape parameters:
  - Area
  - Bending energy
  - Compactness
  - Curvature
  - Eccentricity
  - Extent
  - Orientation
  - Solidity
- Params involving curvature are calculated on smoothed outline.



Red polygon = raw storm object
Blue polygon = smoothed object
Grey dots = radar pixels in storm object

Variable Importance for Damaging Straight-line Storm Winds
Section 5: Calculation of Predictors

- 97 sounding parameters (see Appendix C) were calculated with SHARPpy[17].
- Types of sounding parameters:
  - Wind shear over various layers
  - Mean wind over various layers
  - Mean storm-relative wind over various layers
  - Moisture variables (precipitable water, mean RH over various layers, etc.)
  - Heights of critical levels (LCL, LFC, equilibrium level, etc.)
  - Thermodynamic indices (CAPE, CIN, lapse rates, etc.)
  - Other indices (EHI, microburst index, wind-damage index, etc.)
- 431 predictors overall (after decomposing vectors into magnitude, sin, cos).



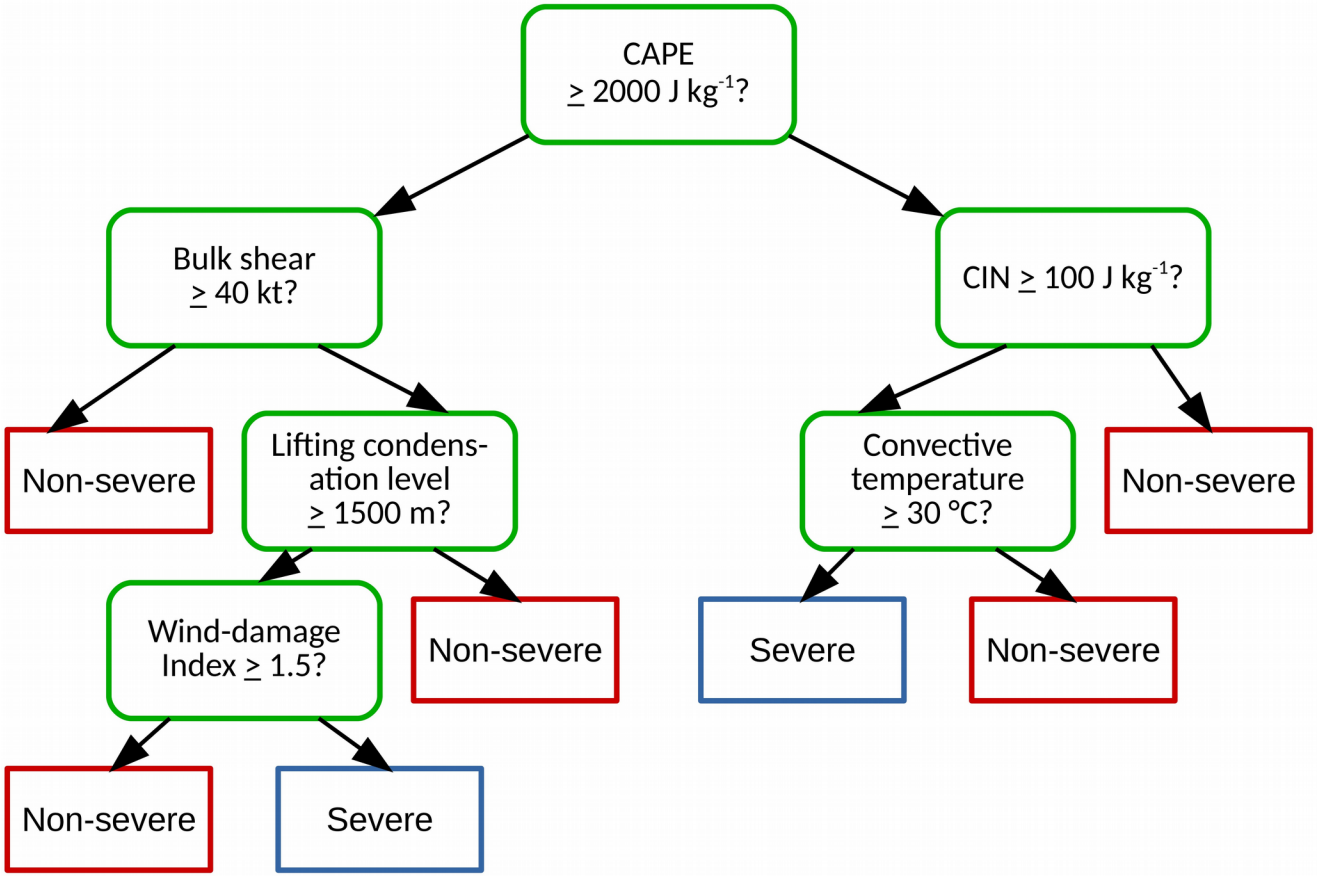
Variable Importance for Damaging Straight-line Storm Winds

Section 6: Calculation of Labels

- Goal is to predict severe winds for each storm cell at 3 buffer distances (0, 5, 10 km) and 5 lead times (0-15, 15-30, 30-45, 45-60, 60-90 minutes).
- Procedure is described below for a single storm object *S*, buffer distance *d*, lead time $[t_{min}, t_{max}]$.
  - Label = 1 if *S* has a wind observation $\geq$ 50 kt within buffer distance *d* and lead time $[t_{min}, t_{max}]$.
  - Label = 0 otherwise.
- Label = predictand for machine learning.
- Thus, machine-learning task is binary classification.

Variable Importance for Damaging Straight-line Storm Winds
Section 7: Machine-learning Methods

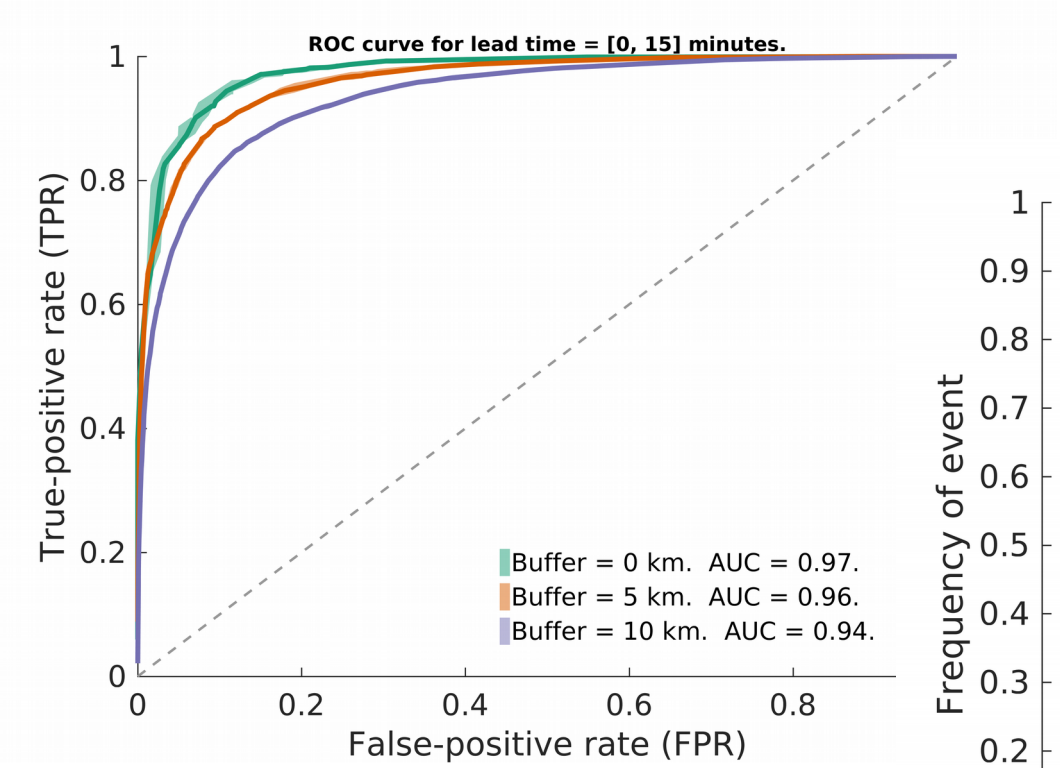
- We used gradient-boosted trees[18] (GBT) as the base model.
- Then used isotonic regression[19] to calibrate GBT probabilities.
- Gradient-boosting is a way of ensembling decision trees, similar to random forest.
- Difference is that GB focuses learning on difficult examples, which usually leads to better performance.

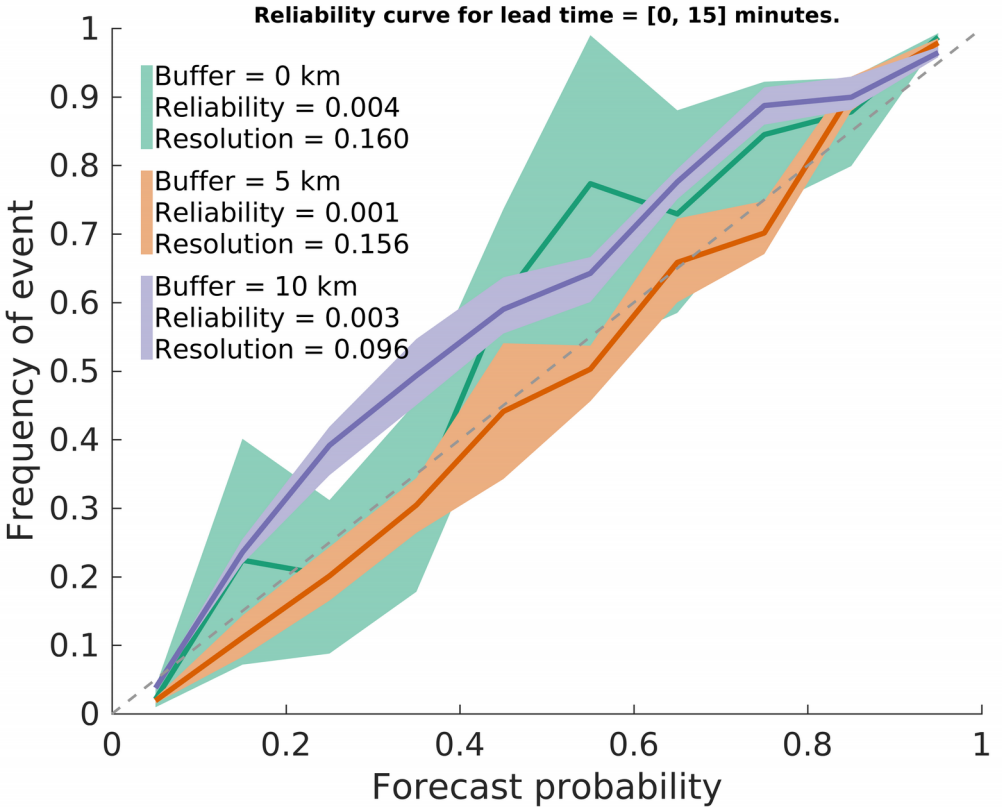


Sample decision tree.
Follow right branch if answer to question is “yes”.

Variable Importance for Damaging Straight-line Storm Winds
Section 7: Machine-learning Methods

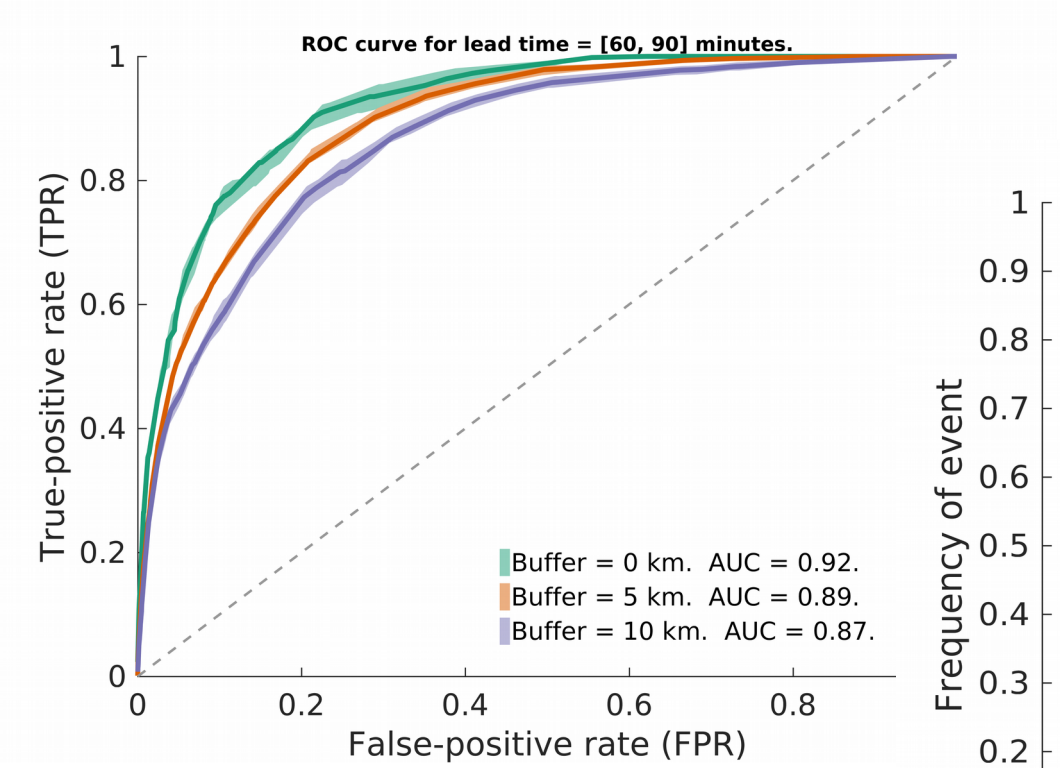
- Receiver operating characteristic (ROC) and reliability curves shown for 0-15 minutes (easiest lead time).
- Areas under curve (AUCs) all > 0.9 (excellent[20, 21, 22]).
- 0-km and 5-km models near perfect reliability ($x = y$); 10-km model underforecasts.

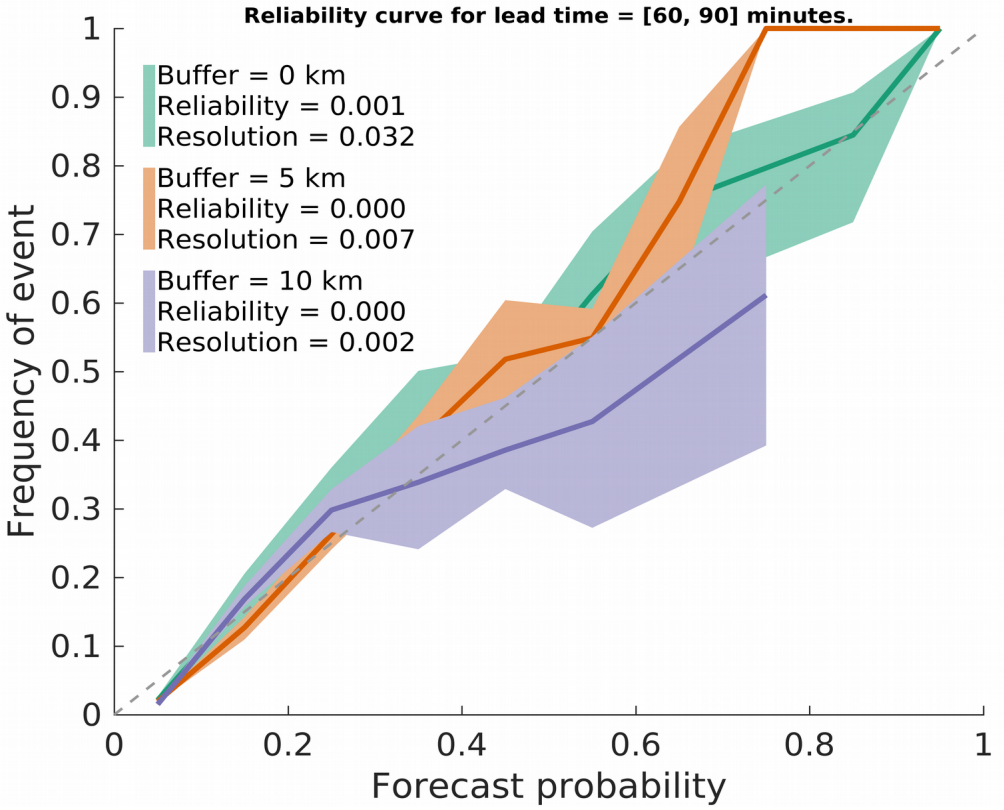




Variable Importance for Damaging Straight-line Storm Winds
Section 7: Machine-learning Methods

- ROC and reliability shown below for 60-90 minutes (hardest lead time).
- AUCs are all > 0.8 (good) or > 0.9 (excellent).
- Near perfect reliability, except 5-km and 10-km at upper end.
- More forecast evaluation in Appendix A.





Variable Importance for Damaging Straight-line Storm Winds
Section 8: Variable Importance

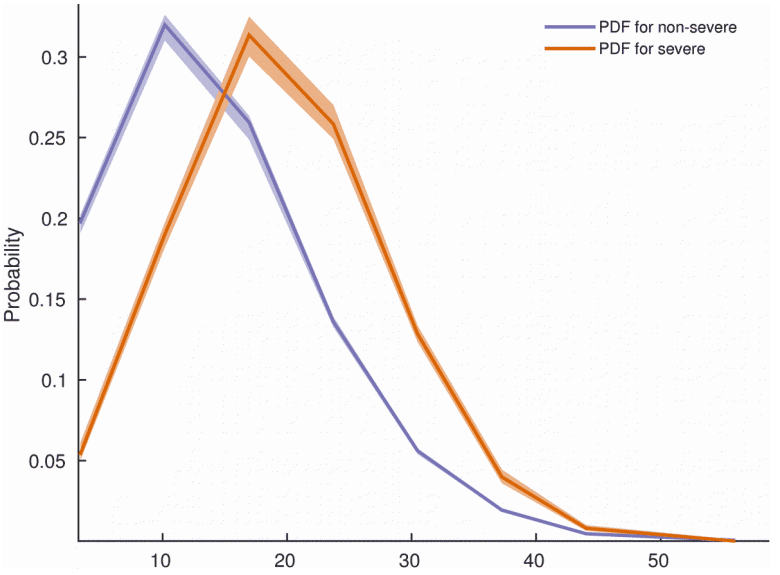
- We used two methods:
  - *J*-measure ranking (JMR[22])
  - Sequential forward selection (SFS[23])
- JMR is a data-driven method (selects variables based on their relationships in the dataset).
- SFS is a model-driven method (selects variables based on how they affect model performance).
- Methods independent of each other.
- Methods independent for each buffer distance and lead time.
- Bootstrapped data 25 times for each buffer distance and lead time.
- We will show results only for 5 km and 30-45 minutes (median values).
- Results were similar for other combinations (see Appendix B).

Variable Importance for Damaging Straight-line Storm Winds
Section 8: Variable Importance

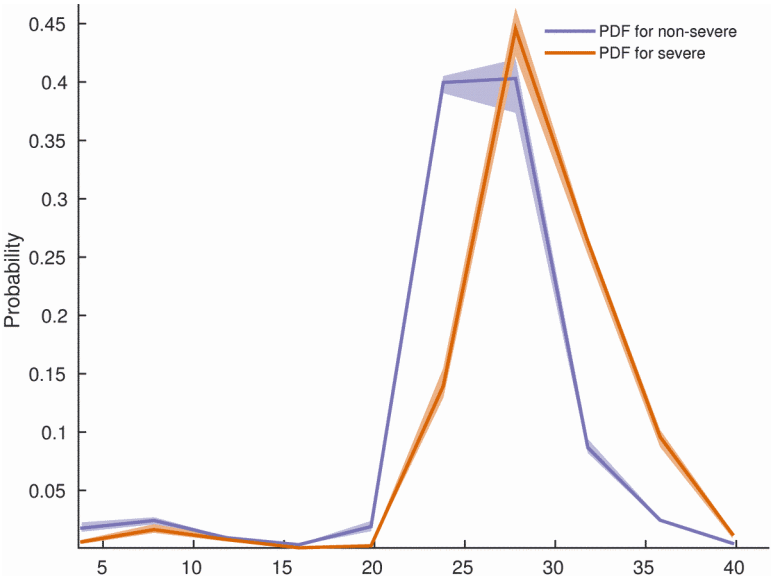
- *J*-measure is basically the difference between two PDFs.
- In our case, difference between PDF of predictor variable when label = 0 and label = 1.

$$J=\sum_{k=1}^{10}[p(x=x_k \mid y=0)-p(x=x_k \mid y=1)]\log_2[\frac{p(x=x_k \mid y=0)}{p(x=x_k \mid y=1)}]$$

- $p(x \mid y=0)$ and $p(x \mid y=1)$ are distributions of predictor variable *x* for non-severe and severe-wind-producing storms, respectively.
- *J* is large when distributions are very different.
- Suggests that *x* is an important predictor.



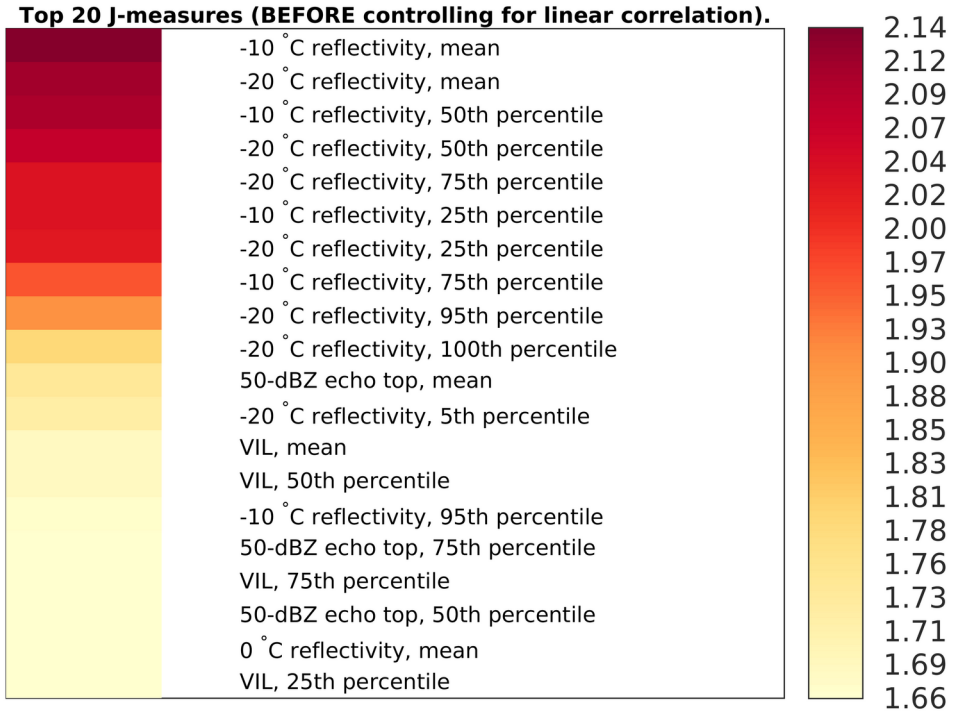
Magnitude of 0—6-km shear (kt)

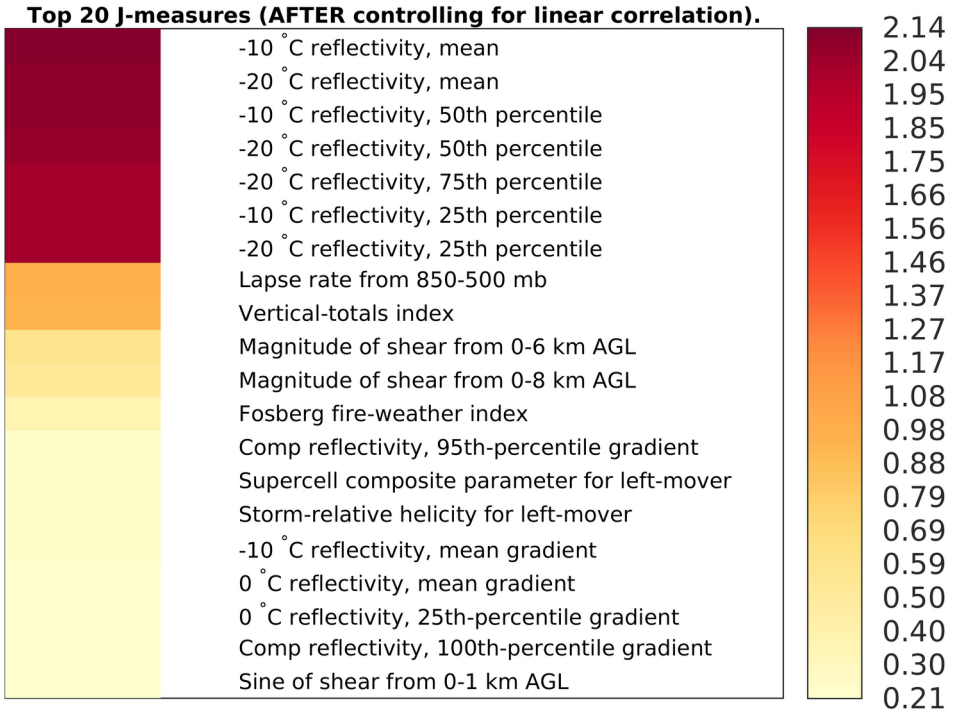


Vertical-totals index

Variable Importance for Damaging Straight-line Storm Winds
Section 8: Variable Importance

- Top 20 of 431 predictors shown below.





- Colour fill is *J*-measure (averaged over 25 bootstrap replicates).
- Even after controlling for linear correlation, top 7 variables are statistics on -10 C° or -20 °C reflectivity.
- Probably an artifact of storm ID (based on -10 °C reflectivity).

Variable Importance for Damaging Straight-line Storm Winds

Section 8: Variable Importance

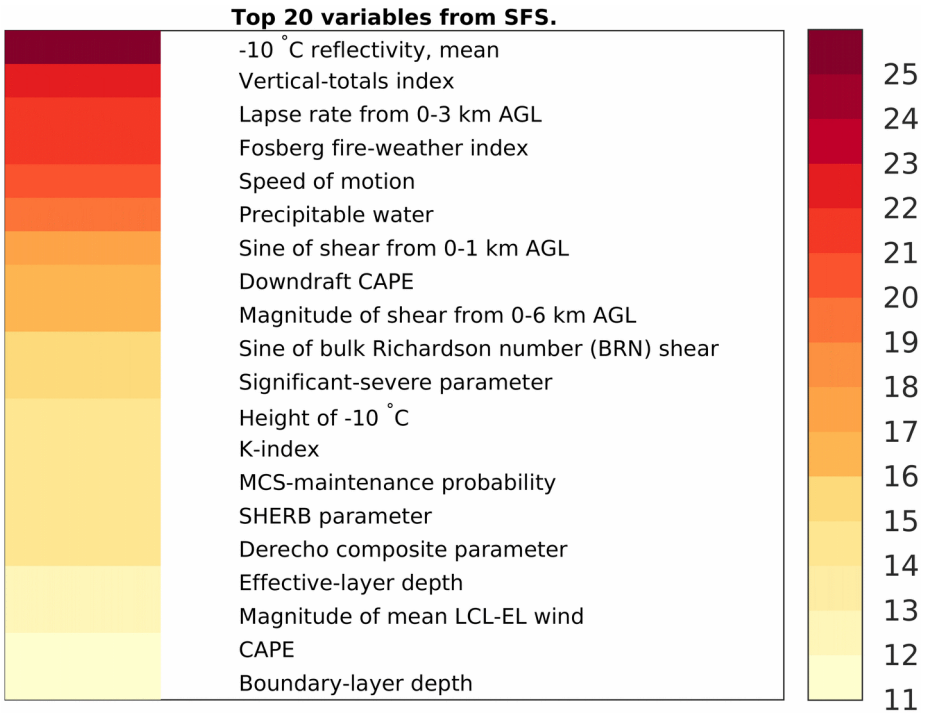
- SFS is a model-driven method, so has underlying machine-learning model.
- We used logistic regression (fast and good for binary classification).
- SFS starts with 0 variables (constant model), then adds one at a time.
- Stops when the last 10 variables have not improved performance (AUC).

Variable Importance for Damaging Straight-line Storm Winds
Section 8: Variable Importance

- Graph below shows how many times (out of 25) each variable was included in the final model.

- Similarities between JMR and SFS:
  - Fosberg fire-weather index
  - Lapse rates
  - Magnitude of deep-layer shear
  - Mean -10 °C reflectivity
  - Sine (related to *v*-component) of 0—1-km shear
  - Vertical-totals index

- Severe-wind probability increases with all of these.



Variable Importance for Damaging Straight-line Storm Winds
Conclusions

- Used machine learning to predict severe winds ($\geq$ 50 kt) for each storm cell.
- Three buffer distances: 0, 5, 10 km around storm cell.
- Five lead times: 0-15, 15-30, 30-45, 45-60, 60-90 minutes.
- Good AUC and reliability achieved for all buffers and lead times.
- *J*-measure ranking and sequential feature selection used to estimate variable importance.
- Methods show strong agreement, except for stats on -10 °C and -20 °C reflectivity (probably artifact of storm ID).
- Variable importance allows straight-line wind threat to be linked to climate models, which do not explicitly resolve convection.

Variable Importance for Damaging Straight-line Storm Winds
References

[1] Insurance Information Institute. “Catastrophes: U.S.” 2015. Webpage. 14 December 2015.

[2] National Severe Storms Laboratory. “Severe Weather 101: Damaging Winds Basics.” n.d. Webpage. 11 August 2015.

[3] Gagne II, David John, *et al.* "Tornadic Supercell Environments Analyzed Using Surface and Reanalysis Data: A Spatiotemporal Relational Data-Mining Approach." *Journal of Applied Meteorology and Climatology* **51.12** (2012): 2203-2217.

[4] Cintineo, John L., *et al.* "An Empirical Model for Assessing the Severe Weather Potential of Developing Convection." *Weather and Forecasting* **29.3** (2014): 639-653.

[5] McGovern, Amy, *et al.* "Enhancing understanding and improving prediction of severe weather through spatiotemporal relational learning." *Machine Learning* **95.1** (2014): 27-50.

[6] Gagne II, David John, *et al.* "Day-Ahead Hail Prediction Integrating Machine Learning with Storm-Scale Numerical Weather Models." *Twenty-Seventh IAAI Conference.* 2015.

[7] Blouin, Karen D., *et al.* "Ensemble lightning prediction models for the province of Alberta, Canada." *International Journal of Wildland Fire* **25.4** (2016): 421-432.

[8] Williams, John K. "Using random forests to diagnose aviation turbulence." *Machine Learning* **95.1** (2014): 51-70.

Variable Importance for Damaging Straight-line Storm Winds
References

[9] Ortega, K. L., *et al.* "The multi-year reanalysis of remotely sensed storms (MYRORSS) project." *26th Conference on Severe Local Storms*, Nashville. American Meteorological Society. Vol. 74. 2012.

[10] Benjamin, Stanley G., *et al.* "An hourly assimilation-forecast cycle: The RUC." *Monthly Weather Review* **132.2** (2004): 495-518.

[11] McNitt, James, Joseph Facundo, and James O'Sullivan. "Meteorological Assimilation Data Ingest System Transition Project Risk Reduction Activity." *24th Conference on IIPS.* 2008.

[12] Brock, Fred V., *et al.* "The Oklahoma Mesonet: a technical overview." *Journal of Atmospheric and Oceanic Technology* **12.1** (1995): 5-19.

[13] National Climatic Data Center. "Data Documentation for Data Set 6406 (DSI-6406): ASOS Surface 1-Minute, Page 2 Data." 12 July 2006. Webpage. 14 December 2015.

[14] Storm Prediction Center. "Storm Prediction Center Severe Weather Summaries." 2016. Webpage. 13 May 2016.

[15] Lakshmanan, Valliappa, and Travis Smith. "Evaluating a Storm Tracking Algorithm." *26th Conference on IIPS.* 2010.

[16] Lakshmanan, Valliappa, Benjamin Herzog, and Darrel Kingfield. "A Method for Extracting Postevent Storm Tracks." *Journal of Applied Meteorology and Climatology* **54.2** (2015): 451-462.

Variable Importance for Damaging Straight-line Storm Winds
References

[17] Halbert, K. T., W. G. Blumberg, and P. T. Marsh, 2015: "SHARPpy: Fueling the Python Cult". Preprints, *5th Symposium on Advances in Modeling and Analysis Using Python*, Phoenix AZ.

[18] Friedman, Jerome H. "Greedy function approximation: a gradient boosting machine." *Annals of Statistics* **29.5** (2001): 1189-1232.

[19] Niculescu-Mizil, Alexandru, and Rich Caruana. "Predicting good probabilities with supervised learning." *Proceedings of the 22nd International Conference on Machine Learning*. ACM, 2005.

[20] Luna-Herrera, J., *et al*. "Use of receiver operating characteristic curves to assess the performance of a microdilution assay for determination of drug susceptibility of clinical isolates of *Mycobacterium tuberculosis*." *European Journal of Clinical Microbiology and Infectious Diseases* **22.1** (2003): 21-27.

[21] Mehdi, Tazhibi. "Kernel smoothing for ROC curve and estimation for thyroid stimulating hormone." *International Journal of Public Health Research* Special Issue (2011): 239-242.

[22] Muller, Matthew P., *et al*. "Can routine laboratory tests discriminate between severe acute respiratory syndrome and other causes of community-acquired pneumonia?" *Clinical Infectious Diseases* **40.8** (2005): 1079-1086.

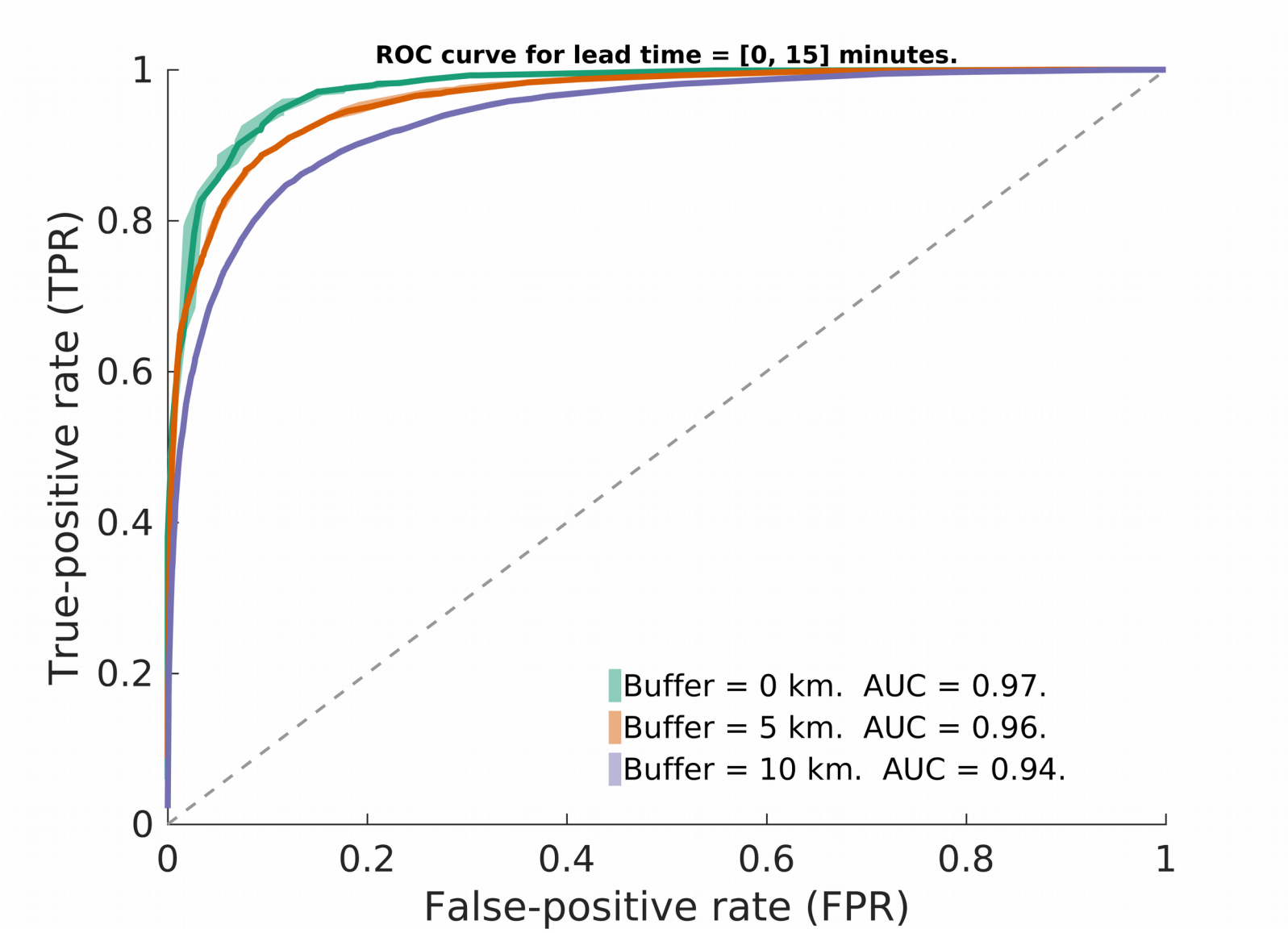
[23] Kohavi, Ron, and George H. John. "Wrappers for feature subset selection." *Artificial Intelligence* **97.1** (1997): 273-324.

Variable Importance for Damaging Straight-line Storm Winds
References

[24] Roebber, Paul J. "Visualizing multiple measures of forecast quality." *Weather and Forecasting* **24.2** (2009): 601-608.

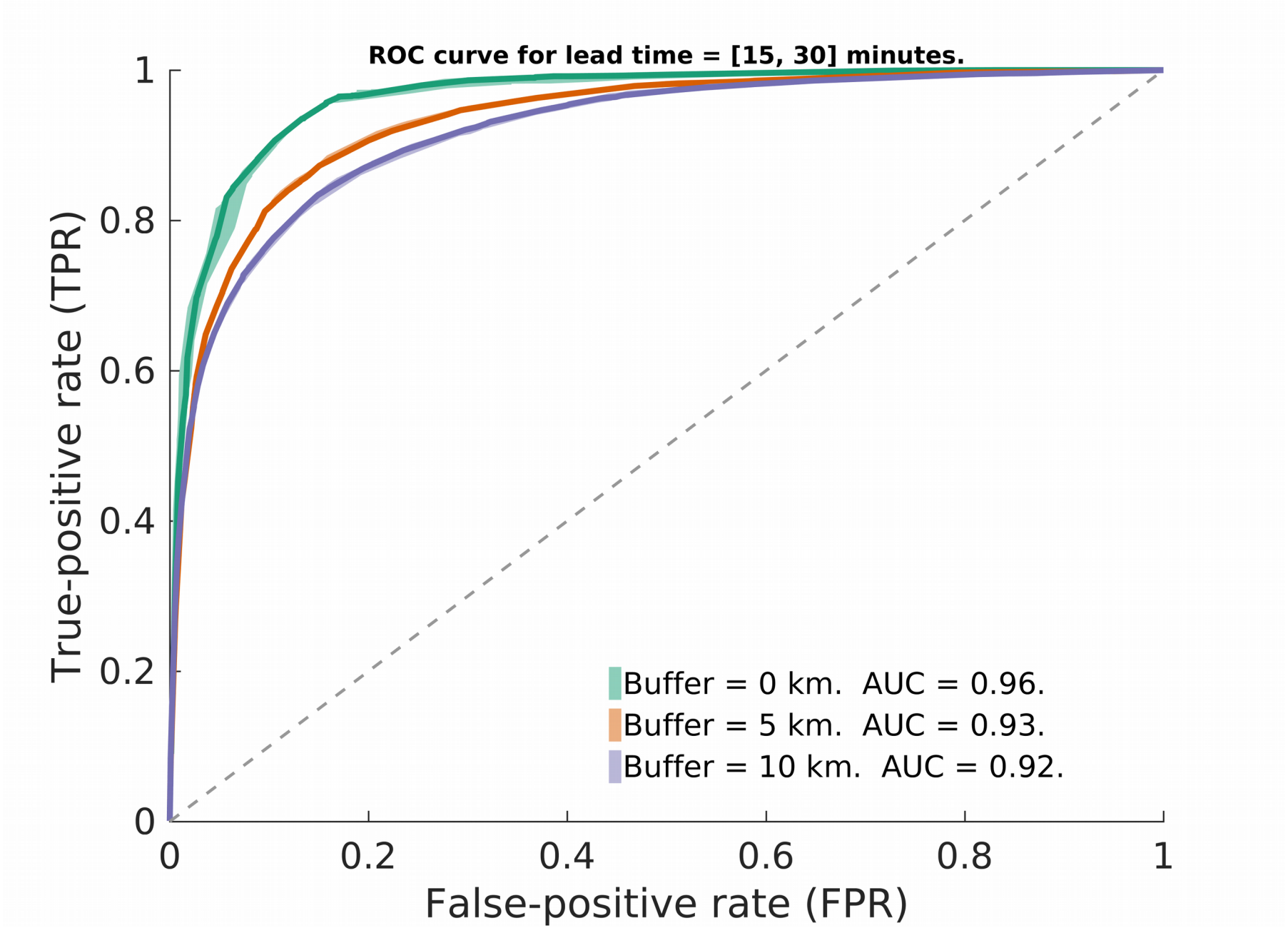
Variable Importance for Damaging Straight-line Storm Winds
Appendix A: Forecast Evaluation

- ROC curves for each lead time are shown on the following pages.



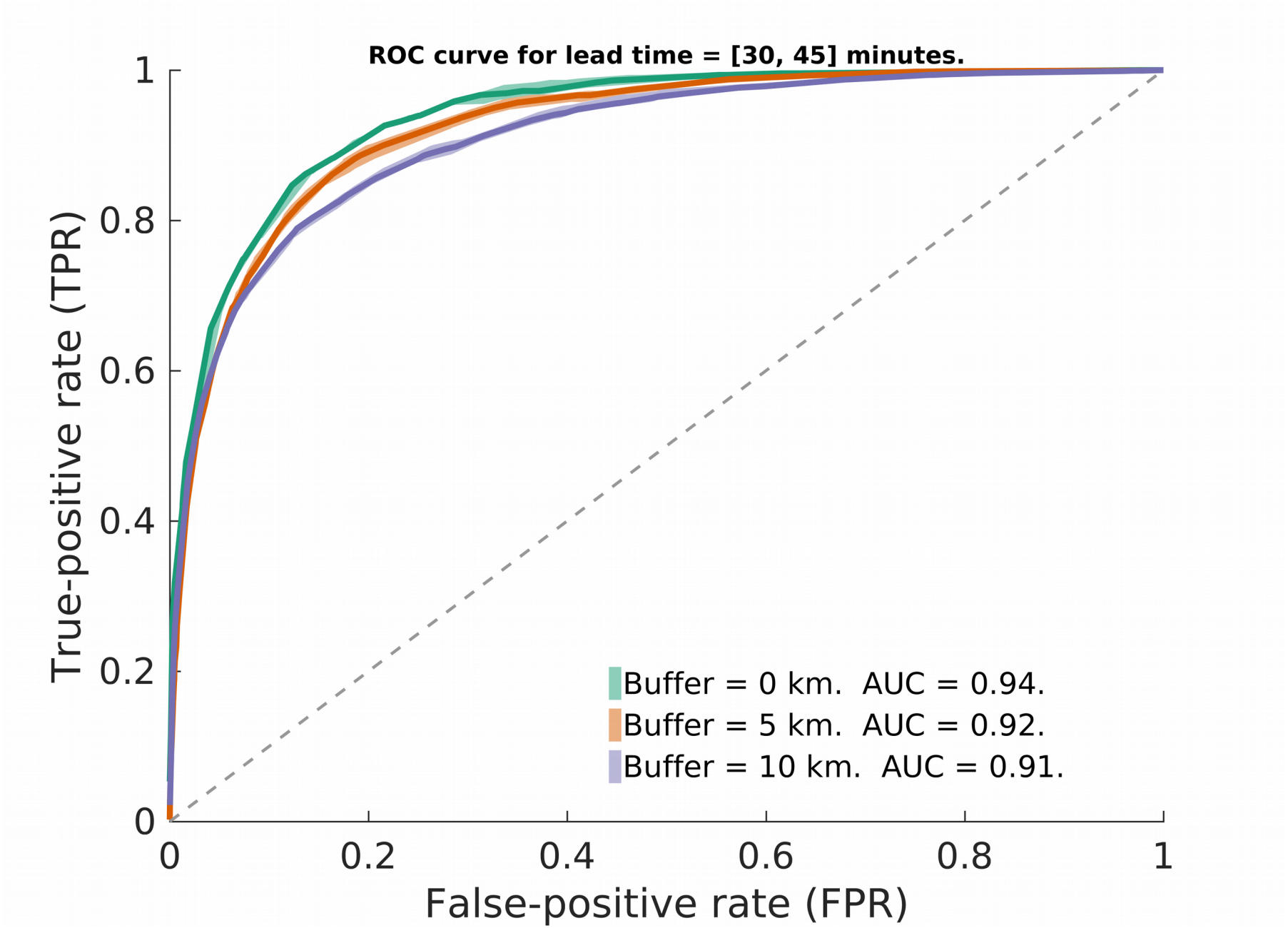
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



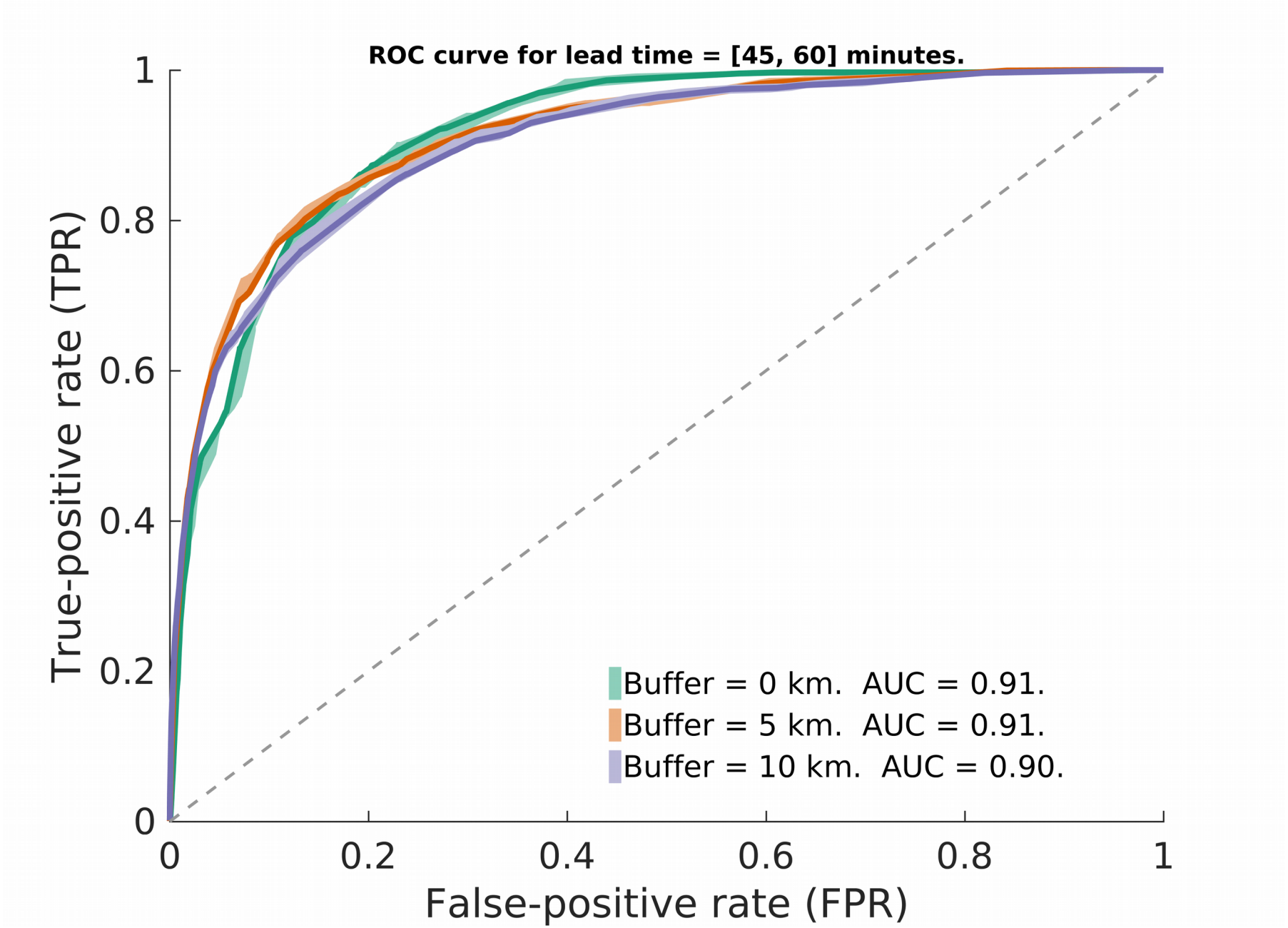
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



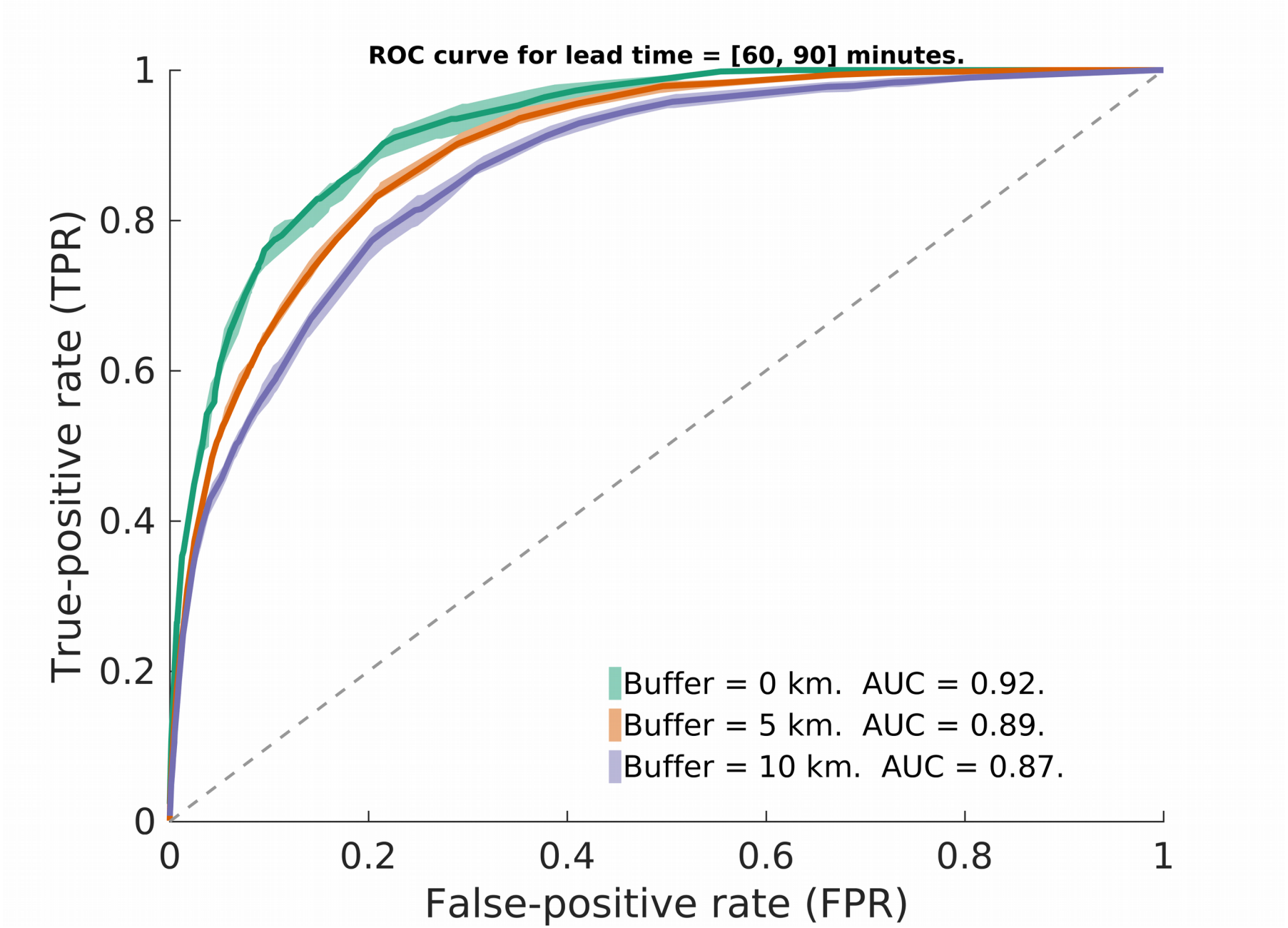
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



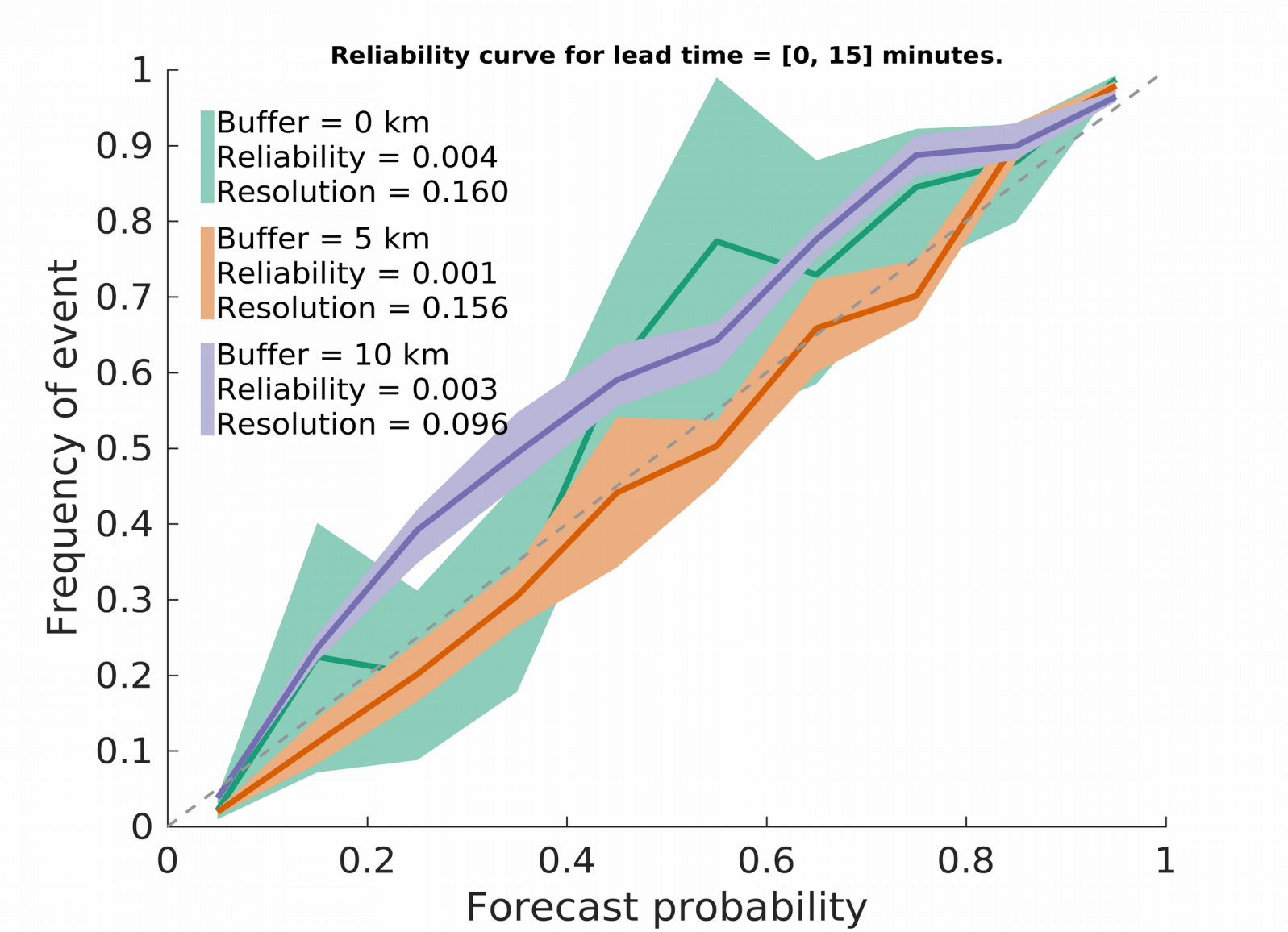
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



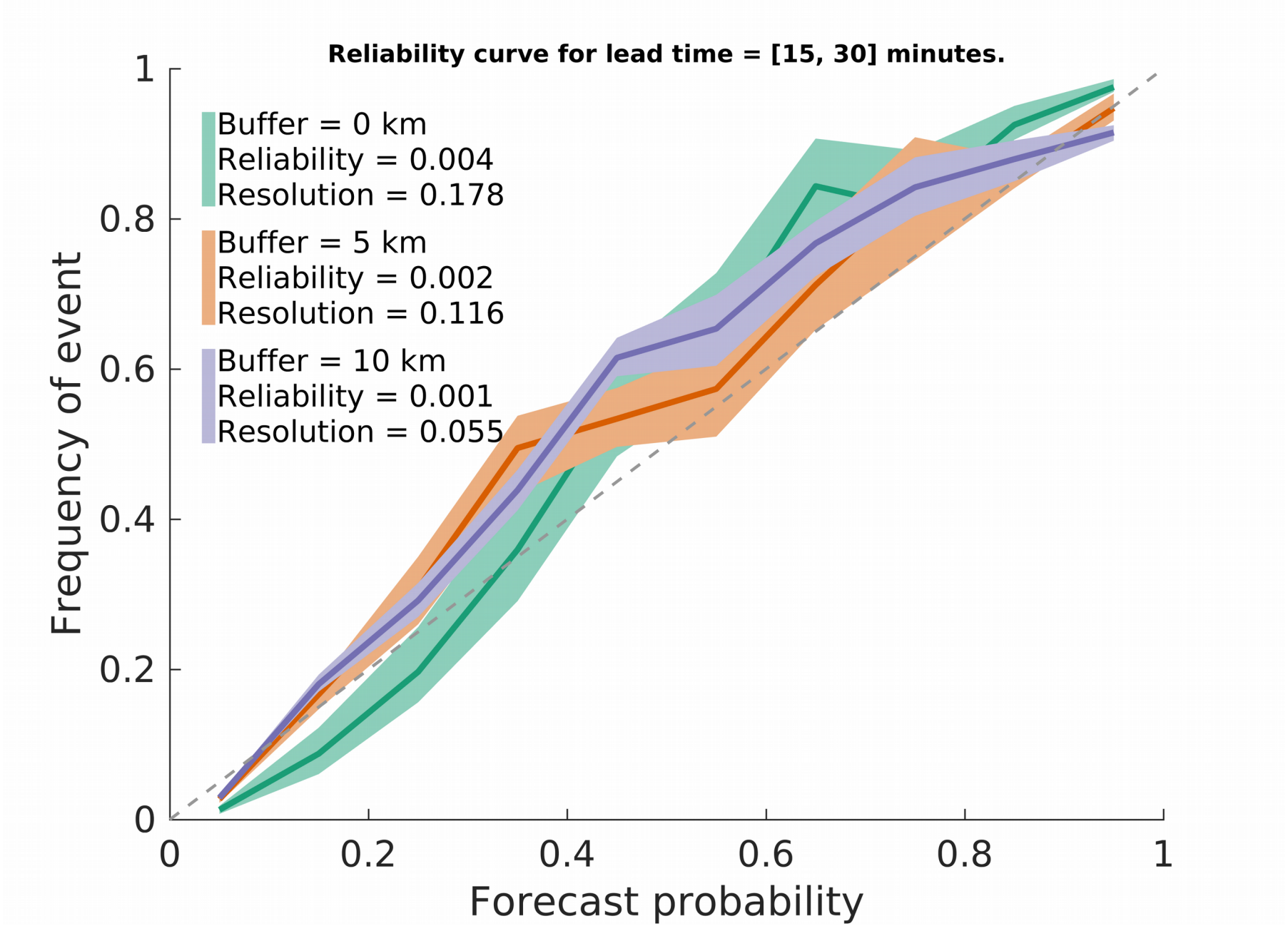
Variable Importance for Damaging Straight-line Storm Winds
Appendix A: Forecast Evaluation

- Reliability curves for each lead time are shown on the following pages.



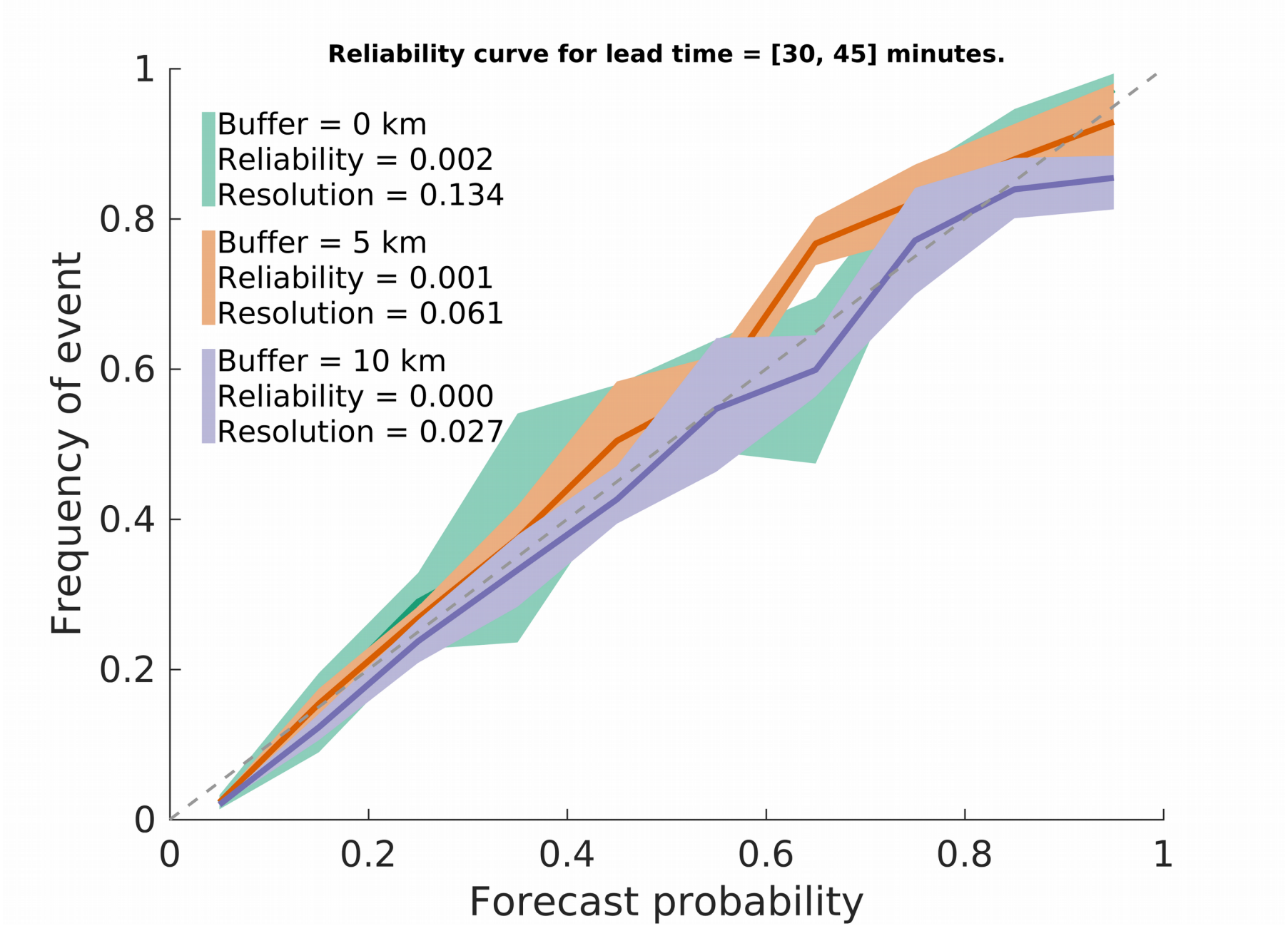
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



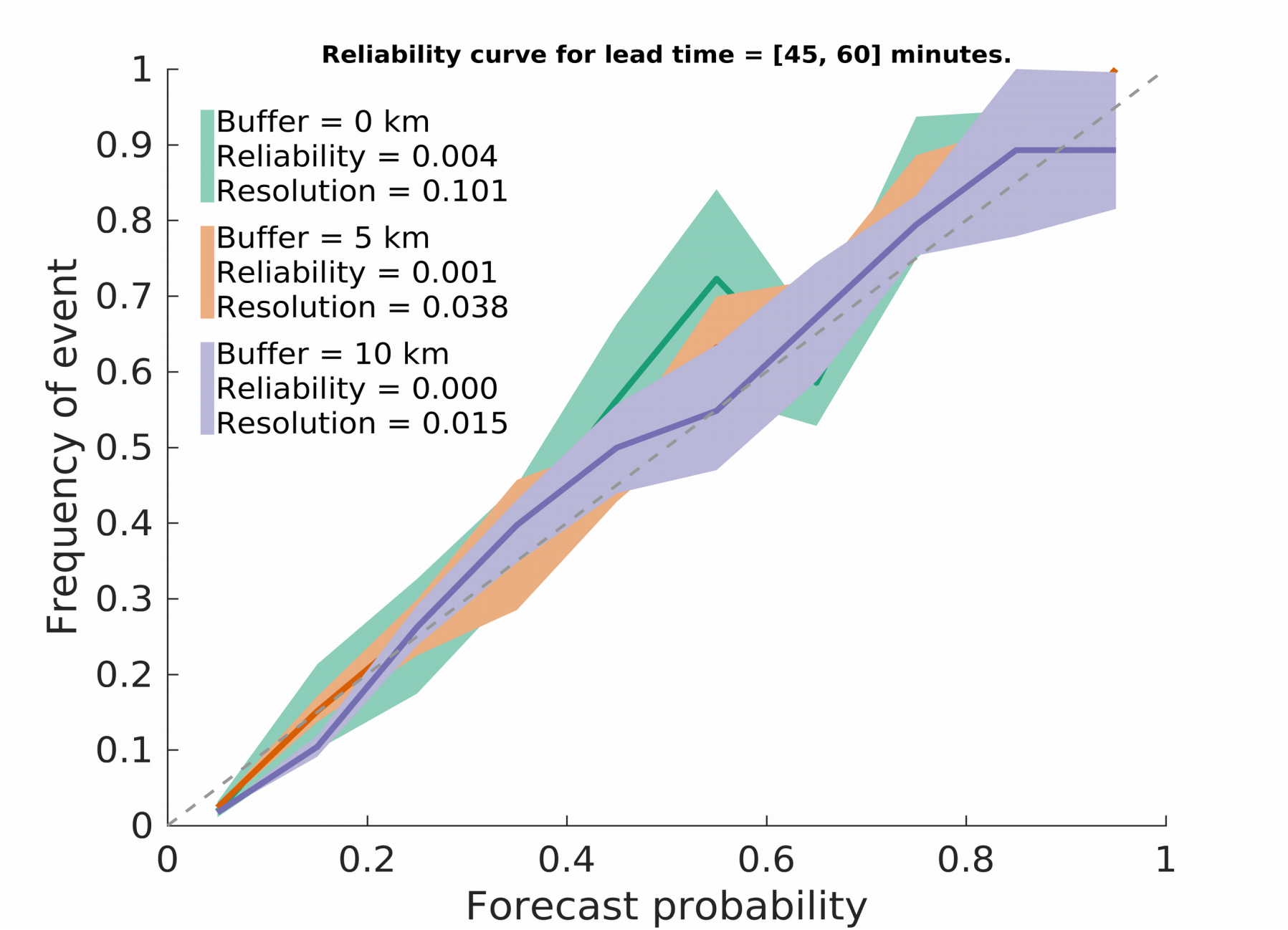
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



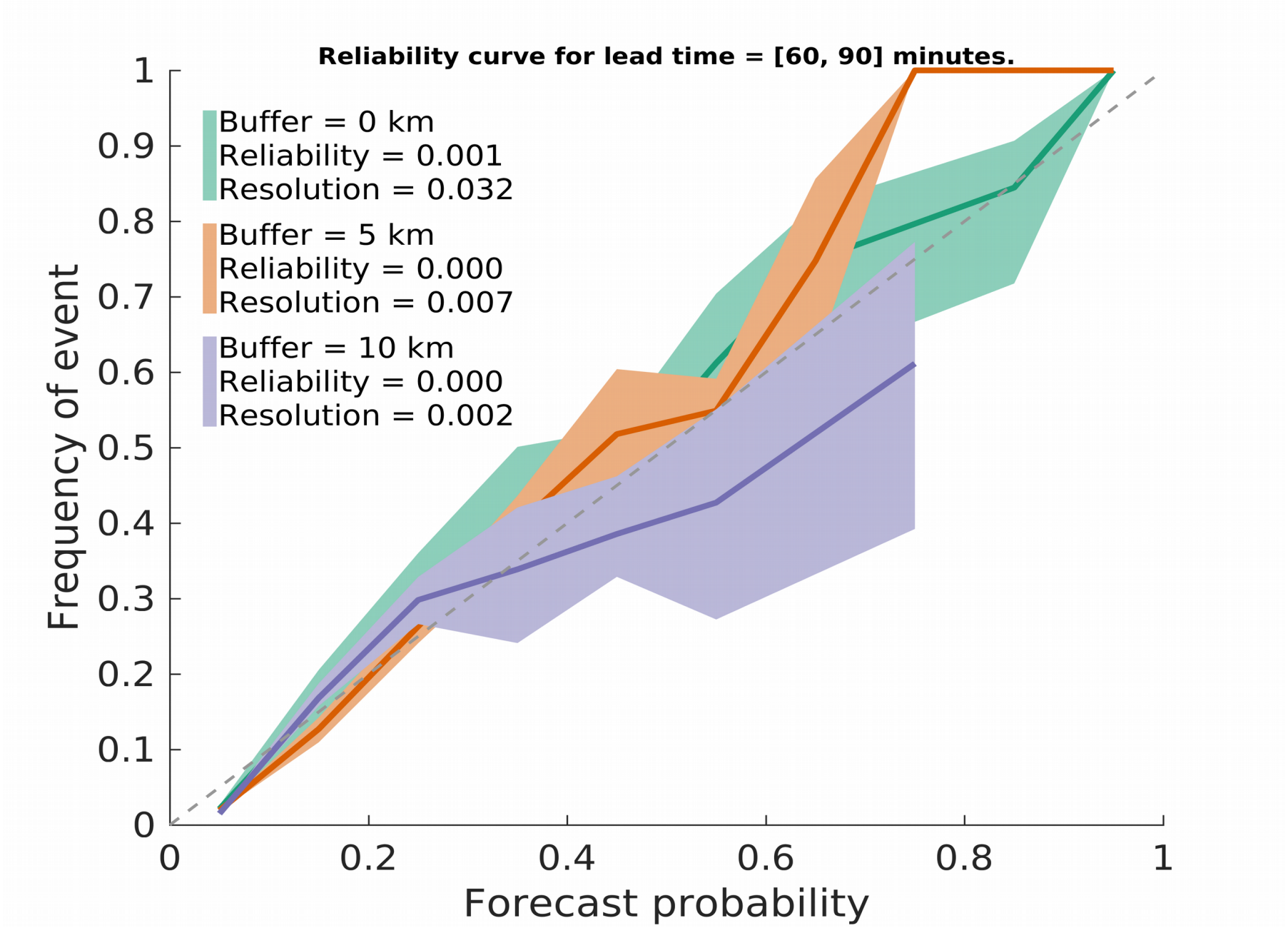
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



Variable Importance for Damaging Straight-line Storm Winds

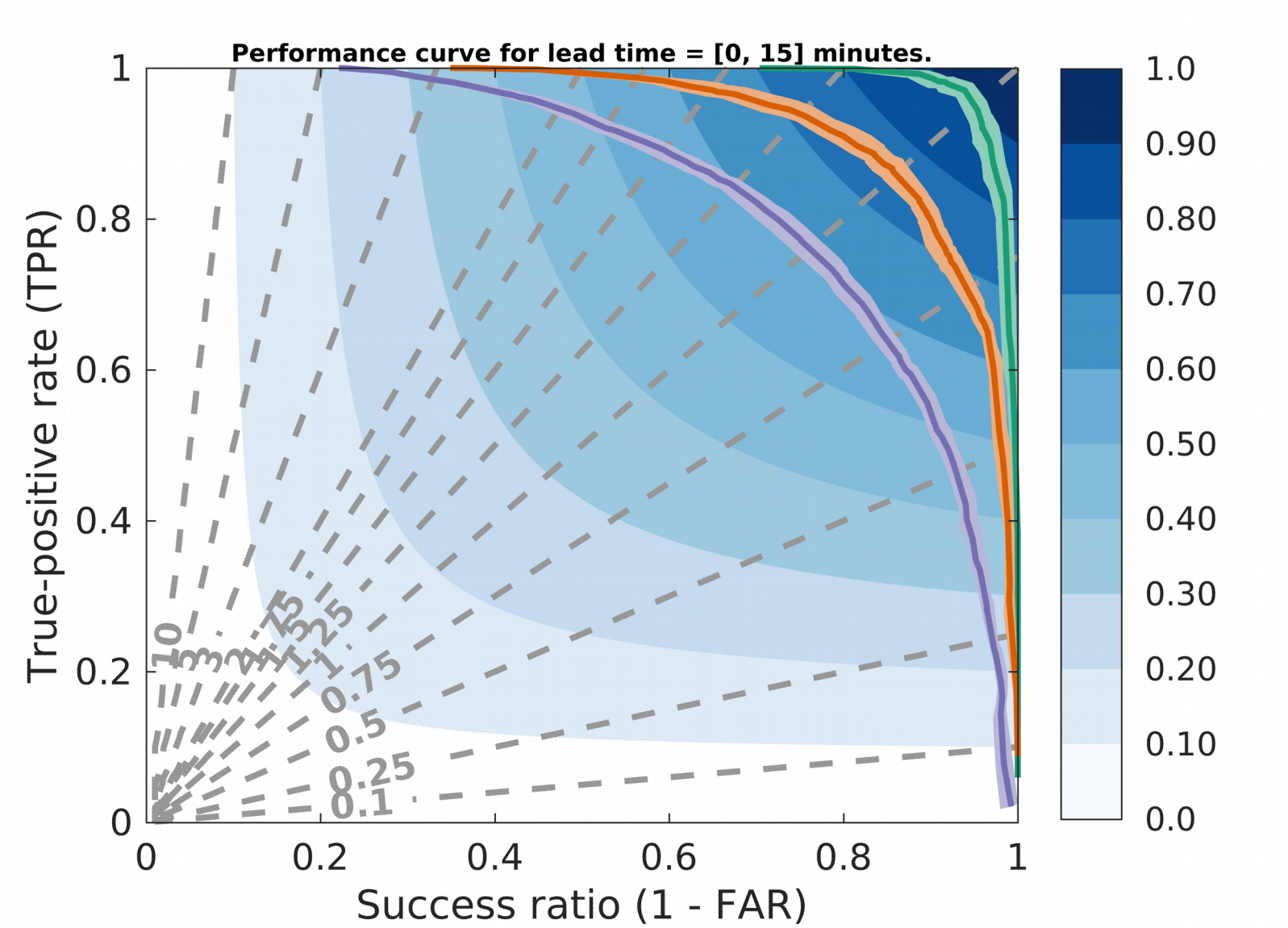
Appendix A: Forecast Evaluation



Variable Importance for Damaging Straight-line Storm Winds

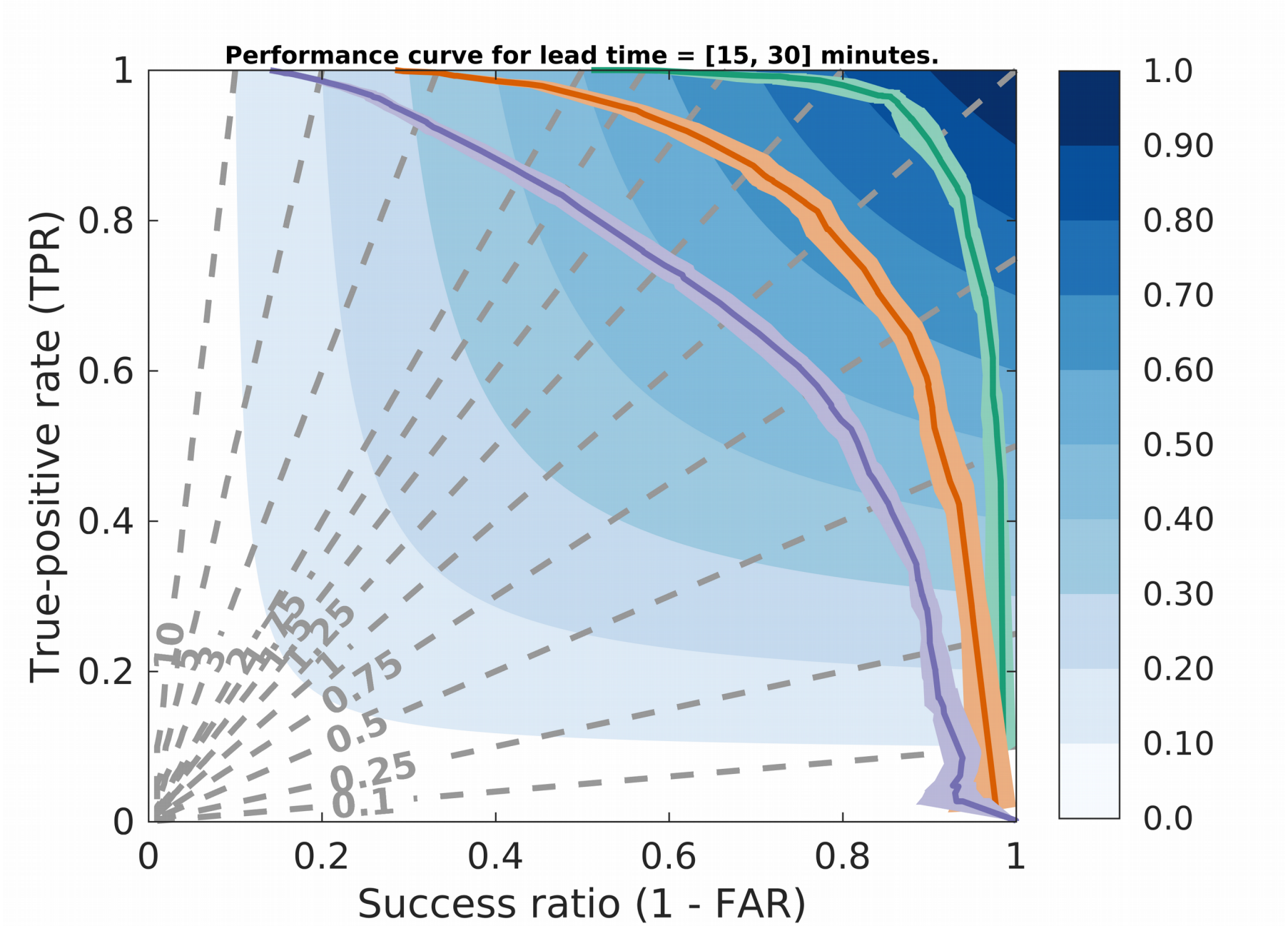
Appendix A: Forecast Evaluation

- Performance diagrams[24] for each lead time are shown on the following pages.
- Colour fill is critical success index (CSI); dashed lines are frequency bias.



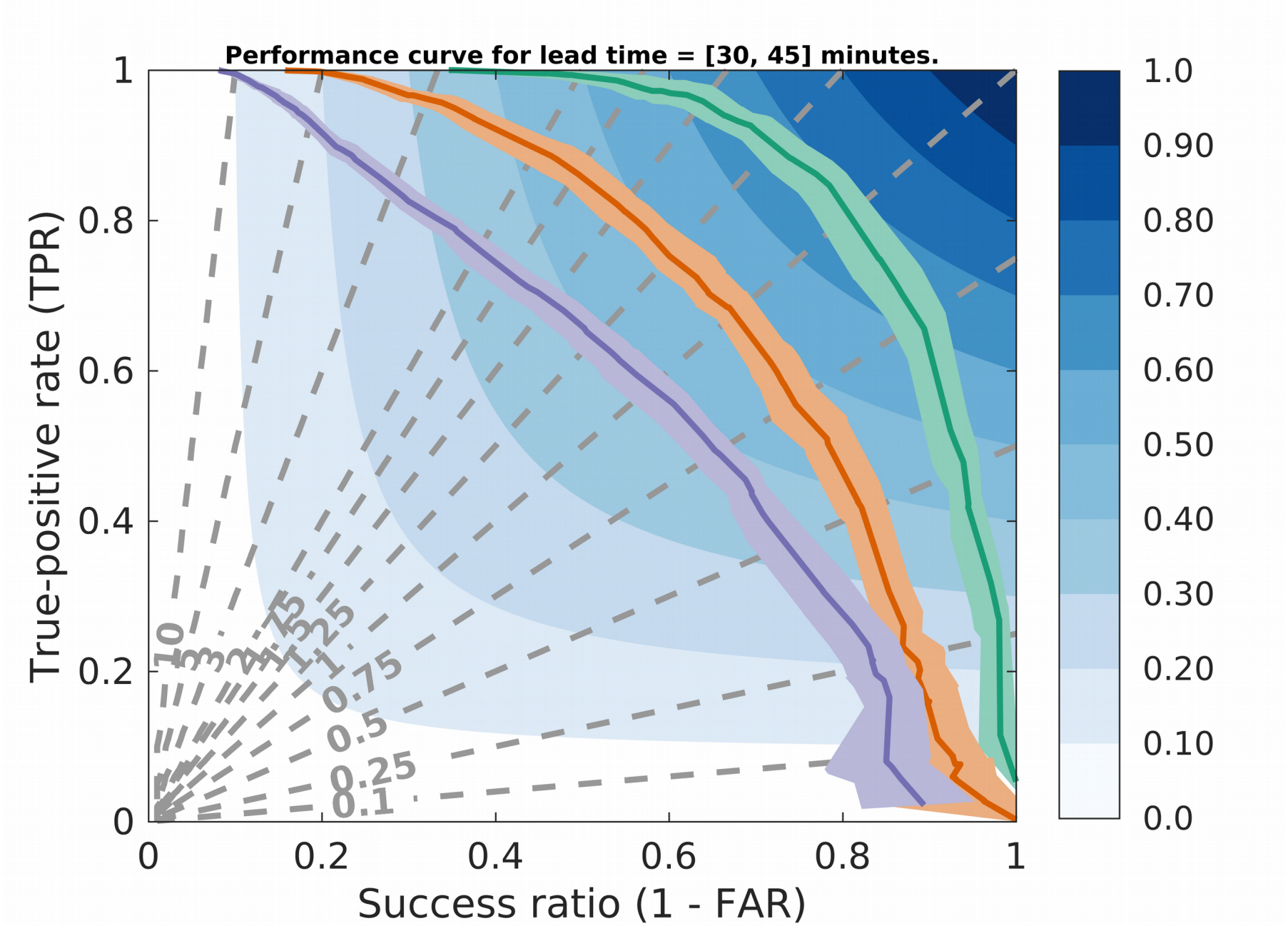
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



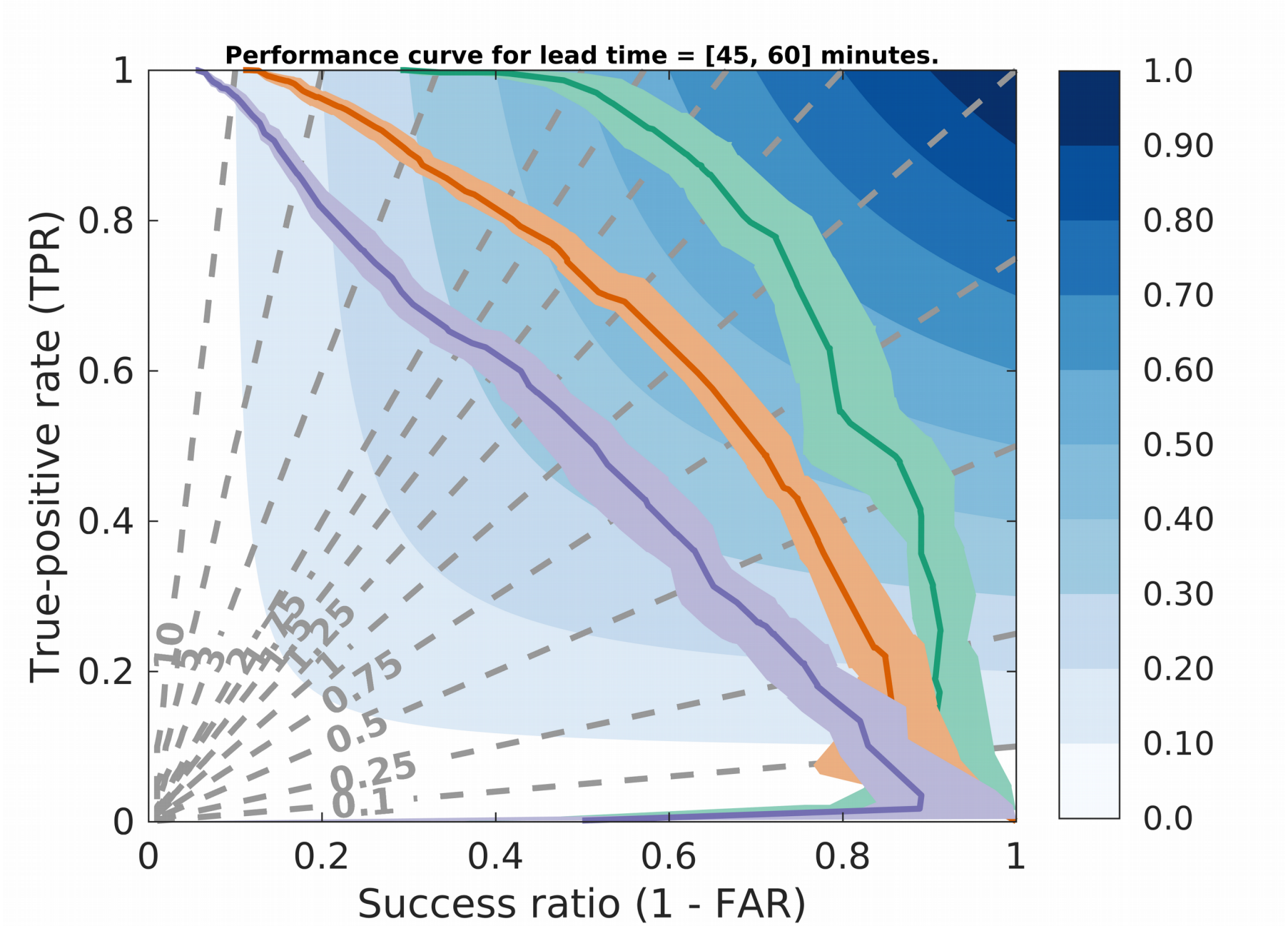
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



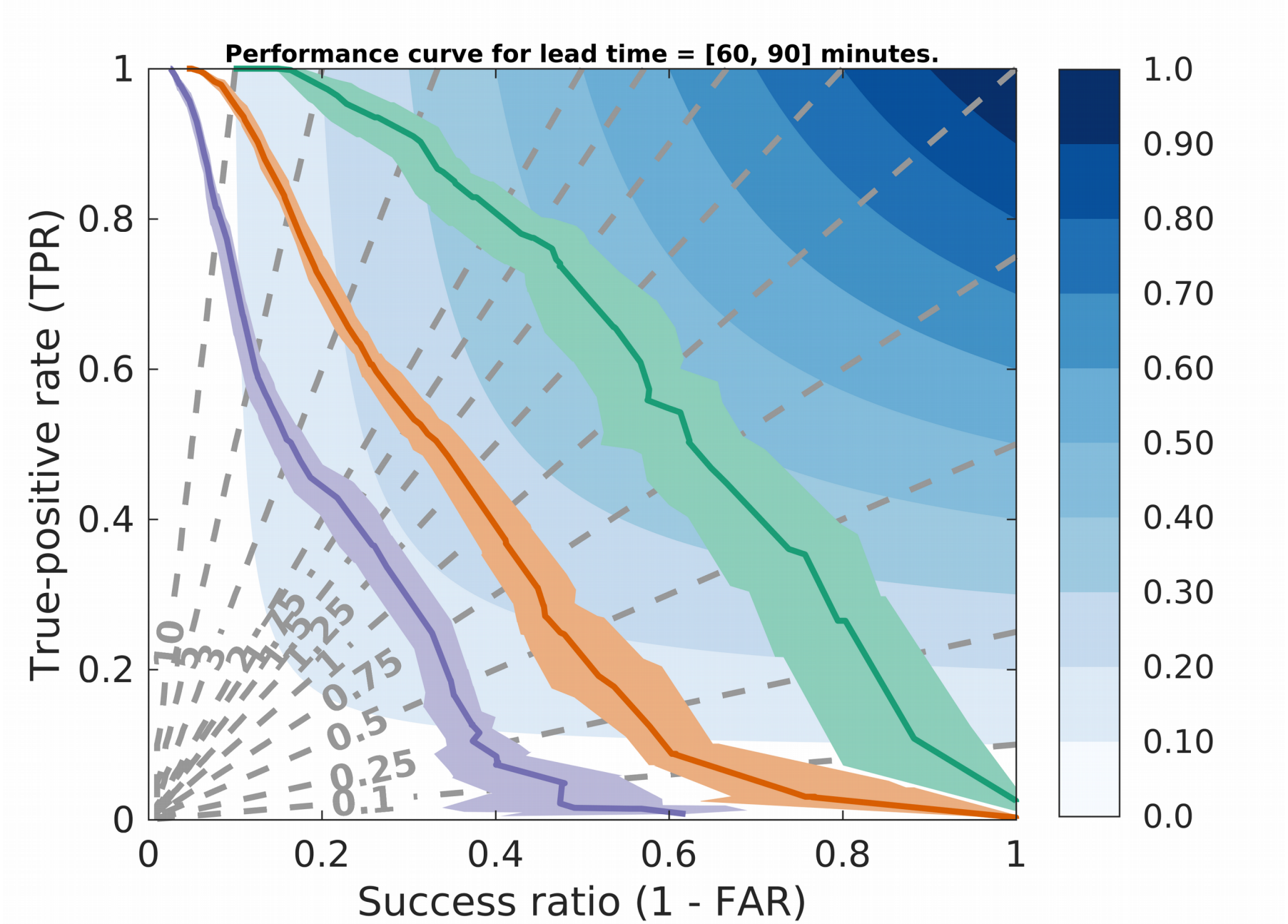
Variable Importance for Damaging Straight-line Storm Winds

Appendix A: Forecast Evaluation



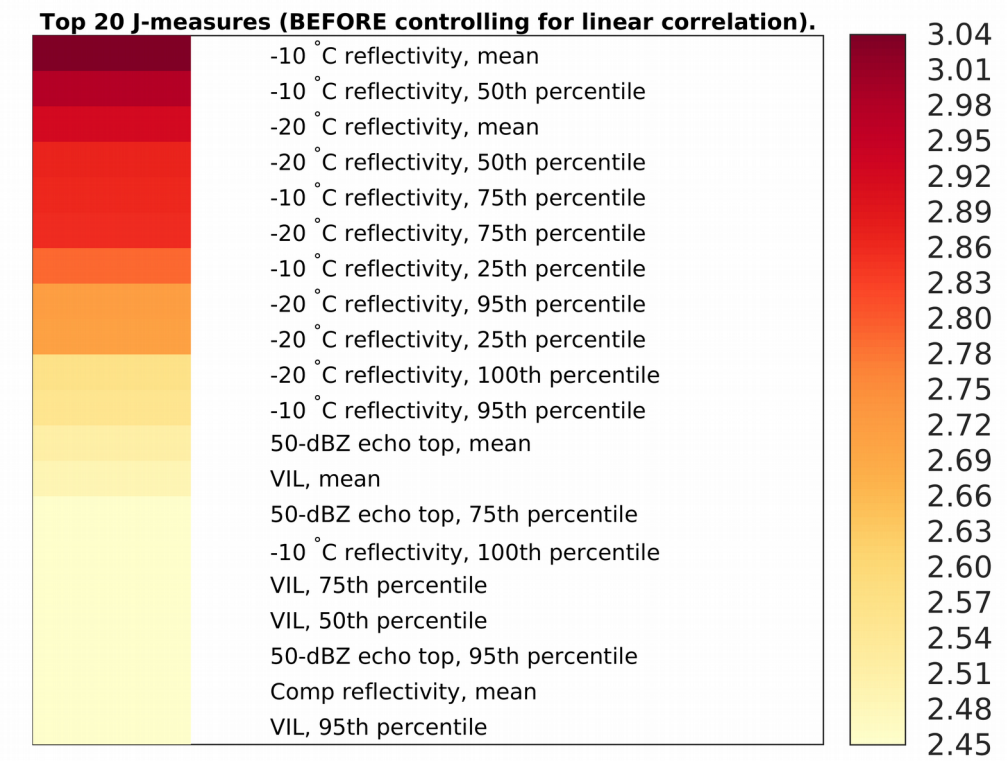
Variable Importance for Damaging Straight-line Storm Winds

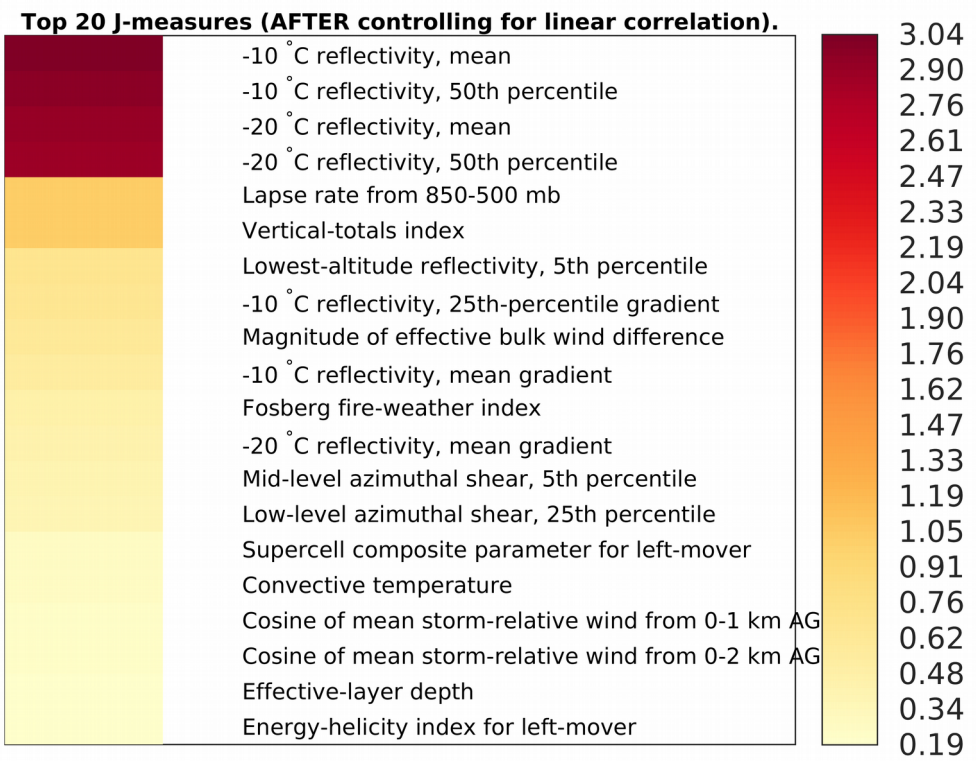
Appendix A: Forecast Evaluation



Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

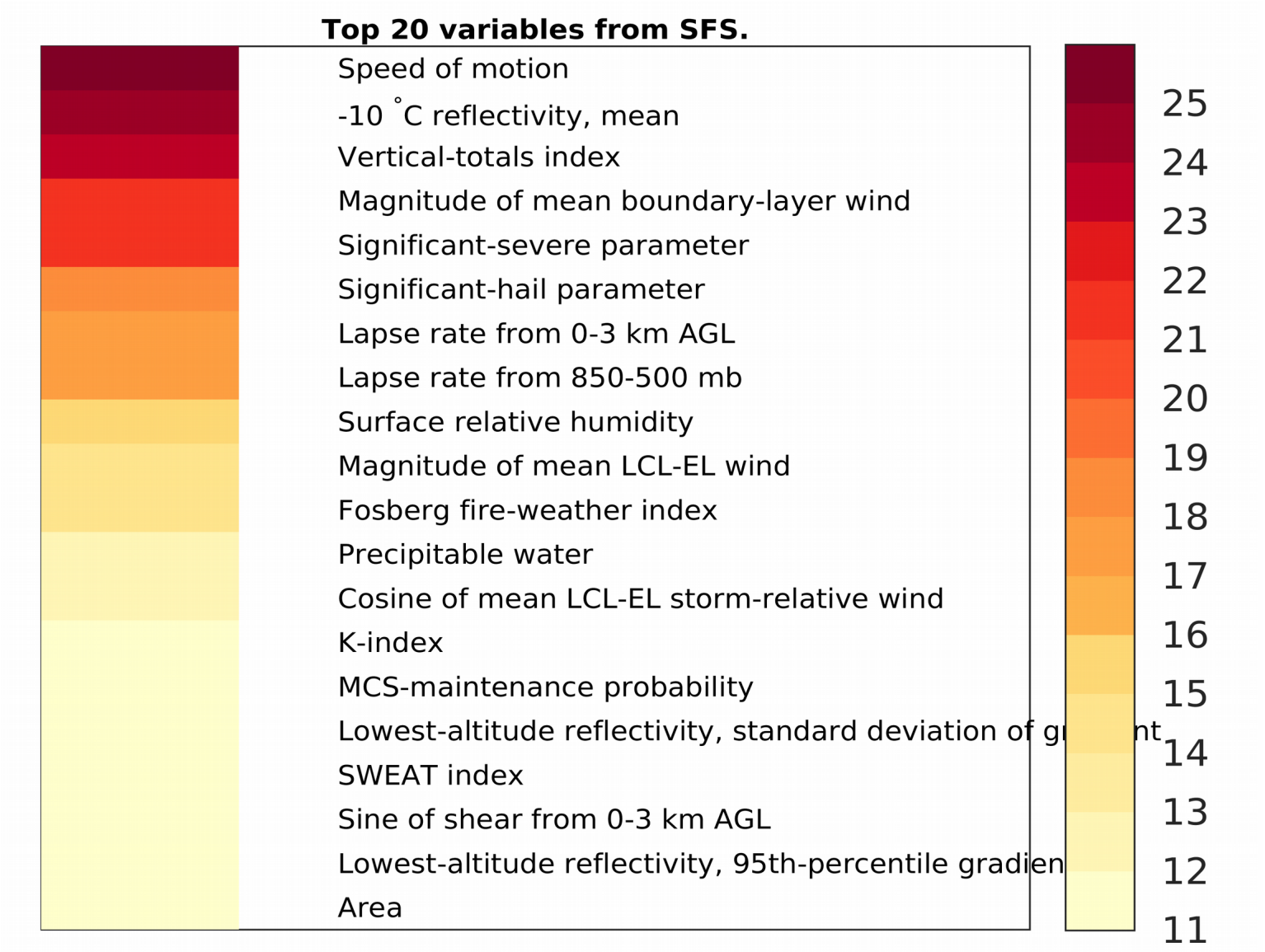
- *J*-measure results for buffer distance of 0 km and lead time of [0, 15] minutes.





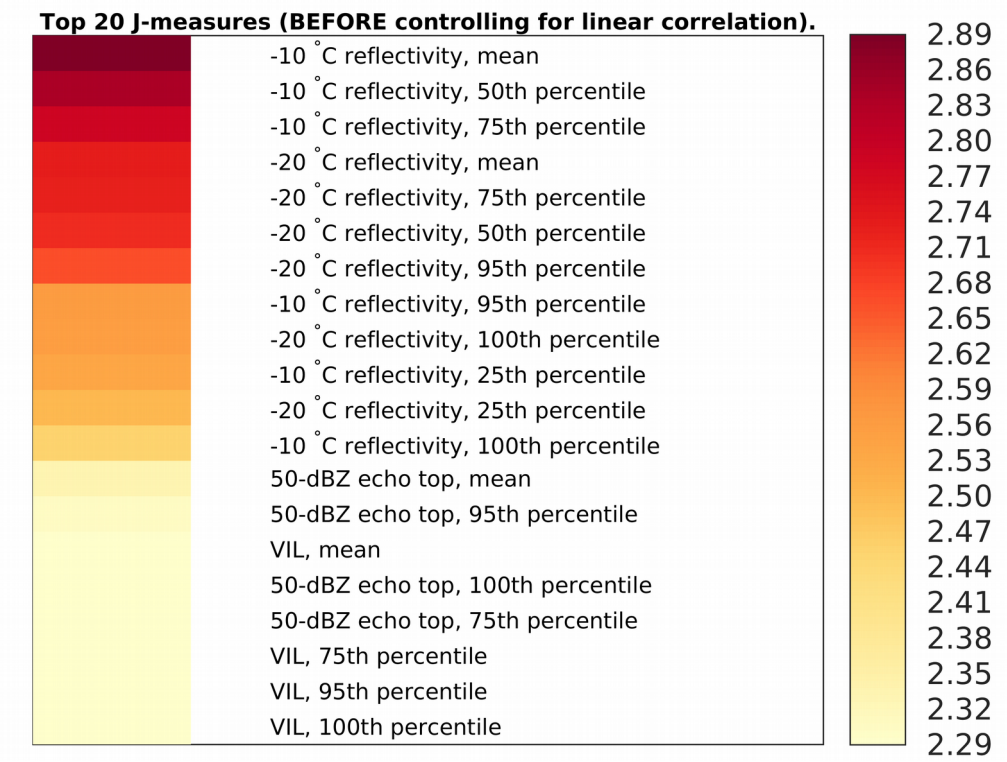
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

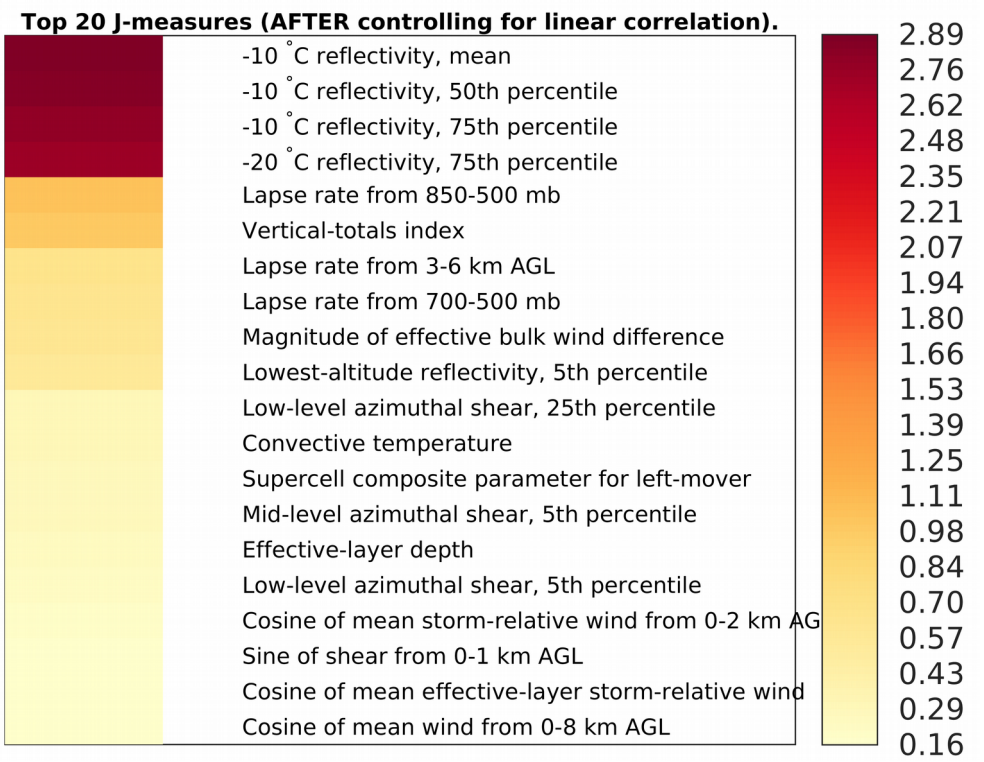
- SFS results for buffer distance of 0 km and lead time of [0, 15] minutes.



Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

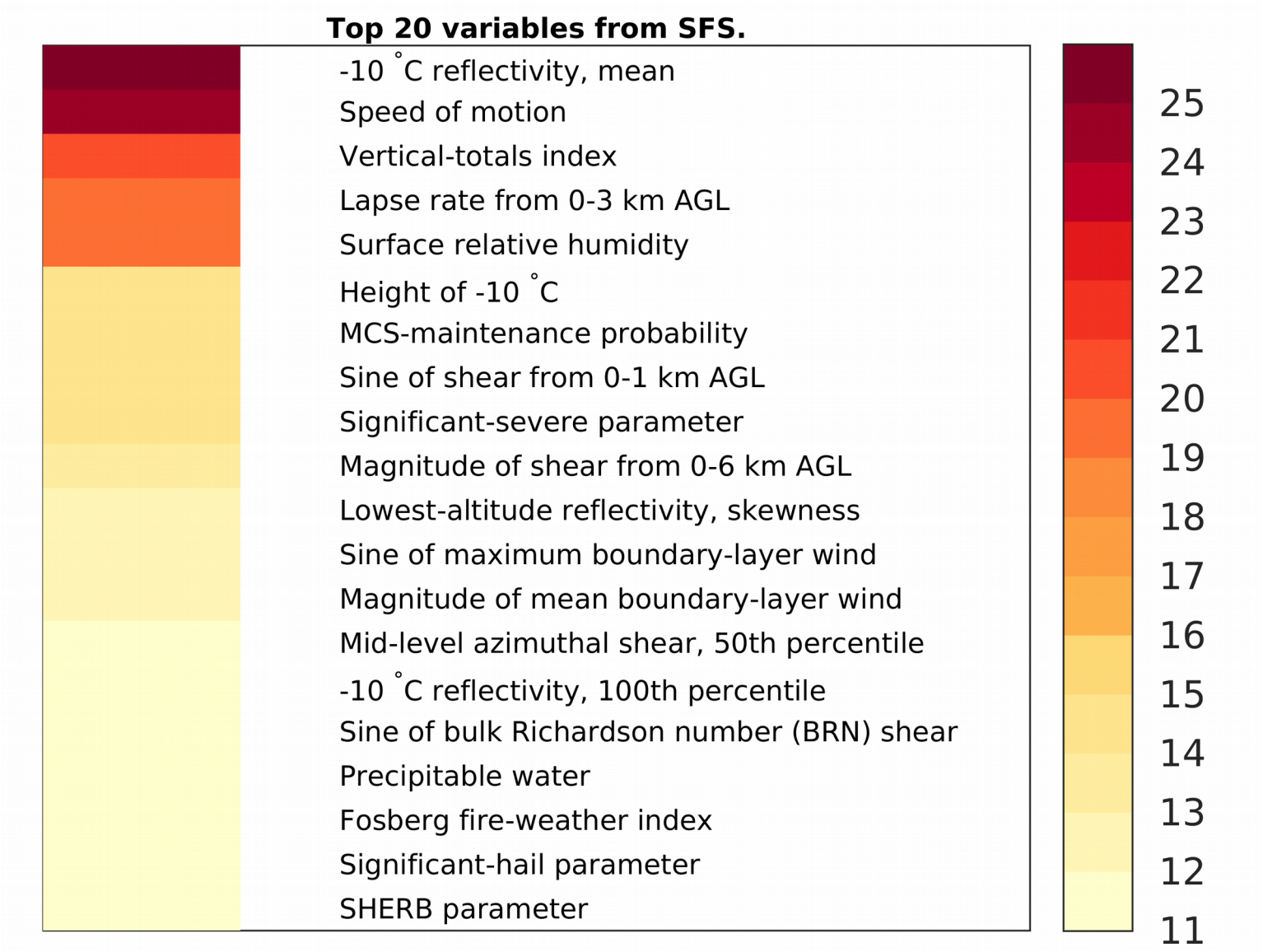
- *J*-measure results for buffer distance of 0 km and lead time of [15, 30] minutes.





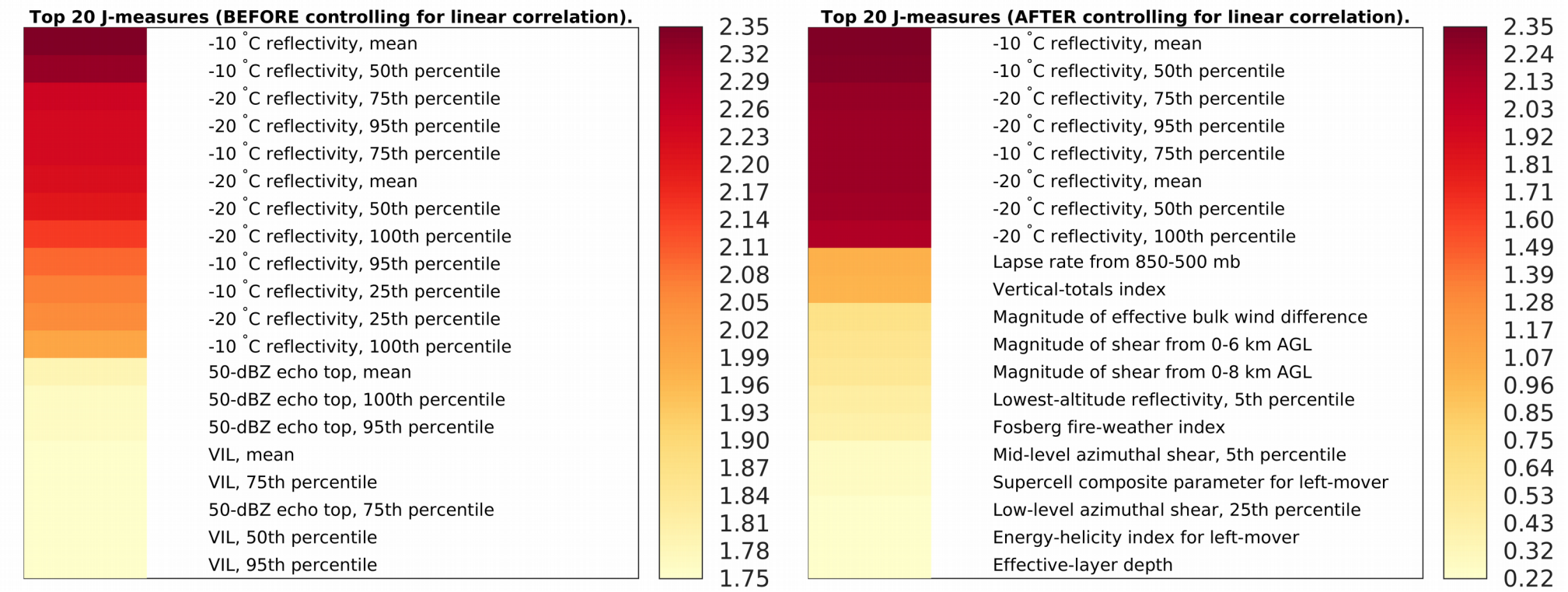
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 0 km and lead time of [15, 30] minutes.



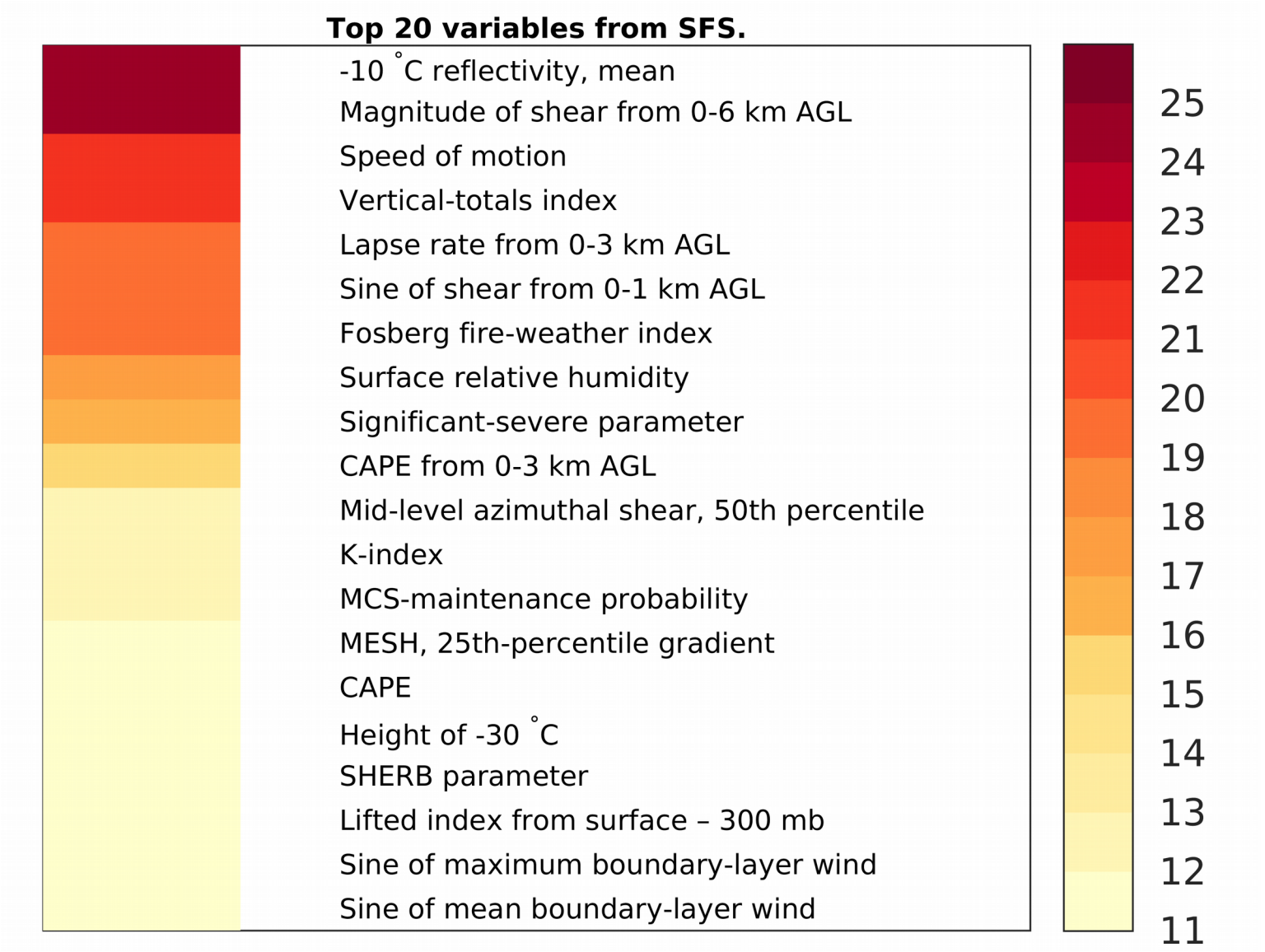
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 0 km and lead time of [30, 45] minutes.



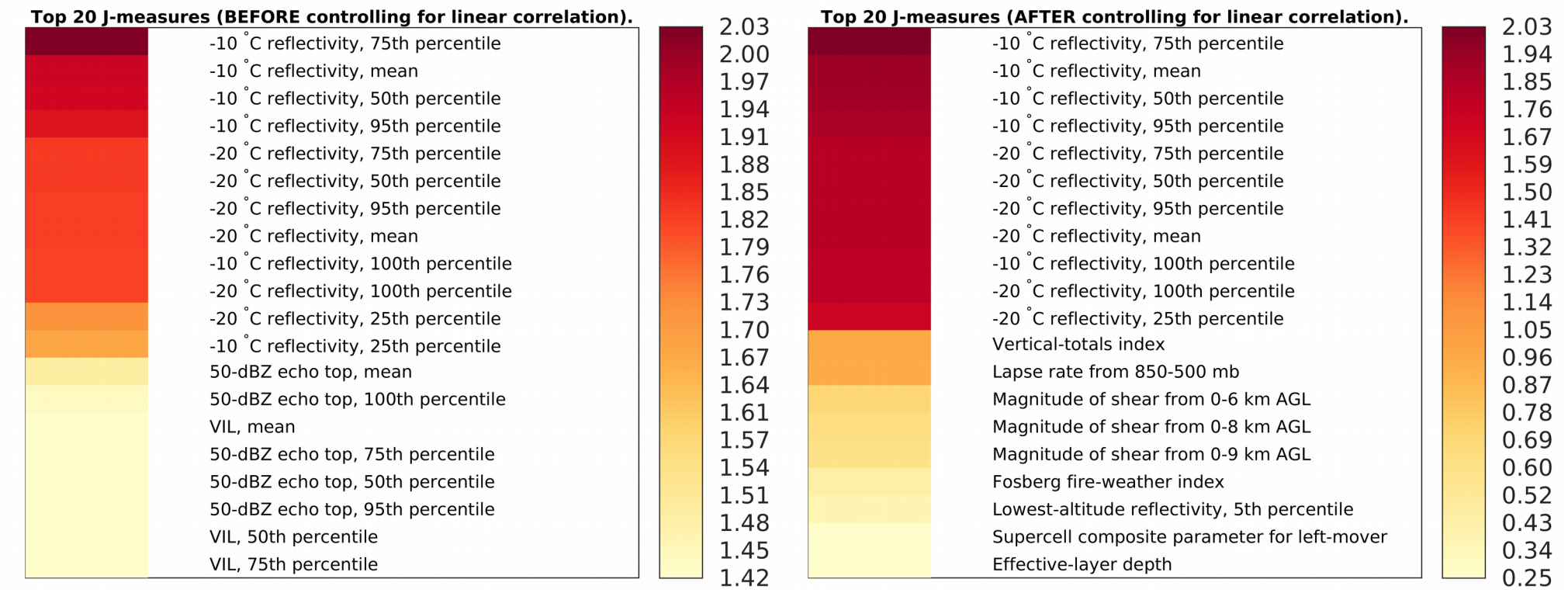
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 0 km and lead time of [30, 45] minutes.



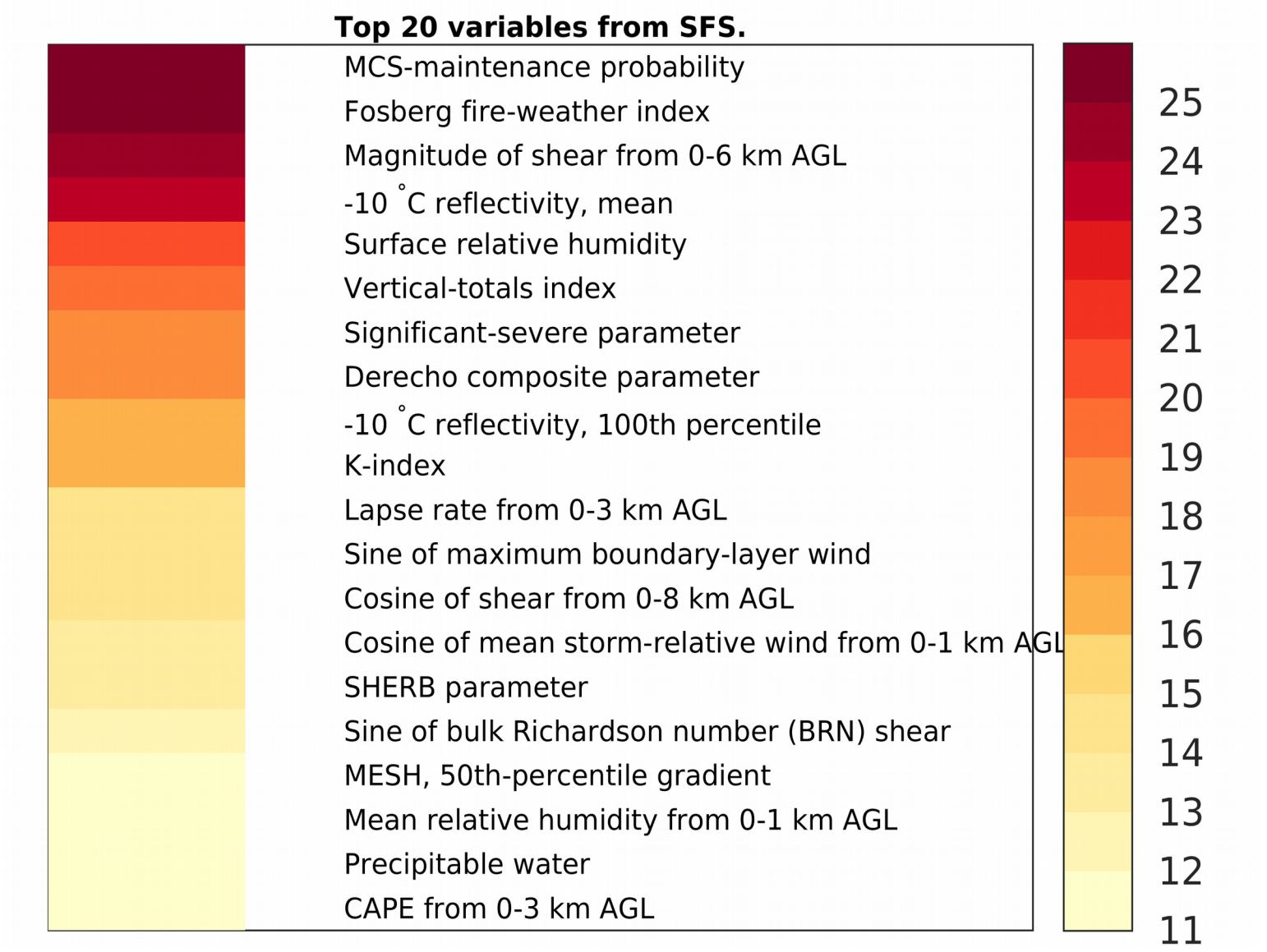
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 0 km and lead time of [45, 60] minutes.



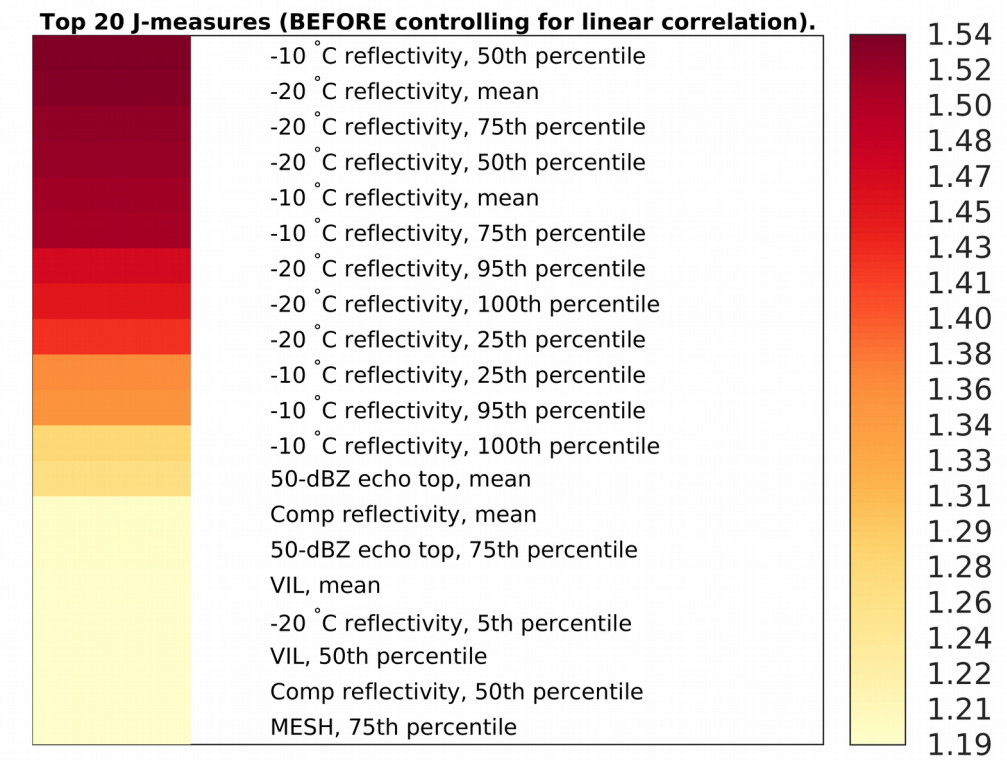
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

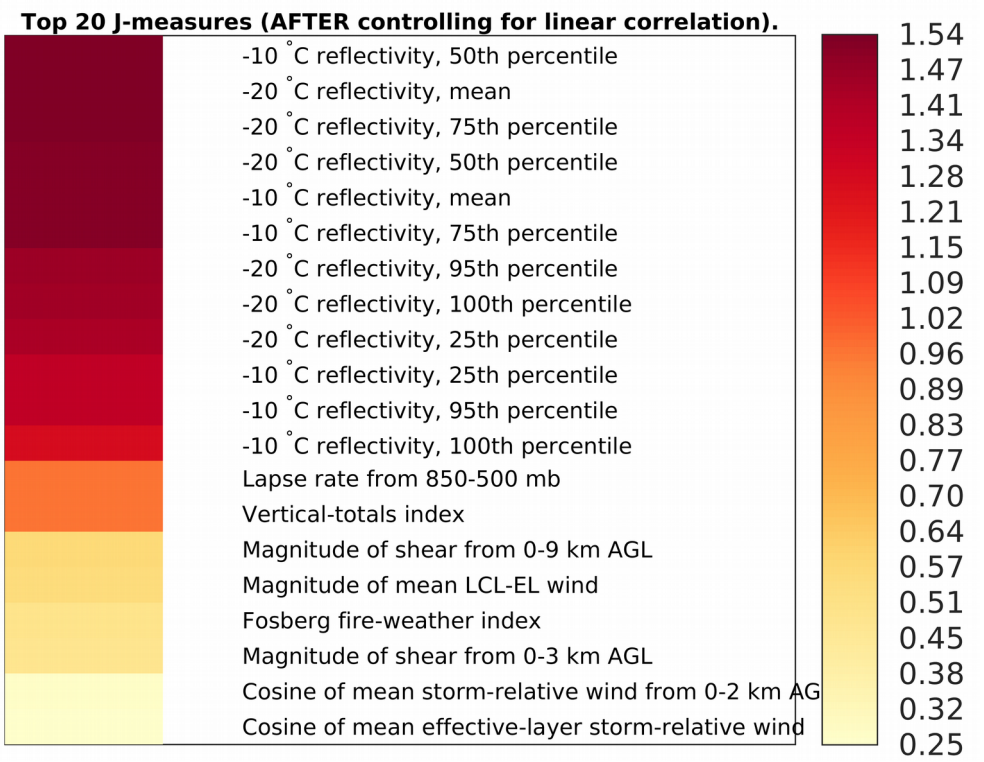
- SFS results for buffer distance of 0 km and lead time of [45, 60] minutes.



Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

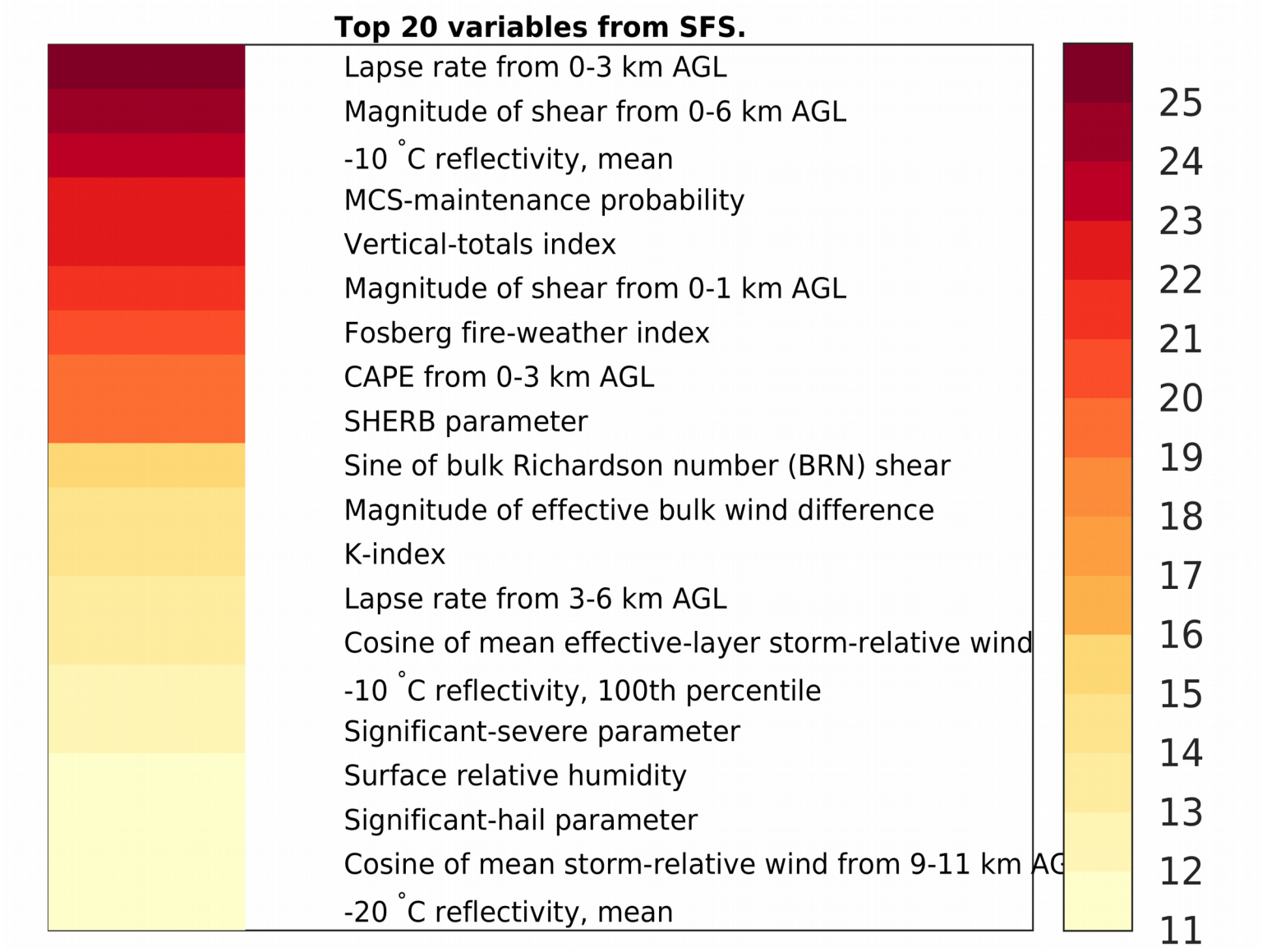
- *J*-measure results for buffer distance of 0 km and lead time of [60, 90] minutes.





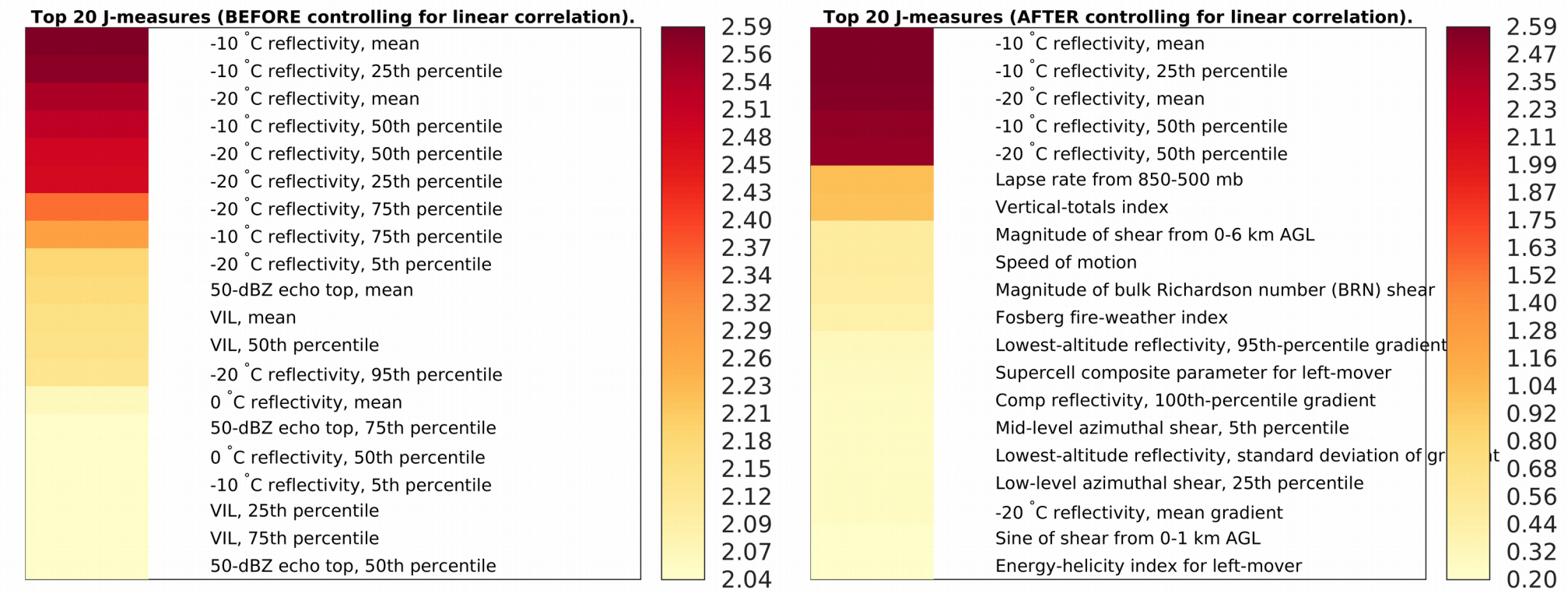
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 0 km and lead time of [60, 90] minutes.



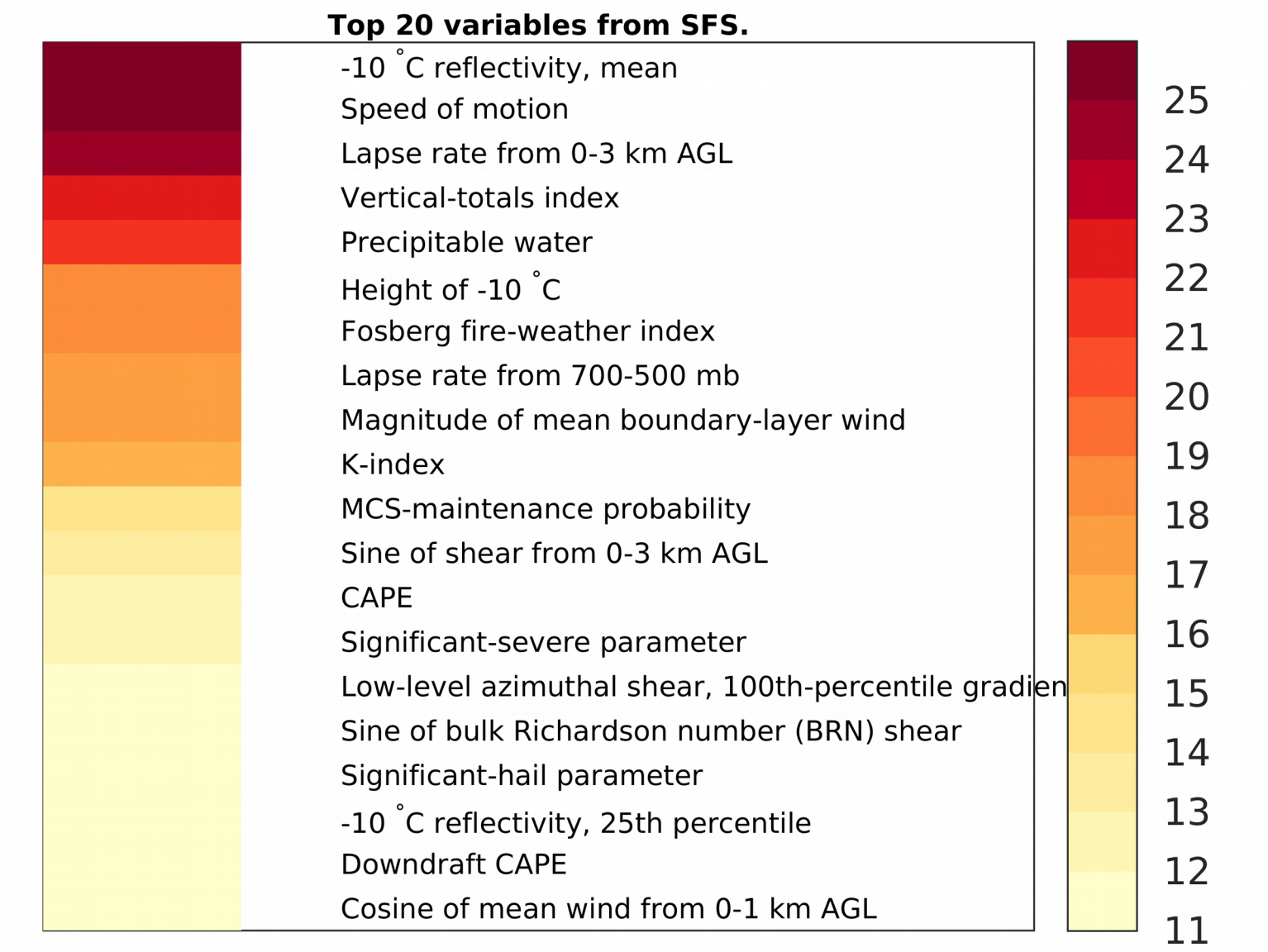
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 5 km and lead time of [0, 15] minutes.



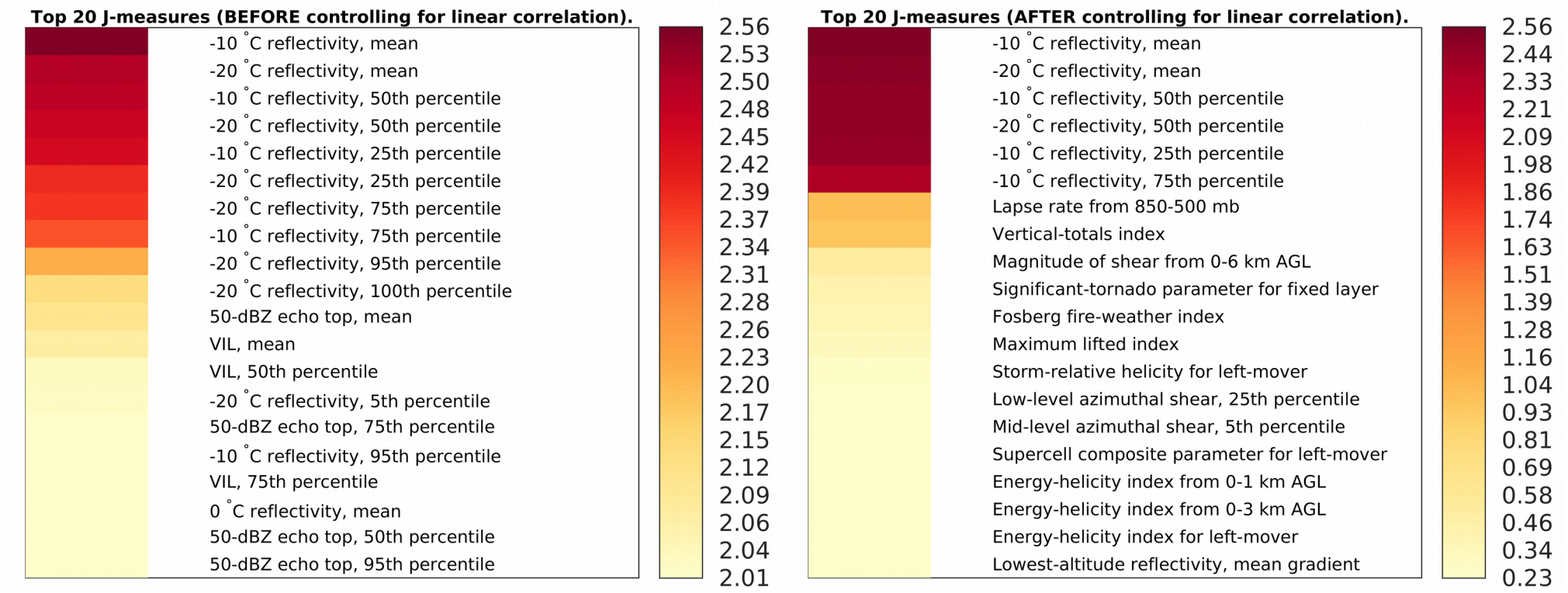
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 5 km and lead time of [0, 15] minutes.



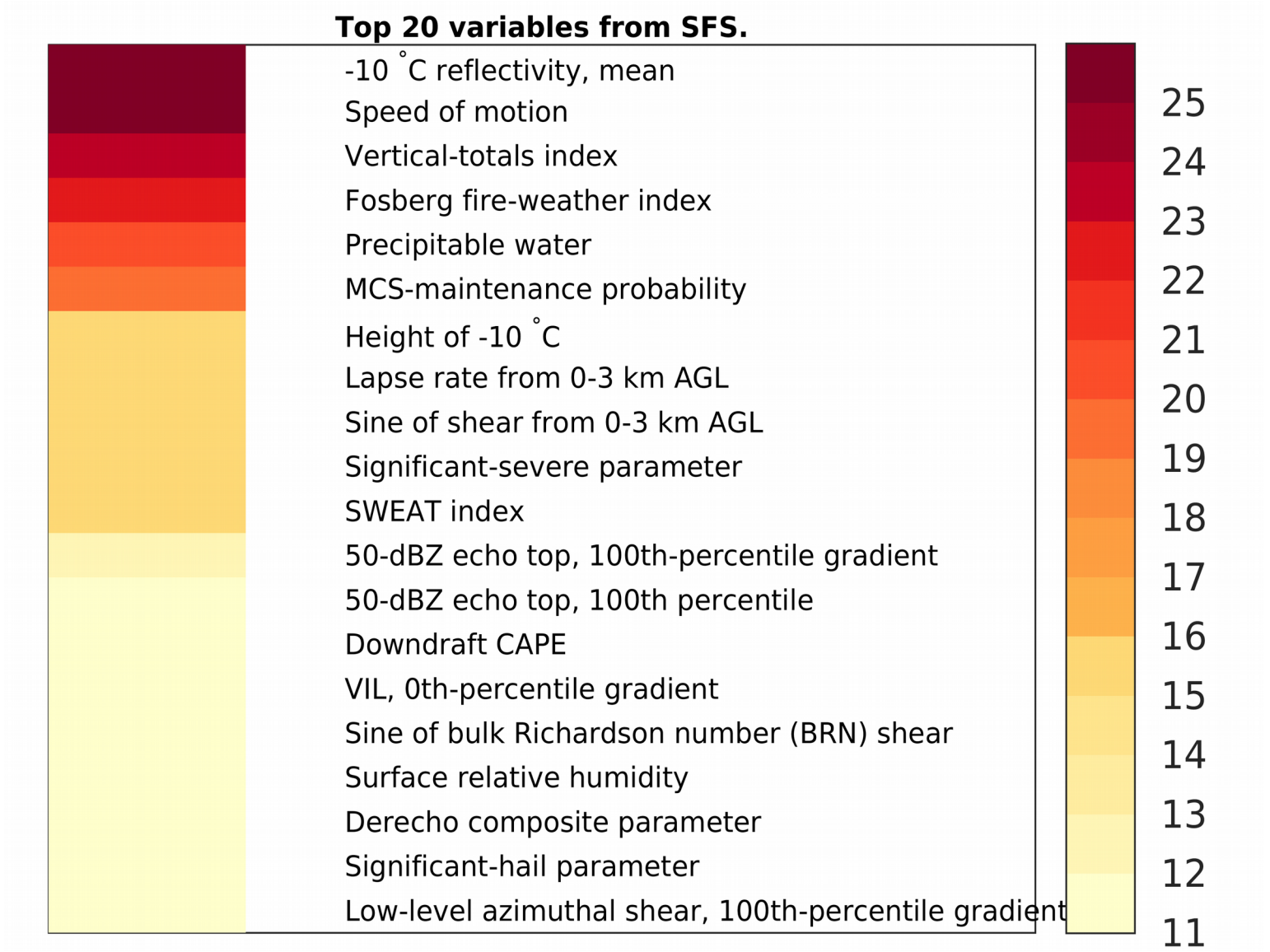
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 5 km and lead time of [15, 30] minutes.



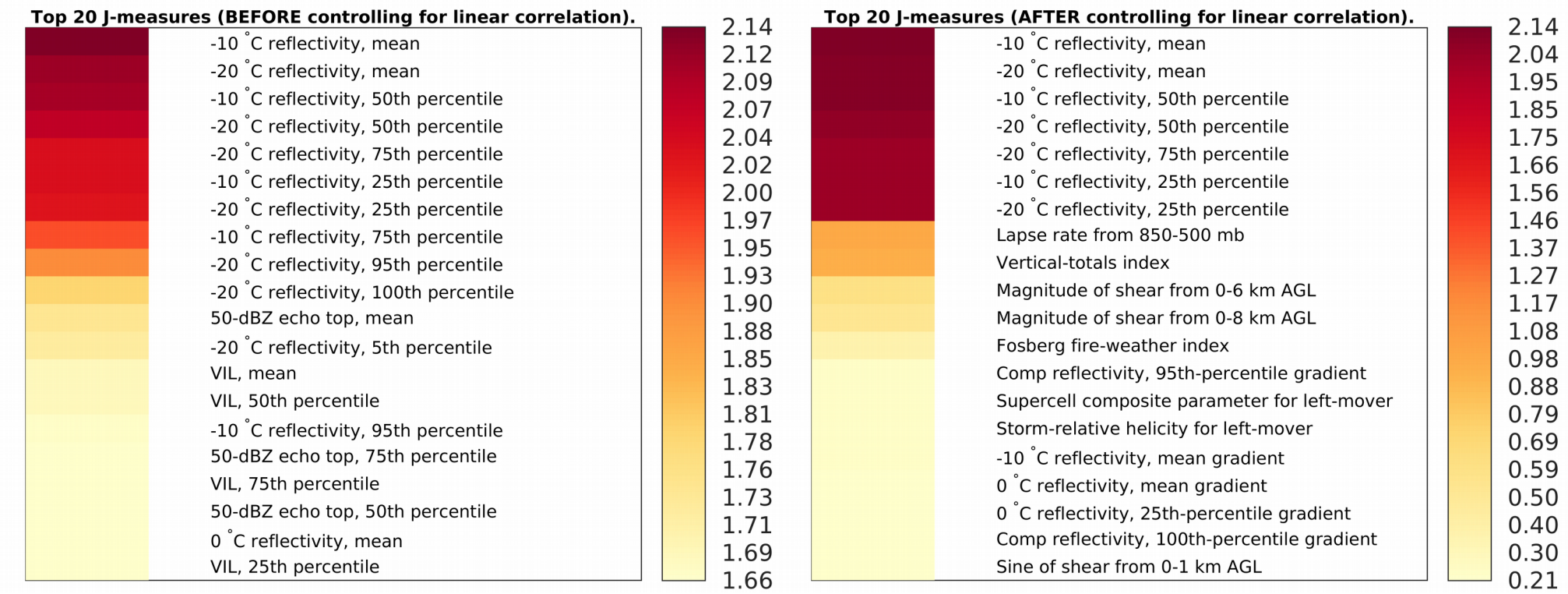
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 5 km and lead time of [15, 30] minutes.



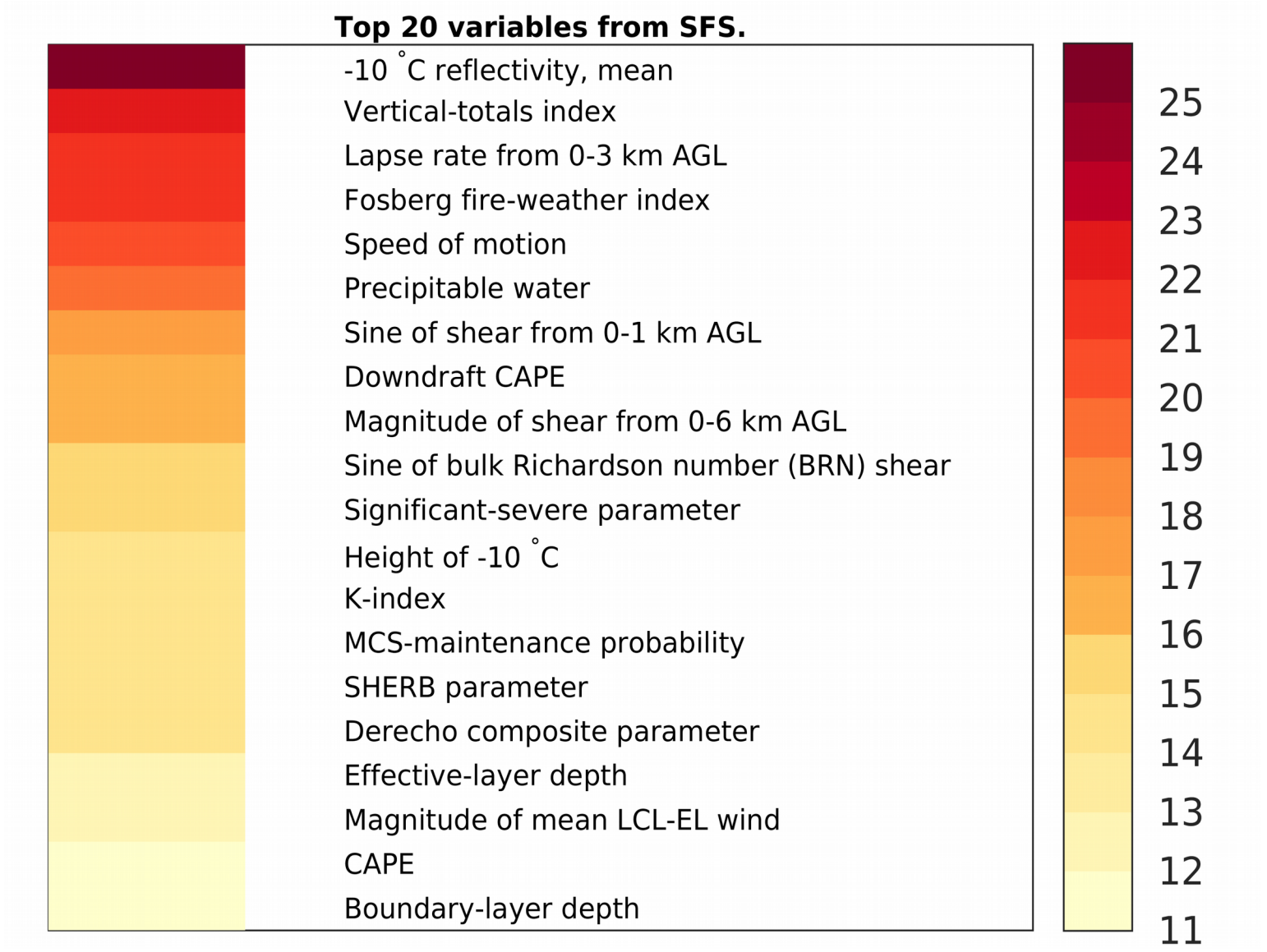
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 5 km and lead time of [30, 45] minutes.



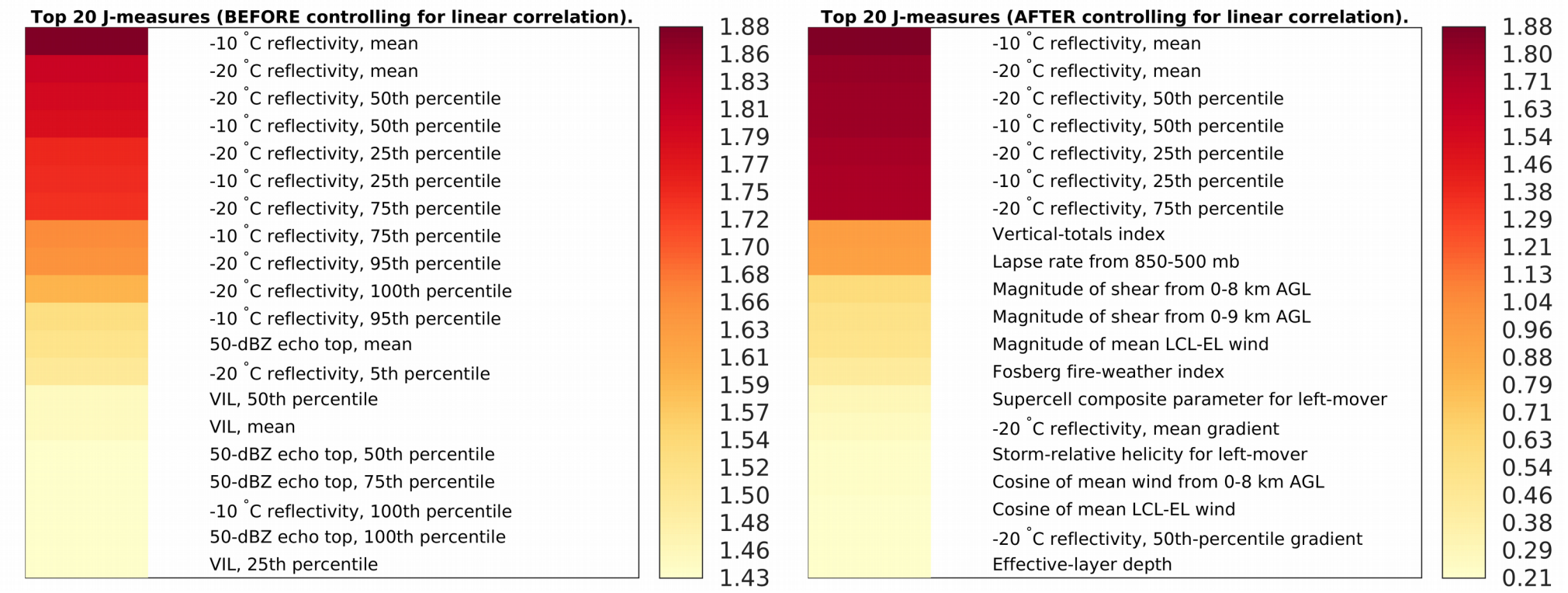
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 5 km and lead time of [30, 45] minutes.



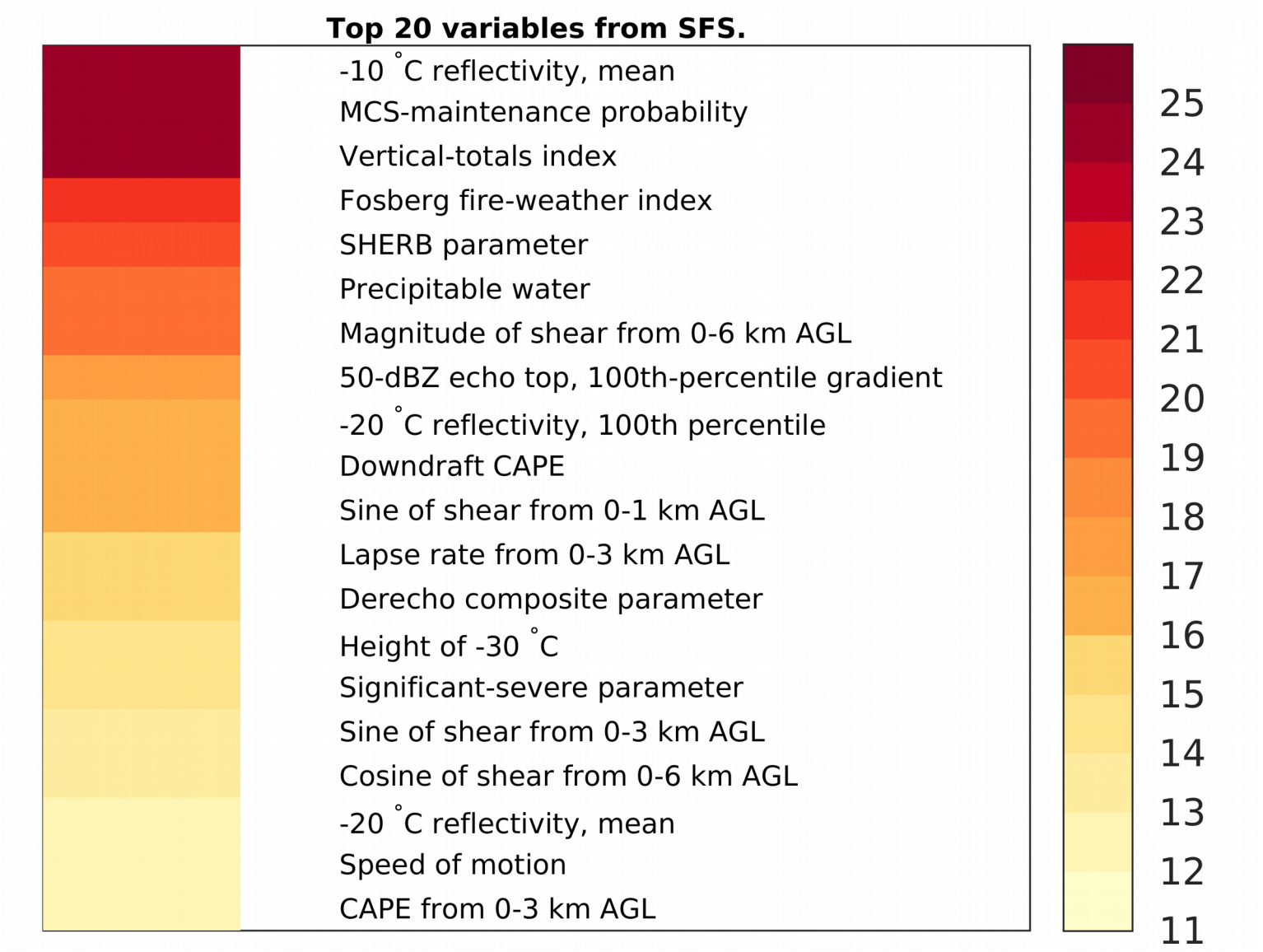
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 5 km and lead time of [45, 60] minutes.



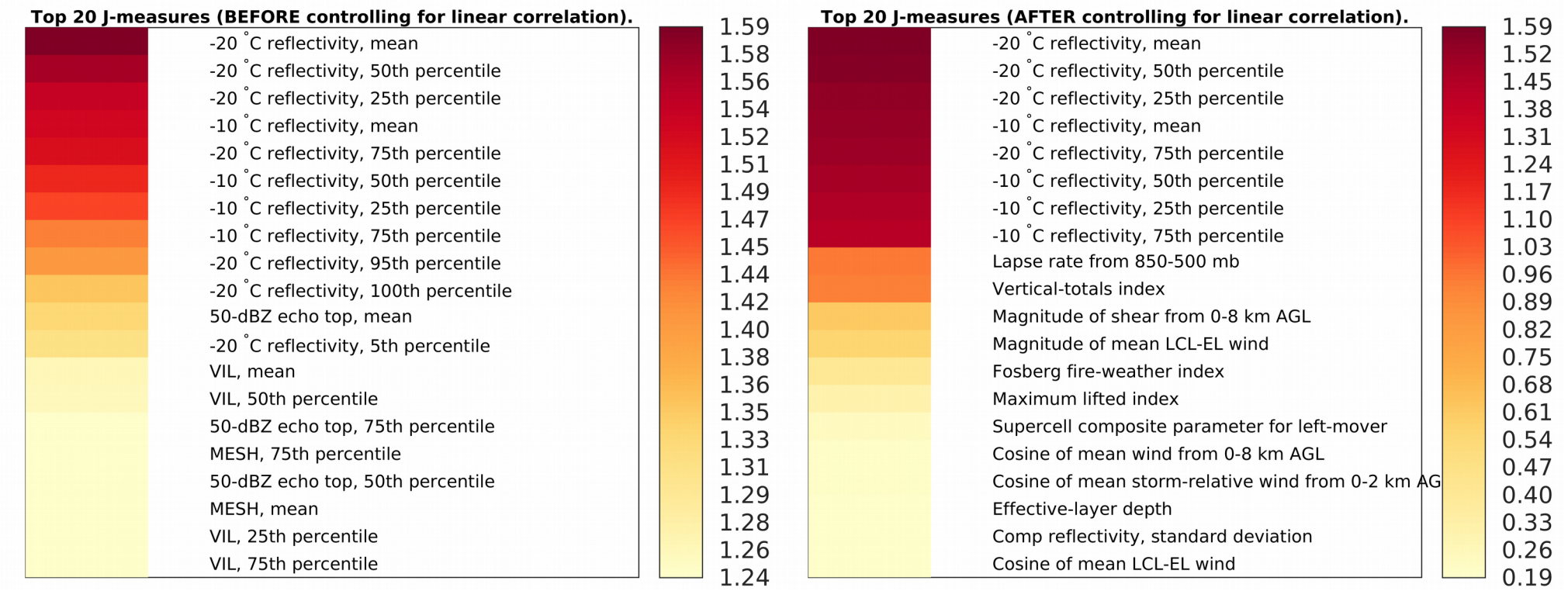
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 5 km and lead time of [45, 60] minutes.



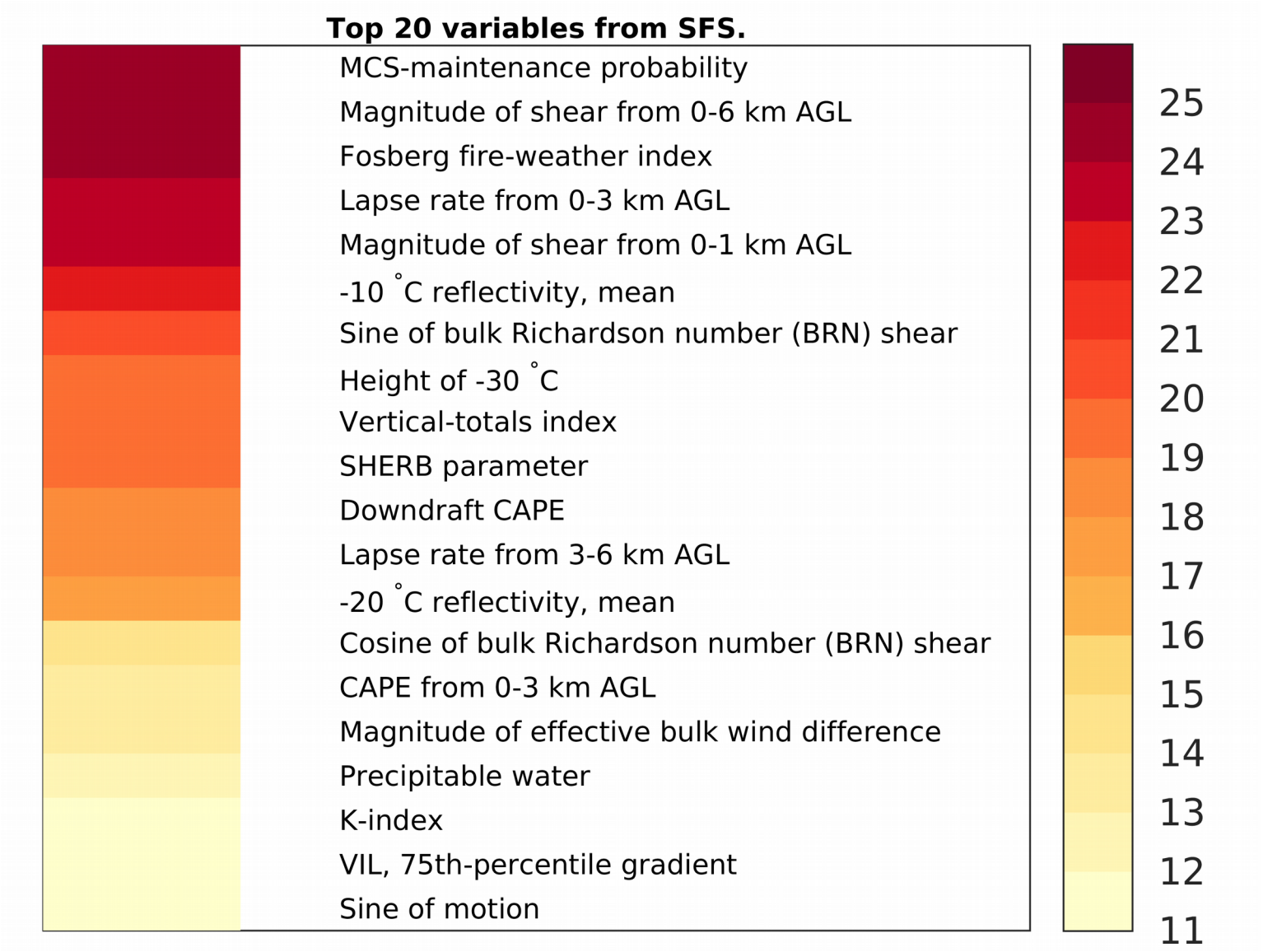
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 5 km and lead time of [60, 90] minutes.



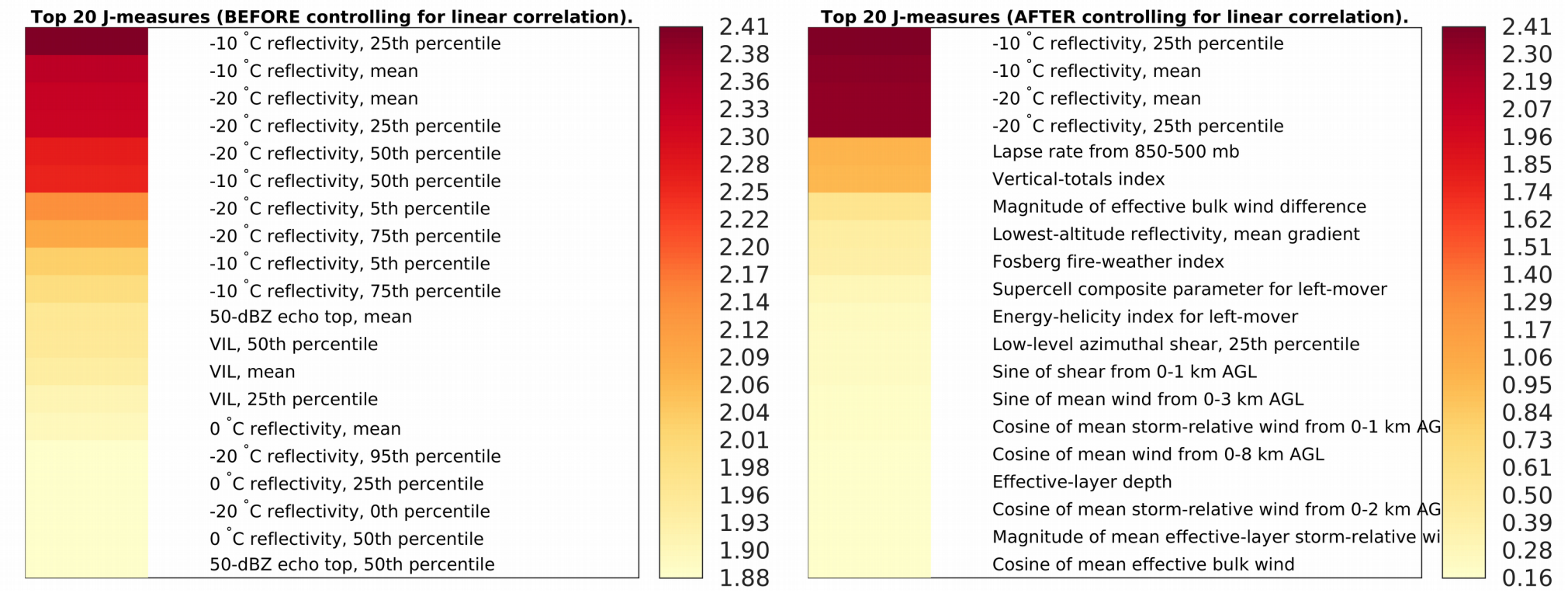
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 5 km and lead time of [60, 90] minutes.



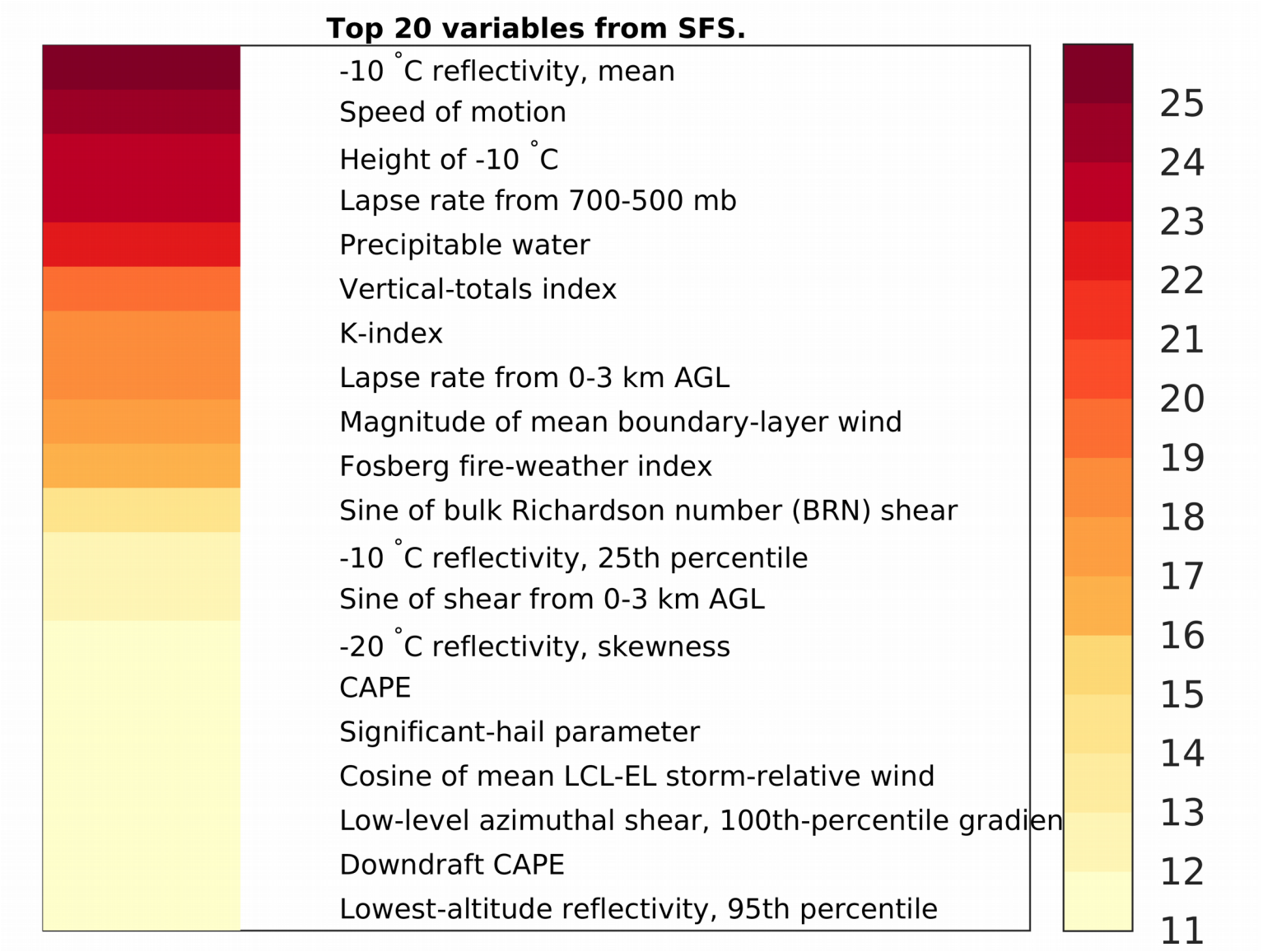
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 10 km and lead time of [0, 15] minutes.



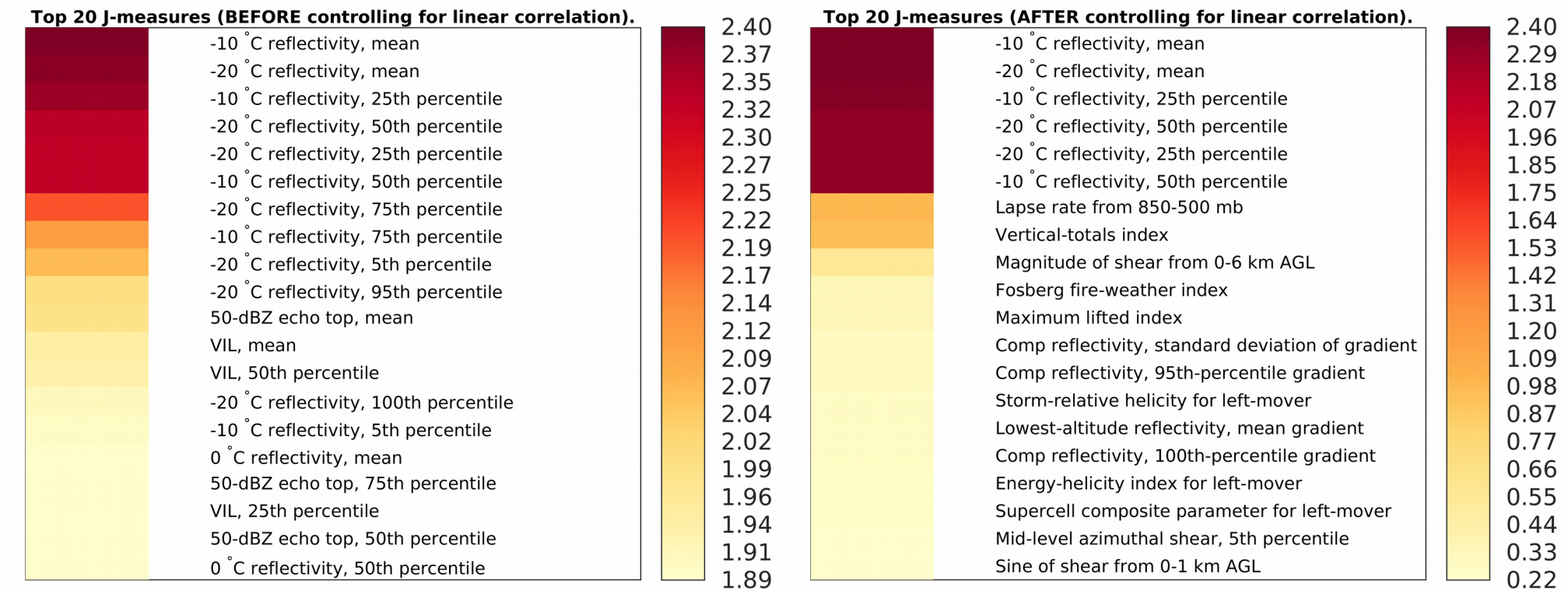
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 10 km and lead time of [0, 15] minutes.



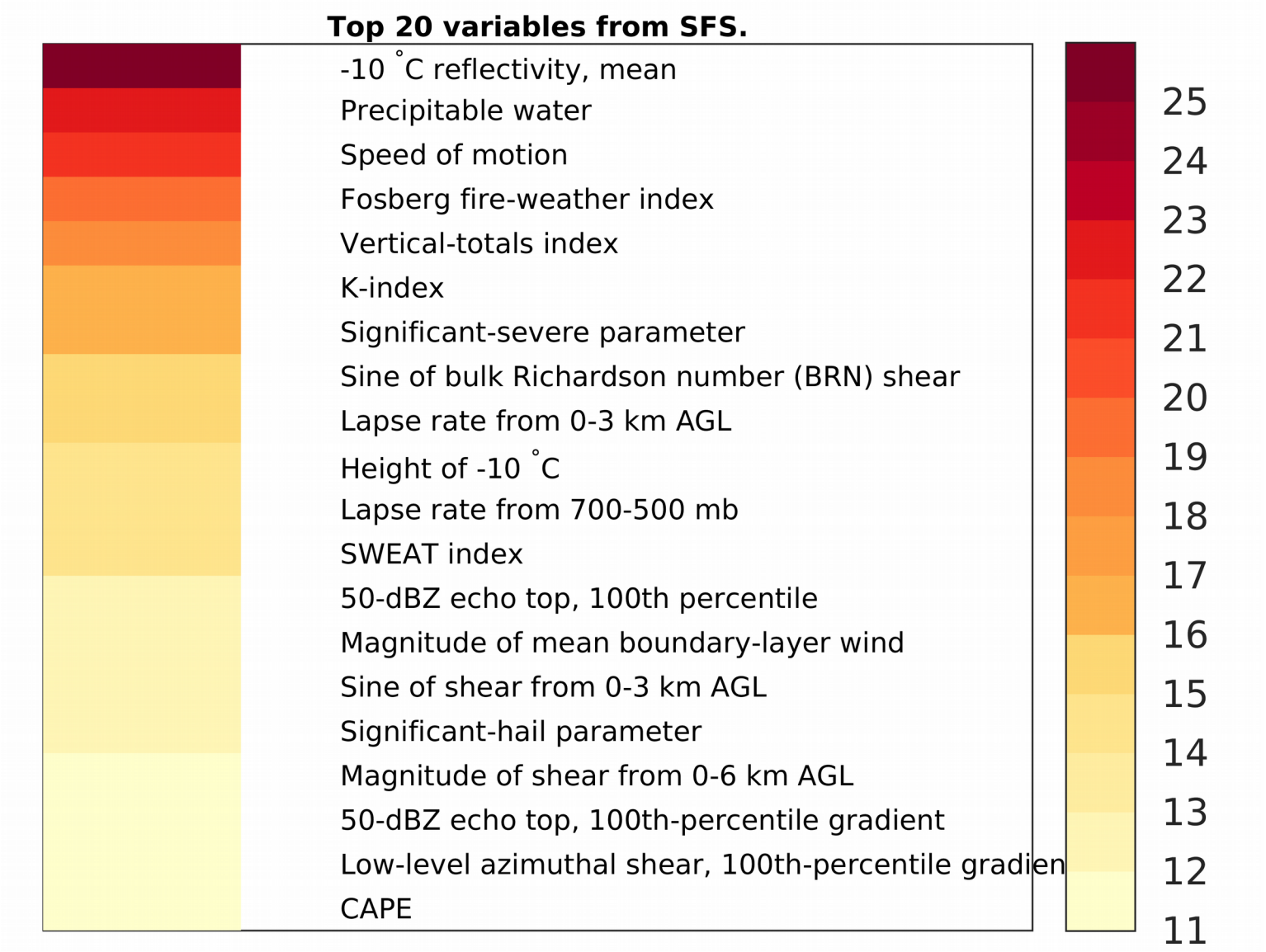
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 10 km and lead time of [15, 30] minutes.



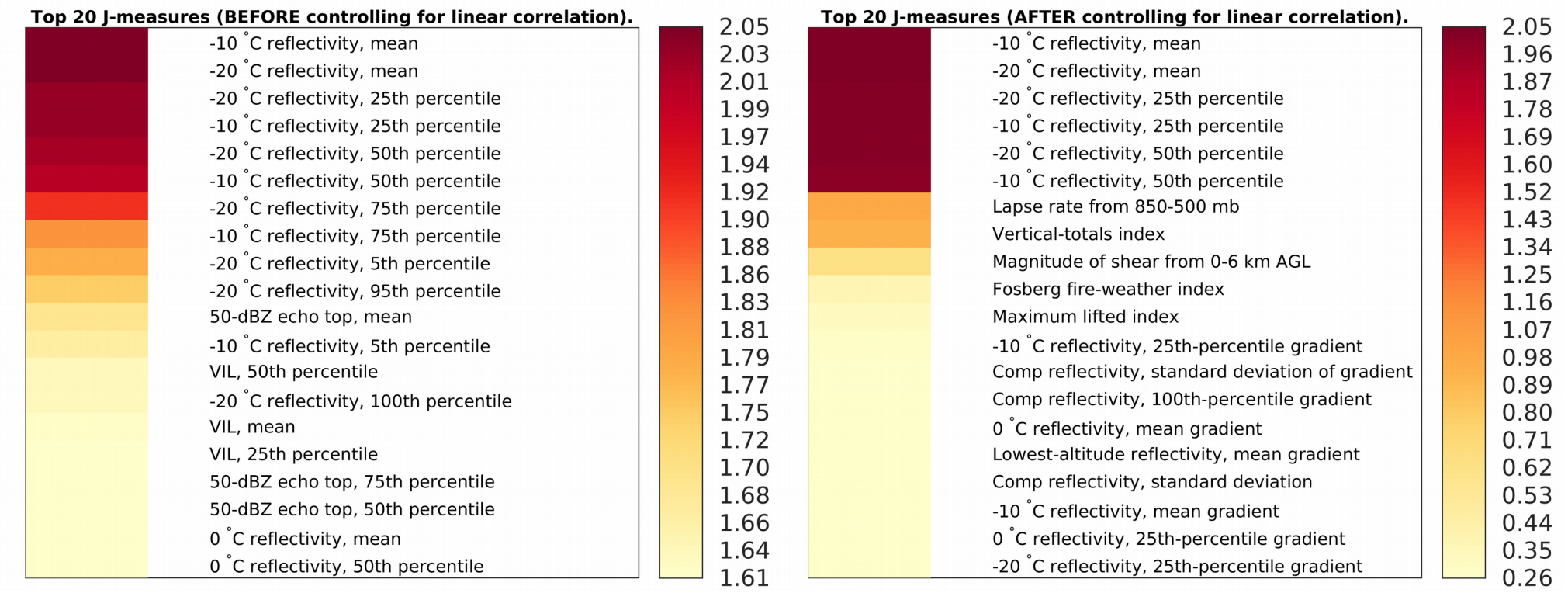
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 10 km and lead time of [15, 30] minutes.



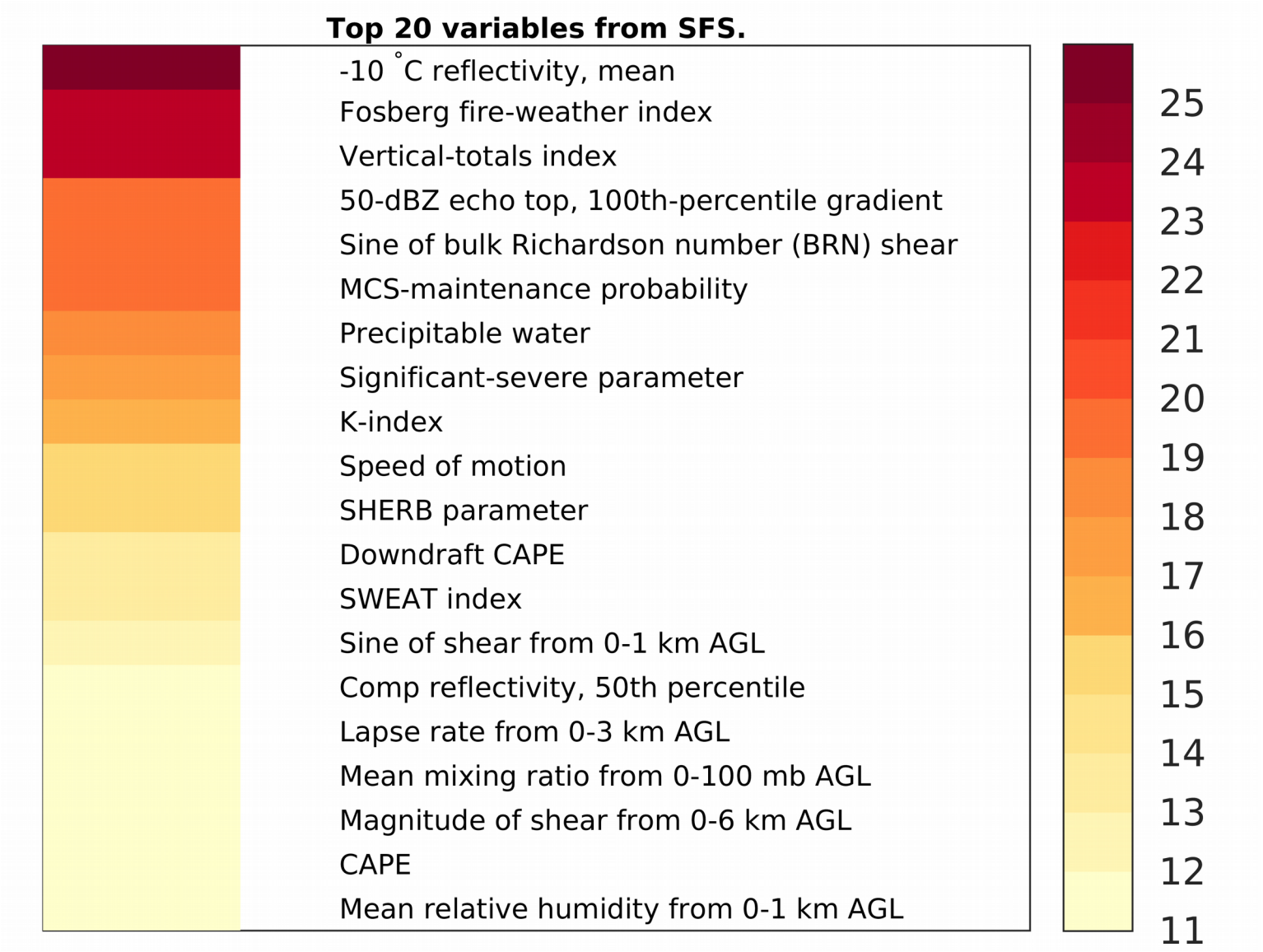
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 10 km and lead time of [30, 45] minutes.



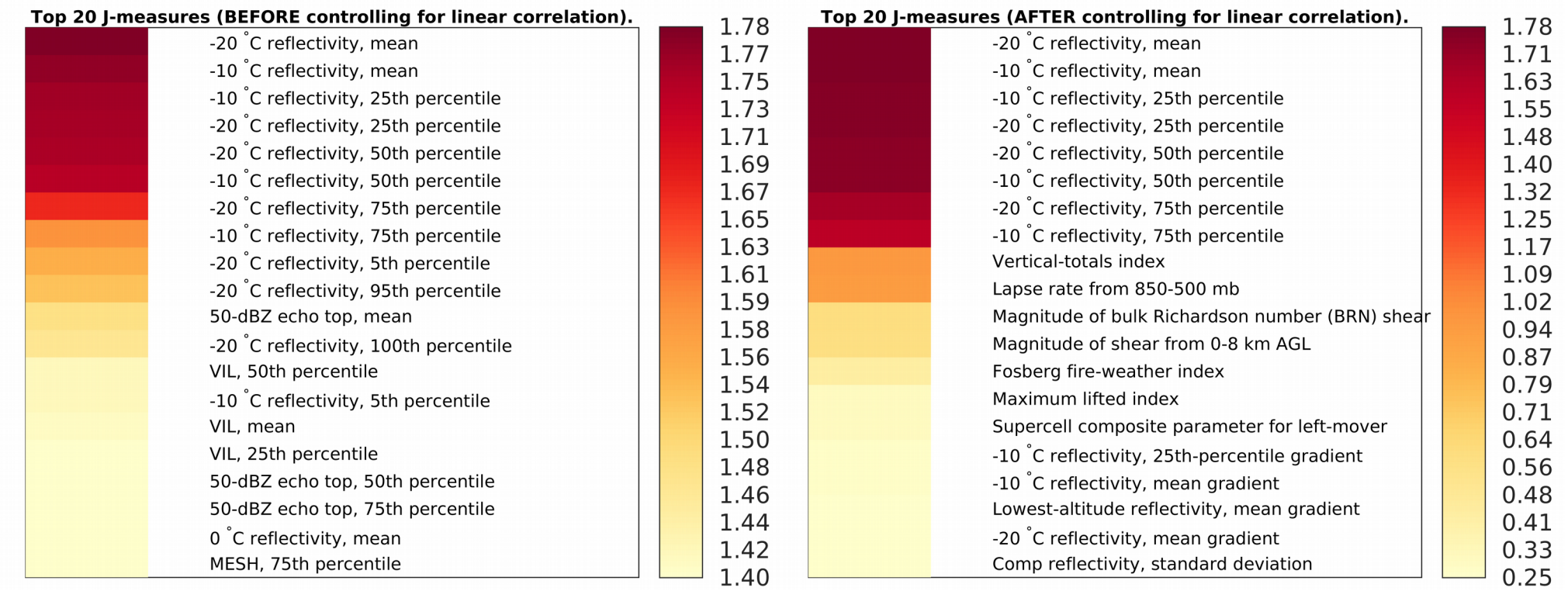
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- SFS results for buffer distance of 10 km and lead time of [30, 45] minutes.



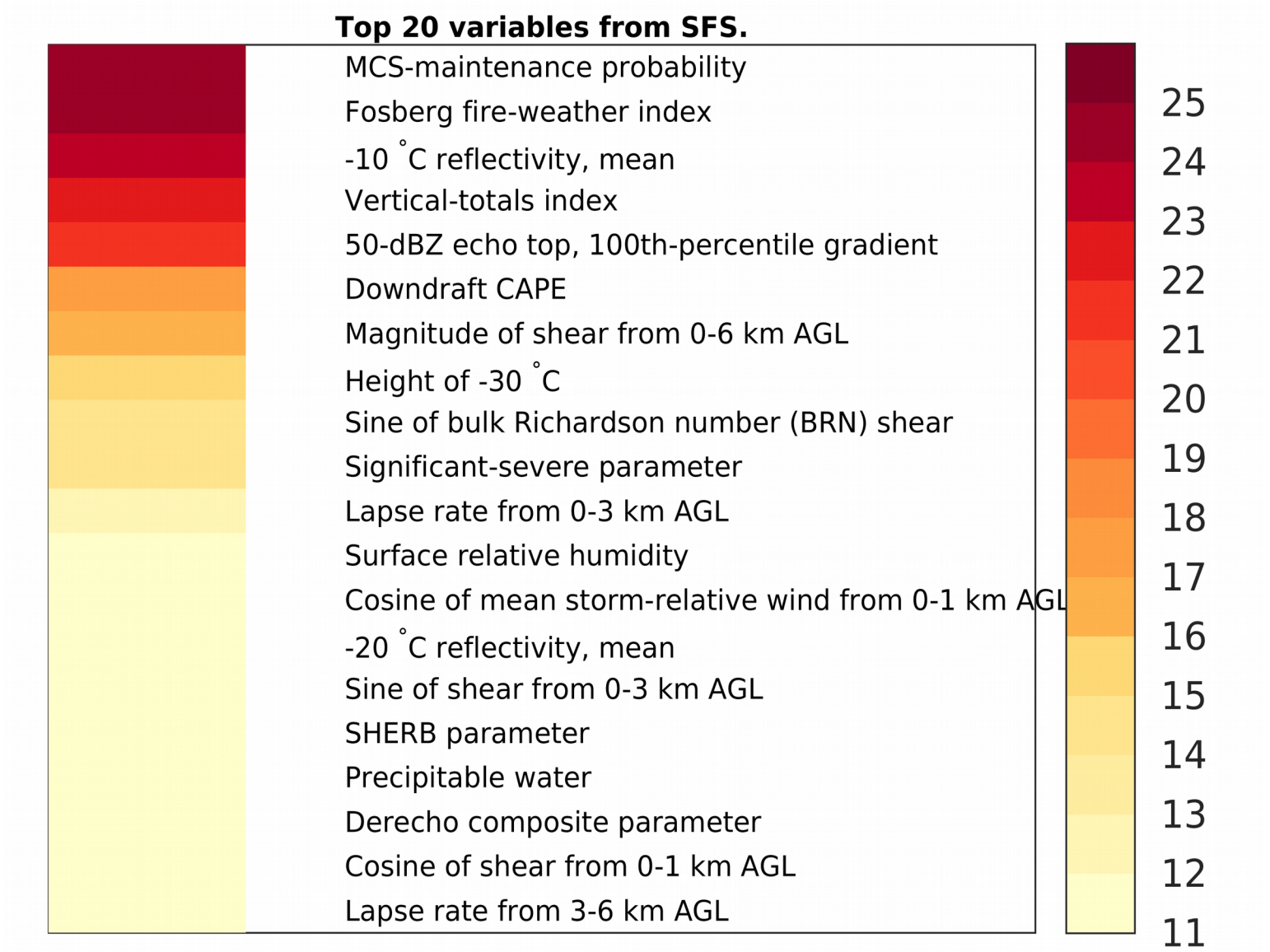
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 10 km and lead time of [45, 60] minutes.



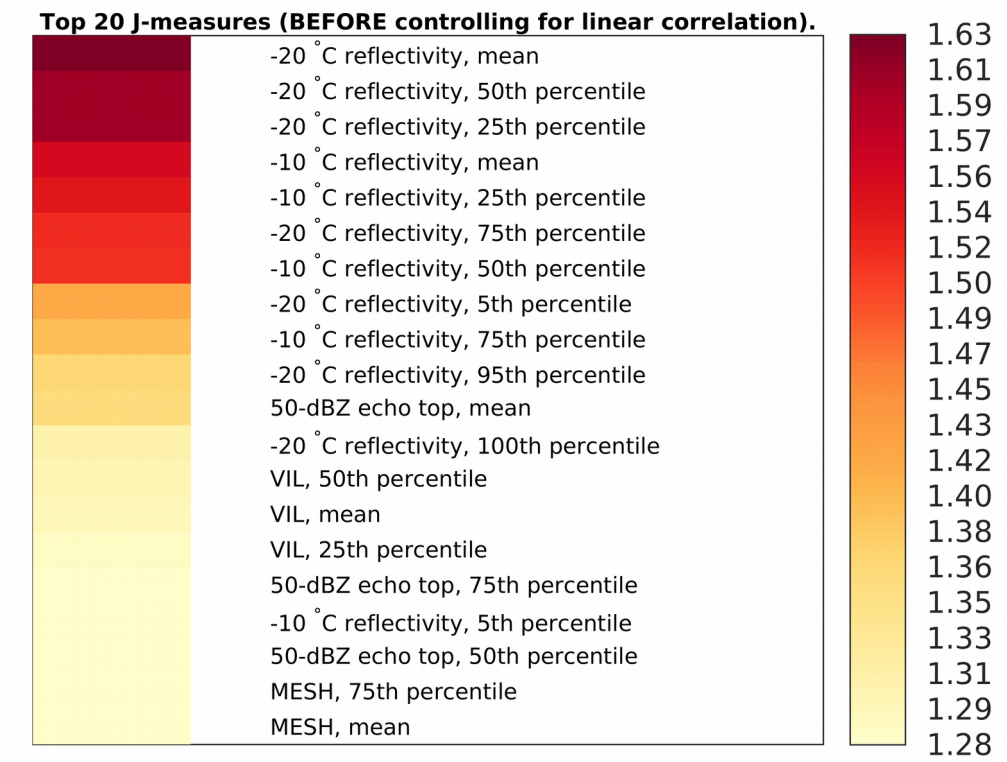
Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

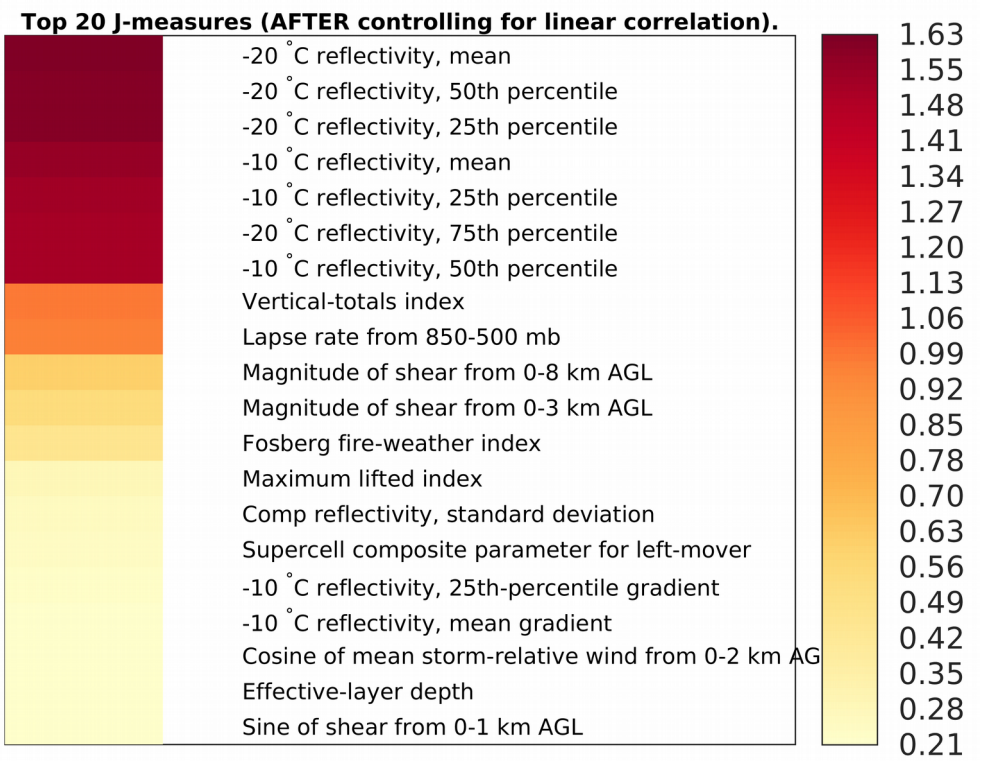
- SFS results for buffer distance of 10 km and lead time of [45, 60] minutes.



Variable Importance for Damaging Straight-line Storm Winds
Appendix B: Variable Importance

- *J*-measure results for buffer distance of 10 km and lead time of [60, 90] minutes.





Variable Importance for Damaging Straight-line Storm Winds

Appendix B: Variable Importance

- SFS results for buffer distance of 10 km and lead time of [60, 90] minutes.

Results forthcoming… sorry!

Variable Importance for Damaging Straight-line Storm Winds
Appendix C: Sounding Parameters

Variable	Description	Units	Vector?
brn	Bulk Richardson number	None	
brnDenom	BRN denominator	$m^2\ s^{-2}$	
brnShear	BRN shear term	$m\ s^{-1}$	Yes
capStrength	Cap strength	K	
cape	Convective available potential energy	$J\ kg^{-1}$	
cape3km	CAPE from 0-3 km above ground level (AGL)	$J\ kg^{-1}$	
cape6km	CAPE from 0-6 km AGL	$J\ kg^{-1}$	
capeFreezing	CAPE from surface – freezing level	$J\ kg^{-1}$	
cin	Convective inhibition	$J\ kg^{-1}$	
convectiveTemp	Convective temperature	K	
critAngle	Critical angle	°	
crossTotals	Cross-totals index	K	
dcape	Downdraft CAPE	$J\ kg^{-1}$	
dcp	Derecho composite parameter	None	
effBwd	Effective bulk wind difference	$m\ s^{-1}$	Yes
effLayerBottom	Effective-layer bottom	m	

Variable Importance for Damaging Straight-line Storm Winds
Appendix C: Sounding Parameters

Variable	Description	Units	Vector?
effLayerDepth	Effective-layer depth	m	
effLayerTop	Effective-layer top	m	
effShear	Effective-layer shear	$m\ s^{-1}$	Yes
ehi1km	Energy helicity index from 0-1 km AGL	$J\ kg^{-1}$	
ehi3km	EHI from 0-3 km AGL	$J\ kg^{-1}$	
ehiLeft	EHI for left-mover	$J\ kg^{-1}$	
ehiRight	EHI for right-mover	$J\ kg^{-1}$	
elHeight	Height AGL of equilibrium level (EL)	m	
esp	Enhanced stretching potential	None	
fosberg	Fosberg fire-weather index	None	
height0C	Height of 0 °C level	m	
height-10C	Height of -10 °C level	m	
height-20C	Height of -20 °C level	m	
height-30C	Height of -30 °C level	m	
kIndex	K-index	K	
lapseRate3km	Lapse rate from 0-3 km AGL	$K\ km^{-1}$	

Variable Importance for Damaging Straight-line Storm Winds
Appendix C: Sounding Parameters

Variable	Description	Units	Vector?
lapseRate3-6km	Lapse rate from 3-6 km AGL	K km^{-1}	
lapseRate700-500mb	Lapse rate from 700-500 mb	K km^{-1}	
lapseRate850-500mb	Lapse rate from 850-500 mb	K km^{-1}	
lclHeight	Height AGL of lifting condensation level (LCL)	m	
lfcHeight	Height AGL of level of free convection	m	
lhp	Large-hail parameter	None	
li300mb	Lifted index from surface – 300 mb	K	
li500mb	Lifted index from surface – 500 mb	K	
liMax	Maximum lifted index (surface – any level)	K	
maxWindPbl	Maximum boundary-layer wind	m s^{-1}	Yes
mburst	Microburst index	None	
meanEffBulkWind	Mean effective bulk wind	m s^{-1}	Yes
meanEffWind	Mean effective-layer wind	m s^{-1}	Yes
meanMixr100mb	Mean mixing ratio from 0-100 mb AGL	g kg^{-1}	
meanRh1km	Mean relative humidity (RH) from 0-1 km AGL	%	
meanRh150mb	Mean RH from 0-150 mb AGL	%	

Variable Importance for Damaging Straight-line Storm Winds
Appendix C: Sounding Parameters

Variable	Description	Units	Vector?
meanRh150-350mb	Mean RH from 150-350 mb AGL	%	
meanRhPbl	Mean boundary-layer RH	%	
meanWind1km	Mean wind from 0-1 km AGL	$m\ s^{-1}$	Yes
meanWind3km	Mean wind from 0-3 km AGL	$m\ s^{-1}$	Yes
meanWind6km	Mean wind from 0-6 km AGL	$m\ s^{-1}$	Yes
meanWind8km	Mean wind from 0-7 km AGL	$m\ s^{-1}$	Yes
meanWindLclEl	Mean wind from LCL-EL	$m\ s^{-1}$	Yes
meanWindPbl	Mean boundary-layer wind	$m\ s^{-1}$	Yes
minBuoyancy	Minimum buoyancy in column	K	
mmp	Mesoscale convective system (MCS)-maintenance probability	%	
mplHeight	Maximum parcel height	m	
pblDepth	Boundary-layer depth	m	
pw	Precipitable water	mm	
rhSurface	Surface RH	%	
scpLeft	Supercell composite parameter (SCP) for left-mover	None	
scpRight	SCP for right-mover	None	

Variable Importance for Damaging Straight-line Storm Winds
Appendix C: Sounding Parameters

Variable	Description	Units	Vector?
shear1km	Wind shear from 0-1 km AGL	m s^{-1}	Yes
shear3km	Wind shear from 0-3 km AGL	m s^{-1}	Yes
shear6km	Wind shear from 0-6 km AGL	m s^{-1}	Yes
shear8km	Wind shear from 0-8 km AGL	m s^{-1}	Yes
shear9km	Wind shear from 0-9 km AGL	m s^{-1}	Yes
shearLclEl	Wind shear from LCL-EL	m s^{-1}	Yes
sherb	Severe hazards in environments with reduced buoyancy (SHERB) parameter	None	
ship	Significant-hail parameter	None	
sigSevere	Significant-severe parameter	None	
srh1km	Storm-relative helicity (SRH) from 0-1 km AGL	J kg^{-1}	
srh3km	SRH from 0-3 km AGL	J kg^{-1}	
srhLeft	SRH for left-mover	J kg^{-1}	
srhRight	SRH for right-mover	J kg^{-1}	
srw1km	Mean storm-relative wind (SRW) from 0-1 km AGL	m s^{-1}	Yes
srw2km	Mean SRW from 0-2 km AGL	m s^{-1}	Yes
srw3km	Mean SRW from 0-3 km AGL	m s^{-1}	Yes

Variable Importance for Damaging Straight-line Storm Winds
Appendix C: Sounding Parameters

Variable	Description	Units	Vector?
srw4-5km	Mean SRW from 4-5 km AGL	$m\ s^{-1}$	Yes
srw4-6km	Mean SRW from 4-6 km AGL	$m\ s^{-1}$	Yes
srw6km	Mean SRW from 0-6 km AGL	$m\ s^{-1}$	Yes
srw8km	Mean SRW from 0-8 km AGL	$m\ s^{-1}$	Yes
srw9-11km	Mean SRW from 9-11 km AGL	$m\ s^{-1}$	Yes
srwBulk	Mean bulk SRW	$m\ s^{-1}$	Yes
srwEff	Mean effective-layer SRW	$m\ s^{-1}$	Yes
srwLclEl	Mean SRW from LCL-EL	$m\ s^{-1}$	Yes
stpEff	Significant-tornado parameter (STP) for effective layer	None	
stpFixed	STP for fixed layer	None	
sweat	SWEAT index	None	
thetaeDiff	Difference between min and max equiv potential temperature (θ_e) from 0-3 km AGL	K	
thetaeIndex	θ_e-index	K	
totalTotals	Total-totals index	K	
updraftTilt	Updraft tilt	°	

Variable Importance for Damaging Straight-line Storm Winds
Appendix C: Sounding Parameters

Variable	Description	Units	Vector?
`verticalTotals`	Vertical-totals index	K	
`wdi`	Wind-damage index	None	