

SDGE Station Level Wind Gust Machine Learning Progress



Update:

1. Remove RH in model

2. Use percentile table generated from SDGE observation data instead of WRF data

- linear interpolate between 95th and 99th percentile

3. Validation plot:

- Add WRF wind gust output (light blue line)
- Add BIAS/RMSE stats
- Replace validation dates (SDGE provide)
- Start with 10 PSPS stations

4. Preliminary study for ens_cfsr_004_d02 data:

- temp_grad_surf, wind_spd_grad_surf, Ri_surf
- wind_dir_500mb
- u_500m, u_1000m, u_1500m, u_2000m, u_2500m, u_3000m
- v_500m, v_1000m, v_1500m, v_2000m, v_2500m, v_3000m
- temp_500m, temp_1000m, temp_1500m, temp_2000m, temp_2500m, temp_3000m

1) temp_grad_surf: Vertical temperature gradient near the surface, which is defined as the ratio of the difference in potential temperature between 2 and 100m to the distance between 2 and 100m (i.e., 98m). Unit: K m⁻¹.

2) wind_spd_grad_surf: Vertical gradient of wind speed near the surface, which is defined as the ratio of the difference in wind speed between 10 and 100m to the distance between 10 and 100m (i.e., 90m). Unit: s⁻¹.

3) Ri_surf: Richardson number near the surface, as defined by eq. (2) in Rob's paper. It is basically a function of temp_grad_surf and wind_spd_grad_surf.

$$Ri = \frac{g_0}{\Theta} \frac{\partial \Theta / \partial z}{(\partial \bar{V} / \partial z)^2},$$

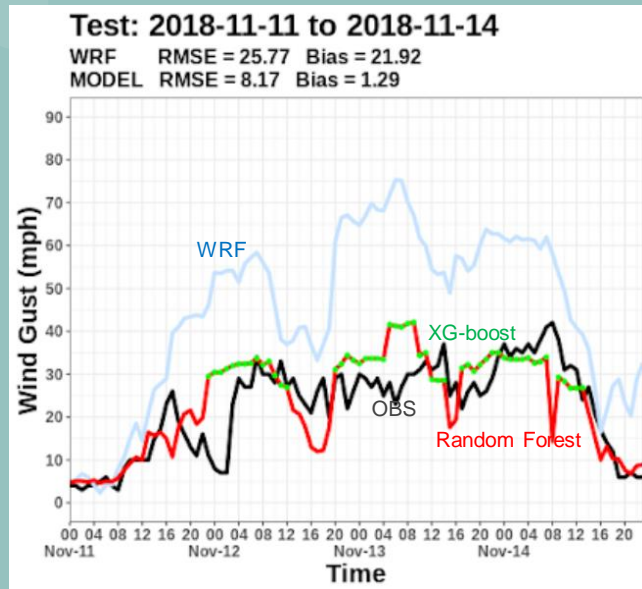
4) wind_dir_500mb: Wind direction at 500 mb. Unit: degree.

5) u_500m, u_1000m, u_1500m, u_2000m, u_2500m, u_3000m: zonal wind at different heights above ground.

6) v_500m, v_1000m, v_1500m, v_2000m, v_2500m, v_3000m: meridional wind at different heights above ground. Unit: m s⁻¹.

7) temp_500m, temp_1000m, temp_1500m, temp_2000m, temp_2500m, temp_3000m: temperature at different heights above ground. Unit: K

Example for new validation plot



Top 10 most likely to result in a PSPS:

Sill Hill (SIL)
 Julian (JUL)
 Pine Hills (PIH)
 Crestline (CLN)
 Ramona (RAM)
 Creelman (CLM)
 Hideaway Lake (HID)
 De Luz (DLZ)
 Highland Valley (HVY)
 Crestwood (CWD)

Top 10 most difficult to forecast:

Hellhole Canyon (HHC)
 Rancho Heights (RHS)
 Cuca Ranch (CCR)
 Rincon Reservation (RCR)
 Rincon Central (RCE)
 Ortega (ORT)
 Otay Mountain (OTM)
 Rancho Santa Fe (RSF)
 Barrett Junction (BRJ)
 Goose Valley (GOS)

Top 10 least difficult to forecast:

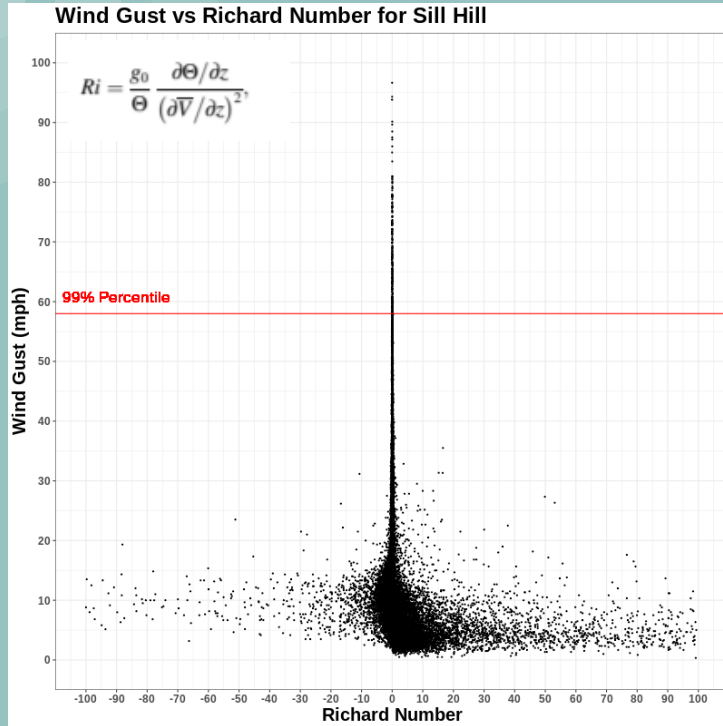
Sill Hill (SIL)
 Volcan Mountain (VCM)
 Hoskings Ranch (HOS)
 West Santa Ysabel (WSY)
 Lucky Five Ranch (LFR)
 East Willows Rd (EWR)
 North Potrero (NPT)
 Boulder Creek (BOC)
 Sherilton Valley (SHV)
 Pine Valley (PIV)

RFW and/or Extreme FPI dates:

Oct. 4-6, 2013
 Nov. 5, 2013
 Dec. 14-15, 2013
 Jan. 13-17, 2014
 Jan. 23-24, 2014
 Apr. 30-May 1, 2014
 May 11-15, 2014
 Nov. 4-6, 2014
 Nov. 16-18, 2014
 Nov. 24-26, 2014
 Apr. 15-16, 2015
 Oct. 9-10, 2017
 Oct. 14-15, 2017
 Oct. 22-25, 2017
 Dec. 4-10, 2017
 Dec. 14-15, 2017
 Jan. 27-29, 2018
 Jul. 6, 2018
 Oct. 15-16, 2018
 Oct. 19, 2018 -- wrf totally missed (other close grid cell)
 Nov. 8-9, 2018
 Nov. 11-14, 2018

Note: OC-only RFWs were only included if the event as a whole was significant (i.e. Apr. 30th)

Example for Importance Rank



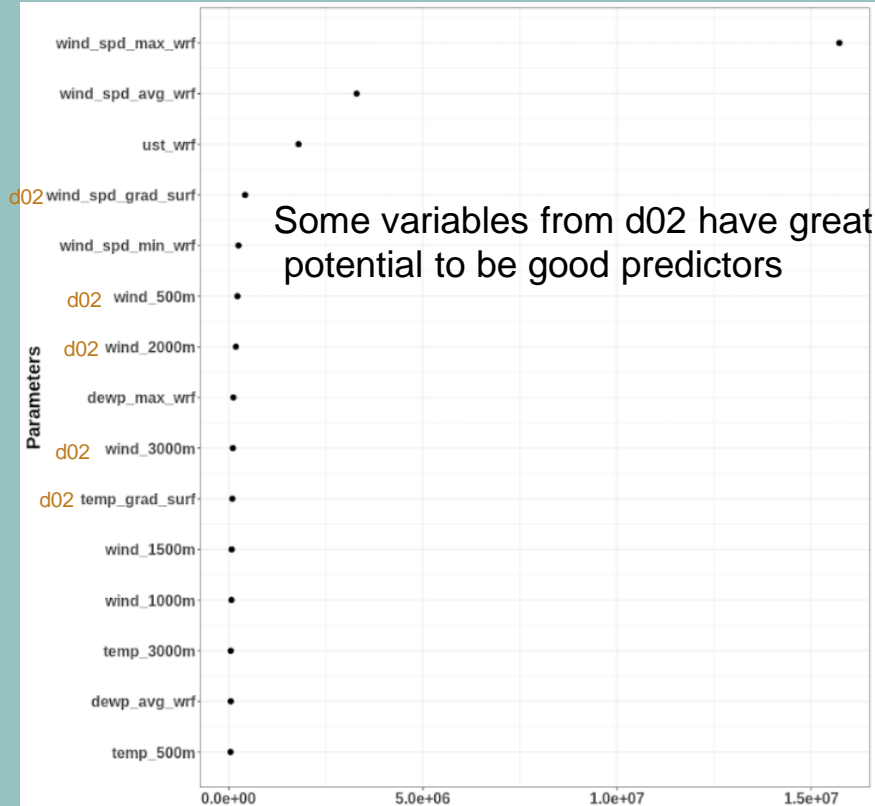
As indicated in Rob's paper, extreme wind event always happens when RN is around zero;

However, a lot of small wind event also happens when RN is around zero;

-> might use RN as a threshold to activate XG-boost model

-> might categorize RN as 1/0 as a new predictor

-> Instead of using RN, it might be better to use temp_grad_surf or wind_spd_grad_surf



SDGE Wind Gust ML Flowchart

Hourly WRF : ens_cfsr_004
ens_cfsr_004_d02
Period: 2013-07-31 to 2015-07-31 &
2017-07-31 to 2019-07-31

Response Variable: observed gust
Period: 2012-10-02 to 2019-09-02

Combined Training Data for Each Station

187 stations in total (164 stations are used)
15 stations with no data; 8 stations with less than 1yr data

Model I: Random Forest for All Wind Conditions
Predictor Selection/Model Parameter Selection
Rank Importance/Bias Pattern Check

Calculate bias with best predictor/parameter combination
Re-train bias with high wind condition in model II

Model II: Gradient Boosting Tree
Retrain Bias on High Wind Time-Points Only
Predictor/Parameter Selection; Resample Strategy

Determine high wind event threshold: wrf gust percentile
Bias correct high wind time-points only

Gust Prediction: bias corrected forecast value
combining model I and model II

Model Construction

SDG&E provides Major Events and
important stations

Check Model Results
Training Performance/Testing Performance
(use the most recent 2 events as testing)

Forecast vs Observed vs WRF Gust
Check Bias/RMSE
Check event/daily max difference
Check under- or over-estimation
Check time-lag issue/timeseries pattern

Model Validation

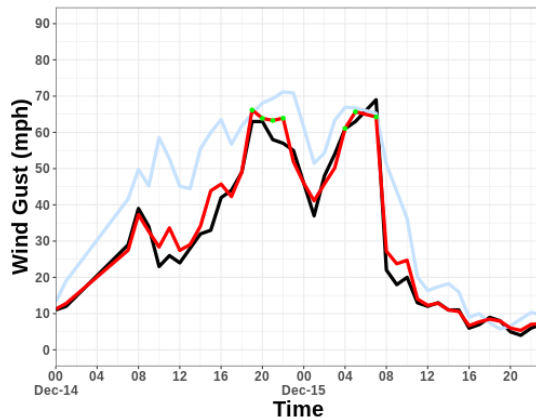
Validation Results for PSPS Stations

- For top 10 PSPS stations, the main WRF issue is overestimate big wind events

Wind Gust ML Validation for Sill Hill

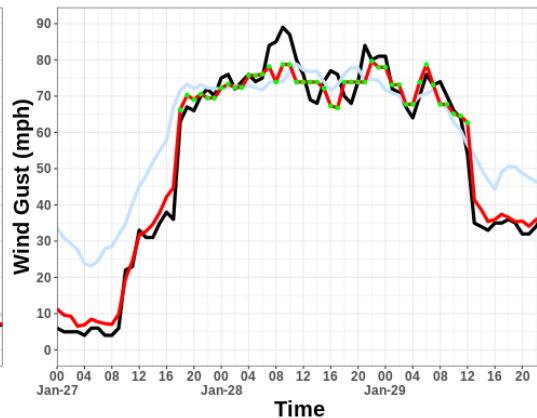
Train: 2017-12-14 to 2017-12-15

WRF RMSE = 14.21 Bias = 10.85
MODEL RMSE = 3.45 Bias = 1.41



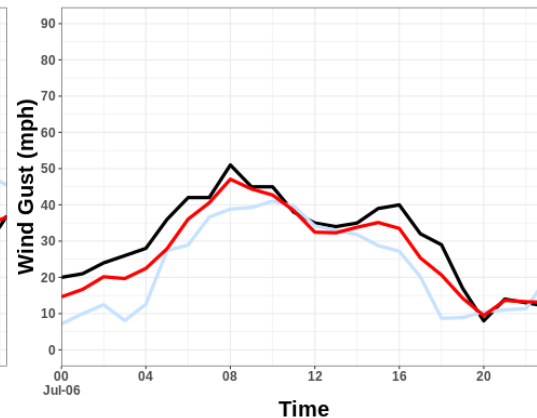
Train: 2018-01-27 to 2018-01-29

WRF RMSE = 13.03 Bias = 6.68
MODEL RMSE = 4.33 Bias = 0.47



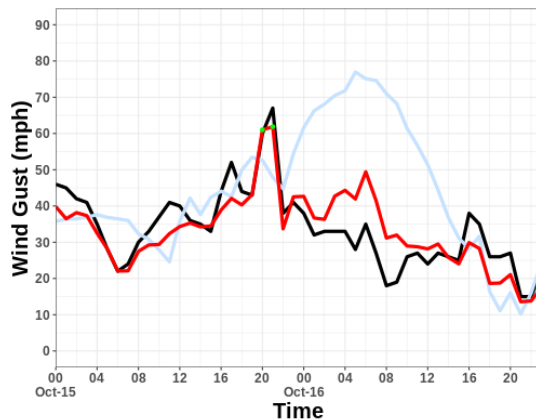
Train: 2018-07-06 to 2018-07-06

WRF RMSE = 10.06 Bias = -7.43
MODEL RMSE = 4.33 Bias = -3.26



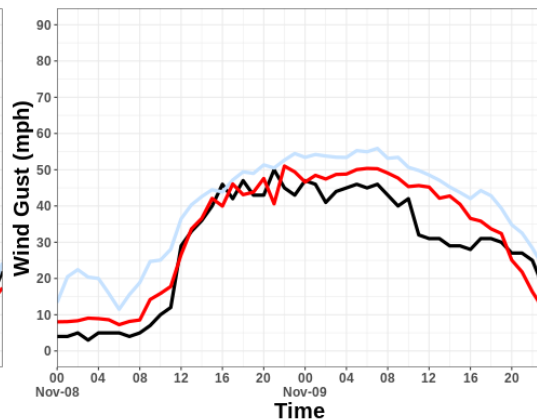
Train: 2018-10-15 to 2018-10-16

WRF RMSE = 22.06 Bias = 10.08
MODEL RMSE = 6.56 Bias = -0.20



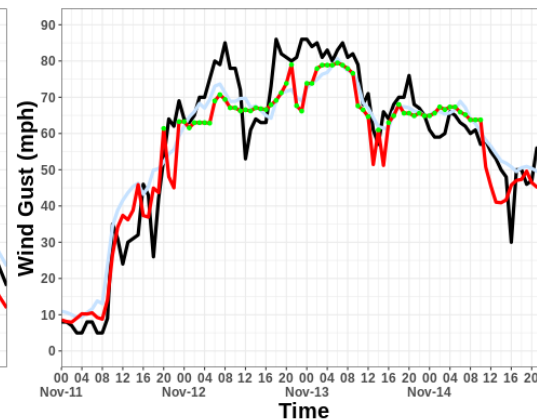
Test: 2018-11-08 to 2018-11-09

WRF RMSE = 11.32 Bias = 10.26
MODEL RMSE = 6.24 Bias = 3.46

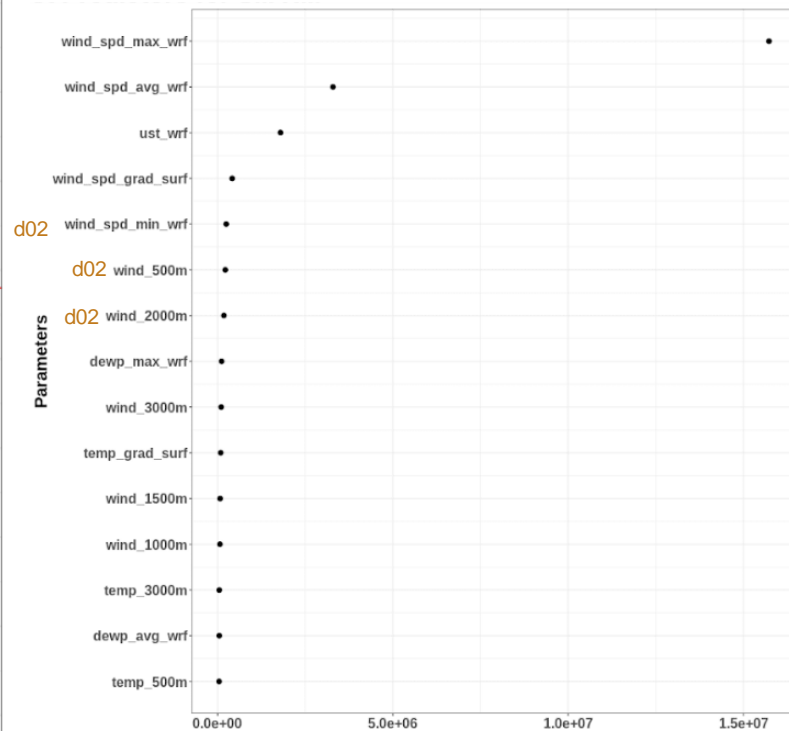
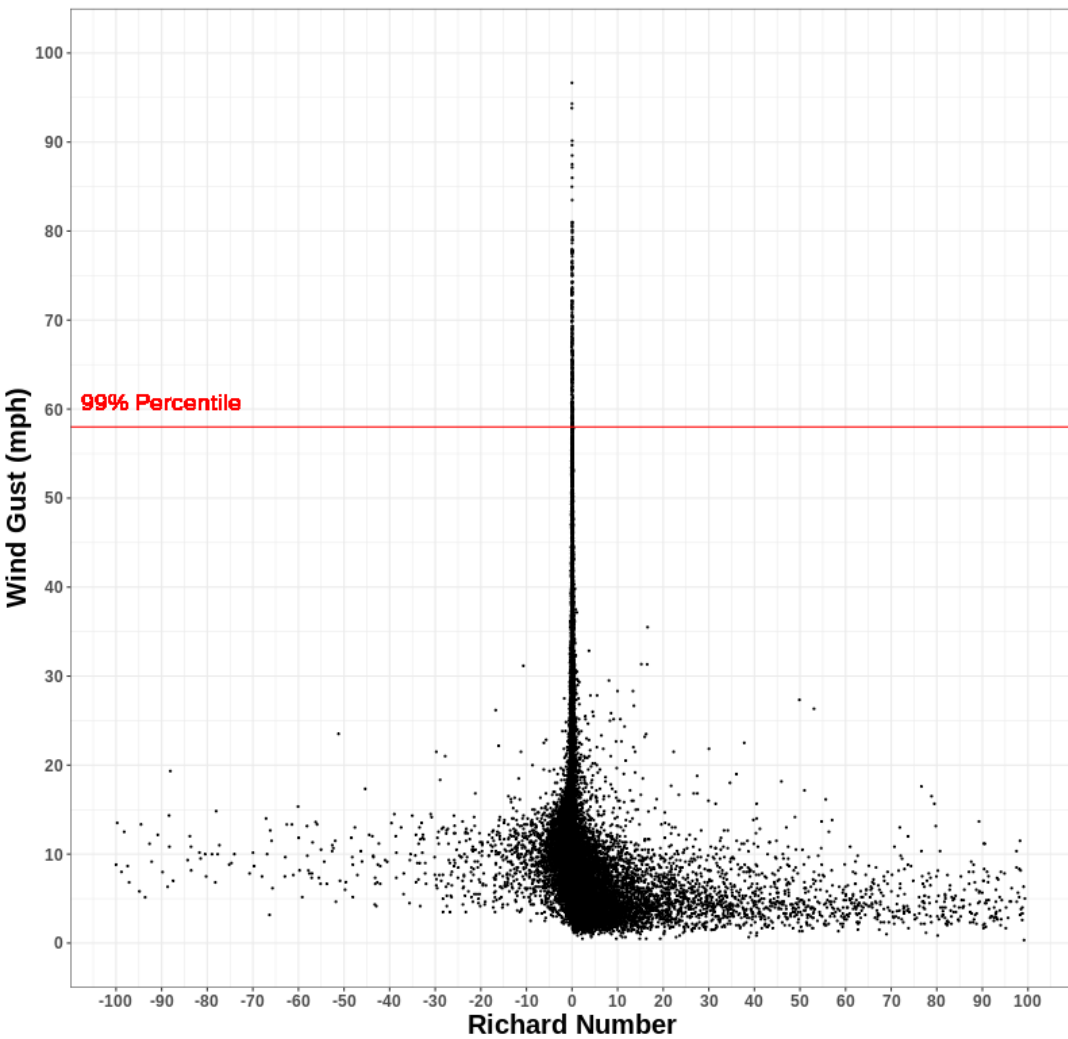


Test: 2018-11-11 to 2018-11-14

WRF RMSE = 7.72 Bias = -0.03
MODEL RMSE = 8.08 Bias = -2.34



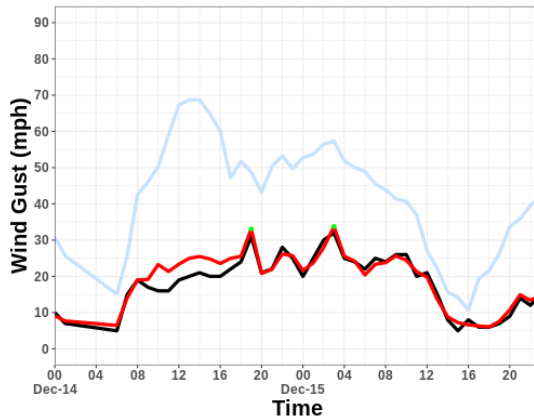
Wind Gust vs Richard Number for Sill Hill



Wind Gust ML Validation for Julian

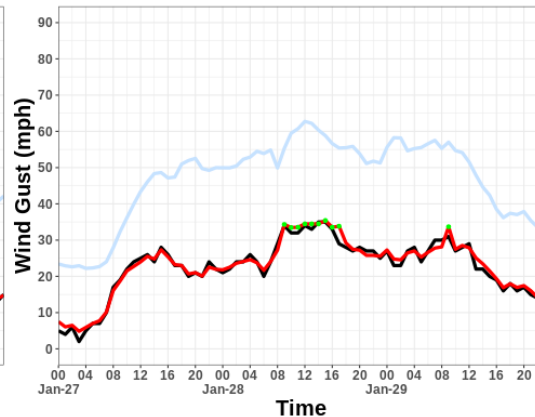
Train: 2017-12-14 to 2017-12-15

WRF RMSE = 26.46 Bias = 23.94
MODEL RMSE = 2.37 Bias = 0.98



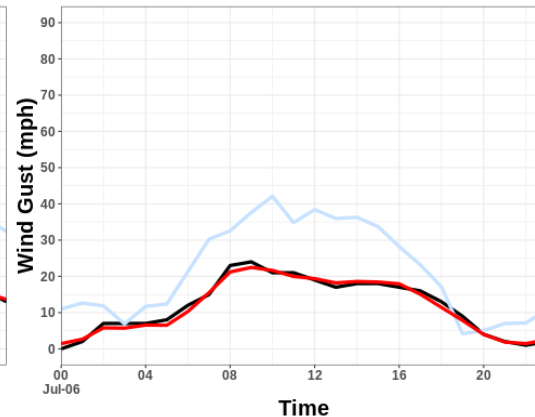
Train: 2018-01-27 to 2018-01-29

WRF RMSE = 24.80 Bias = 24.21
MODEL RMSE = 1.28 Bias = 0.35



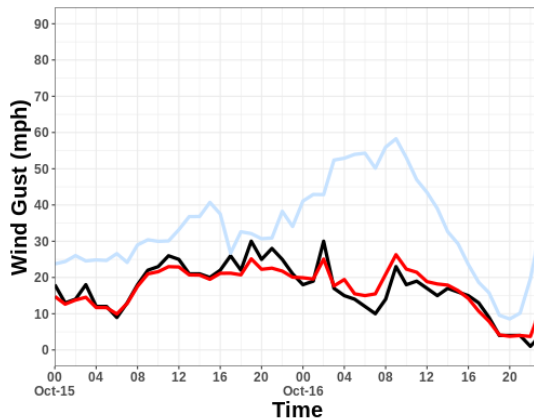
Train: 2018-07-06 to 2018-07-06

WRF RMSE = 11.57 Bias = 9.55
MODEL RMSE = 1.04 Bias = -0.27



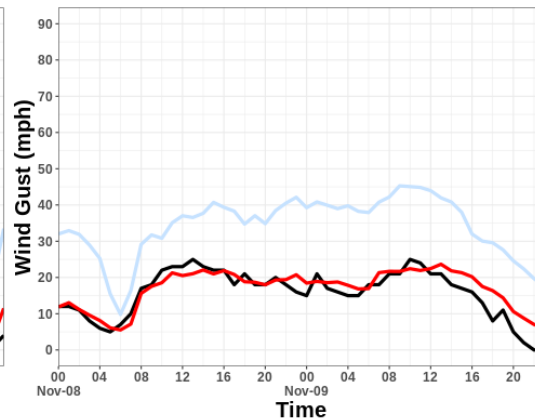
Train: 2018-10-15 to 2018-10-16

WRF RMSE = 20.53 Bias = 16.78
MODEL RMSE = 2.87 Bias = 0.03



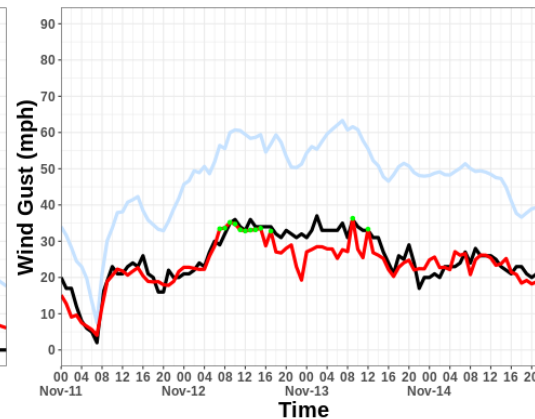
Test: 2018-11-08 to 2018-11-09

WRF RMSE = 19.08 Bias = 18.44
MODEL RMSE = 3.18 Bias = 1.26

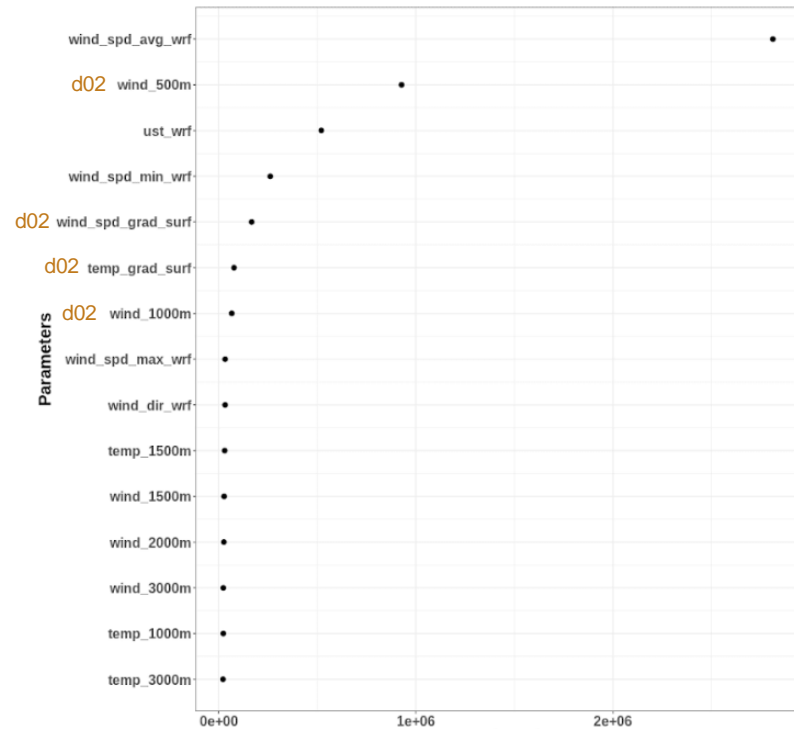
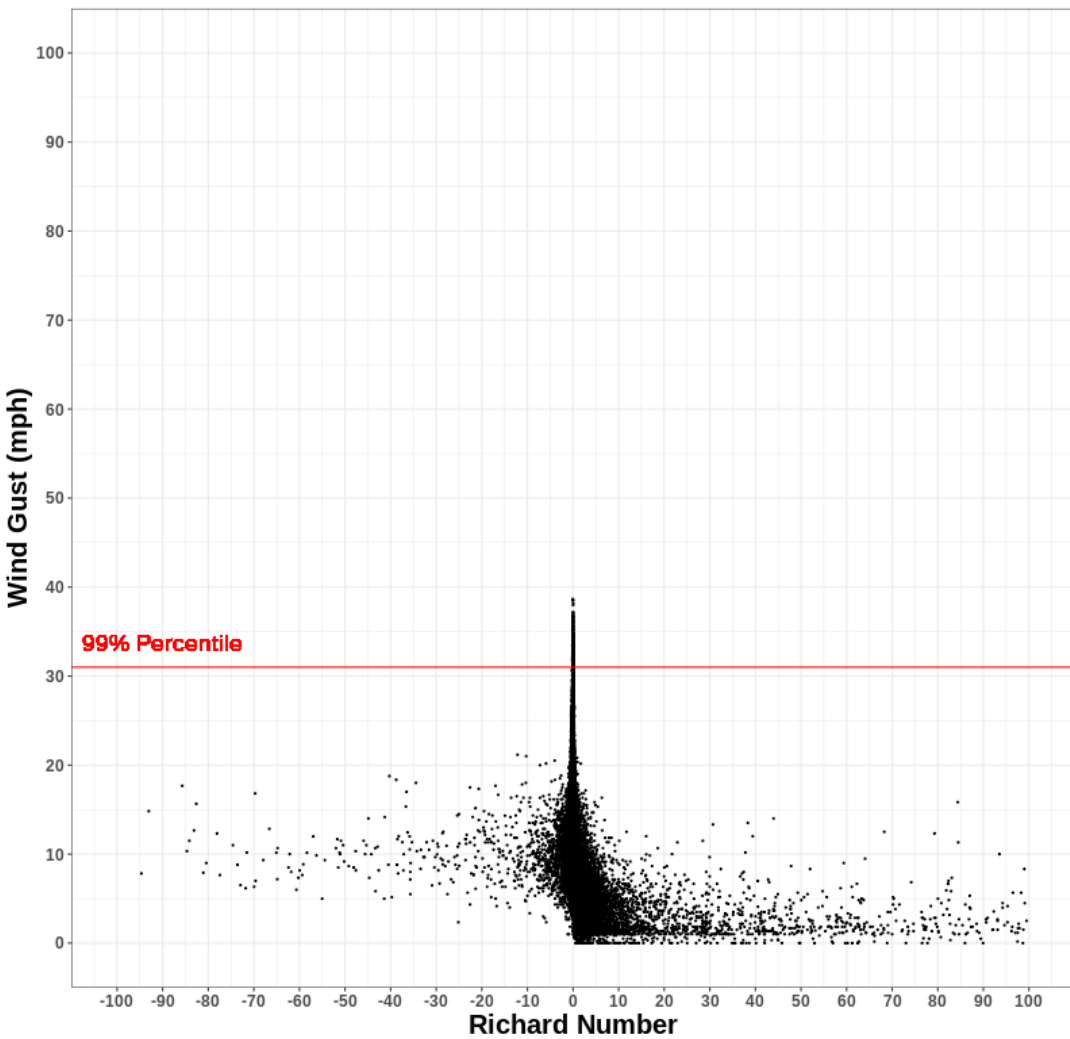


Test: 2018-11-11 to 2018-11-14

WRF RMSE = 22.03 Bias = 21.31
MODEL RMSE = 3.73 Bias = -1.76



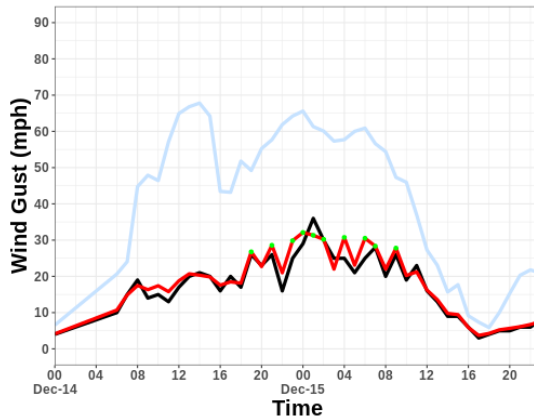
Wind Gust vs Richard Number for Julian



Wind Gust ML Validation for Pine Hills

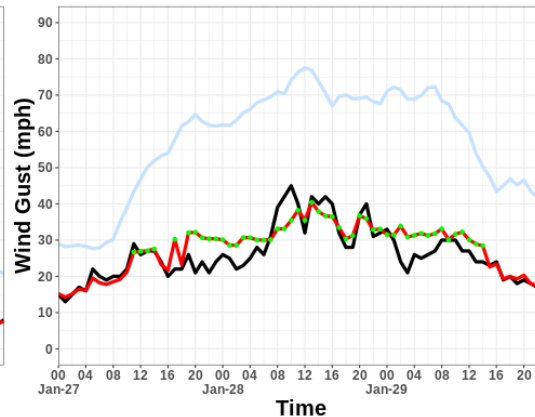
Train: 2017-12-14 to 2017-12-15

WRF RMSE = 28.38 Bias = 24.71
MODEL RMSE = 2.22 Bias = 0.91



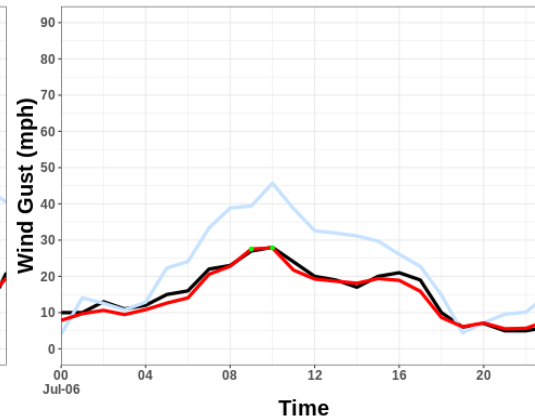
Train: 2018-01-27 to 2018-01-29

WRF RMSE = 32.80 Bias = 30.97
MODEL RMSE = 4.32 Bias = 1.53



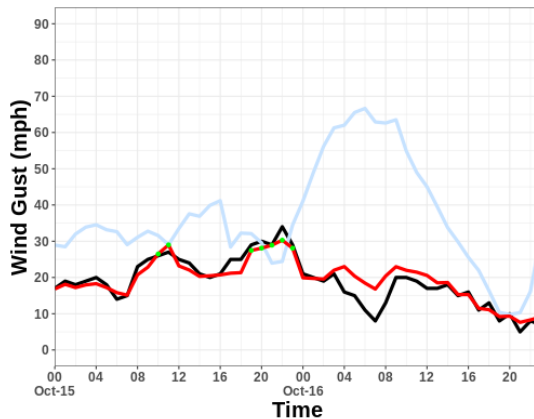
Train: 2018-07-06 to 2018-07-06

WRF RMSE = 9.24 Bias = 6.92
MODEL RMSE = 1.49 Bias = -0.81



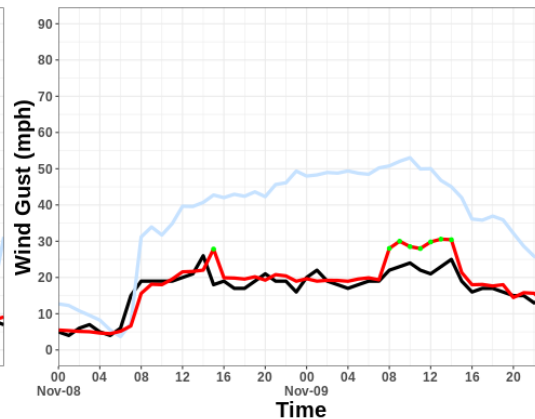
Train: 2018-10-15 to 2018-10-16

WRF RMSE = 23.93 Bias = 17.80
MODEL RMSE = 2.89 Bias = 0.58



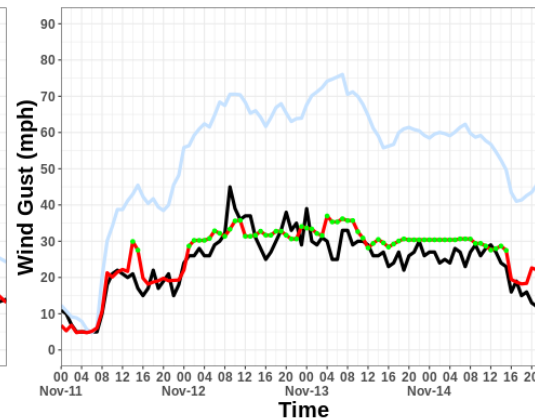
Test: 2018-11-08 to 2018-11-09

WRF RMSE = 21.96 Bias = 19.61
MODEL RMSE = 3.50 Bias = 1.33

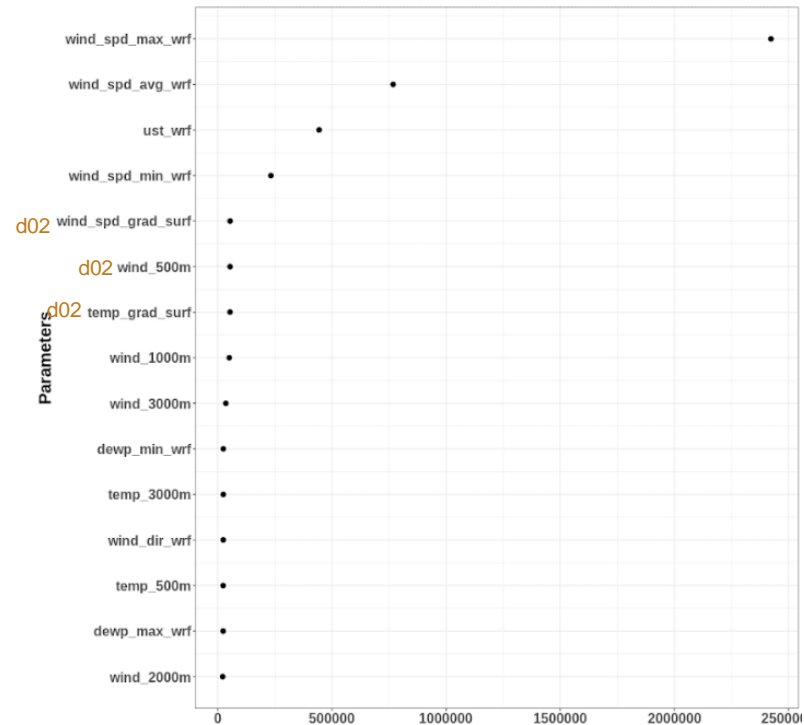
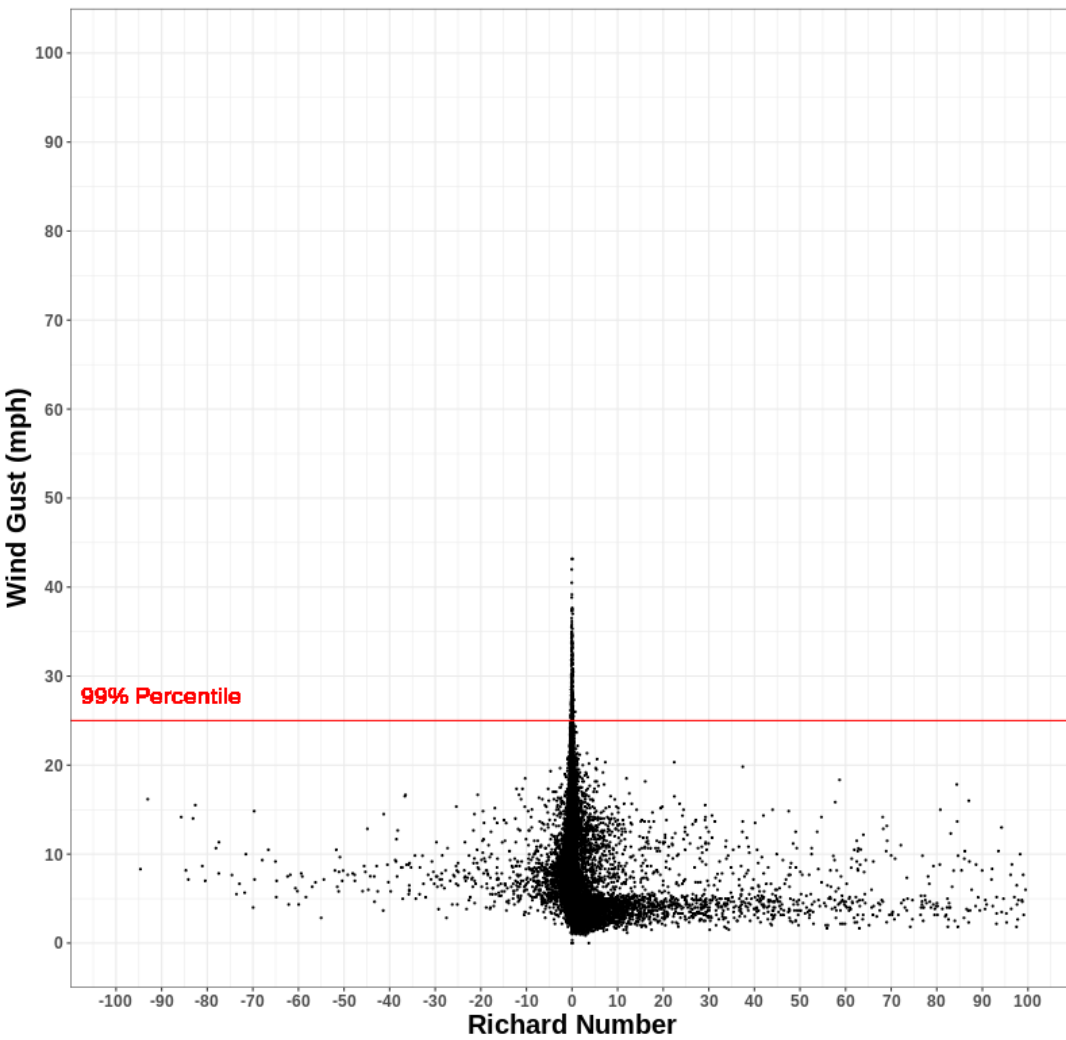


Test: 2018-11-11 to 2018-11-14

WRF RMSE = 30.99 Bias = 29.02
MODEL RMSE = 4.58 Bias = 2.28



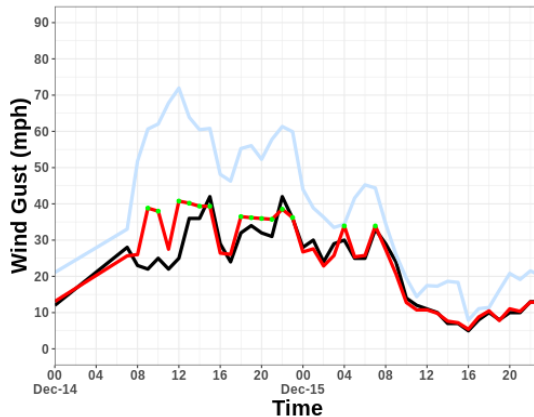
Wind Gust vs Richard Number for Pine Hills



Wind Gust ML Validation for Crestline

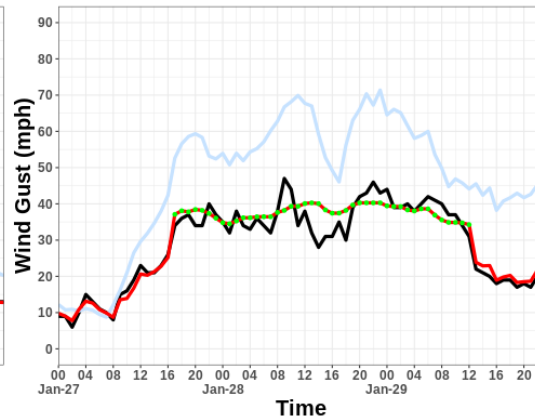
Train: 2017-12-14 to 2017-12-15

WRF RMSE = 19.51 Bias = 15.65
MODEL RMSE = 4.69 Bias = 1.50



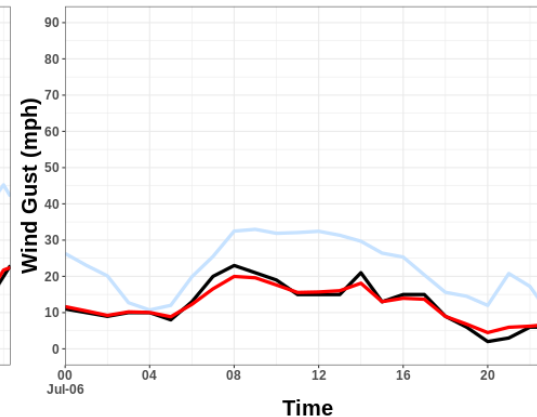
Train: 2018-01-27 to 2018-01-29

WRF RMSE = 19.75 Bias = 17.47
MODEL RMSE = 3.39 Bias = 0.52



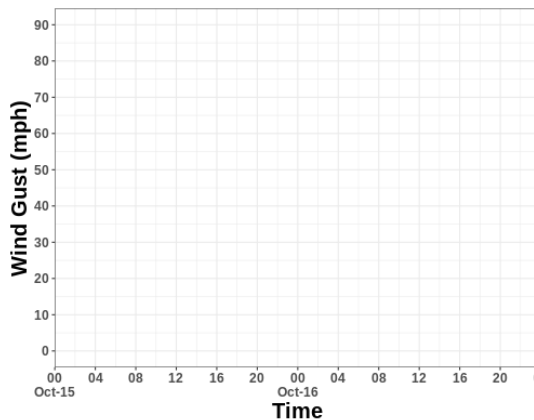
Train: 2018-07-06 to 2018-07-06

WRF RMSE = 11.09 Bias = 9.99
MODEL RMSE = 1.53 Bias = -0.13



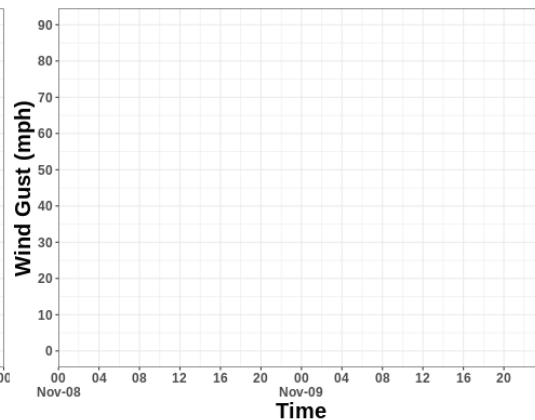
Train: 2018-10-15 to 2018-10-16

WRF RMSE = NaN Bias = NaN
MODEL RMSE = NaN Bias = NaN



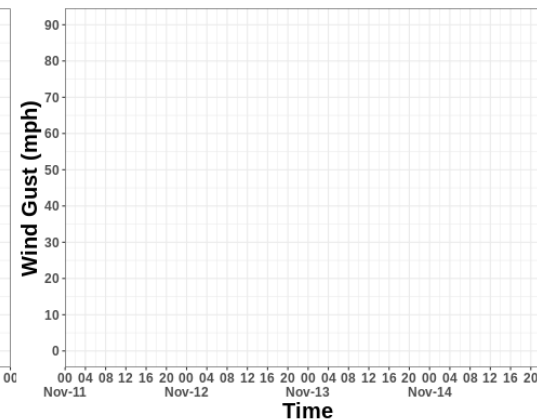
Test: 2018-11-08 to 2018-11-09

WRF RMSE = NaN Bias = NaN
MODEL RMSE = NaN Bias = NaN

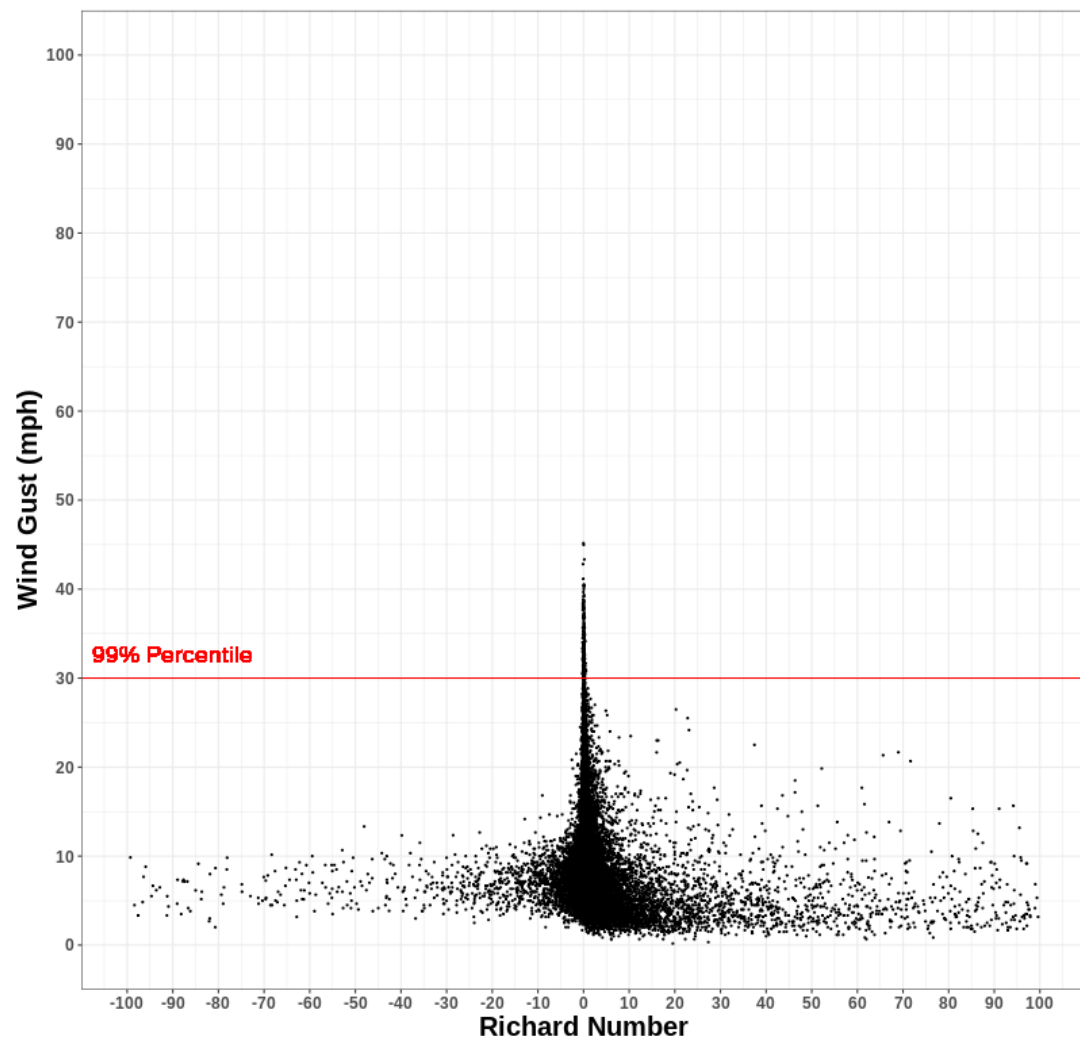


Test: 2018-11-11 to 2018-11-14

WRF RMSE = NaN Bias = NaN
MODEL RMSE = NaN Bias = NaN



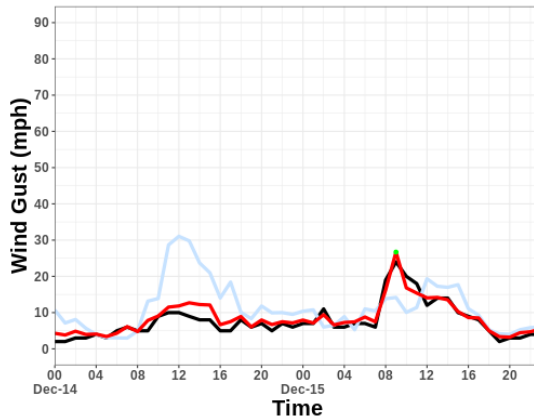
Wind Gust vs Richard Number for Crestline



Wind Gust ML Validation for Ramona

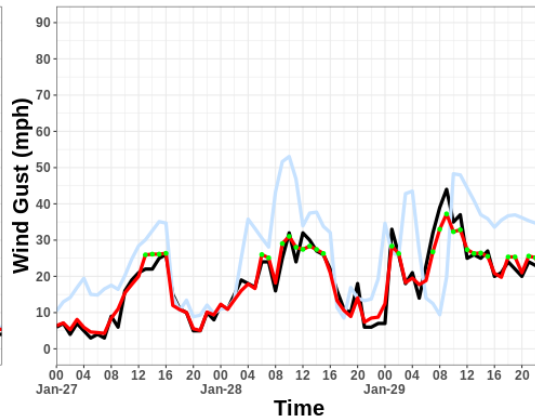
Train: 2017-12-14 to 2017-12-15

WRF RMSE = 7.56 Bias = 3.69
MODEL RMSE = 1.74 Bias = 0.86



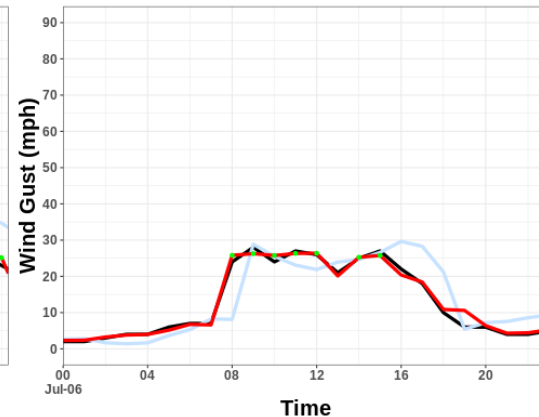
Train: 2018-01-27 to 2018-01-29

WRF RMSE = 12.90 Bias = 7.81
MODEL RMSE = 2.50 Bias = 0.02



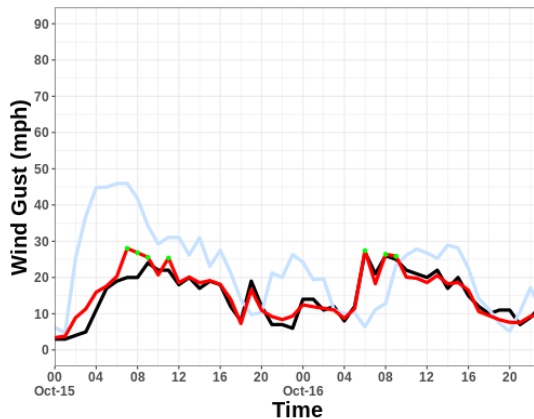
Train: 2018-07-06 to 2018-07-06

WRF RMSE = 5.27 Bias = 0.64
MODEL RMSE = 1.27 Bias = 0.19



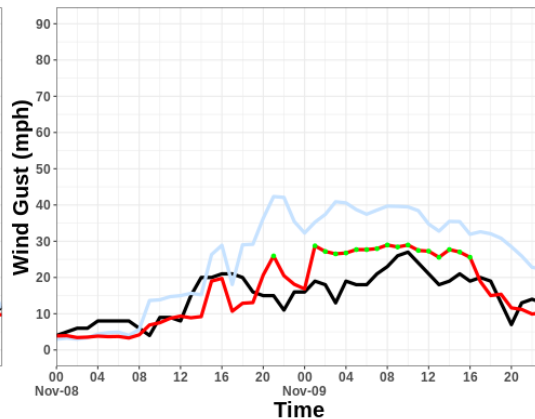
Train: 2018-10-15 to 2018-10-16

WRF RMSE = 13.28 Bias = 7.33
MODEL RMSE = 2.53 Bias = 0.55



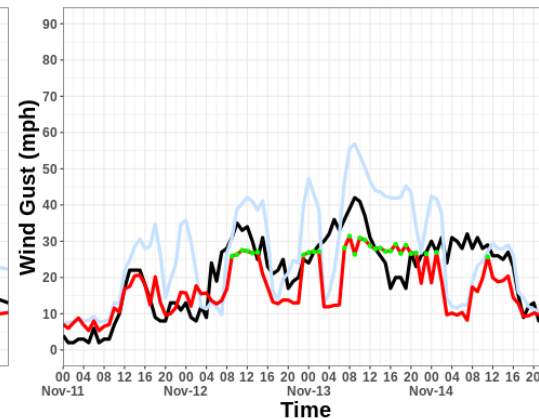
Test: 2018-11-08 to 2018-11-09

WRF RMSE = 14.28 Bias = 10.79
MODEL RMSE = 5.97 Bias = 1.57

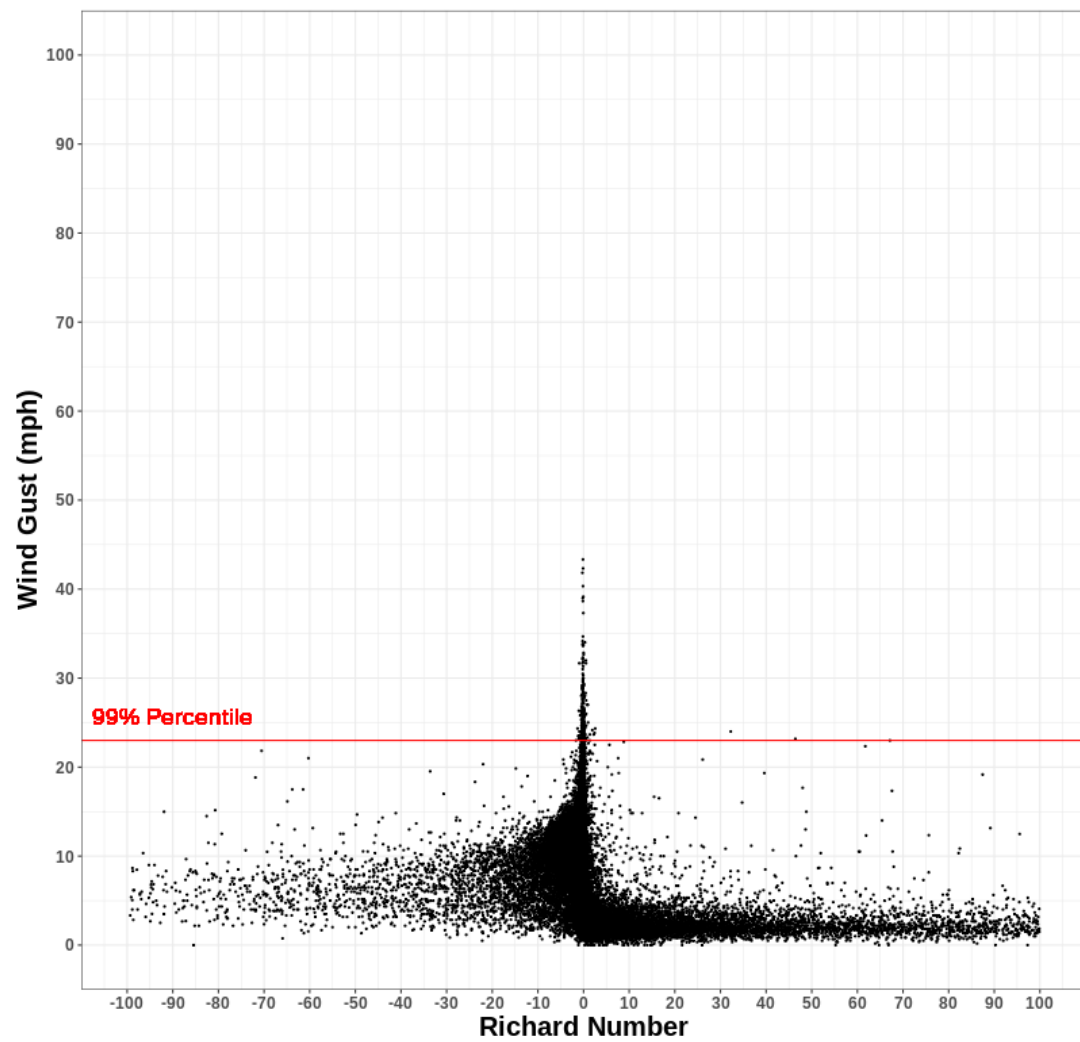


Test: 2018-11-11 to 2018-11-14

WRF RMSE = 12.01 Bias = 5.33
MODEL RMSE = 8.72 Bias = -3.51



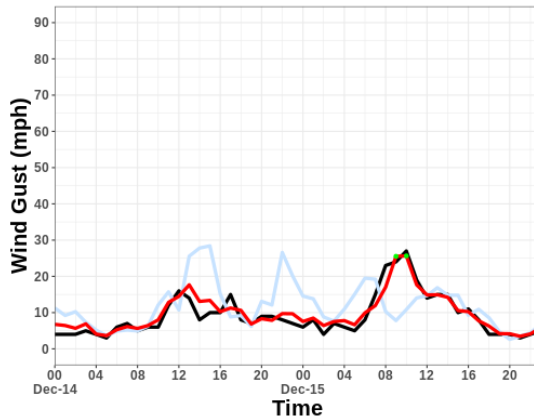
Wind Gust vs Richard Number for Ramona



Wind Gust ML Validation for Creelman

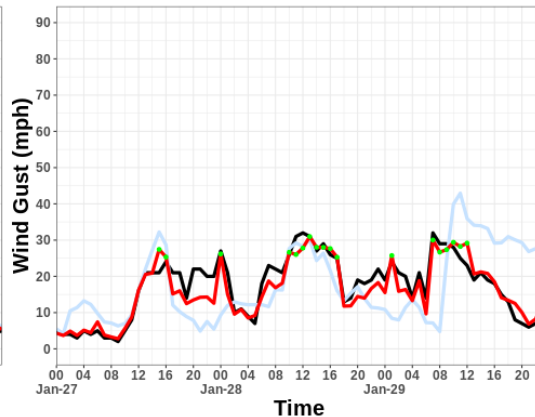
Train: 2017-12-14 to 2017-12-15

WRF RMSE = 7.73 Bias = 2.67
MODEL RMSE = 1.99 Bias = 0.58



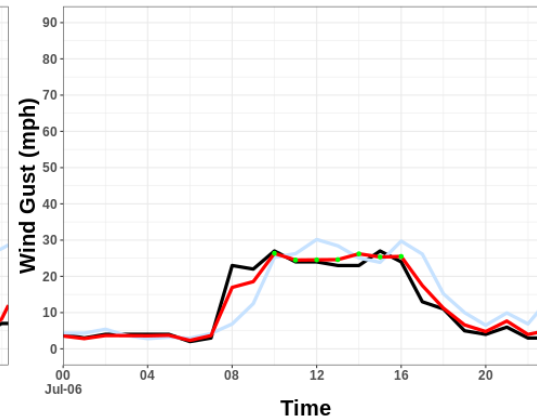
Train: 2018-01-27 to 2018-01-29

WRF RMSE = 10.99 Bias = 0.36
MODEL RMSE = 3.25 Bias = -0.94



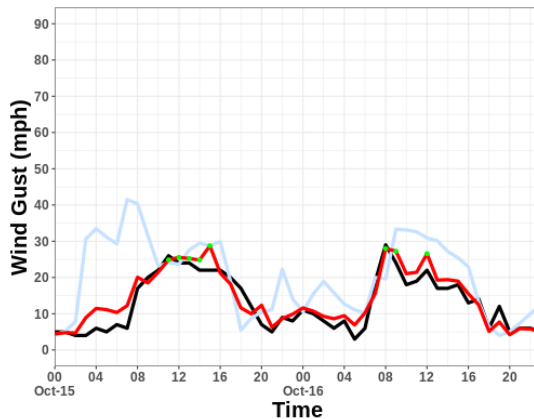
Train: 2018-07-06 to 2018-07-06

WRF RMSE = 5.95 Bias = 1.56
MODEL RMSE = 2.05 Bias = 0.27



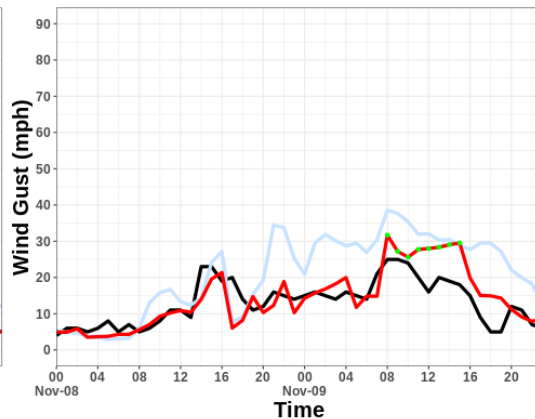
Train: 2018-10-15 to 2018-10-16

WRF RMSE = 11.88 Bias = 7.09
MODEL RMSE = 2.98 Bias = 1.26



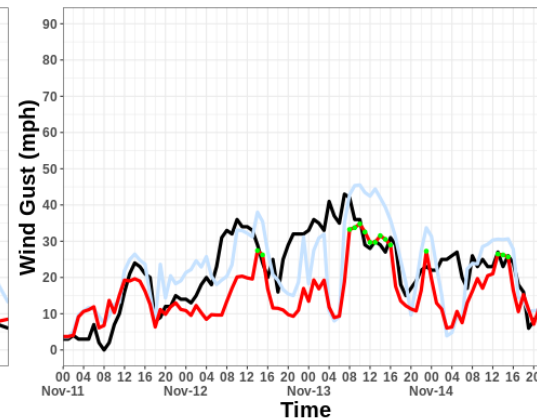
Test: 2018-11-08 to 2018-11-09

WRF RMSE = 11.25 Bias = 7.76
MODEL RMSE = 5.27 Bias = 1.08

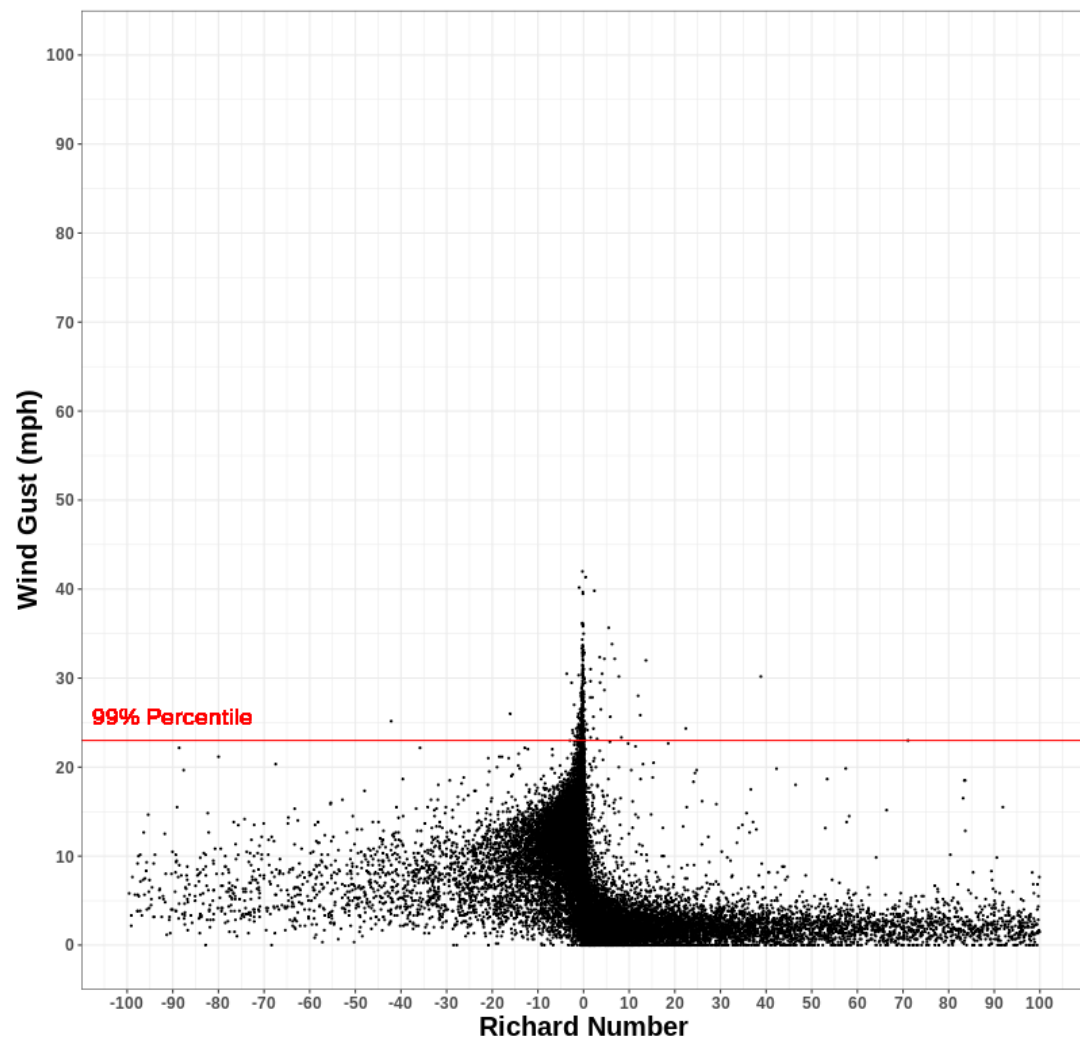


Test: 2018-11-11 to 2018-11-14

WRF RMSE = 9.45 Bias = 0.40
MODEL RMSE = 10.80 Bias = -6.11



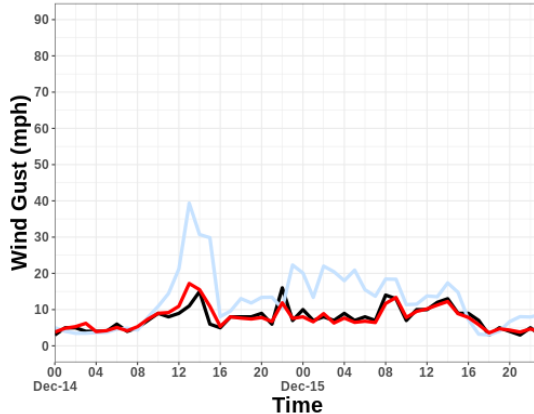
Wind Gust vs Richard Number for Creelman



Wind Gust ML Validation for Hideaway Lake

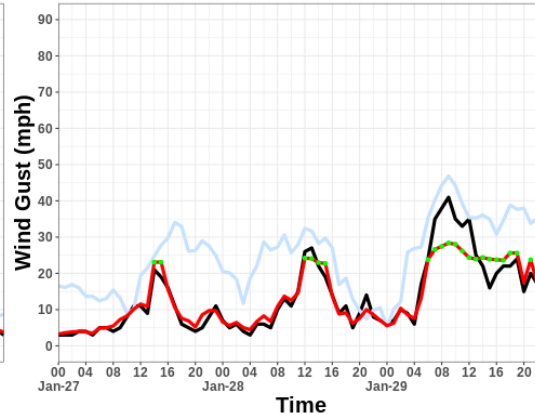
Train: 2017-12-14 to 2017-12-15

WRF RMSE = 8.35 Bias = 5.08
MODEL RMSE = 1.55 Bias = 0.09



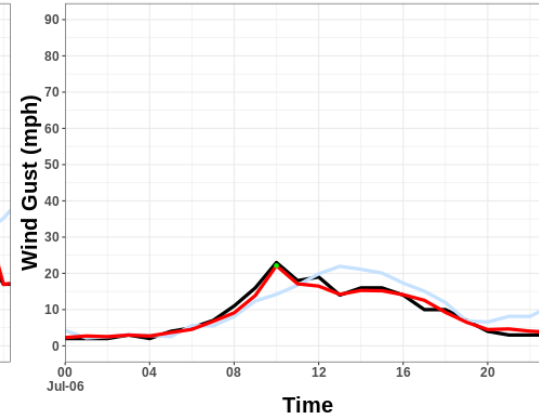
Train: 2018-01-27 to 2018-01-29

WRF RMSE = 13.37 Bias = 11.48
MODEL RMSE = 3.36 Bias = -0.10



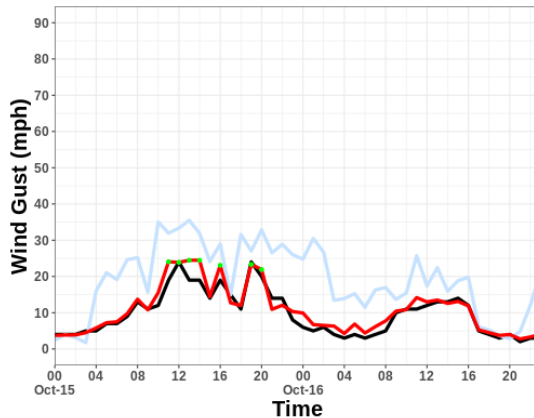
Train: 2018-07-06 to 2018-07-06

WRF RMSE = 3.99 Bias = 1.42
MODEL RMSE = 1.13 Bias = -0.13



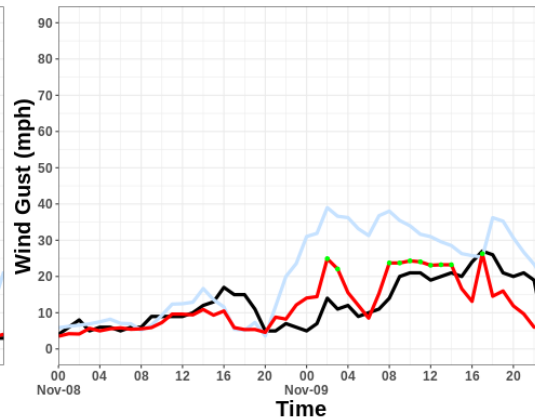
Train: 2018-10-15 to 2018-10-16

WRF RMSE = 11.77 Bias = 9.48
MODEL RMSE = 2.09 Bias = 1.02



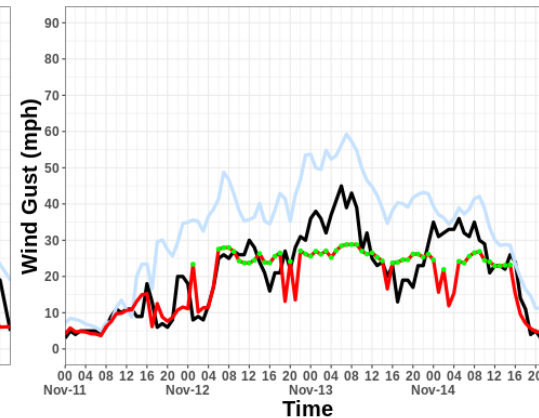
Test: 2018-11-08 to 2018-11-09

WRF RMSE = 12.76 Bias = 8.00
MODEL RMSE = 5.74 Bias = -0.45

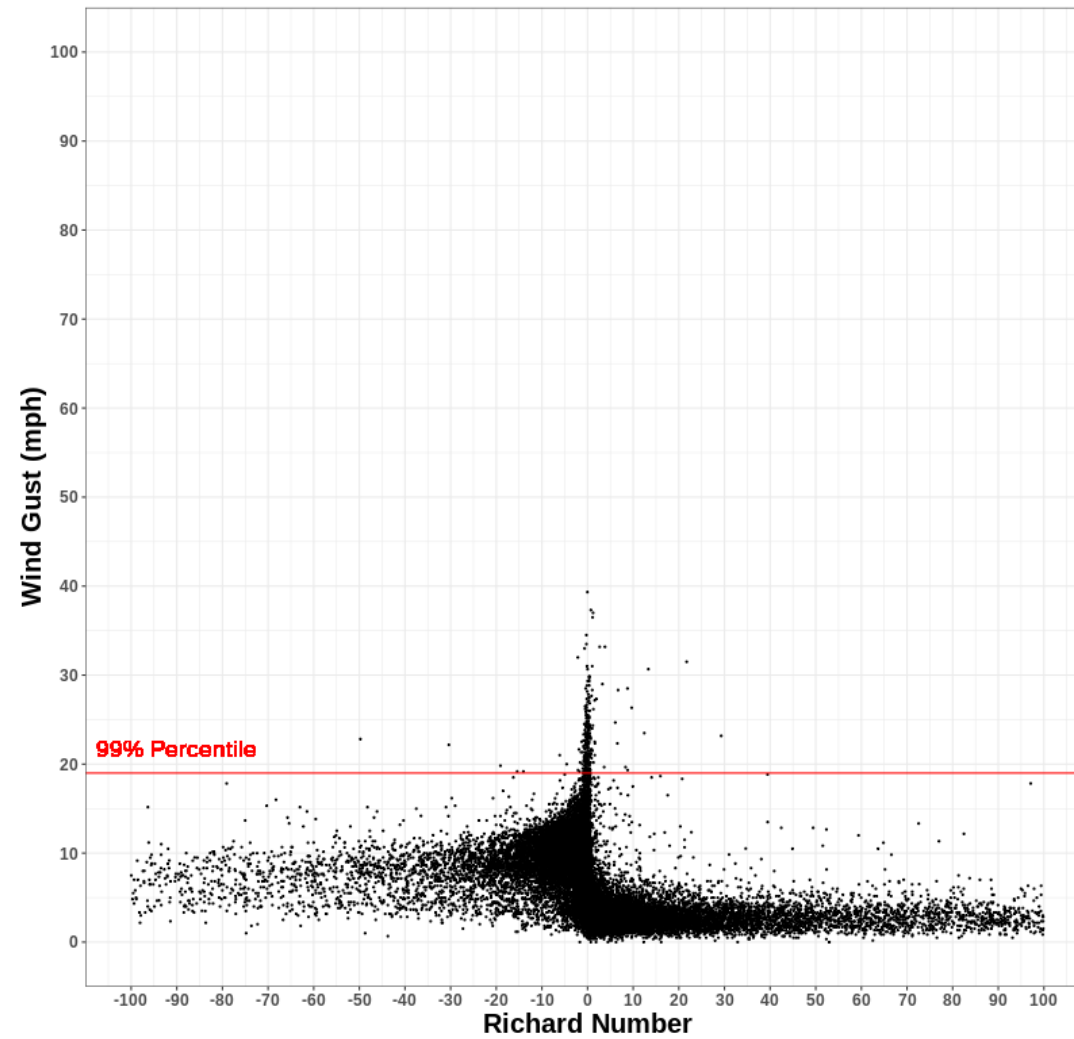


Test: 2018-11-11 to 2018-11-14

WRF RMSE = 14.39 Bias = 12.09
MODEL RMSE = 6.96 Bias = -2.34



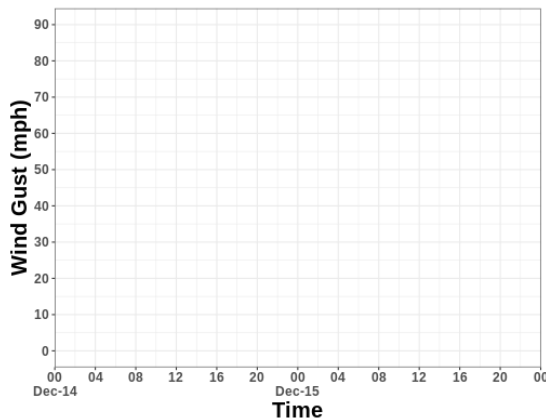
Wind Gust vs Richard Number for Hideaway Lake



Wind Gust ML Validation for Deluz

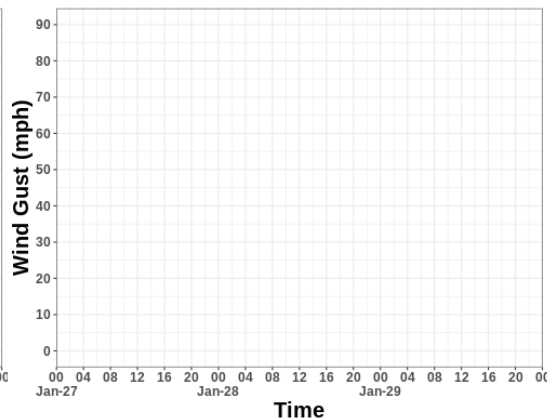
Train: 2017-12-14 to 2017-12-15

WRF RMSE = NaN Bias = NaN
MODEL RMSE = NaN Bias = NaN



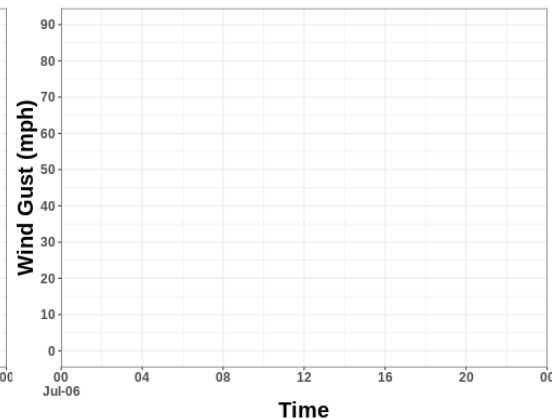
Train: 2018-01-27 to 2018-01-29

WRF RMSE = NaN Bias = NaN
MODEL RMSE = NaN Bias = NaN



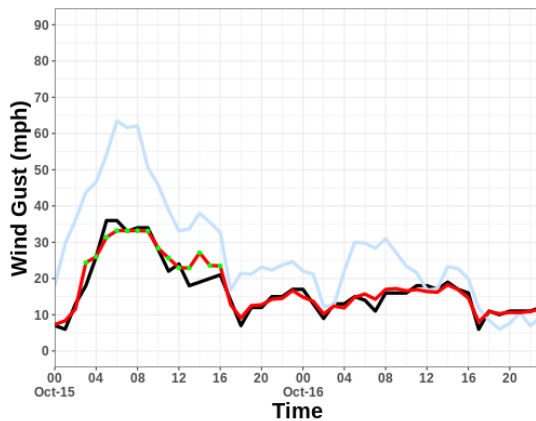
Train: 2018-07-06 to 2018-07-06

WRF RMSE = NaN Bias = NaN
MODEL RMSE = NaN Bias = NaN



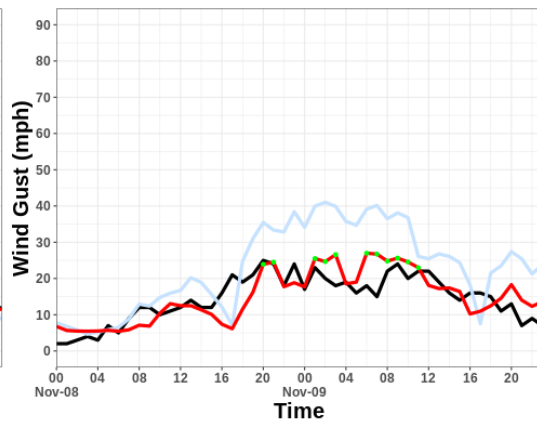
Train: 2018-10-15 to 2018-10-16

WRF RMSE = 13.75 Bias = 10.45
MODEL RMSE = 2.26 Bias = 0.46



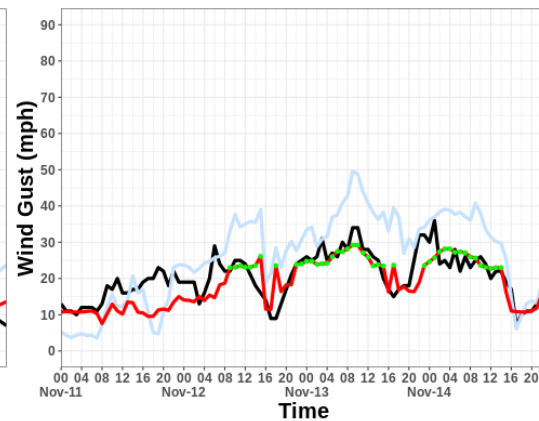
Test: 2018-11-08 to 2018-11-09

WRF RMSE = 11.78 Bias = 8.55
MODEL RMSE = 4.84 Bias = 0.38

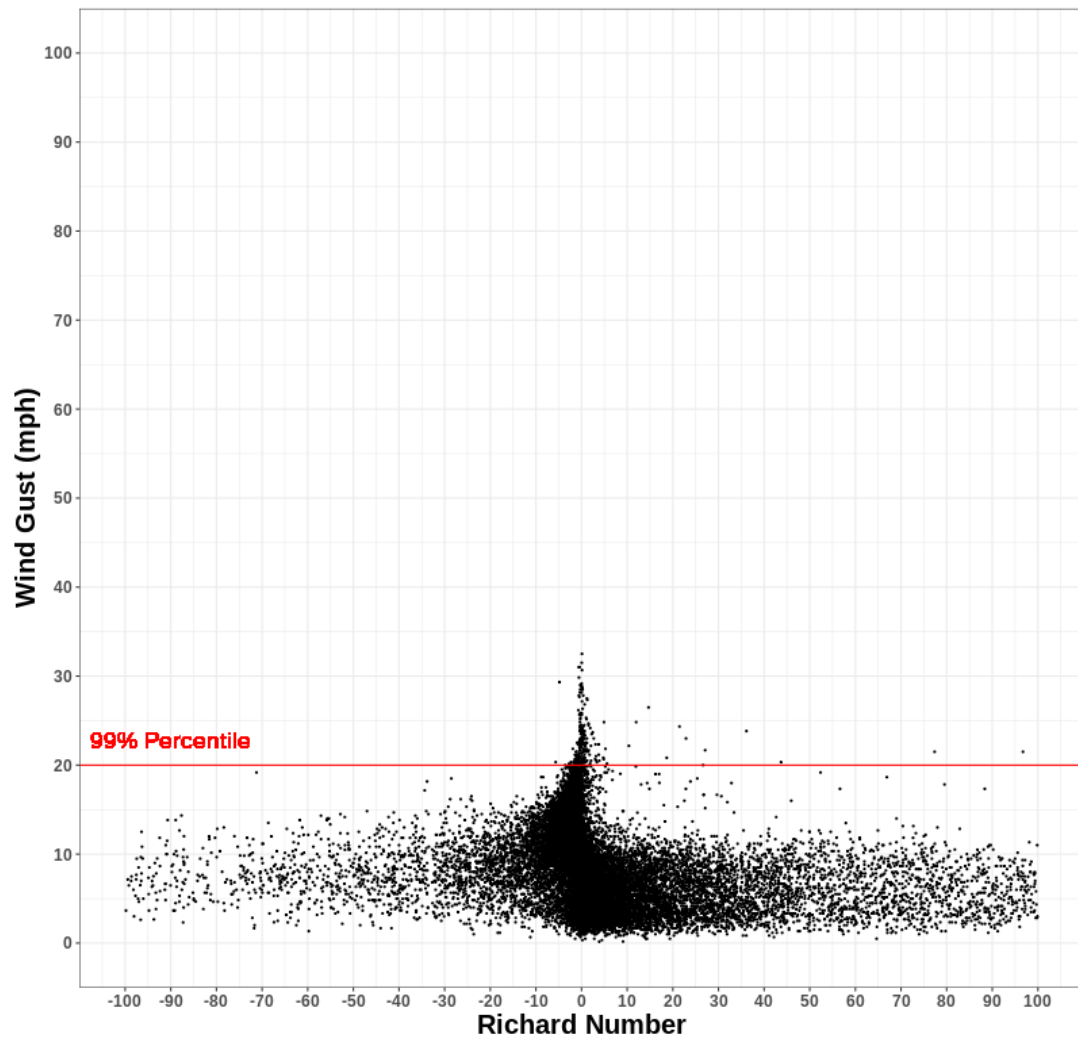


Test: 2018-11-11 to 2018-11-14

WRF RMSE = 10.33 Bias = 5.57
MODEL RMSE = 5.21 Bias = -1.80



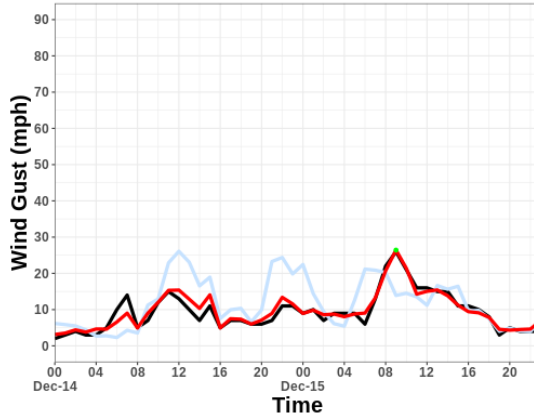
Wind Gust vs Richard Number for Deluz



Wind Gust ML Validation for Highland Valley

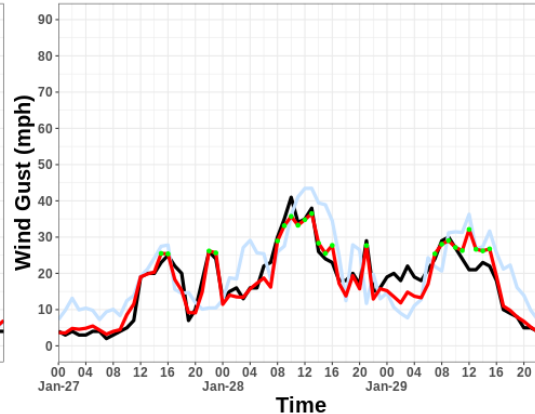
Train: 2017-12-14 to 2017-12-15

WRF RMSE = 6.66 Bias = 2.49
MODEL RMSE = 1.63 Bias = 0.39



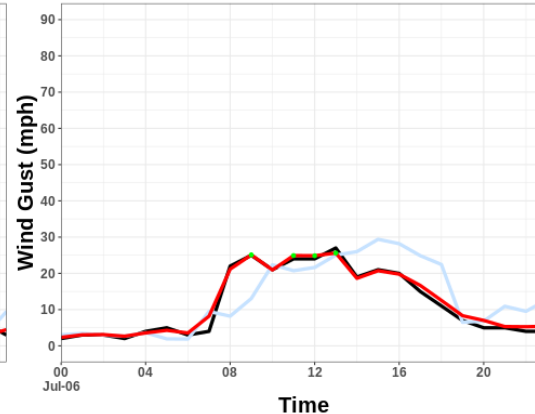
Train: 2018-01-27 to 2018-01-29

WRF RMSE = 8.03 Bias = 2.69
MODEL RMSE = 3.08 Bias = -0.27



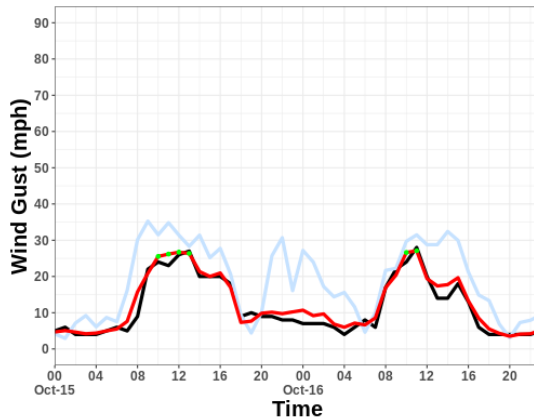
Train: 2018-07-06 to 2018-07-06

WRF RMSE = 6.32 Bias = 1.56
MODEL RMSE = 1.26 Bias = 0.54



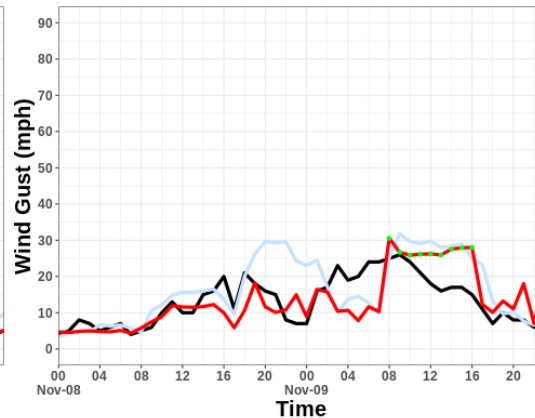
Train: 2018-10-15 to 2018-10-16

WRF RMSE = 9.53 Bias = 6.99
MODEL RMSE = 1.93 Bias = 0.88



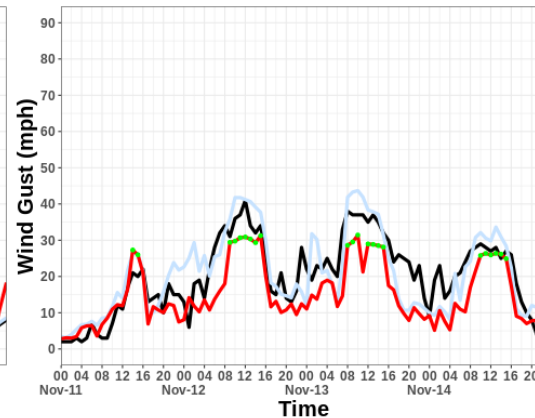
Test: 2018-11-08 to 2018-11-09

WRF RMSE = 8.22 Bias = 3.28
MODEL RMSE = 6.44 Bias = 0.09

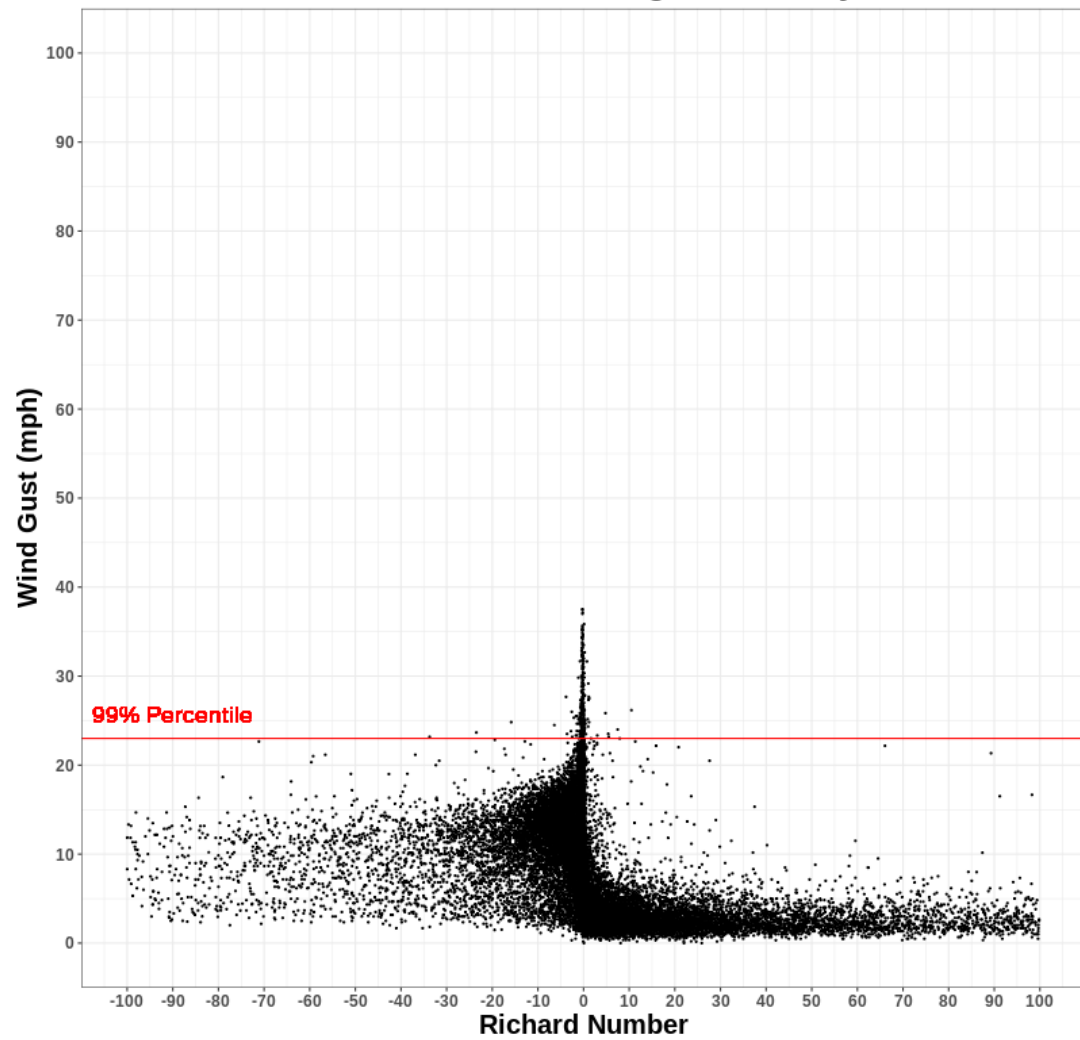


Test: 2018-11-11 to 2018-11-14

WRF RMSE = 6.21 Bias = 0.81
MODEL RMSE = 7.75 Bias = -4.98



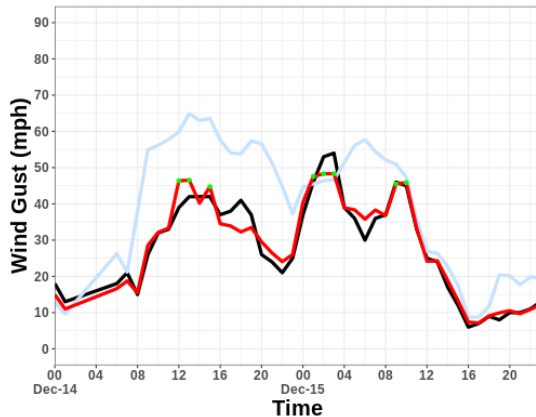
Wind Gust vs Richard Number for Highland Valley



Wind Gust ML Validation for Crestwood

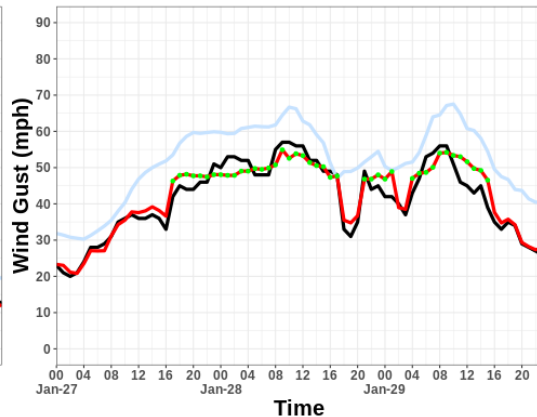
Train: 2017-12-14 to 2017-12-15

WRF RMSE = 15.68 Bias = 11.73
MODEL RMSE = 2.94 Bias = 0.22



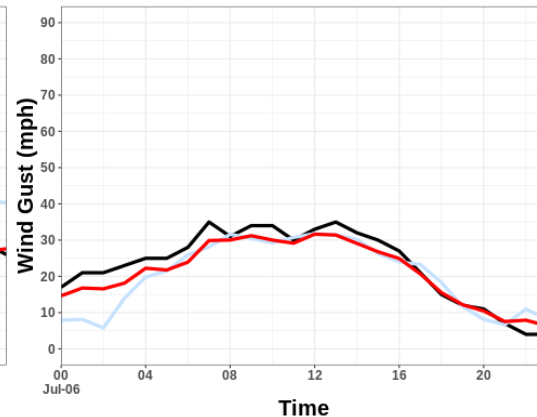
Train: 2018-01-27 to 2018-01-29

WRF RMSE = 11.24 Bias = 10.41
MODEL RMSE = 3.05 Bias = 0.76



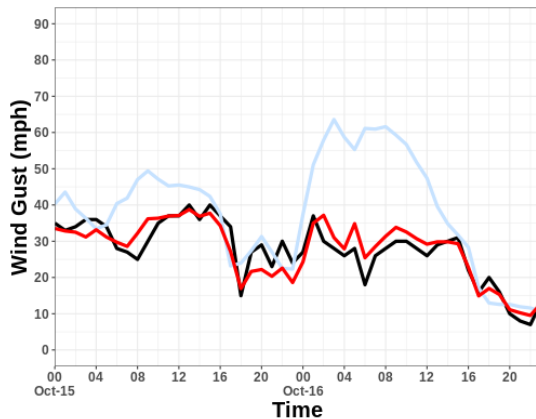
Train: 2018-07-06 to 2018-07-06

WRF RMSE = 5.83 Bias = -2.96
MODEL RMSE = 2.97 Bias = -1.93



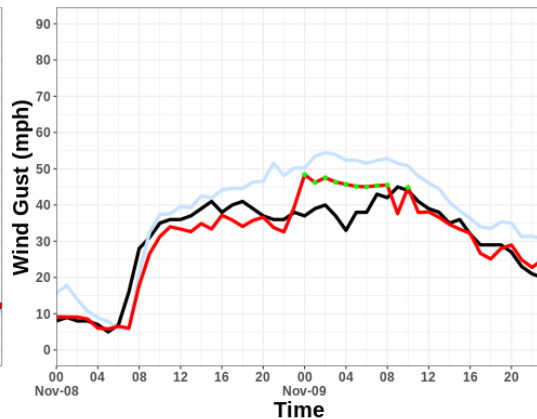
Train: 2018-10-15 to 2018-10-16

WRF RMSE = 16.66 Bias = 10.65
MODEL RMSE = 3.64 Bias = 0.15



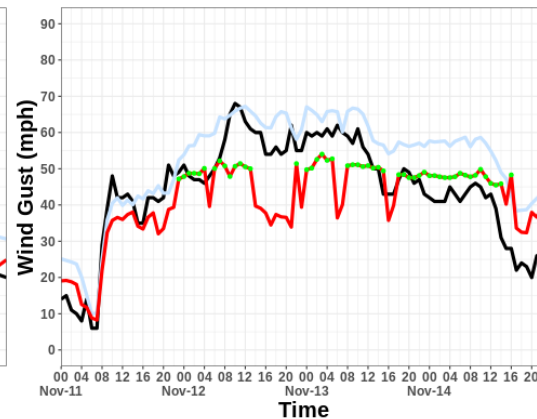
Test: 2018-11-08 to 2018-11-09

WRF RMSE = 8.80 Bias = 6.77
MODEL RMSE = 4.95 Bias = -0.20

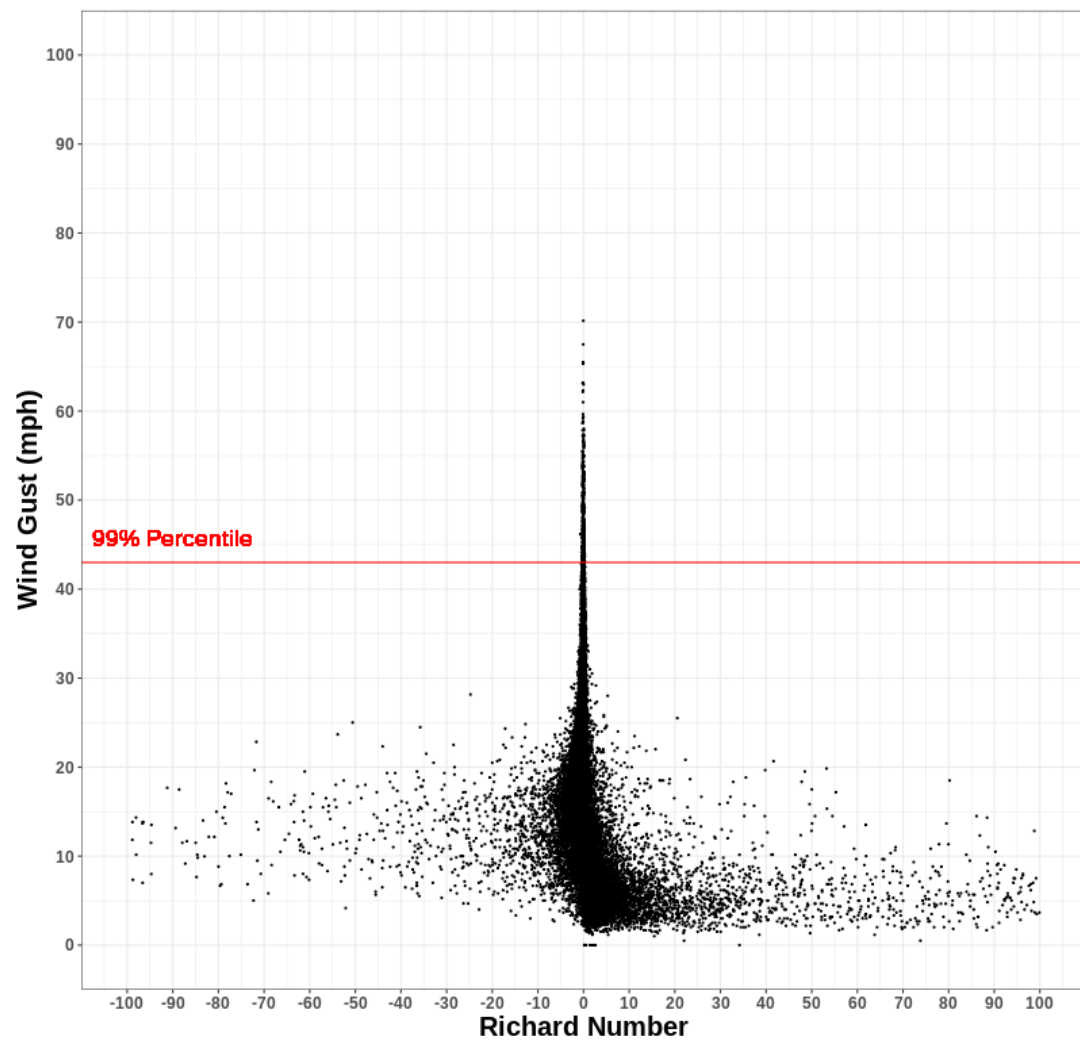


Test: 2018-11-11 to 2018-11-14

WRF RMSE = 10.11 Bias = 8.01
MODEL RMSE = 10.16 Bias = -2.54



Wind Gust vs Richard Number for Crestwood



Future Plan:

- 1. Use surrounding Grid Cell Data when available (especially when WRF totally miss the event)**
- 2. Use percentile table +/- RN as threshold to activate Bias correction model (XG-boost)**
- 3. Validation plot for more stations and dates (provided by SDG&E)**
- 4. Use d02 data (ens_cfsr_004_d02 data) to add values in current model**
 - temp_grad_surf, wind_spd_grad_surf, wind_speed_500m