Southern California Edison
2022-WMPs – 2022 Wildfire Mitigation Plan Updates

DATA REQUEST SET M G R A - S C E - 0 0 2

To: MGRA
Prepared by: David Michael Siuta
Job Title: Senior Advisor, Meteorology
Received Date: 3/7/2022

Response Date: 3/10/2022

Question 08:
Please provide analysis describing and supporting SCE's use of machine learning to enhance its weather station predictions, including verification statistics. (p. 30)

Response to Question 08:
Weather model forecasts are subject to error from imperfect initial conditions sources, incomplete representation of the underlying terrain, and scientific unknowns affecting small-scale meteorological processes. These sources of error can be random or systematic (repeatable) and are some of the primary limitations of weather model-based forecasts. To overcome these challenges, meteorologists can employ statistical methods such as machine learning (ML) to remove forecast biases from weather models resulting in improved forecasts when evaluated against observed values.

ML modeling is a statistical approach which trains a set of predictor variables against a known set of outcomes to result in a model that minimizes the error in the prediction. Once these models are trained, they are applied to future scenarios (forecasts) given the same predictors as input. For SCE, the predictors for the ML modeling are derived from SCE's in-house 2-KM deterministic WRF model. Table 1, below, is a comparative verification aggregated over six case studies between the 2-KM deterministic WRF model forecast and the ML model forecast at the 64 weather station point locations currently operational at SCE. The verification demonstrates the improvements the ML modeling approach provides over the 2-KM deterministic WRF model forecast with respect to the ability to differentiate between locations that will exceed set sustained wind speed or gust wind speed thresholds from those that will not exceed thresholds.

A forecast hit is a forecast and corresponding observation that exceeds a set threshold. A forecast miss occurs when an observed value exceeds the threshold, but the forecast does not. Finally, a false positive occurs when the forecast exceeds threshold, but the corresponding observed value remains below threshold. For this analysis, SCE used set thresholds of 31 MPH for the sustained wind speed and 46 MPH for the gust wind speed to determine the hit, miss, and false positive percentage. Over the test cases, the ML model improved the hit percentage for gust wind speeds by 33.5%, reduced misses by 33.5%, and reduced false positives by 8.7% (Table 1). For sustained winds, the hit percentage was improved by 49.8%, the miss percentage reduced by 49.8%, and false positives reduced by 29.7%.

Table 1: Hit, Miss, and False Positive Percentage statistics derived over six test cases prior to implementing the machine learning (ML) models at SCE. WRF represents the 2-KM deterministic WRF model forecast. Forecasts from both systems were derived at equivalent weather station point locations prior to performing the comparison. A total of 8,965 forecast-observation pairs were evaluated in the comparison.

Threshold	46 MPH Gust		31 MPH Sustained	
(n=8965)	**WRF**	**ML**	**WRF**	**ML**
Hit %	34.2	67.7	10.8	60.6
Miss %	65.8	32.3	89.2	39.4
False Positive (%)	44.1	35.4	65.8	36.1

Since the 2-KM deterministic WRF model is used as input into the ML model, the changes in verification scores between the two model sources provided in Table 1 can be interpreted directly as improvements gained by reducing bias in a weather model forecast. A more direct demonstration of this effect is provided in Figure 1, below, which compares the ML forecasts (red) and the 2-KM deterministic WRF model (blue) to the observations (black) at three separate case studies for the SCE Oat Mountain weather station location. In each case in Figure 1, the 2-KM deterministic WRF model forecast (blue) is consistently lower in magnitude than the ML forecast (red). The ML forecast (red) provides a closer fit to the observations (black). Additionally, the horizontal black line represents a threshold of 31 MPH sustained. In two of the three cases shown, the ML forecast correctly predicted winds to exceed the 31 MPH threshold, while in the third case, the ML forecast was much closer to the threshold than the raw weather model.

Figure 1: Forecast timeseries comparison at the SCE Oat Mountain weather station location for the machine learning forecast (red) and 2-KM deterministic WRF model (blue). Observations are provided in the black curve.