2022A10 STANDARD DEVIATION

Level 1: To calculate the mean (μ) of a set of data, sum all of the data points and then divide by the number of data points. For the wind speed data, this would be:

$$\mu_{WS} = \frac{5.24 + 5.46 + 5.32 + 4.93 + 4.82 + 4.65 + 4.72 + 4.80 + 4.86 + 4.58 + 4.16 + 3.91}{12}$$

$$\mu_{WS} = 4.78\ \frac{m}{s}.$$

For the power production data, this would be:

$$\mu_{PP} = \frac{32.7 + 44.0 + 44.0 + 32.7 + 32.7 + 23.7 + 23.7 + 32.7 + 32.7 + 23.7 + 15.8 + 15.8}{12}$$

$$\mu_{PP} = 29.5\ kW.$$

Level 2: To find the standard deviation of these sets of data, we'll use the mean wind speed and power production we found in the Level 1 question. Let's first recall the equation for standard deviation:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}.$$

The standard deviation for the wind speed becomes:

$$\sigma_{WS} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu_{WS})^2}{N}}$$

$$\sigma_{WS} = \sqrt{\frac{(5.24 - 4.78)^2 + (5.46 - 4.78)^2 + (5.32 - 4.78)^2 + \cdots + (3.91 - 4.78)^2}{12}}$$

$$\sigma_{WS} = 0.43\ \frac{m}{s}.$$

The standard deviation for the power production becomes:

$$\sigma_{PP} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu_{PP})^2}{N}}$$

$$\sigma_{PP} = \sqrt{\frac{(32.7 - 29.5)^2 + (44.0 - 29.5)^2 + (44.0 - 29.5)^2 + \cdots + (15.8 - 29.5)^2}{12}}$$

$$\sigma_{PP} = 8.89\ kW.$$

What do these two standard deviations tell us about our data sets? Remember, the standard deviation tells us how a data set is dispersed around the data set mean. In the case of our wind speed, the standard deviation of 0.43 m/s indicates that the majority of the data is within plus or minus 0.43 m/s of the mean. This is a low standard deviation in terms of the data we were measuring. Our wind speed data set is

relatively closely dispersed around the mean. The wind speed did not vary much over the two-hour period we were looking at.

In the case of our power production data, the standard deviation of 8.89 kW tells us that the majority of our data is within plus or minus 8.89 kW of our mean. Our power production data set is not as closely dispersed around our mean because the standard deviation is higher in terms of our power production data set.