Response to Reviewer 2

We thank the reviewer for their thorough examination of this manuscript.

In response to the general comment: the data from this field campaign is targeted in better understanding the role of lake breeze on shoreline ozone concentrations. The vertical structure obtained from the measurements will allow investigations into the dimensions of the stable boundary layer, the height of which can be modeled using a Richardson Number. Such investigations can be used in comparison to current weather forecasting models to test the validity in a shoreline environment. The vertical profile of ozone and meteorological variables allows for a nuanced understanding of the changes to stratification and vertical mixing occur as an air mass moves from over-water to over-land and how the coastal interface plays a role in pollutant mixing within the marine layer and lake breeze movement. Understanding the vertical and horizontal dimensions of the lake breeze phenomena can also inform model resolution improvements in this area impacted by high ozone, where 1.3 km grid scale models are only now starting to capture the lake breeze adequately. See revised manuscript P4-5, lines 111-124.

All specific comments have been addressed in the revised manuscript.

P 2, Line 46, (now lines 58-59): “volatile organic compounds (VOCs) and nitrogen oxides (NOx)” is added

P 2, line 50, (now P2 line 63): Deleted comma.

P 3, line 89-90, (now P 4-5 lines 111-125): Additional definitions given along with how the measurements can be used.

P 4, line 101, (now P16 lines 283-285): The 8-hour design value for the site is given along with the 8-hour federal ozone standard.

P 6: line 117-119: (Figure 5 has been added, now on P19): The shoreline monitor ozone was not modeled well, although the model predicted high ozone over the lake and the meteorological forecast suggested the possibility of a shallow lake breeze on May 22nd and 24th. Lines 304-313 on P18 accompany this new figure.
P 6, Table 1 (now P 19): language in the caption indicates patterns A and B are shown in Figure 2. Figure 2 caption also gives the color schemes for Pattern A and B.

Page 8, Figure 2 (Now Figure 6, P21): Figures 6 and 7 have “normalized probabilities” that are constructed so that the integral curve is equal to 1. Here, the value provided for a given bin is the number of elements in a given bin divided by the total number of elements in the input data. See figure caption.

Page 9 Figure 3 (Now Figure 7, P23): The figure has been updated to include the letters, as recommended.

Page 10, line 161 (now P5 line 134): units are corrected.

Page 15, Figure 5 (Now Figure 2, P12): Changed.

Page 17, Figure 6 (Now Figure 3, P15): Changed.

Page 18, Line 333 (Now P 24, lines 384-387): This text describes the procedure applied, not a single equation or relationship for the entire dataset. This procedure is applied independently to each flight to account for differences resulting from weather conditions and location-specific GPS offsets for a given flight that may impact the accuracy of the altitude estimate. Therefore, it is not practical to show a single figure or single equation, as there will be slight differences from one flight to the next.

Page 20, Figure 7 (Now Figure 8, P 26): We have updated the figure and associated caption to make the one-to-one line red.

Page 21, Line 383-384 (Now P 27 lines 439-445): The relative humidity values from the multihole pressure probe (MHP) are significantly impacted by the exposure of that sensor to sunlight and the associated impact on sensor temperature. This is not corrected for, resulting in large fluctuations in the RH values at times. As a result, this measurement (from the MHP) only provides a reality check to ensure that the RSS-421 is reporting something that is in line with the general reality, and therefore such a large offset (15%) is allowed. The more important comparison is between the two RSS-421 sensors, which should agree much more closely, as they are the same sensor type, and are mounted within close proximity of one another. Historical comparisons between the sensors across a wide variety of environmental regimes have led us to understand that 6% is an acceptable threshold for this comparison.

Page 23, Line 454 (now P 29 line 514): 2 flights did not recover iMET data because of a loss in battery power.