Associate Editor Evaluations:
Accurate Key Points: Yes

Reviewer #1 Evaluations:
Recommendation: Return to author for **major revisions**
Significant: The paper has some unclear or incomplete reasoning but will likely be a significant contribution with revision and clarification.
Supported: Mostly yes, but some further information and/or data are needed.
Referencing: Yes
Quality: The organization of the manuscript and presentation of the data and results need some improvement.
Data: Yes
Accurate Key Points: Yes

Reviewer #1 (Formal Review for Author (shown to authors)):

Review attached as PDF file and pasted below.

1. General comments: After revision, this paper should be a significant contribution to the literature on sea breeze impacts on temperature an ozone in the NY area, as it uses an advanced cluster technique. Its introduction and methodology are generally sound, but both require many clarifications. Its results section is hard to follow and needs careful revision, particularly of many of the graphics. Numerous minor JGR format issues also need to be addressed, and a review of their English usage is needed. Details below.

Recommendation: Major revision, with a focus on the following points, which will greatly clarify the presentation. I do need to see the revised Ms.

We sincerely thank the reviewer for the overall positive assessment and comprehensive review comments to help improve the presentation of this study. We find the major comments to be very constructive. Please see our responses below and the corresponding updates in the manuscript and figures. Thanks again for your professional comments and the time/effort you put into the review!

2. Major points to address

Major Comment 1 for Section 1 (Introduction): The literature review is well structured, as its review of previous studies goes from modeling (unstated as such), to analyses, and then to cluster analysis (the technique currently used). The discussion of the work of these efforts, however, is too superficial, as the authors do not explain what these studies contribute to our understanding of the physics involved and what was left undone/unexplained (to justify further study). They must end this section with something like the following: Whereas all previous studies have only…., the current study will…

Thanks for the great recommendations. We have reorganized the section following the logic below: Ozone problem in the NYC region → Case studies (contributions detailed) across some coastal regions including NYC; however, comprehensive study is lacking for the region → analyses/group-based comparisons (contributions detailed) such as sea breeze days vs non-SB days; however, temperature is overlooked → introduce what we do (multi-year comprehensive study with temperature considered for the NYC region) including method, data, and organization of the paper. Please see detailed revisions in the marked-up version.

Major Comment 2 for Section 2.1 (Data): The description of the meteorological observations should be further clarified, e.g., hourly average definition (as stations collect data at slightly different time intervals, values within how many minutes of hour were averaged); what was the filter for calm winds; and QA/QC procedures. Description of the wind lidar data should be clarified, such as equipment used, scanning frequency, and time averaging. The QA/QC procedures for NOx, O3, HCHO, and NO2 should also be clarified.

Thanks for the comments! We have double-checked the data QA/QC and averaging for the revision. "The hour for all hourly data indicates the end of the averaging window in EST (UTC - 05:00, EST is used throughout this study), e.g., 11:01 to 12:00 is averaged for the 12:00 condition for data measured at 1 min frequency." Averaging and QA/QC for quantitatively used data are described below. The wind lidar description has been added in the revision.

NYSM:

> Many QA/QC procedures (both automated and manual ones) are applied to flag erroneous data by NYS Mesonet. We used pre-generated hourly surface observations with time marked at the end of the average window of xx:00 and averaged lidar data to the same period.

AQS:

> AQS pre-generated hourly data has a time stamp at the sample beginning time xx:00 (sources below) and is now shifted one hour to be compatible with other data.
>
> "The time stamp on sample data is always the sample begin time." https://aqs.epa.gov/aqsweb/documents/about_aqs_data.html
>
> "The time corresponds to the begin time of the sampling period." https://aqs.epa.gov/aqsweb/documents/AQCSV_Format.html
>
> AQS QA/QC: "Ozone, NOx, and NO2 observations used in this study are measured by either a Federal Reference Method (FRM) or a Federal Equivalent Method (FEM) as defined in the Code of Federal Regulations for Ambient Air Monitoring Reference and Equivalent Methods (https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-53)." Specifically: Ozone in NY, NJ, and CT by FEM; NO2 in NY, NJ, and CT by FEM or FRM methods; NOx at 2 NYC sites by FRM.

Buoy data retrieved has various temporal resolutions from 6 minutes to 1 hour and are also reported to the end-of-acquisition time (https://www.ndbc.noaa.gov/faq/acq.shtml). Only quality-controlled data is used and averaged for temporal resolutions higher than one hour to xx:00. For the two sites used quantitatively, they are reported to end at xx:50 and assigned to (xx+1):00 with a ten minutes shift.

ASOS one-minute data is retrieved and averaged to the end hour following NYSM.

Major Comment 3 for Section 2.2 (Clustering): As k-means clustering is the core analysis method of the paper, two things should be further clarified: (1) reason why you choose data from the QUEE station instead of other stations and (2) advantages, limitations, and mathematical analysis principles of k-means clustering.

Thanks for the suggestions. The reasons for choosing QUEE have been added in this section as follows:

"QUEE is selected because 1) as part of the newly established NYSM, this site delivers quality controlled hourly average meteorological observations with extraordinary temporal data coverage; 2) it is "collocated" with an active AQS site reporting both ozone and NOx observations; 3) observations from AQS and NYSM at this site are representative in both physical and chemical conditions around the NYC where high ozone episodes occur frequently."

We have improved the method description in the revision by adding mathematical analysis principles, spelling out its input (number of clusters and feature vectors of the samples), and discussing its limitations and how we dealt with them. The advantages are that it is "an effective and efficient technique to separate features of interest into distinctive groups." Please see updates in the marked-up version of this section.

Major Comment 4 for Section 3 (Results): There is a great deal of good science results presented, but the presentation is difficult to follow (and a more detailed analyses of the results could not be done), as (1) the quality of many figures must be increased and (2) the use of the many case abbreviations needs to be locally defined regularly in the text so that readers will not get lost/confused while reading. With respect to the first point many figures are (i) too small, (ii) have poor color choices, (iii) lack details of what we are looking at, and/or (iv) have incomplete legends. With respect to the second point, it would also be helpful if the authors included a table that defined the 12 cases in Fig. 5. Specific instances are given below.

Thanks for the comments. We have improved the presentation of the results by:

1) Improving all figures: enlarging all figured elements (vector figures are also provided so that readers can enlarge all to suit their needs with the digital copy), deleting unnecessary elements or subplots, filling in more details such as statistics (min, med, and max) of all map composites and necessary legends and units, and improving the colormaps used for clarity.
2) Including necessary tables: Table 1 to summarize the wind and temperature clusters and associated ozone conditions for the 10 cases presented in Fig 5; Table 2 to conclude temperature and ozone conditions for the three regions studied; Table S1 to include details of sea breeze features and other meteorological and ozone conditions at Westport and Queens during all *SB* days.

Major Point 5, for Conclusions (or should it be Conclusion)? This section should summarize what was done (needs more); what was found, including specific quantitative results (not enough given); why results are important (needs more); and what future efforts should be carried out (done).

Thanks for the comments. We have expanded the section by adding more to what was done and specific quantitative results and filling in the importance and limitations of our work. Please see the marked-up version for details. For the section title, we would keep it plural to refer to the points we want to make instead of the act of concluding. The plural form is also consistent with other section names.

Major Comment 6: Pls note JGR requires consistent (C vs K) SI units/formats for wind, T, time, etc. In addition, many other notational/format errors also exist. Specific examples are given below. A careful review of the English usage in the paper is necessary to polish he Ms.

Many thanks for the comments. We have considered all specific comments seriously and addressed them below and in the manuscript. We have also checked the English writing before revision submission and made appropriate changes when necessary.

2. Specific points

For title: As this is only a three year study, how about: Sea breeze impacts on coastal urban pollution in the New York metropolitan area during three summers.

Thanks for the suggestion. However, we argue that the original title describes the study more accurately. Specifically:

Sea breeze (In addition to sea breeze, synoptic wind conditions, as indicated by "Local Circulation" in the title, play an important role by controlling the overall ozone distribution and magnitude and interacting with the development and penetration of sea breeze throughout the study. These are explicitly discussed in Section 3.1 for all meteorological clusters; Section 3.2 for the three "Hot" day clusters when sea breeze all developed under different synoptic wind conditions; and Section 3.3 for how early morning wind conditions impact sea breeze type and thus ozone exceedances in the region, etc. Therefore, we think "Local Circulation" would be better.) impacts on coastal urban pollution (We only covered ozone and a few of its precursors in the study. Urban pollution would overstate the study subject; it could include particulate matter and other air pollutants or even other coastal-related environmental issues other than air pollution. Thus, we argue that "Coastal Ozone Pollution" would be more precise and accurate.) in the New York metropolitan area during three summers (Three summers accurately described the study period. However, we would like to emphasize the generalization of our conclusions. "Evidence from Multi-year Observations" better help convey this notion and, in the meantime, acknowledge the study period.).

For key point No: (2) the westerlies delay the sea breeze onset (thus change order) and (3) What specific change in wind direction causes this

Thanks for your suggestions for the key points.

(2) We have thought about rewriting the second point seriously and could not come up with a better revision with the 140-character JGR limit for key points:

1) We want to focus more on the dramatic ozone change during hot summer days by the presence of the westerlies; thus, it is more reasonable to place the ozone change part in the beginning, even though the changed order would follow the event sequence logically and chronologically.

2) The westerlies not only delay the sea breeze onset but also impact how far north the sea breeze front can penetrate and the subsequent impact on ozone. Making this clear with the changed order would take up more space, exceeding the character limit.

(3) We think the original "meridional wind direction" specifies the change in wind direction as "meridional" indicates the v wind direction either from the north or south.

For abstract: add some additional specific numerical results

Thanks for the comment! We would love to include more quantitative results, but the 250-word limitation forces us to be picky. We added the regional DMA8 decrease for NYC and coastal CT from Hot SB to Hot W_sb days as this is the most critical numerical change from the impact of local circulation.

For PLS: Fine as is

Thanks!

For Section 1. Introduction:

Line 61: define "moderate"

The sentence has been rewritten as follows to make the definition more obvious.

"The region is classified as a moderate 8-hour ozone nonattainment area (area that has a design value of 81 up to but not including 93 ppb, US EPA, 2018) with the Ozone National Ambient Air Quality Standards (NAAQS) level at 70 ppb."

Lines 64-9: What effect does T have on PBL height and how does that effect ozone?

Thanks for bringing the PBL height impact into the discussion. We have added the following with appropriate citations to the paragraph.

"In addition, higher temperature facilitates the growth of a deeper planetary boundary layer (PBL) due to increased convection driven by surface heating. However, the impacts of PBL height on surface ozone are often coupled with influences from other physical and chemical factors, including the absolute height of PBL and ozone vertical profile (Haman et al., 2014; Zhang et al., 2023). The overall effect of PBL height on surface ozone is less certain. In general, the highest surface ozone most likely occurs when PBL height is moderate (between 1-2 km for

Beijing) and decreases as PBL becomes higher (dilution) or lower (decreased photochemistry because of the availability of sunlight) (Zhang et al., 2023)."

Line 70: "This" should be "NY area"

Thanks. We have spelled out this as "the NY metropolitan area".

Lines 77-82: Should clearly state the two urban effects on sea breeze circulations: (1) UHI acceleration and (2) roughness/building deceleration and frontal stalling

Thanks for the comment. We have clarified these two urban effects on sea breeze front penetration as follows:

"In addition, the large-scale urbanization in the region alters the local circulation with its excessive heat flux and increased surface roughness. Urban heat island (UHI) generated by the heat contrast between the urban center and surrounding area produces convergence flow into the urban center. The convergent flow would accelerate the sea breeze front towards the city center (Freitas et al., 2007; Hu et al., 2022). However, sea breeze front penetration over the city center would likely be more difficult and can be stalled for a few hours (Freitas at al., 2007; Hu et al., 2022) due to the convergent flow and urban frictional retardation and would be slower overall than that over less-urbanized surrounding areas (Ferdiansyah et al., 2020; Han et al., 2022)."

Line 81: clarify "favorable"

Thanks for your comments to help improve the clarity of the statement here. We have decided to drop the discussion on background wind conditions because of its complexity and limited relevance in the context since we aim to discuss the impact of UHI and surface roughness here. Please see updates in the response above.

Lines 85-92: Were these modeling studies?

Thanks! All Clarified in the revised manuscript.

Line 99: "been" should be "also been," as this is an important transition

Thanks for the comment. However, the cluster method is used in one of the studies mentioned above (Li et al., 2020). Since we have greatly improved and restructured the introduction, this has been changed to the following when introducing the cluster method.

"With the addition of temperature as a contributing factor, this study adapted the k-means clustering technique used in Li et al. (2020) to cluster NYC wind conditions and diurnal temperature profiles."

Line 103: where was work of Li carried out

We have added their study region of Houston in the revised manuscript.

Line 99-103: What type of cluster analysis (we know that it is not k-type) were used in these studies?

They incorporated k-means as part of the approach after the standard principal component analysis (PCA) procedure. Please refer to the short summary below:

Ngan & Byun (2011) clustered 2005-2006 ozone season (May-September) local weather patterns in the Houston–Galveston–Brazoria area based on the 850 hPa u and v components of winds from the gridded analysis products at 06:00 local time with the help of a standard principal component analysis (PCA) procedure. Classification of resultant principal components then involves two steps: a hierarchical method to determine the initial number of clusters and a nonhierarchical iterative method (k-means clustering) to generate the final results.

Line 106: New paragraph, as this begins the discussion of what you will do.

It has been revised accordingly.

Line 108: Clarify DMA8

We think the full name "daily maximum 8-hour average ozone" before DMA8 in the sentence is adequately clear.

Lines 109-113: Just tell us what you will do, without use of Section numbers.

Thanks for the suggestion! We have improved the writing of this paragraph. However, we believe the spelled-out section numbers will help guide readers better. It has been revised as follows: "The contents are organized as follows. Details of data and methodologies are described in Section 2. Section 3.1 addresses the first question by presenting the cluster results and examining the overall impact of local circulation and temperature scenarios on daily maximum 8-hour average ozone (DMA8) spatial distribution in the New York metropolitan area during the summertime of 2017-2019. Then, Section 3.2 focuses on hot summer days, during which ozone levels are more elevated and harmful to human health and other living beings, and summarizes how background wind conditions and sea breeze impact ozone diurnal cycles and spatial distributions, answering the second question. Lastly, critical meteorological characteristics that can help modulate ozone exceedances during extreme hot sea breeze days in ozone hot spots of NYC and coastal CT are identified and illustrated with representative case studies during LISTOS in Section 3.3 (the third question). A comprehensive table detailing all related meteorological characteristics and ozone levels during hot sea breeze days is also included to serve as a reference for future case studies."

For Sections 2. Data and Methods:

At start of Section 2.1, you could add a short list of the science questions you will answer. These should follow from your "Whereas statement (see Major Comment 1)" and will guide the reader through your results and conclusions.

Thanks for the comment. It is a good idea to list the scientific questions to help guide readers. However, we think the end of the introduction would be a better place for them. The added questions are as follows, and we cited these questions here at the beginning of Section 2.1.

1. What is the general relationship between surface ozone and meteorological conditions in the region during summertime?
2. What are the temporal and spatial characteristics of ozone during hot summer days, when elevated ozone levels are more prevalent, and how does local circulation contribute to these patterns?
3. Can local circulation help modulate ozone exceedances during extreme heat events, and if so, how?

Line 119: are the hourly data centered on the hour?

Thanks for the question again. Please see reply in the major comment on data.

Lines 124-140: To much specific details about sites. This info is in figure and its legend. Just name the sites and add "(Figure 1)".

Thanks for the comments. We would prefer to keep the site details, such as additional links, descriptions, and site ID, in this section as they could be helpful to some readers; moving them to the caption would make it too long and difficult to screen through. We have also improved Figure 1 using annotations on the map.

Line 142: five scans per hour?

About 20 seconds. We have added it to the manuscript.

"NYSM wind Doppler lidar (WDL, Leosphere scanning Doppler Windcube 100S) collects data using the Doppler Beam Swinging (DBS) scan mode, which consists of five scans in four cardinal directions (north, east, south, and west) and nadir with a cycle of 20 seconds (Shrestha et al., 2021)."

Line 152: What wind characteristics were used and how frequently was this done?

Please see explanation below: "In addition, we employed the gradient in HRRR hourly *(how frequently: hourly, but maps for only two hours around noon are shown)* surface wind fields *(wind characteristics: surface u and v wind)* to locate the sea breeze fronts in the case studies in Section 3.3."

Lines 174-5: Move to Section 1

We think it would be better to keep the original Line 174-5 here to provide a more detailed introduction and description of the methodology in this section.

Lines 179-80: What criteria were used?

Thanks for the question and the general comment on improving the description of k-means. We have added the rationale, how it works, limitations, etc., in this section. The number of clusters is the only "criteria" that needs to be preselected, and the algorithm will cluster the samples by minimizing the inertia or within-cluster sum-of-squares: $\sum_{i=1}^{I}\sum_{j=1}^{J} w_{ij}\left\|x_i-\mu_j\right\|^2$. "The initial cluster centroids sampling method was set to be based on an empirical probability distribution to speed up convergence and guarantee the inertia converging to a global minimum (Pedregosa et al., 2011)." However, we did manually choose the suitable wind cluster number. The best would be the one with: "small inertia and reasonable wind clusters by visual inspection of the clustering results for all summer days."

Line 186: To be clearer and more consistent, the first four features should also be marked with numbers (first feature, second feature, … etc.) like the fifth "average early morning (4:00-6:00) wind speed (fifth feature)" and sixth characteristics "the ratio of morning and afternoon mean wind speed (sixth feature)".

First four features marked.

Lines 188-9: Other studies have found land breezes in pre-sea breeze hours. Clarify.

Thanks for the comment. We assume the reviewer refers to the land breeze in pre-sea breeze hours in general, but not for the NYC region. Land breeze can be found in some of the coastal regions. However, as in the citation (figure pasted below) and in this study (Figure 9), there is no clear diurnal cycle of comparable land breeze (from the north) and sea breeze (from the south) caused by the land-sea temperature difference in the region. This is primarily due to the large-scale wind exerted by the extensive Bermuda High system (southwest wind in the study region, Figure 7). It prevented the development of land breeze, as shown in Figure 9.

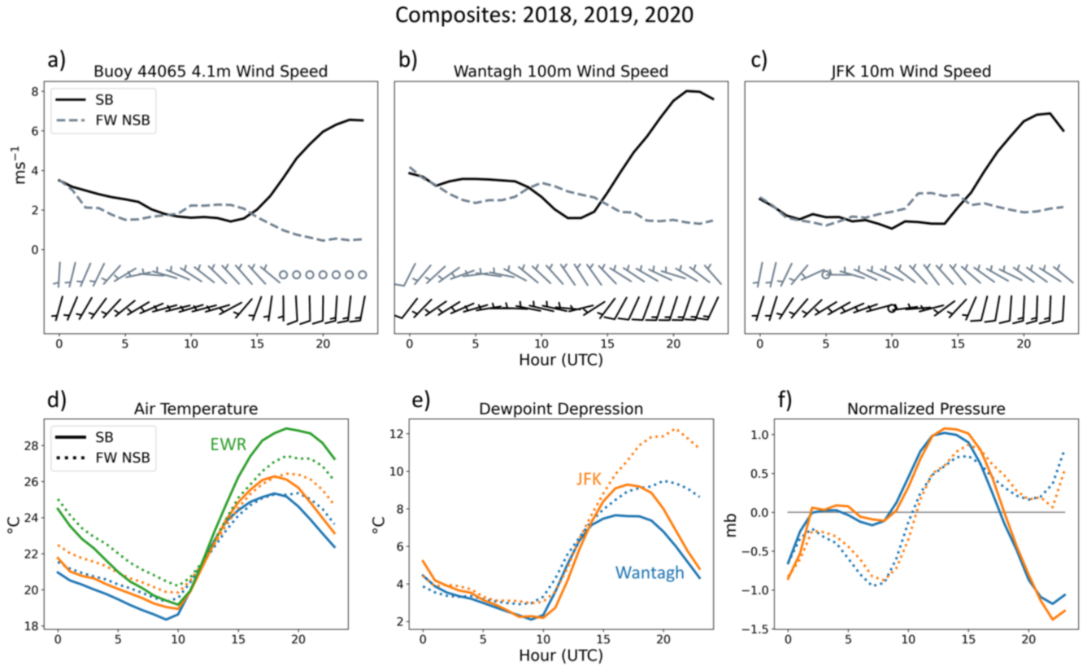

Figure 13. Composites over a 3-year period from 2018 – 2020 on 81 sea breeze days (SB; solid lines) and 60 fair weather non-sea breeze days (FW NSB; dashed lines). Top: Shown are composites of wind speed (m s^{-1}) and direction (1 full barb = 5 m s^{-1}) at a) Buoy 44065 (4.1 m measurements); b) Wantagh, NY (100 m measurements); and c) JFK (10 m measurements). Bottom: Shown are measurements from 3 sites EWR (green), JFK (orange) and Wantagh (blue) of d) air temperature (°C); e) dewpoint depression (°C); and f) normalized pressure (mb).

Line 198: Best in what way?

Thanks for the question. We have added the info: "The number of clusters was tested between 3 and 6, and four clusters yielded the best wind condition separations (small inertia and reasonable wind clusters by visual inspection of the clustering results for all summer days) were adopted."

Lines 202-4: Need to discuss the lidar and its data in Section 2.1.

We have added more details of the lidar and its data in Section 2.1. Please see details in the marked-up version.

Line 204: Figure 2 has an (a) and (b). You need to tell us which we are looking at throughout the discussion.

We have specified a or b in the revision and improved the figure accordingly.

Line 211: what are units of 0.5?

It is a unitless correlation.

Line 225: remove “in Section 3/.3” and add “All” at start of sentence.

Thanks for noticing this. However, in addition to the approach described here, we also located sea breeze fronts manually, as shown and described in Figure 10. The original statement resolves the confusion.

Lines 228-33: Do we really need these equations. Couldn’t you say the averages between two adjacent grid cells.

Thanks for the comment. They are actually the addition of the 2-D wind gradients in x and y directions. We prefer to keep these equations and numbered them accordingly. They would help 1) explain how we calculated the two-dimensional gradient and located their grid cell centers; 2) make it easier for readers who want to replicate the calculations.

For Section 3. Results and Discussion:

Line 238: Shorten section title to: Clustering results

We think it would be better to keep the title as is because there’s also an important overview of ozone conditions under these clustered conditions.

Line 240: The two missing days should be addressed in Section 2.1 and delete phase in parentheses

NYSM data coverage is discussed in Section 2.1 and details are deleted here.

Discussion of Fig. 3: well dome

Thanks!

Line 253: change arrows to vectors

Changed here and all other places.

Line 259: why do they not penetrate into NYC?

We suspect the main reasons for that could be: 1) sound breeze is weaker in general as the surface area of LIS, especially its left corner close to NYC, is much smaller than that of the New York Bight/the Atlantic Ocean; 2) the sound breeze direction towards NYC would be opposite of the sea breeze from the Atlantic Ocean. The land-sea temperature contrast for the LIS and the

New York Bight/the Atlantic Ocean might not be the reason as shown in the figure and caption below.

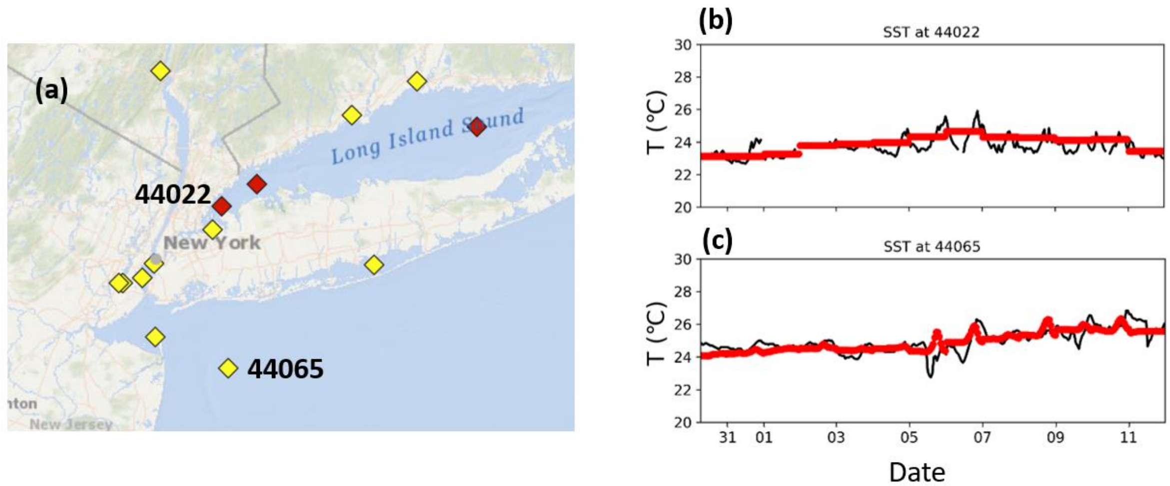

Figure Line259. Sea surface temperature (SST) comparison between LIS (NDBC Site 44022, black line in b, lower SST, higher temperature difference with land) and the New York Bight (NDBC Site 44064, black line in c) from Jul. 30 to Aug. 11, 2018. Among them, Aug 5, 6, 8, and 10 are hot sea breeze days. Please ignore the red lines, as they are Hourly SST products from the UK Met Office.

Line 280: what is Phil.-NYC south shore?

We apologize for the confusion. The “Philadelphia-NYC-south shore of CT corridor” describes the high ozone corridor seen in conditions such as Hot SB days. We renamed it “Philadelphia-NYC-LIS corridor” and visualized its location in updated Figure 5a. Locations of all three places are identified in updated Figure 1.

Line 296: Shorten section title to: Local circulation impacts

Thanks for the comment. However, as explained by the progression of Section 3 in the major comments, we are limiting the days to higher and higher temperature levels from Section 3.1 to 3.3. Thus, their titles should reflect the temperature segment talked about in each subsection. We have also highlighted the meaning of hot days and extreme heat in the manuscript and added transitional discussion between subsections to make it clear.

Line 311: Be specific as to map type

We have added details of map as: “Weather maps based on composites of HRRR reanalysis of 850hPa geopotential height and wind fields during Hot day clusters (SB, W_sb, and S) are shown in Figure 7.”

Line 391: Shorten section title to: Extreme heat case studies

In addition to case studies, this subsection includes an important overview of extreme heat conditions in the beginning.

Line 395: All geographic references in test should appear in Fig. 1, or in an insert (e.g., for the NYC boroughs)

Thanks for all your comments on Figure 1. We have added an insert to show details around NYC and included names for all locations mentioned in the manuscript. However, the three regions of "…NYC (squares in Figure 9h), CTcoastal (coastal region of western CT, triangles in Figure 9h), and LICTe (LI and coastal region of eastern CT, reverse triangles in Figure 9h) are selected and grouped manually based on their spatial proximity and similarity in the characteristics of their ozone diurnal cycles during SB, W_sb, and S days..." (later data analysis). Figure 9h would be a better place for these regional references.

Line 430: why define "T" here; it must have been used before

The full name has been deleted accordingly.

Line 462: Shorten section title to: abrupt sea breeze onset

The title now reads "Case 1: Abrupt strong sea breeze onset" as the case number would help readers locate these case studies; the "strong" is also an important feature ("strength") of the sea breeze, in addition to "abrupt".

Line 525: Shorten section title to: weak sea breeze onset

The title now reads "Case 2: Gradual weak sea breeze onset".

Line 557: change to Conclusion

Thanks for the comment. We would keep it plural to refer to the points we want to make instead of the act of concluding. The plural form also is consistent with other section names.

Line 559: Define DMA8 again, just to remind us

To be consistent with the comment on "T", we decided not to repeat the definition here. Also, this is a widely known concept for the ozone study.

Line 561: certain wind pattern has is poor English and too vague

It has been updated to: "…specific wind patterns exhibit consistent effects on the spatial distribution of ozone regardless of temperature levels. For example, sea breeze days favor high ozone levels along the Philadelphia-NYC-LIS corridor at both hot and moderate temperatures."

Line 580: Do not cite “cases” and Figs. in the conclusion: explain what the case was and the results of the Figs.

Thanks for the comment. We have improved the last section accordingly. Please see marked-up version for details.

For Reference list: The correct JGR format was generally followed, one error was found, i.e., all words in titles were incorrectly capitalized.

Thanks for noticing. All such mistakes have been corrected accordingly. However, these changes made using the *Change Case* in *Microsoft Word* is not tracked.

For Figures:

Figure overall reply: Thanks for all the detailed comments. We have updated all figures accordingly. All figures and texts in figures are in adequately large sizes. If further enlargement is needed, the digital copy can be zoomed in to suit the readers’ needs as the digital vector figures in pdf are submitted for publication. Thanks again for helping improve the presentation.

Fig. 1: Needs an insert to show details of NYC. Needs color bar scale to show topo heights. The station location symbols are too close in color and shape to identify.
Fig. 1, Legend: Should not have to identify each symbol if the symbols were clearer.

Thanks for the comments. We have greatly improved Figure 1 with an insert of the NYC region and decreased the number of symbols and legends to increase the readability of the figure. Specific sites are now noted directly on the map. The corresponding main text and figure captions have been updated accordingly.

Fig. 2a: Vectors are too small, enlarge figure. Must use m/s and not knots. Should use 1 full barb as 1 or 5 (not 10) kt (or m/) so we can see details of low wind speeds. Why are 800 m and 20180805 on fig., as they are in legend? Is height in m AGL or m MSL (should make general statement in text and then be consistent throughout the text). Can give vales at every 400 or 500 m, only.
Fig. 2b: Here you use m/s, while in (a) you used kts. Remove triangle from top of fig, and define it in the legend

Thanks for the comments on Figure 2. We have updated Figure 2 accordingly. The wind speed in (a) is now indicated by color in the legend in m/s. Low wind direction is now clearly defined by the vector directions. Height definition (above ground) and label have been updated as well.

Fig. 4 a and b: Enlarge both for clarity

Done as in the figure overall reply above.

Figs. 5 and 8: Add ozone to color bar. Can you add some more colors so lower panels do not look the quite so similar

Thanks for the comments. Names of color bars are now added to all figures. The color map of the second row in Figure 8 has been changed to be different from those in the first row.

Fig. 6: Too small. Are these 10 m speeds? Define subsets in legend, as it is hard to keep track of all the abbreviations.

Please see responses for size in the overall reply. They are now both clarified in the figure caption: "Diurnal cycles of mean 10 m wind at WANT during Hot W wind subsets of W_w (5 days) and W_sb (17 days). The lowercase w and sb in the subset names represent coastal wind conditions of westerly and sea breeze at WANT."

Fig. 7: Use more colors. Confusion: are colors T or gpz values? Wind vectors: impossible to see. Add some summary vectors and height contours to illustrate main pressure and wind patterns

Thanks for the comments. We have greatly improved Figure 7 by making wind vectors and geopotential heights (color) more visible and removing the less relevant temperature. In addition, major pressure systems impacting the NYC regional wind conditions have been highlighted on the map.

Fig. 9: A bit larger, please. Add units to color bar & identify symbols (can say: same as in Fig. x)

Thanks for the comments. Please see responses for size in the overall reply and symbols now identified in the added legend.

Fig. 10: Bigger, pls, e.g., hard to see geographic lines. Do not see "thinner dash line".

Please see responses for size in the overall reply. To make each subplot larger in the printed version, we have deleted the less relevant condition at 09:00. "Thinner dash line" indicating the secondary sound breeze was only shown in some of the subplots and was indeed hard to see. Also, given that only one site on the north central shore of LI (Stony Brook) is available, we decided to remove those lines and added/changed the figure caption accordingly.

Fig. 11: Again, bigger please. Define box graph edges and lines

Thanks! We have rearranged Figure 11 and improved its readability. Definitions of the lines are added to the figure caption as: "Boxplots show interquartile range (IQR) with the median value in black; the whiskers spread to the farthest point if within 1.5 times the IQR from the box; data points outside the whisker limits are shown as black open circles."

Fig 12 b and d: Too small

Thanks for the comment again. We have removed the less relevant times and enlarged the figures for the two hours of interest as follows.

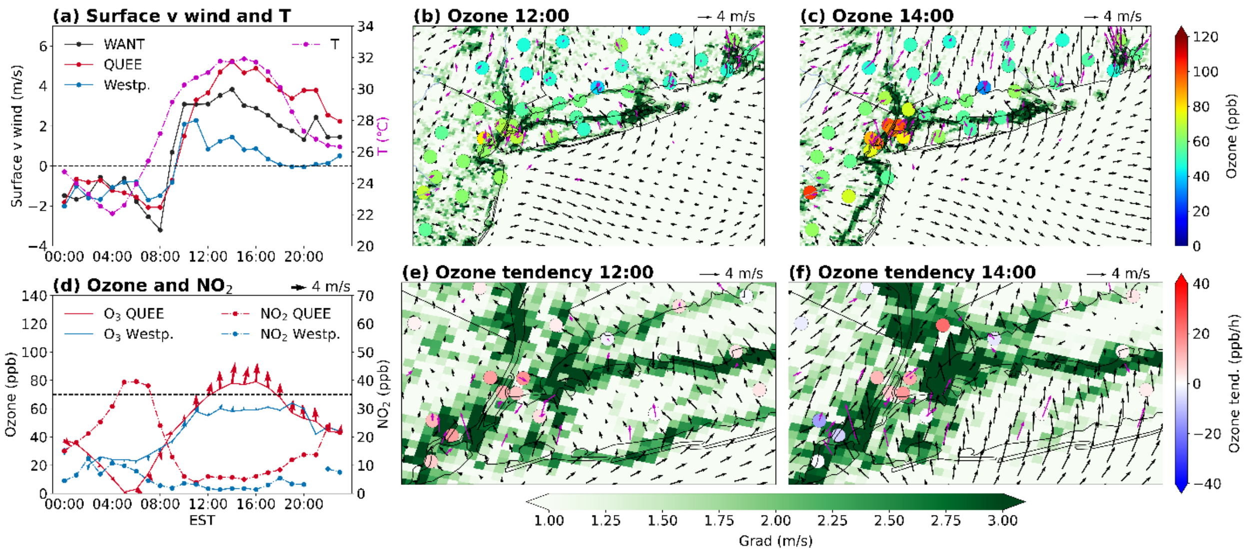

Figure 12…

JGR Technical Format errors and typos: only one example given for each one

Line 61:8-hour should be 8-h or eight-hour (in general: integers <10 are spelled out unless followed by a correct unit abbreviation

Thank the reviewer for the suggestion. However, the daily maximum 8-hour average ozone (DMA8) is a notion used by the regulatory agencies such as US EPA (https://www.epa.gov/green-book/green-book-8-hour-ozone-2015-area-information) and adopted widely. We decided to keep this instance as is.

Line 65: Ref in a list should be in chronological order

They are sorted alphabetically by first author family name according to the guidance from AGU website:

“If a parenthetical citation includes two or more papers, separate the citations with a semicolon and list alphabetically by first author name: (Forbes et al., 1999; Hausler & Wu, 2001).”

Source: https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Grammar-Style-Guide#reference

Line 73: Since Long Island has already been abbreviated as LI, use LI.

Revised.

Line 101: U and V should be u and v

Revised throughout the manuscript.

Line 207: “10m” should be “10 m”

Corrected throughout the manuscript.

Line 210: 4:00-6:00 should be 0400-0600 GMT or LT (or EST), where LT (of EST) is defined at first usage as = GMT- x h

All time references in the manuscript have been changed to 08:00 (EST) throughout. It is easier to read and in a commonly used format in studies such as in Li et al. (2020), following the guideline pasted below. We defined EST based on the UTC at first use and stated it would be used throughout this study “The hour for all hourly data indicates the end of the averaging window in EST (UTC - 05:00, EST is used throughout this study) …”

“For time, use the accepted time standard among your scientific community.”

Source: https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Grammar-Style-Guide#reference

Line 242: 24.2 0C should be 24.20C

We remove the space before “°C” throughout the manuscript.

Line 281: change Figure to Figures

Corrected; also updated for many other instances throughout the manuscript.

Line 424: avoid too-long paragraphs

Thanks for the comment. We double-checked the manuscript and have broken long paragraphs down if possible.

Line 466: change ppp respectively to ppb, respectively

Added “,” after “ppb”; also corrected in original Line 529 for the same mistake.

##

Reviewer #2 Evaluations:
Recommendation: Return to author for **major revisions**
Significant: The paper has some unclear or incomplete reasoning but will likely be a significant contribution with revision and clarification.
Supported: Yes
Referencing: Yes

Quality: The organization of the manuscript and presentation of the data and results need some improvement.
Data: Yes
Accurate Key Points: Yes

Reviewer #2 (Formal Review for Author (shown to authors)):

Review comments on the manuscript "Investigating impacts of local circulation on coastal ozone pollution in the New York Metropolitan area: Evidence from multi-year observations" by Luo and Lu.
General comments:
This manuscript presents an extensive observational analysis that investigates the influence of local wind circulations on ambient ozone levels in the New York Metropolitan area. The study spans three years and incorporates meteorological and chemical observations, including conventional and field campaign data. It focuses on elucidating the association between ozone exceedance events and sea breezes occurring on hot days with weak background winds. The authors identify the change in the v-component of winds as the most critical meteorological factor determining sea breeze onset type and modulating ozone exceedances during extreme hot sea breeze days. The study also incorporates satellite retrieval data to illustrate the spatial variations in O3 formation regime in this complex landscape. The study contributes to our understanding of ozone exceedance events and the driving factors in the region. However, to meet the criteria for acceptance by JGR-Atmosphere, the authors need to further demonstrate the novelty of their study and establish the sufficiency of their findings. Additionally, a deeper analysis is required to enhance our understanding of the mechanisms through which local circulations and high temperatures contribute to ozone exceedance events. Therefore, I recommend a major revision of the manuscript to address the following comments before considering it for publication in JGR.

We thank the reviewer for the overall positive assessment of our manuscript and for providing helpful feedback to help improve the quality of our manuscript. We have addressed all questions and suggestions in our response below as well as in the text and figures, as necessary. Please see detailed responses below and the marked-up version of the revised manuscript.

Major comments:

1. Novelty and significance: The authors should clearly highlight the novelty and significance of their study and the scientific questions they want to address. They should discuss how their findings contribute to the existing literature and what new insights they bring to the field. Additionally, the authors should clearly state why their study is important for the understanding of ozone exceedance events in the New York Metropolitan area specifically.

Thanks for the great comments. We have reorganized the introduction following the logic below and highlighted the research gap on this topic: Ozone problem in the NYC region → Case studies (contributions detailed) across some coastal regions including NYC; however, a comprehensive study is lacking for the region → analyses/group-based comparisons (contributions detailed) such as sea breeze days vs non-SB days; however, contributing temperature is overlooked → introduce what we do (multi-year comprehensive study with temperature considered for the NYC region) including method, data, and organization of the paper. We also highlighted the scientific questions addressed in the introduction, which will help guide the readers through the study. Please see detailed revision in the marked-up version.

2. Methodology: The application of the K-means method is critical for this study. However, the manuscript lacks specific details on how this method was used to classify the data into different groups. It is crucial for the authors to provide a comprehensive explanation of the criteria used when classifying the winds into four clusters: Sea Breeze (SB), Oscillation (O), Southerly (S), and Westerly (W) at QUESS (refer to Figures 4 and the corresponding text at Lines 253-266). The authors should outline the specific parameters and conditions employed to define each category. Merely providing the number of data points within each category is insufficient; a detailed description of the classification methodology is necessary.

Thanks for the great comments. The reviewer might have overlooked our description of the methodology details in Section 2.2. To improve the presentation of the methodology, we have expanded Section 2.2 by adding its mathematical analysis principles, spelling out its input (number of clusters and feature vector of the samples), discussing on its limitations and how we dealt with them. Please see changes in the marked-up version.

3. The structure and organization of the manuscript require some improvements to better represent the new findings and key points. Currently, both key sections (i.e., Sections 2 and 3) of the manuscript focus on the impact of local circulation on ambient levels of O3 on hot days. However, the latter section appears to be more concerned with the differences in O3 and corresponding meteorological conditions between O3 exceedance and non-exceedance events.

To enhance clarity and coherence, I recommend the following revisions:

a) Reorganize the two sections: Consider rearranging the sections to provide a more logical flow and emphasize the main findings. It may be helpful to create a distinct section for the case studies, perhaps titled "Case Studies" or a similar descriptive subtitle.

b) Refine subtitles: Ensure that the subtitles clearly differentiate the two sections and accuratyely reflect their content. The existing difference between the subtitles is not significant, so revising them would help improve the structure.

c) Contextualize the sections: Provide a clear introduction and contextualization for each section to help readers understand the purpose, scope, and significance of the analysis conducted in each part.

Thanks for the comments! Data and methodology in Section 2 and results and discussions in Section 3 is a common way of organizing manuscripts. We assume the reviewer is commenting on the three subsections in Section 3. Please see details below.

a) We aim to study the impacts of local circulation on surface ozone in the NYC area. The logic behind Subsections 1 to 3 is:

Section 3.1 (summertime all-weather) presents a general picture of ozone level and its spatial distribution under ten different meteorological conditions clustered by wind and temperature;

Section 3.2 (hot summer days) focuses on hot days (not moderate nor cool) when ozone is higher and more of concern to human health, etc., and limits the impact of temperature;

Section 3.3 (extreme heat) further zooms into hot sea breeze days under extreme heat with regional daytime temperature higher than 29°C. Since DMA8 ozone is likely to exceed the NAAQS 70 ppb, we checked whether/how local circulation can help modulate ozone exceedances.

b) We think the subtitles are accurate if sufficient context is given. See response to a) for details and c) for exact revisions made.

c) Thanks for this great suggestion! We understand the logic behind the progression of Section 3 might be hard to follow. To this end, we had greatly improved the Introduction to provide a better overview of the study and added more contextualization at the beginning of each subsection (Sections 3.1-3.3) to guide the readers through.

To sum up, we have made significant improvements in the Introduction and transition of the subsections and respectively disagree with the reorganization of the manuscript. Please see improvements in the marked-up version of the revised manuscript.

4. Lines 191-194: The authors state that they replaced certain features from Li et al. (2020) with two new features in order to distinguish between two specific conditions. These new features are the average early morning (4:00-6:00) wind speed (fifth feature) and the ratio of morning and afternoon mean wind speed (sixth feature), obtained by dividing the average early morning wind speed by the average afternoon (14:00-16:00) wind speed. However, it is important for the authors to provide evidence or conduct a sensitivity study to demonstrate how these changes improve the analysis.

Thanks for the comments! This is important to address in the manuscript. We have done a lot of sensitivity studies to find the most appropriate set of features to group wind conditions with minimal similarities among different groups and the highest uniformity inside each group. The ones presented in the manuscript yield the best results. We have added a new Figure S1 to demonstrate the results using the original features. As stated in the manuscript, “days with strong southerly wind and sea breeze days (calm in the early morning and southernly during sea breeze) would not be distinguished using the original recirculation-based features (Figure S1) as these last three features focus on the temporal variation and summation of wind directions.” In addition, mixed wind speed and directions in (b), compared with the approach in the manuscript, demonstrated the effectiveness of the new features used for clustering.

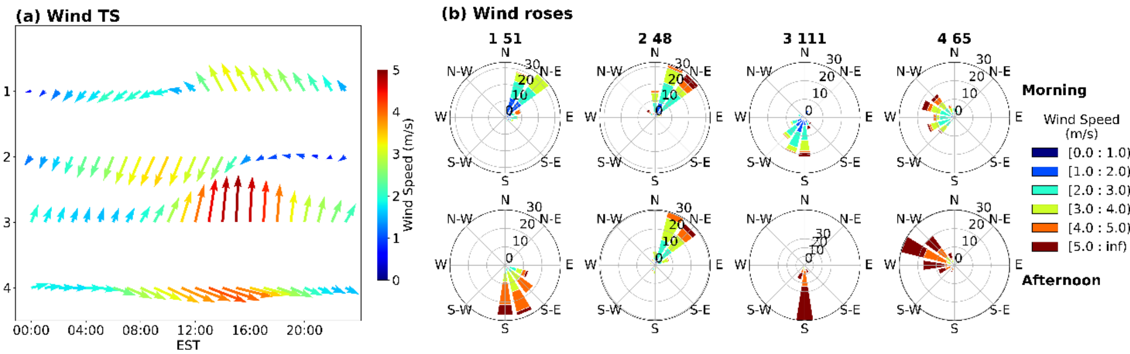

Figure S1. Wind clusters at QUEE based on features in Li et al. (2020). (a) diurnal cycles of mean wind for wind clusters 1-4; (b) early morning (04:00-06:00, top panels) and afternoon (14:00-16:00, bottom panels) wind roses for each wind cluster. The number of days in each cluster is shown in (b). The mixed wind speed and directions in (b), compared with the results in Figure 4, demonstrated the effectiveness of the new features used for wind clustering.

5. Figure 2, Line 220-224: How do you determine the depth of sea breeze? I did not see a large change in wind vectors above and below the depth of sea breeze as indicated by a thick black line in Figure 2.a.

As in the manuscript, "SB_Depth is the height above ground at which the correlation of v wind with 10 m v wind drops below 0.5 starting from 100 m to 2000 m." For example, during this day the correlation of 100 m, 200 m, … v wind and surface v wind are 0.97 (100 m), 0.87, 0.81, 0.78, 0.65 (500 m), 0.54, 0.51 (700 m), 0.41 (800 m), …. 800 m will be the estimated sea breeze depth. The method and the cutoff correlation of 0.5 are somewhat arbitrary as it is hard to make a clean cut between the sea breeze structure and the background wind fields. It is even more challenging for readers to eyeball from the wind vector profile time series in Figure 2a. We wish this could be made easier. In summary, this is one approach to estimate the depth of sea breeze fount, and we acknowledged and emphasized its uncertainty by stating "In addition, stronger upper-level flow and lidar data availability during some of the Hot SB days might challenge the quality of the identified SB_Depth." at the end of this paragraph.

6. Lines 239-246: The authors define three categories of days: hot, moderate, and cold, based on air temperature measurements taken at QUEE. However, it is necessary for the authors to provide further clarification on how these values were calculated and whether a statistical significance test was performed to validate the categorization. Here hot days are clustered with the maximum temperature around 30.7 ºC. I assume 30.7ºC is a threshold value used to define a hot day when the maximum air temperature exceeds 30.7ºC. A similar question for other two threshold values for moderate and cold days. Please clarify them.

Thanks for the comments. This is different from how k-means works. The exact temperatures listed comes from the statistics of the cluster results and thus presented in the result Section 3. We do not set any specific temperature threshold, but just input features (24 hourly temperature) and number of clusters we want. The algorithm determines which cluster certain day belongs to by minimizing the overall variation within clusters. We have updated the descriptions of the approach significantly in Section 2.2. Please refer to the marked-up version for details.

7. Figure 7 and pages 311-322: It is important to expand the weather maps in Figure 7 to include the location of the Bermuda High for each case to support the discussion presented on Lines 311-322. In addition, the quality of these weather maps require further improvement.

Thanks for the comments. We have greatly improved Figure 7 by making wind vectors and geopotential heights (color) more visible and removing the less relevant surface temperature. In addition, relative locations of the major pressure systems (center of those systems might possibly outside of the HRRR domain) impacting the NYC region have been highlighted on the map.

8. Figure 10: The inland penetration distance is an important parameter indicating the strength of the sea breeze. Generally, a longer penetration distance signifies a stronger sea breeze. However, in Fig. 10d, h, and l, the penetration distance for case (l) is much longer than that of cases (d) and (h). The O3 level for case (l) is lower than the other two cases (d and h) over LI and South Coastal CT. This is not consistent with the authors' analyses in other places, where stronger sea breezes are associated with higher O3 levels near coastal areas.

Thanks for the comments. Please see our clarifications below. Penetration distance is one possible parameter to indicate sea breeze strength (not the SB_Str defined in this study as in Section 2.3 and Figure 2, which we meant the change in v wind because of the sea breeze onset) if the background wind is comparable. For instance, the long penetration of sea breeze in Figure 10l is because of the favorable background southerly wind, but not necessarily the sea breeze induced by the land-sea temperature contrast. Thus, we used SB_Str instead of penetration distance to indicate the strength of the sea breeze. Also as in the manuscript, "lowest ozone levels are observed during Southerly days for all three regions because of the presence of marine air (warmer during nighttime and colder during daytime) throughout the day brought in by the steady southerly flow."

We assume by "in other places" the reviewer is referring to Section 3.3 as it is the only section in which we talked about some associations between sea breeze strength and ozone levels. Only some of Hot SB days are discussed in Section 3.3. Abrupt stronger sea breeze cases (higher SB_Str) are associated with higher ozone at NYC city center (updated Figure 12c), such as Queens (updated Figure 11i) under extreme high heat (Hot cluster with regional T > 29°C) and weak background flow (SB cluster); In contrast, gradual weaker sea breeze (lower SB_Str) is associated with higher ozone along LIS, especially coastal CT (updated Figure 14c), such as Westport (updated Figure 11c).

In order to make this more evident to the readers, we have improved our sectional description in the introduction and updated the transitional description before each section with your comments in mind. Please see details in the marked-up version, and many thanks for your comments!

9. Lines 411-412: The statement "we limited our analysis to 16 Hot SB days with regional daytime temperature (T) higher than 29.0ºC, defined as extreme heat/hot days" seems to be slightly contradictory to the statement provided on Line 241, where 30.7ºC is used to define hot days. Could you please clarify this discrepancy?

Thanks for the comments. Please see the clarification below:

1. 30.7ºC is the maximum temperature of hot day hourly median temperature at QUEE and is a statistical value based on the cluster results (not a selected threshold).

   To clarify, we modified "As a result, three summers of temperature diurnal cycles are grouped into Hot, Moderate, and Cool days." to "The resultant clusters are named Hot, Moderate, and Cool days based on their diurnal temperature profile, and their statistics will be discussed in Section 3.1." in Section 2.2.

2. 29.0ºC is a threshold selected for regional average daytime mean temperature during 07:00-17:00 (T) at the eight NYSM sites (the corresponding maximum hourly temperature at QUEE mean would likely be higher than 30.7ºC on the 16 selected days).

10. Figure 11 and Lines 430-435: There are minimal differences observed in several key features, such as breeze depth, onset time (SB_OnT), arrival time (SB_ArrT), regional mean temperature (T), etc., between days with ozone (O3) exceedances and non-exceedance days. However, it remains unclear why some days experience O3 exceedances while others do not. Understanding the underlying factors contributing to this phenomenon should be a priority for the authors, as it requires further investigation on the mechanism rather than solely focusing on establishing statistical relationships.

Thanks for the comments. It is unclear till this portion, but the analysis of the underlying factors follows. To make it more evident to the readers, we broke this paragraph into two right after this part. The new paragraph explains the contributing factors and leads up to the two case studies that detailed the progress/mechanism of these two different scenarios.

11. To facilitate comparison and understanding, it is recommended to include one or two tables that summarize the relationship between wind clusters, temperature, ambient O3 levels, and exceedance events (number or frequency). These tables will effectively highlight the key features depicted in Figures 4, 5, 8, 9, and 10.

Many thanks for the great suggestion! We have added two tables (Tables 1 and 2) for direct quantitative comparison, one for those all-weather 10 clusters, as in Figure 5, and one for those three hot day clusters, as in Figure 9. Please see marked-up version of the manuscript for details.

12. Section 2.3: It will be helpful if the authors create a new table to summarize the criteria that are used to define sea breezes?

As in the responses to Comment 2, we adopted the six features aiming to cluster wind conditions based on the presence of sea breeze and the condition of synoptic flow. We did not specify certain criteria to define sea breeze, but the cluster Sea Breeze is separated from cluster Southerly mostly by the fifth and sixth feature: calm morning wind and large scale of speed increase in the afternoon. The other two cluster (W and O) have completely different wind directions in the morning and afternoon (first-fourth features). We have clarified the general term sea breeze and the SB cluster as follows in the revision:

"The Southerly cluster represents 36 days with consistent southerly wind throughout the day. Its daytime southerly wind tends to be strengthened by sea breeze due to the land-sea temperature contrast, and its wind speed peaks early around noontime compared with 15:00 in the SB cluster with light synoptic flow in the region. Note that sea breeze, any wind that blows from a large body of water toward or onto a landmass, as a physical phenomenon might happen under many synoptic-scale wind conditions, such as in the Southerly cluster; the italicized cluster Sea Breeze (SB) only refers to the days in that cluster with relatively calm background wind fields."

13. It is useful by creating another table to compare the sea breeze features, ozone levels, and exceedance probability among different wind clusters or local circulations.

We have incorporated the comparison of ozone levels, and exceedance probability among different wind clusters or local circulations into summary tables and the main text (Pex) in response to Comment 11. However, we only have "reliable" and "comparable" sea breeze feature estimates (Section 2.3) during hot sea breeze days. Detailed meteorological and ozone conditions for all 37 hot sea breeze days are now included in the supplemental Table S1.

14. The conclusion section requires further refinement to accurately represent the key points derived from the discussion and analysis presented in Section 3.

Thanks for the comment! We have refined the section by adding more to what was done and specific quantitative results and filling in the importance and limitations of our work. Please see the marked-up version for details.

Minor comments:

1. Figure 3: Y-axis label, T[C] should be T[ºC].

Corrected for all figures.

2. Lines 228-236: It would improve clarity if the authors assigned numbers to each equation or formula. Additionally, it is helpful by providing the specific threshold values that are utilized to define the sea breeze front locations.

Thanks for the suggestion. Equations are now numbered. However, we are unable to pinpoint a specific threshold as there is generally large variability in the wind gradient on different days (Take Figures 12 and 14 for example). Any visible Lines formed by connected higher Grad, with a minimum of 1 m/s in the color bar, can be interpreted as a (sea breeze) front.

3. What are the units presented in Panels a-b, and d-e of Figs. 13 and 15?

Many thanks for catching the missing units here. It should be molecules/cm^2 and has been added to the figures.

4. L105: Please spell out "PCA" and make sure all abbreviations are defined throughout the manuscript.

Thanks! It should be principal component analysis (PCA). We have also cleared the manuscript for all abbreviation definitions. However, this instance of PCA is deleted in the revision as the contents are less relevant.

5. L493: The sentence "Sound breeze ... are ..." contains a grammatical error that needs to be corrected. Please double-check the entire manuscript to ensure that all grammar errors are avoided.

Corrected. We have also carefully screened the entire manuscript and corrected any grammar errors we found.

##

Reviewer #3 Evaluations:
Recommendation: Return to author for **minor revisions**
Significant: The paper has some unclear or incomplete reasoning but will likely be a significant contribution with revision and clarification.
Supported: Mostly yes, but some further information and/or data are needed.
Referencing: Yes
Quality: The organization of the manuscript and presentation of the data and results need some improvement.
Data: Yes
Accurate Key Points: Yes

Reviewer #3 (Formal Review for Author (shown to authors)):

Review attached as PDF file and pasted below.

Review comments

The manuscript presents an interesting study of the sea breeze impacts on the high ozone events in New York metropolitan area and nearby down-wind locations during the summers of 2017-2019.
A K-mean cluster method with the temperature and wind parameters is used to classify the meteorological fields and sea-breeze onset. The paper is generally well written and well organized. However, the K-mean cluster method and the definition or criteria for a sea-breeze event are not clearly described.

We thank the reviewer for the positive assessment of our manuscript and for providing helpful feedback to help improve the quality of it. We added more details for the k-means in the methodology section and cleared out the sea breeze criteria in the responses below and the manuscript. We have also addressed all other questions and suggestions in our response below. Please see detailed responses below and the marked-up version of the revised manuscript.

Details as follows.

1. The paper needs a better descriptions of how the K-mean cluster method is utilized and how it works. What are the input and feature parameters used for classifying the cluster number?

Thanks for the comments. We have added a more comprehensive description of the K-mean cluster by adding its mathematical analysis principles, spelling out its input (number of clusters and feature vector of the samples), discussing on its limitations and how we dealt with them. Specifically, features for temperature are "24 hourly mean temperatures at QUEE from 00:00 to 23:00"; six features for wind are described as follows:

"The first four features of morning and afternoon wind components remain similar: morning zonal (u, first feature) and meridional (v, second feature) winds were averaged during the early morning of 04:00-06:00 and afternoon u (third feature) and v (fourth feature) winds were averaged during the early afternoon of 14:00-16:00, respectively. … Thus, we replaced these last three features with two new ones designed to separate southerly from sea breeze days: average early morning (04:00-06:00) wind speed (fifth feature) and the ratio of morning and afternoon mean wind speed (sixth feature) by dividing the average early morning wind speed by the average afternoon (14:00-16:00) wind speed."

2. A sea breeze event is not clearly defined in this paper. How are the local sea-breeze events isolated or distinct from the synoptic-scale wind fields?

Thanks for the comment. Wind conditions are clustered based on the presence of sea breeze and impact from background synoptic scale flow. For example, to separate S and SB cluster around the city center, we replaced some of the original features (Li et al., 2020) used in the k-means clustering with early morning wind speed and its ratio with afternoon wind speed as their summed diurnal wind vectors are both southerlies. Later on, we figured out that many of the W days at the city center are not strong enough to prevent sea breeze forming at the LI south coast; we further divided W clusters into W_w (strong westerlies in the region with no sea breeze developed at all) and W_sb (light westerlies around city center and sea breeze developed at the LI south coast). To sum up, sea breeze developed in all three hot day clusters (SB, W_sb, and S) as discussed in Section 3.2 with different synoptic flow impact as illustrated in Figure 7. Please see updates in the marked-up version and key clarifications are pasted below:

"Southerly cluster represents 36 days with consistent southerly wind throughout the day. Its daytime southerly wind tends to be **strengthened by sea breeze as a result of the land-sea temperature contrast** and its wind speed peaks early around noontime compared with 15:00 in the SB cluster with light synoptic flow in the region. **Note that sea breeze, any wind that blows from a large body of water toward or onto a landmass, as a physical phenomenon might happen under many synoptic-scale wind conditions, such as in the Southerly cluster; the**

italicized cluster Sea Breeze (SB) only refers to the days in that cluster with relatively calm background wind fields."

"Weather maps based on composites of HRRR reanalysis of 850 hPa geopotential height and wind fields during Hot day clusters (SB, W_sb, and S) are shown in Figure 7. **It should be noted that sea breeze developed from the Atlantic Ocean in all three clusters. However, the characteristics of the sea breeze vary in geographical extent, magnitude, and timing.** The Bermuda High pressure system ("H" in Figure 7) sits near the Southeast Coast of contiguous United States (CONUS) during summertime, generating light southwesterlies in the study region during Hot SB days (Figures 4a and 7a). When the low-pressure system in southeastern Canada ("L" in Figure 7) moves eastwards and pushes the north part of Bermuda High south during Hot W_sb days, westerly components of the local wind around the NY metropolitan area strengthen, delaying the sea breeze development near the coast and preventing its penetration further inland to the city center (Figures 4a, 6, and 7b). Strong southerly wind in the New York Bight and its coastal region are generated during Hot S days as center of the Bermuda High advances closer to the East Coast of CONUS with its large-scale clockwise rotation (Figure 7c). The southerly wind strengthens during daytime as the Atlantic sea breeze develops (Figure 4a)."

3. The paper uses the NYS-Mesonet surface meteorological parameters and wind lidar data at Wantagh to identify the sea-breezes from NY Bight and assess their impacts on the near-surface O3 events in NY metro area. However, according to its geolocation, the other NYS-Mesonet site at Staten Island is located on the path of Sea-breezes from NY Bight, why not use the winds data at this site or other nearby NYSM sites? In addition, authors compare the O3 variation at QUEE and Westport. But, the Long Island Sound-breeze information at Westport is not discussed as those at WANT.

Thanks for the comments. The reason for choosing the Mesonet site at Wantagh, instead of Staten Island, to study the feature of sea breeze is: sea breeze fronts passing by NYC, especially the AQS ozone site at Queens, are likely from the south (Figure 4a), similar to those at Wantagh with west-east coastline; while Staton Island is nested much "inland" and surrounded by multiple water bodies such as Upper Bay, and many of the sea breeze fronts arrive at NYC, especially the AQS ozone site at Queens, might not arrive at Staton Island given it is location and the southwesterly background flow in the region (Figures 7, 10, and 14). Even when sea breeze arrives at Staton Island, it is often later during the day with direction from south east (e.g. Figure S5). Please refer to the updated zoomed-in map in Figure 1 for locations of the sites.

Thanks for bringing up the Long Island Sound breeze at Westport. Normally, sound breeze develops not long after sunrise (Figures 12a and 14a, as well as in S4-S7) and remains southerly flow throughout the day. We have added the following to the case study discussion.

Case 1: "LIS sound breeze towards coastal CT developed at 10:00 as the land surface temperature rose. It remained a southerly flow throughout the day, making ozone over coastal CT sensitive to ozone and its precursors over the LIS."

Case 2: "In the meantime, the LIS sound breeze towards coastal CT developed and remained a southerly flow throughout the day as in the abrupt strong sea breeze case. … Finally, the sea breeze front arrived at Westport in the late afternoon (around 17:00), and the decreasing of ozone slowed down with its progression towards coastal CT (Figures 14, S6, and S7)."

4. The paper utilizes the winds and temperatures for the cluster analysis and shows promising results. The cloud covers and PBL-height are not discussed for the two case studies. Do they potentially affect the spatial patterns of the ground O3 in the region of study?

Thanks for the comments. We have incorporated a discussion of cloud cover and PBL height in the introduction, as pasted below. We suspect these two factors are highly coupled with temperature and have an overall impact on regional ozone levels but not the spatial patterns of ground ozone, as reflected in the three temperature clusters in Figure 5 and S1. This can be explored further when we get access to high-quality cloud cover and PBL height data at relevant spatial scales.

"Higher temperatures often indicate clear sky conditions with higher levels of solar radiation reaching the lower atmosphere. Together, they favor higher surface ozone concentration by 1) accelerating photochemical reaction rates and 2) boosting VOCs and NOx emissions from both natural and anthropogenic sources through various processes, such as accelerated biogenic emission of VOCs and increased NOx emissions from higher energy demand for air conditioning (Coates et al., 2016; Guenther et al., 1995; Porter & Heald, 2019). In addition, higher temperature facilitates the growth of a deeper planetary boundary layer (PBL) due to increased convection driven by surface heating. However, the impacts of PBL height on surface ozone are often coupled with influences from other physical and chemical factors, including the absolute height of PBL and ozone vertical profile (Haman et al., 2014; Zhang et al., 2023). The overall effect of PBL height on surface ozone is less certain. In general, the highest surface ozone most likely occurs when PBL height is moderate (between 1-2 km for Beijing) and decreases as PBL becomes higher (dilution) or lower (decreased photochemistry because of the availability of sunlight) (Zhang et al., 2023)."

5. There is some confusion in Fig.4 (QUEE-site) and Fig.2 (WANT site) that show the seabreeze wind information at Queens College and Wantagh sites. Fig.2 shows the seabreeze information with the surface winds at Wantagh. Do the surface horizontal winds show similar directions at Queens College and Wantagh sites?

In general, yes, since they are located only ~30 km away (updated Figure 1); except for when sea breeze develops along the coast, but not penetrated into the city center during W_w days in the context of this study.

Wind conditions at these two sites are used in conjunction to reflect the regional conditions. WANT is used mainly to derive sea breeze features developed from the Atlantic Ocean because of its proximity to the coast, far from other local impact such as UHI. While QUEE is in the much-polluted city center with the highest percentage of days with DMA8 exceeding NAAQS of 70 ppb, its wind condition is chosen to cluster the local wind condition first. For Hot W days, the secondary cluster of W_w and W_sb is used to separate out the days when W is so strong that

SB won't be able to develop at all along the coast. During those days, the area is occupied by the cleaner air from northwest.

6. Line 297-298, how to identify that the sea-breezes cannot penetrate into NYC urban area? Can you give the front positions of the sea-breezes?

Thanks for the questions. This sentence is a summary of the paragraph. The way to identify those cases is detailed in the following sentences. In short, they are determined by the v wind direction at the south shore of LI (WANT). Figure 6 shows the wind condition at WANT for sea breeze developed at the coast but not penetrated to NYC (W_sb), and no sea breeze developed at all (W_w). Sea breeze front positions for hot SB, W_sb, and S are identified by the dash lines in Figure 10 based on surface wind observation composites. Note the stalled sea breeze fronts along the south shore of LI (not penetrate into NYC urban area) in the W_sb cluster because of the strong westerly flow.

7. Line 371 and Fig.9, what area does "LICTe" represent? How many ground sites of the ozone are available for the LICTe?

"LICTe" compose of 6 sites in "LI and coastal region of eastern CT", which is defined in the 4th paragraph of Section 3.2. Number of sites in all three regions are now included in the legend of Figure 9h.

8. Fig.10, for the Sound breezes marked along the CT coast, are you using the HRRR model data or the surface observation data?

Thanks for the comment. As in the original figure caption: "Sea breeze front locations are manually marked with dashed gray lines (…) based on the observed surface wind (…)."

9. Fig.S5, authors compare the O3 variation and wind component V at Queen-College (QUEE) and Westport, and indicate the effects from sea breezes. However, WANT and Westport are located at the different coastal areas. In particular, Westport might be more affected from the Sound breezes or sea –breezes from Long Island Sound. A plot of the wind component U to identify the Sound breeze onset at Westport will be very useful.

We agree that additional information on sound breeze would benefit the discussion. However, v wind component would be more representative of it given the LIS/CT coastline geometry. Please see v wind examples in case studies (Panel a in Figures 12, 14, and S4-S7) for the onset of early morning sound sea breeze and later sea breeze arrival (v wind increase with a likely secondary ozone increase period) at Westport.

Please refer to responses in 3 for the added discussion on LIS sound breeze at Westport.

10. Line 542-545, this study does not show the evidences on the shallow marine boundary layer and active ozone production at Westport.

Right, this is not a direct observation but a speculation based on the sharp hourly increase of ozone. We have revised it accordingly by adding “likely” in the statement.

Some minors:

1. Fig.2, the negative/positive values of the wind speeds represent different wind direction. Please describe it.

Thanks for the suggestion to make it more apparent in the caption. The negative/positive values are of the meridional wind direction. We have added “negative values indicate wind from the north and positive values indicate wind from the south; the same convention applies throughout this study” to the caption.

2. Line 335 and Fig. 8, Pex definition?

It was defined in Original Line 332 as “the probability daily DMA8 exceeding the NAAQS level of 70 ppb” and also repeated in the original figure caption as “percentage of days with DMA8 exceeding NAAQS of 70 ppb”.

3. Fig.8 caption, describe what the “W-sb” and “S” represent.

Thanks for helping make the figure captions more reader-friendly. However, we would argue this one is unnecessary: wind clusters are the backbones of this study and are used throughout the manuscript. They are well defined and demonstrated in Section 3.1 and 3.2. Adding the long full description here in the figure caption would be too distracting.

4. Fig.S5, the X-axis needs a label.

Thanks for the comment! “EST” has been added in the revision.