Reviewer #1 (Comments to the Author):

Title: On the dynamics of ozone depletion events at Villum Research Station in the High Arctic

This manuscript presents a statistical analysis of near surface ozone observations over a 23 year period at Villum Research Station in the high Arctic, utilizing local meteorological observations, backward air mass trajectory modeling, and statistical analysis to elucidate mechanisms controlling observed ODEs. The dataset and analysis are interesting, and the majority of the discussion section is really well done. However the data analysis suffers from some major issues that need to be rectified.

> We thank the reviewer for their comments and suggestions. We have addressed each comment below with review comments in black, author response in blue, and additions to the original text in red. We have indented the author's response for clarity. Lines numbers given in the author's response refer to lines in the revised manuscript.

The way this paper is written suggests a fundamental misunderstanding of the role of sea ice regions in halogen activation and ozone depletion chemistry. **Sea ice has snow on it!** I'm sure the authors are aware of this fact, but the analysis and discussion give the impression that they believe snow only exists on land. The physical surface of the sea ice itself does not have a pH conducive to halogen activation chemistry (Abbatt et al 2012, Wren et al 2013, Pratt et al 2013). It is the snow in sea ice regions that drives the halogen chemistry. Your analysis and discussion of snow vs sea ice needs to be completely reworked to reflect the complexity of sea ice regions. An analysis of the surface temperature along the back trajectory would potentially help with determining the potential for halogen activation along the back trajectory.

> The authors are fully aware that sea ice has snow on top of it. The satellite products used in this study (National Oceanic and Atmospheric Association/National Environmental Satellite, Data, and Information Service (NOAA/NESDIS) Interactive Multisensor Snow and Ice Mapping System (IMS)) provides information for each grid cell about the underlying surface, where each grid cell can belong to only one of five categories (0: Outside the coverage area, 1: Sea, 2: Land (without snow), 3: Sea ice, 4: Snow covered land). The product does not indicate if the sea ice is covered by snow or not. This is due to the similar spectral signatures of sea ice and snow which makes differentiation difficult, although IMS integrates different data sources, ancillary data, and advanced algorithms for surface mapping, they do not always provide clear delineation between sea ice and snow (U. S. National Ice Center, 2008). In essence, the data source we used cannot definitively discern if there is snow on top of sea ice or not. Therefore, we chose to keep the original labels from the satellite products in our analysis to remain true to the original data product. We do admit this issue with snow vs sea ice detection was not described adequately in the Methods section. We have amended the text to indicate this. We have also addressed this shortcoming in the satellite product in the Results and in the Discussion sections.
>
> Lines 195-213: For each trajectory, a surface-type footprint analysis was performed. The underlying surface types used for the surface footprint type analysis were produced by the National Oceanic and Atmospheric Association/National Environmental Satellite, Data, and Information Service (NOAA/NESDIS) Interactive Multisensor Snow and Ice Mapping System (IMS) developed under the direction of the Interactive Processing Branch (IPB) of the Satellite Services Division (SSD). The altitude at each step along the trajectory was compared to the height of the mixed layer. Steps were classified as being above the mixed layer (AML) if the

trajectory altitude was above this height. If the trajectory altitude was below this height, then the underlying surface type (land without snow, sea, sea ice, or snow on land) was recorded using a polar stereographic map of the Northern Hemisphere classified into 1024×1024 24 km grid cells. It is important to note that grid cells classified as sea ice likely contain snow on the surface, although the satellite products used in this study does not differentiate between bare sea ice and snow-covered sea ice, likely due to the similar spectral signatures between sea ice and snow (U. S. National Ice Center, 2008). We opted to keep the original labels from the satellite product for this analysis, as we cannot make any definitive statements about the presence of snow on top of sea ice. The reader should keep this in mind when interpreting the results. The time spent over different surfaces is expressed as a percentage of the total trajectory length.

Lines 380-381: It should be kept in mind that the air mass history variable, time spent over sea ice, does not give information about the presence of snow cover and only if the underlying surface was classified as sea ice or not.

Lines 502-504: The air mass history variable, time spent over sea ice, does not give information about the presence of snow cover and only if the underlying surface was classified as sea ice or not.

Lines 895-897: It should be noted that the snowpack on top of sea ice is the likely source of these halogens, given that the surface of sea ice is not conducive for halogen activation (Abbatt et al., 2012), although the satellite product used in this study cannot differentiate between snow covered sea ice and bare sea ice (see Methods).

An analysis of the temperature along the trajectories would indeed be an interesting analysis method. We have extensively examined the air mass history and source regions for ODEs within the mixed layer, which agree with previous literature indicating the central Arctic Ocean is the primary source region for ODEs and enhanced halogen levels (Ahmed et al., 2023; Begoin et al., 2010; Bognar et al., 2020; Bottenheim and Chan, 2006; Bougoudis et al., 2020; Oltmans et al., 2012; Seo et al., 2020). Halogen activation requires a frozen, acidic heterogeneous surface (Burd et al., 2017; Jeong et al., 2022; Sander et al., 2006) and given that the temperature in the springtime central Arctic Ocean is usually below freezing, such an analysis would likely give very similar results to the source region analysis already performed in this study thus likely not yielding new information. We feel we have adequately and comprehensively analyzed the source regions of ODEs observed at Villum with our current analysis utilizing frequency maps for trajectory steps below the mixed layer, time over different surface types, their geographic distribution, and their temporal dependences, all of which agrees well with previous studies.

The selection of the time period for further analysis seems arbitrary, as ODEs don't necessarily follow a clear Mar-May pattern as seen in Fig 2, particularly at these high latitudes. The paper would be strengthened if the time period analyzed were empirically defined utilizing the first to last ODE day. You could choose the earliest and latest over the whole study period to have a consistent time frame across years. It might end up being March to May still but at least you would have a better justification for the choice.

We selected the March-May period for an in-depth analysis as this is the main occurrence of ODEs throughout the Arctic as demonstrated by numerous previous studies (Barrie et al., 1988; Bottenheim and Chan, 2006; Simpson et al., 2007; Whaley et al., 2023) and from analyzing the results displayed in Fig. 2, which shows ODEs *are* mainly confined to the spring (March, April, and May) season. Furthermore, from Fig. 2, it is evident that no observations below 10 ppbv occurred in February (likely due to the absence of sunlight) and only a few occurred during the first part of June. Indeed, the ODE frequency for June is 2.37 % which is slightly less than March (3.88 %) and ODEs in June only occurring during the first few days of the month (Fig. 2a). Different environmental conditions during the summer compared to spring also contributed to this decision. June (and other summer months) also regularly experiences temperatures near or above freezing, therefore, the mechanisms behind ozone depletion during the summer are likely different from the spring as halogen propagation needs an acidic frozen, heterogeneous surface (Burd et al., 2017; Jeong et al., 2022; Sander et al., 2006). During the summer months, there is limited transport of ozone and its precursors from the mid-latitudes (diminished sources), the low absolute humidity and low NO_x levels limit in situ photochemical production, and increased areas of open water and bare land increases dry deposition compared to ice covered surfaces (increased sinks) (AMAP, 2015; Barten et al., 2021, 2023). For these reasons, we explicitly selected the months of March to May for a comprehensive and systematic analysis of ODEs and excluded June due to the low frequency of ODEs and different environmental conditions affecting ozone variability. Therefore, our selection is not arbitrary but based on the main occurrence of ODEs across the Arctic as noted in the numerous previous studies and the initial results from this study. If we understand the reviewer's suggestion correctly, using the first and last ODE day to define the ODE season would result in losing the monthly information provided in this study as the analysis would be limited to only ODEs vs Non-ODEs over this period and would not examine any temporal dependencies. Using the first and last ODE days and analyzing each month individually (as done in this study) would only result in a few additional ODEs for the first few days of June. We are also planning a separate publication which analyzes the dynamics of ozone during the summer months (June, July, and August) as described below.

Ozone seems to be persistently below background through the summer months, this is an interesting finding that merits more discussion/analysis. In my view this is a big missed opportunity by the authors particularly given the low number of ozone observations at this latitude and the discussion of the potential role of iodine motivated by the MOSAIC papers (e.g. Benavent et al 2023).

We agree the dynamics of ozone during the summertime is an interesting topic that warrants further discussion. The focus of this paper was confined to the springtime ODEs for reasons described above. Therefore, we were already planning a follow up publication on the dynamics of ozone in the High Arctic summertime. This paper will examine the observations below 10 ppbv in the first weeks of June to determine if they are caused by the same mechanisms as ozone depletion in spring, the role of IO in summertime ozone destruction, the role of entrainment from aloft on ozone levels (preliminary findings indicate that subsidence of dry, ozone rich air from above the mixed layer contributes to enhanced ozone levels). This work is currently under preparation and will be submitted to ACP in due time.

The description and utility of a SHAP value needs to be in the main text as the whole ML discussion relies on the reader having an understanding of those values and being able to interpret them. Additionally, Section 3.4 needs to be revised for clarity, I've read it a few times and I'm not entirely sure what I am supposed to be taking away from this section, especially figure 10. Maybe folks with a background in machine learning will find value here, but the broader community I think is going to be lost.

We have moved the description of the ML methodology including the description of the SHAP methodology to the main text for readers unfamiliar with these concepts. We originally included them in the SI for brevity.

Sect. 3.4 of the Results section describes the results of the ML model and we discuss these results in the Discussion Sect 4.2. We begin by highlighting why we utilized ML and its added benefits over statistical analysis, which were both performed in this study. We then detail the accuracy and applicability of our ML model through an evaluation of its predictive performance using robust and comprehensive evaluation metrics for a classification ML model. The most important features in each month are then described (which is not evident using classical statistical analysis). The relationships between the input features and their contribution to the model prediction are then analyzed. From Fig. 10, the relationship between the input features and their contribution to the model output is displayed, this gives information about how certain levels (and threshold ranges) of the input features affects the model's prediction of an observation being an ODE or not. For example, this is especially evident for solar radiation (Fig. 10d), which shows that after the 112 to 153 W m^{-2} bin range solar radiation starts to make a positive contribution to the model prediction an ODE (i.e., the model is more likely to predict a positive label or an ODE) and below this range it makes a negative contribution (the model is more likely to predict a negative label or Non-ODE). Such a threshold range is not evident from Fig. 4e. Overall, this section shows the ML brings added value to our analysis, our ML model is robust and accurate, the input features that are most important to modeling ODEs, and how the features affect the model prediction.

We have added the following lines to make this clearer in the text:

Lines 566-582: The evaluation metrics of the ML for all spring months combined and individual months are displayed in Table 1. We use three common metrics for evaluating a binary classification ML model: accuracy, recall, and AUC ROC (Area Under Curve Receiver Operating Characteristics). Briefly, accuracy is the fraction of correctly predicted observations regardless of label (ODE vs Non-ODE), recall is the fraction of ODEs correctly predicted and AUC ROC evaluates how well a model can discriminate between positive and negative labels across all decision thresholds for binary classification. In general, the ML model can accurately reproduce ODEs over all spring months combined as evidenced by how all three metrics are close to unity (their maximum value). However, when evaluating the results on an individual monthly basis, there is an increase in the recall metric and decrease in the accuracy and AUC ROC (see Sect. 2.6 for a detailed description of the evaluation metrics) from March to May (Table 1), which is likely connected to the increasing frequency of ODEs from March to May. With increased ODE occurrence, the recall metrics would increase as positive labels (ODEs) are more likely to be identified when they occur more often and the accuracy and AUC ROC metrics would decrease with the increased occurrence of positive labels due to a concurrent increase in number of incorrectly labeled ODEs. The ML model is also free from over-fitting given the close agreement between the train and test sets. Overall, this ML model is sufficiently accurate, robust, and suitable for the investigation of ODEs.

Caption of Table 1: The accuracy gives an overview of the model performance for both labels (ODEs vs Non-ODEs), recall gives the model performance only for positive labels (ODEs), and AUC ROC evaluates the model performance over different decision thresholds, together, these three metrics give a comprehensive view of the model's performance. The three metrics range from 0 (worst) to 1 (best).

Lines 593-605: The SHAP approach is designed to estimate the importance of each input variable to the model output based on coalitional game theory (Molnar, 2022) (see Sect. 2.6 for a more detailed description). SHAP values represent the marginal contribution of each input variable to the model output, or in other words: how each observation for each variable affects the model's prediction. SHAP values can be positive or negative, with positive values indicating a variable is more likely to contribute to an observation being predicted as an ODE while negative values mean a variable is more likely to contribute to an observation being labeled as a Non-ODE. SHAP can produce both local and global explanations. The global importance gives an overview of the most important variables to the model output. The local importance of each observation can give information about the relationship between the SHAP and input values (positive or negative relationship, linear or non-linear), or in other words how does the model output vary over the range of input values.

Minor points:

Line 284: Given that high wind speed enhances vertical mixing it is not surprising that ozone would not be depleted during those conditions.

As noted in the literature, both low and high wind speed can have an effect on ozone variability (Blechschmidt et al., 2016; Choi et al., 2012; Jones et al., 2009; Zhao et al., 2016), therefore while maybe not surprising this observation is of importance. We have noted the dual effect of wind speed on ODEs in the Results and Discussion and we will add this insight to our discussion.

Lines 841-846: High wind speeds can also enhance vertical mixing of ozone enriched air masses from aloft, which could mask the contribution of halogen activation from blowing snow. Only during May does high wind speeds regularly make a positive contribution to the model output, and the magnitude of this contribution is small (Fig. 10b). Overall, the rare occurrence of high wind speeds (Fig. S4b) hinders any definitive conclusions about their effect on ODEs.

Abbatt, J. P. D., Thomas, J. L., Abrahamsson, K., Boxe, C., Granfors, A., Jones, A. E., King, M. D., Saiz-Lopez, A., Shepson, P. B., Sodeau, J., Toohey, D. W., Toubin, C., von Glasow, R., Wren, S. N., and Yang, X.: Halogen activation via interactions with environmental ice and snow in the polar lower troposphere and other regions, Atmos. Chem. Phys., 12, 6237–6271, https://doi.org/10.5194/acp-12-6237-2012, 2012.

Ahmed, S., Thomas, J. L., Angot, H., Dommergue, A., Archer, S. D., Bariteau, L., Beck, I., Benavent, N., Blechschmidt, A.-M., Blomquist, B., Boyer, M., Christensen, J. H., Dahlke, S., Dastoor, A., Helmig, D., Howard, D., Jacobi, H.-W., Jokinen, T., Lapere, R., Laurila, T., Quéléver, L. L. J., Richter, A., Ryjkov, A., Mahajan, A. S., Marelle, L., Pfaffhuber, K. A., Posman, K., Rinke, A., Saiz-Lopez, A., Schmale, J., Skov, H., Steffen, A., Stupple, G., Stutz, J., Travnikov, O., and Zilker, B.: Modelling the coupled mercury-halogen-ozone cycle in the central Arctic during spring, Elementa: Science of the Anthropocene, 11, 00129, https://doi.org/10.1525/elementa.2022.00129, 2023.

AMAP: AMAP Assessment 2015: Black carbon and ozone as Arctic climate forcers., Arctic Monitoring and Assessment Programme (AMAP), 116, 2015.

Barrie, L. A., Bottenheim, J. W., Schnell, R. C., Crutzen, P. J., and Rasmussen, R. A.: Ozone destruction and photochemical reactions at polar sunrise in the lower Arctic atmosphere, Nature, 334, 138–141, https://doi.org/10.1038/334138a0, 1988.

Barten, J. G. M., Ganzeveld, L. N., Steeneveld, G.-J., and Krol, M. C.: Role of oceanic ozone deposition in explaining temporal variability in surface ozone at High Arctic sites, Atmos. Chem. Phys., 21, 10229–10248, https://doi.org/10.5194/acp-21-10229-2021, 2021.

Barten, J. G. M., Ganzeveld, L. N., Steeneveld, G.-J., Blomquist, B. W., Angot, H., Archer, S. D., Bariteau, L., Beck, I., Boyer, M., von der Gathen, P., Helmig, D., Howard, D., Hueber, J., Jacobi, H.-W., Jokinen, T., Laurila, T., Posman, K. M., Quéléver, L., Schmale, J., Shupe, M. D., and Krol, M. C.: Low ozone dry deposition rates to sea ice during the MOSAiC field campaign: Implications for the Arctic boundary layer ozone budget, Elementa: Science of the Anthropocene, 11, 00086, https://doi.org/10.1525/elementa.2022.00086, 2023.

Begoin, M., Richter, A., Weber, M., Kaleschke, L., Tian-Kunze, X., Stohl, A., Theys, N., and Burrows, J. P.: Satellite observations of long range transport of a large BrO plume in the Arctic, Atmospheric Chemistry and Physics, 10, 6515–6526, https://doi.org/10.5194/acp-10-6515-2010, 2010.

Blechschmidt, A.-M., Richter, A., Burrows, J. P., Kaleschke, L., Strong, K., Theys, N., Weber, M., Zhao, X., and Zien, A.: An exemplary case of a bromine explosion event linked to cyclone development in the Arctic, Atmos. Chem. Phys., 16, 1773–1788, https://doi.org/10.5194/acp-16-1773-2016, 2016.

Bognar, K., Zhao, X., Strong, K., Chang, R. Y.-W., Frieß, U., Hayes, P. L., McClure-Begley, A., Morris, S., Tremblay, S., and Vicente-Luis, A.: Measurements of Tropospheric Bromine Monoxide Over Four Halogen Activation Seasons in the Canadian High Arctic, J. Geophys. Res. Atmos., 125, e2020JD033015, https://doi.org/10.1029/2020jd033015, 2020.

Bottenheim, J. W. and Chan, E.: A trajectory study into the origin of spring time Arctic boundary layer ozone depletion, J. Geophys. Res. Atmos., 111, https://doi.org/10.1029/2006JD007055, 2006.

Bougoudis, I., Blechschmidt, A.-M., Richter, A., Seo, S., Burrows, J. P., Theys, N., and Rinke, A.: Long-term time series of Arctic tropospheric BrO derived from UV–VIS satellite remote sensing and its relation to first-year sea ice, Atmos. Chem. Phys., 20, 11869–11892, https://doi.org/10.5194/acp-20-11869-2020, 2020.

Burd, J. A., Peterson, P. K., Nghiem, S. V., Perovich, D. K., and Simpson, W. R.: Snowmelt onset hinders bromine monoxide heterogeneous recycling in the Arctic, J. Geophys. Res. Atmos., 122, 8297–8309, https://doi.org/10.1002/2017jd026906, 2017.

Choi, S., Wang, Y., Salawitch, R. J., Canty, T., Joiner, J., Zeng, T., Kurosu, T. P., Chance, K., Richter, A., Huey, L. G., Liao, J., Neuman, J. A., Nowak, J. B., Dibb, J. E., Weinheimer, A. J., Diskin, G., Ryerson, T. B., da Silva, A., Curry, J., Kinnison, D., Tilmes, S., and Levelt, P. F.: Analysis of satellite-derived Arctic tropospheric BrO columns in conjunction with aircraft measurements during ARCTAS and ARCPAC, Atmospheric Chemistry and Physics, 12, 1255–1285, https://doi.org/10.5194/acp-12-1255-2012, 2012.

Jeong, D., McNamara, S. M., Barget, A. J., Raso, A. R. W., Upchurch, L. M., Thanekar, S., Quinn, P. K., Simpson, W. R., Fuentes, J. D., Shepson, P. B., and Pratt, K. A.: Multiphase Reactive Bromine Chemistry during Late Spring in the Arctic: Measurements of Gases, Particles, and Snow, ACS Earth Space Chem., 6, 2877–2887, https://doi.org/10.1021/acsearthspacechem.2c00189, 2022.

Jones, A. E., Anderson, P. S., Begoin, M., Brough, N., Hutterli, M. A., Marshall, G. J., Richter, A., Roscoe, H. K., and Wolff, E. W.: BrO, blizzards, and drivers of polar tropospheric ozone depletion events, Atmos. Chem. Phys., 9, 4639–4652, https://doi.org/10.5194/acp-9-4639-2009, 2009.

Molnar, C.: Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, 2nd ed., 2022.

Oltmans, S. J., Johnson, B. J., and Harris, J. M.: Springtime boundary layer ozone depletion at Barrow, Alaska: Meteorological influence, year-to-year variation, and long-term change, J. Geophys. Res. Atmos., 117, https://doi.org/10.1029/2011JD016889, 2012.

Sander, R., Burrows, J., and Kaleschke, L.: Carbonate precipitation in brine – a potential trigger for tropospheric ozone depletion events, Atmospheric Chemistry and Physics, 6, 4653–4658, https://doi.org/10.5194/acp-6-4653-2006, 2006.

Seo, S., Richter, A., Blechschmidt, A.-M., Bougoudis, I., and Burrows, J. P.: Spatial distribution of enhanced BrO and its relation to meteorological parameters in Arctic and Antarctic sea ice regions, Atmos. Chem. Phys., 20, 12285–12312, https://doi.org/10.5194/acp-20-12285-2020, 2020.

Simpson, W. R., von Glasow, R., Riedel, K., Anderson, P., Ariya, P., Bottenheim, J., Burrows, J., Carpenter, L. J., Frieß, U., Goodsite, M. E., Heard, D., Hutterli, M., Jacobi, H. W., Kaleschke, L., Neff, B., Plane, J., Platt, U., Richter, A., Roscoe, H., Sander, R., Shepson, P., Sodeau, J., Steffen, A., Wagner, T., and Wolff, E.: Halogens and their role in polar boundary-layer ozone depletion, Atmos. Chem. Phys., 7, 4375–4418, https://doi.org/10.5194/acp-7-4375-2007, 2007.

U. S. National Ice Center: IMS Daily Northern Hemisphere Snow and Ice Analysis at 1 km, 4 km, and 24 km Resolutions, Version 1, , https://doi.org/10.7265/N52R3PMC, 2008.

Whaley, C. H., Law, K. S., Hjorth, J. L., Skov, H., Arnold, S. R., Langner, J., Pernov, J. B., Bergeron, G., Bourgeois, I., Christensen, J. H., Chien, R.-Y., Deushi, M., Dong, X., Effertz, P., Faluvegi, G., Flanner, M., Fu, J. S., Gauss, M., Huey, G., Im, U., Kivi, R., Marelle, L., Onishi, T., Oshima, N., Petropavlovskikh, I., Peischl, J., Plummer, D. A., Pozzoli, L., Raut, J.-C., Ryerson, T., Skeie, R., Solberg, S., Thomas, M. A., Thompson, C., Tsigaridis, K., Tsyro, S., Turnock, S. T., von Salzen, K., and Tarasick, D. W.: Arctic tropospheric ozone: assessment of current knowledge and model performance, Atmos. Chem. Phys., 23, 637–661, https://doi.org/10.5194/acp-23-637-2023, 2023.

Zhao, X., Strong, K., Adams, C., Schofield, R., Yang, X., Richter, A., Friess, U., Blechschmidt, A.-M., and Koo, J.-H.: A case study of a transported bromine explosion event in the Canadian high arctic, J. Geophys. Res. Atmos., 121, 457–477, https://doi.org/10.1002/2015JD023711, 2016.