Responses to Editors and Reviewers

We sincerely appreciate the reviewers for their constructive and insightful comments, which are of great benefit to improve the quality of the manuscript. In response, we have carefully revised the manuscript and addressed each comment in a point-by-point manner. For clarity, the reviewers' comments are presented in black, our responses in blue, and the added or revised sections of the manuscript are highlighted in red.

RC2: 'Comment on egusphere-2024-1638', Anonymous Referee #3, 03 Jan 2025

In this manuscript, the authors present a study that investigates key factors driving the production of $ClNO_2$ based on field observations and XGBoost-SHAP model. Furthermore, the authors evaluated the potential impact of $ClNO_2$ photolysis on the formation of RO_2 and hence, the atmospheric oxidative capacity.

Overall, I found this manuscript interesting and well-constructed. Although the conclusion drawn for the nighttime $ClNO_2$ formation has been well recognized for two decades, the contribution of NO_3^- photolysis to daytime $ClNO_2$ is confirmed by the authors, which brings sufficient novelty to this manuscript.

Despite this, I do have some comments, particularly on the interpretation of the machine learning results, which need to be fully addressed before this manuscript can be accepted for publication.

Response: Thank you for your valuable and thoughtful comments. Your comments and suggestions have greatly enhanced the overall quality and readability of the manuscript. We have made the necessary revisions and provided detailed responses to each point below for your consideration.

General comments:

1. Machine learning, especially SHAP value, starts to be widely used in atmospheric research very recently, but many readers may not be sufficiently familiar with it. To improve the readability, I believe the way of interpreting SHAP values must be fully informed in the manuscript. E.g., what do the negative and positive SHAP values

stand for? Should the contribution be evaluated by the true value or absolute value.

Response: Thank you for your comment. We have added a detailed introduction to SHAP values in the revised manuscript.

Added/rewritten: "The SHAP model is an interpretability tool designed to analyze the contributions of individual features to model predictions. It employs an additive explanatory framework that considers all features as contributors, drawing inspiration from cooperative game theory. For each predicted instance, SHAP assigns a Shapley value, representing the cumulative contribution of each feature. Positive SHAP values indicate that a feature increases the model's predicted outcome, signifying a positive contribution. Conversely, negative SHAP values suggest that the feature reduces the predicted value, reflecting a negative contribution. The absolute value of the SHAP score reflects the magnitude of the contribution, regardless of direction, offering insight into the overall importance of the feature. The true value, on the other hand, reveals the direction of the contribution (positive or negative), facilitating a clearer understanding of the relationship between the feature and the prediction."

2. I am not fully convinced by the way of performing SHAP model and its interpretation.

1) why does the aerosol surface, as a known important factor for N_2O_5 uptake, not used as an input of SHAP model?

Response: Thank you for your valuable comment. We agree that aerosol surface area is a crucial factor influencing the heterogeneous uptake of N_2O_5. Initially, we had included particle surface area concentrations (S_a) in the XGBoost-SHAP model to assess its significance in $ClNO_2$ formation. However, the results indicated that S_a did not play a prominent role (Figure R1). Furthermore, it is found that R^2 values of the training and testing sets slightly improved from 0.963 and 0.861 to 0.965 and 0.891, respectively, when S_a was not used as an input of a machine learning model. Given that $PM_{2.5}$ and its inorganic compositions serve as representative indicators of aerosol conditions to some extent, we chose not to include aerosol surface area as a dependent variable in the machine learning model to avoid redundancy.

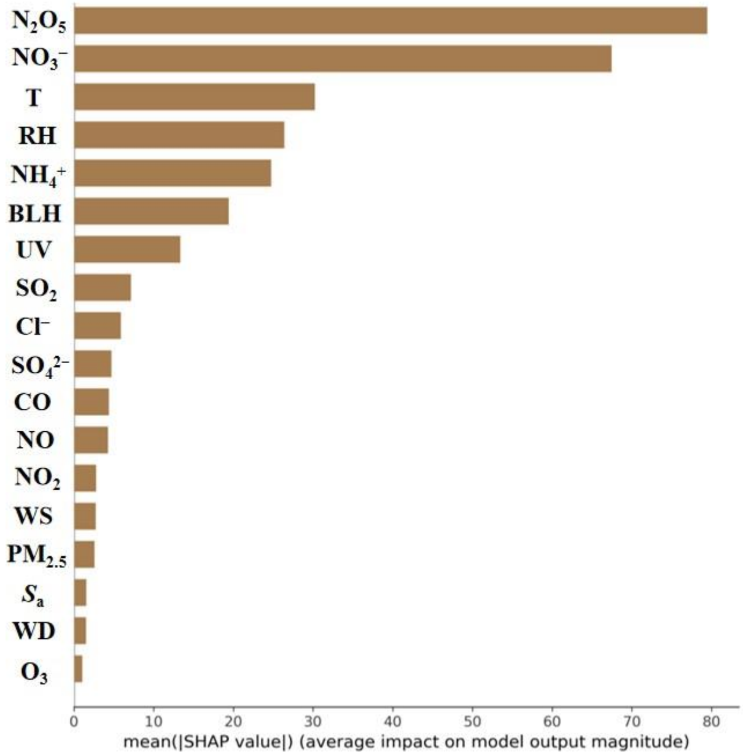

Figure R1. Relative importance of each feature to $ClNO_2$ using the XGBoost-SHAP model during the autumn observation period, with S_a included as an additional variable in the model.

2) $ClNO_2$ has a rather long nighttime lifetime, which means $ClNO_2$ could be accumulated during airmass transport. Meanwhile, N_2O_5 could both form and loss through the transport, leading to varying patterns of its concentration. In fact, this can be testified by calculating the maximal $ClNO_2$ production through N_2O_5 uptake by, e.g., assuming gamma = 0.1 and $ClNO_2$ yield = 1. Given this assumption, I didn't see any model input that could represent the influence of airmass transport. I suggest to reconsider their model input and incorporate certain transport parameters.

Response: Thank you for your thoughtful comment. I fully agree with your opinion that $ClNO_2$ tends to accumulate at night. We had indeed considered the impact of air mass transport in our analysis. In this study, trace gases (SO_2, CO, NO_2, NO, O_3, and N_2O_5), $PM_{2.5}$ and its inorganic compositions (NO_3^-, SO_4^{2-}, NH_4^+, and Cl^-), along with meteorological parameters (T, RH, UV, WS, WD, and BLH) were selected as independent variables. Typically, WS and WD effectively reflect the influence of air masses and play a significant role in the transport, dispersion, and accumulation of atmospheric pollutants. However, results from the XGBoost-SHAP model indicate that WS and WD have a minimal impact on $ClNO_2$ concentrations (Figure R2). Notably, previous observations indicating that $ClNO_2$ is easily influenced by air mass transport were primarily conducted in clean rural areas or under background atmospheric conditions (Niu et al., 2022; Tan et al., 2022). Given that our study site located in a typical urban area surrounded by shopping malls, residential zones, and major traffic arteries, it is highly affected by fresh anthropogenic emissions. Therefore, these results suggest that $ClNO_2$ concentrations are primarily driven by local processes, rather than by air mass transport during our study period.

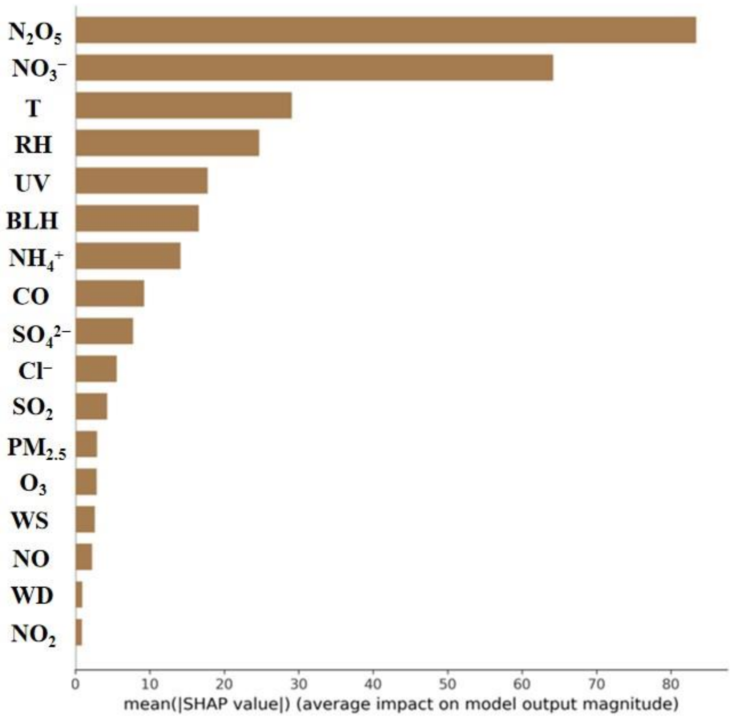

Figure R2. Relative importance of each feature to $ClNO_2$ using the XGBoost-SHAP model during the autumn observation period.

3) As this study suggested, daytime and nighttime $ClNO_2$ are driven by different processes, which however, were affected by similar parameters (in different ways). For instance, NO_3^- is a co-product with $ClNO_2$ at nighttime, but a precursor of

$ClNO_2$ in the daytime. I suggest to consider conducting SHAP models daytime and nighttime data sets separately, so that the exact role of these parameters can be better revealed.

Response: Thanks for your constructive comment. We fully agree with your insightful perspective. Through our in-depth analysis, we found that $ClNO_2$ exhibits distinctly different influence pathways during the daytime and nighttime, with certain parameters potentially playing different roles in these two periods. To investigate this further, we integrated all daytime and nighttime data into a unified machine learning model, resulting in a high-performing model. Using SHAP analysis, we were able to effectively distinguish the roles of key influencing factors between daytime and nighttime.

While the primary formation mechanisms of $ClNO_2$ differ between daytime and nighttime, there is a clear interconnection between daytime and nighttime $ClNO_2$ concentrations. Especially, the elevated nighttime $ClNO_2$ concentrations can significantly affect its concentrations in the early morning. Machine learning models trained exclusively on daytime data show poor performance, with R^2 values for the testing sets dropping below 0.6, thereby constraining further analysis of factor importance. As a result, separating daytime and nighttime data for independent machine learning analyses may risk overlooking the intrinsic linkages between these periods.

We believe that a comprehensive analysis, incorporating both daytime and nighttime data, is crucial for a complete and accurate assessment of $ClNO_2$ production and loss processes. Although we did not segregate the data into daytime and nighttime subsets for machine learning, SHAP analysis enabled us to clearly identify the relative importance of various factors during the daytime and nighttime, providing deeper insights into their respective mechanisms across these two periods.

For example, we used SHAP analysis to evaluate the key influencing factors of daytime $ClNO_2$. The simulated concentrations of $ClNO_2$, based on the XGBoost-SHAP model, were significantly elevated when NO_3^- concentrations were higher than 3.7 $\mu g \cdot m^{-3}$. Consequently, the average daily concentrations of NO_3^- were classified as high ($>$ 3.7 $\mu g \cdot m^{-3}$) and low ($<$ 3.7 $\mu g \cdot m^{-3}$) to further elucidate the impacts of NO_3^- on the formation of $ClNO_2$. Fig. R3 presents the diurnal variations in the relative importance of the most critical influencing factors based on the SHAP values under high and low NO_3^- concentrations. Unexpectedly, daytime NO_3^- was the dominant influencing factors for daytime $ClNO_2$ (Fig. R3a). High concentrations of daytime NO_3^- positively affected the daytime concentrations of $ClNO_2$, independent of N_2O_5 uptake processes. As depicted in Fig. R3a, daytime N_2O_5 did not promote the elevation of daytime $ClNO_2$. Negative SHAP values for N_2O_5 during the daytime indicate that the contribution of N_2O_5 chemistry to daytime $ClNO_2$ levels was limited. Therefore, it is very likely that high concentrations of daytime NO_3^- participated in daytime $ClNO_2$ production.

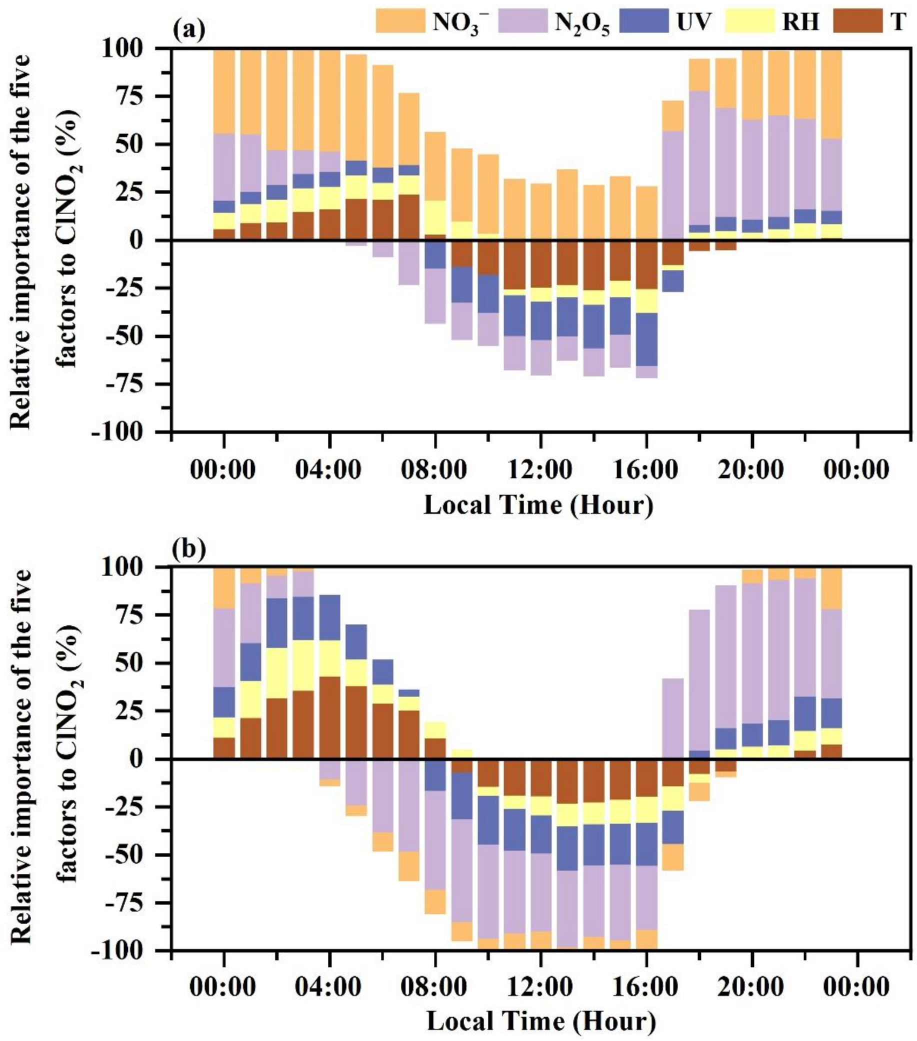

Figure R3. The diurnal variations of the relative importance of factors to $ClNO_2$ based on the SHAP values under the high (> 3.7 $\mu g \cdot m^{-3}$) (a) and low (< 3.7 $\mu g \cdot m^{-3}$) (b) $ClNO_2$ concentrations.

Detailed comments:

Line 64 "were" could be replaced by "are", as this is common case.

Response: Thanks for your comment. We have revised it.

Added/rewritten: "The reaction rates between Cl radical and some alkanes are several orders of magnitude faster than those involving OH radical."

Line 99-100 "our research integrated…." This sentence has grammatic error, please rephrase.

Response: Thanks for your comment. This sentence has been rephrased.

Added/rewritten: "Field observations, combined with a machine learning model, were used to reveal the key driving factors of $ClNO_2$ formation. Furthermore, we further investigated the potential mechanisms driving daytime $ClNO_2$ generation."

Line 141-143. The statement of $JClNO_2$ calculation is not clear, please consider to rephrase.

Response: Thanks for your comment. The statement of $JClNO_2$ calculation has been rephrased.

Added/rewritten: "The Tropospheric Ultraviolet and Visible Radiation (TUV) model was used to calculate $ClNO_2$ photolysis rates ($JClNO_2$) under clear-sky conditions. The simulated $JClNO_2$ values were then scaled based on field-measured JNO_2 values."

Line 167-168 "Simultaneously, …" I think the high correlation between $ClNO_2$ and N_2O_5 (and NO_3^-) does not mean simultaneous peaking. From Fig.1, I can clearly see that their concentrations do not reach the maxima at exactly the same time.

Response: Thanks for your valuable comment. We agree with your opinion that the concentrations of $ClNO_2$, N_2O_5, and NO_3^- did not reach their maxima simultaneously. We intended to convey that their peak concentrations were observed during the night of November 27th. The sentences have been revised accordingly.

Added/rewritten: "The highest concentrations of $ClNO_2$ were detected during the night of November 27th, with a maximum hourly average of 3.4 ppb. Peak concentrations of N_2O_5 and NO_3^- were also observed on that night."

Line 203-204 the authors first indicate NO_3^- could affect the formation of $ClNO_2$; but afterwards, the authors say that the high NO_3^- and $ClNO_2$ together were caused by the simultaneous formation. Please improve the logic of this part.

Response: Thanks for your comment. We have improved the logic of this part.

Added/rewritten: "Differently, the relative importance of NO_3^- derived from the XGBoost-SHAP result indicated that elevated $ClNO_2$ concentrations were associated with high concentrations of NO_3^- besides N_2O_5. According to Fig. 5b, high NO_3^- concentrations (> 3.7 $\mu g\cdot m^{-3}$) are accompanied by the elevation of $ClNO_2$, especially its concentrations reaching 6.2 $\mu g\cdot m^{-3}$. The importance of nighttime NO_3^- for $ClNO_2$ levels is that they are co-products from the processes of N_2O_5 heterogeneous uptake. As shown in Fig. 1, compared to low NO_3^- conditions, $ClNO_2$ production was enhanced in high NO_3^- conditions."

Line 221 "did not promoted…" should be "did not promote".

Response: Thanks for your comment. We have revised it.

Added/rewritten: "As depicted in Fig. 5a, daytime N_2O_5 did not promote the elevation of daytime $ClNO_2$."

Line 222 "A recent study declared that…". Please use "suggested" or "argued" instead of "declared".

Response: Thanks for your comment. We have revised it.

Added/rewritten: "A recent study suggested that nitrate photolysis produced $ClNO_2$ in addition to Cl_2 (Dalton et al., 2023), while it has been not verified by field observations."

Line 236-237. I am not convinced by the discussion about the role of temperature. The authors suggested that N_2O_5 is not important for $ClNO_2$ in the daytime. Then how can temperature affect $ClNO_2$ through the thermal equilibrium of N_2O_5? Also, N_2O_5 is a measured quantity. Such a temperature impact should be already reflected by the connection between daytime N_2O_5 and $ClNO_2$.

Response: Thank you for your comments. We believe that N_2O_5 plays a critical role in the formation of $ClNO_2$, as $ClNO_2$ is generated through the heterogeneous uptake of N_2O_5 on chloride-containing aerosols. In this study, we emphasized that limited contribution of heterogeneous N_2O_5 uptake to daytime $ClNO_2$ concentrations was primarily due to very low daytime N_2O_5 levels, which are largely associated with its thermal decomposition. In other words, the thermal decomposition process affects $ClNO_2$ generation by reducing the availability of N_2O_5 in the daytime. Specifically, the elevated ambient temperature from nighttime to daytime reduced N_2O_5 concentrations through enhanced thermal decomposition. During the entire observation period from October to November, the overall drop in ambient temperature facilitated $ClNO_2$ production by reducing the thermal decomposition of N_2O_5, thereby increasing its availability for heterogeneous uptake.

Added/rewritten: "The impact of ambient temperature on $ClNO_2$ was probably reflected in its thermal equilibrium with N_2O_5. Elevated daytime ambient temperature suppressed the formation of N_2O_5, resulting in low N_2O_5 concentrations, which further limited the contribution of heterogeneous N_2O_5 uptake to daytime $ClNO_2$ generation. During the whole observation period from October to November, the drop in ambient temperature facilitated $ClNO_2$ production by decreasing the thermal decomposition process."

Line 243 I suggest the subtitle of "Impact of $ClNO_2$ photolysis on RO_x budget"

Response: Thanks for your suggestion. We have revised it.

Added/rewritten: "3.3 Impact of $ClNO_2$ photolysis on RO_x budget."

Figure 2: the N_2O_5 in the lowest panel is barely seen. Please consider to show the pattern by perhaps N_2O_5*5.

Response: Thank you for your suggestion. We have revised Figure 2 to update the presentation of N_2O_5 accordingly.

Added/rewritten:

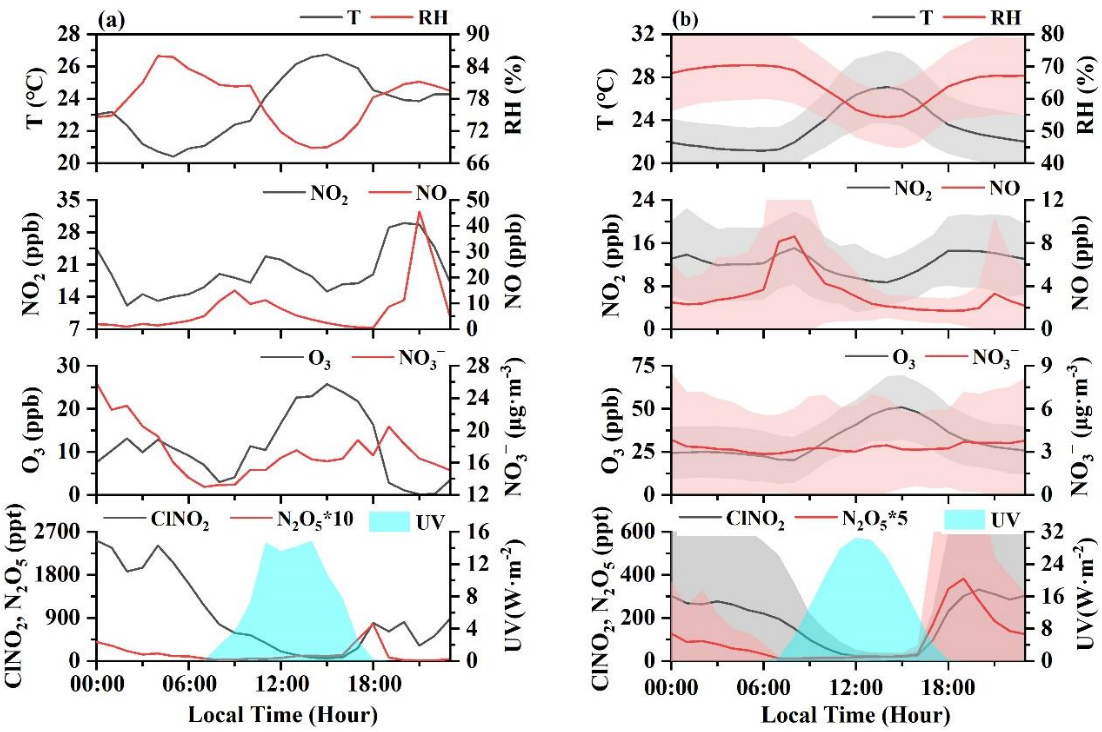

Figure 2. Diurnal variations of $ClNO_2$ and other related parameters for the highest concentrations of $ClNO_2$ (case) on November 28th (a) and the observation-average condition (from 9 October to 5 December) (b).

Figure 4. the division of x ticks looks strange. Please modify.

Response: Thanks for your comment. We have modified Figure 4.

Added/rewritten:

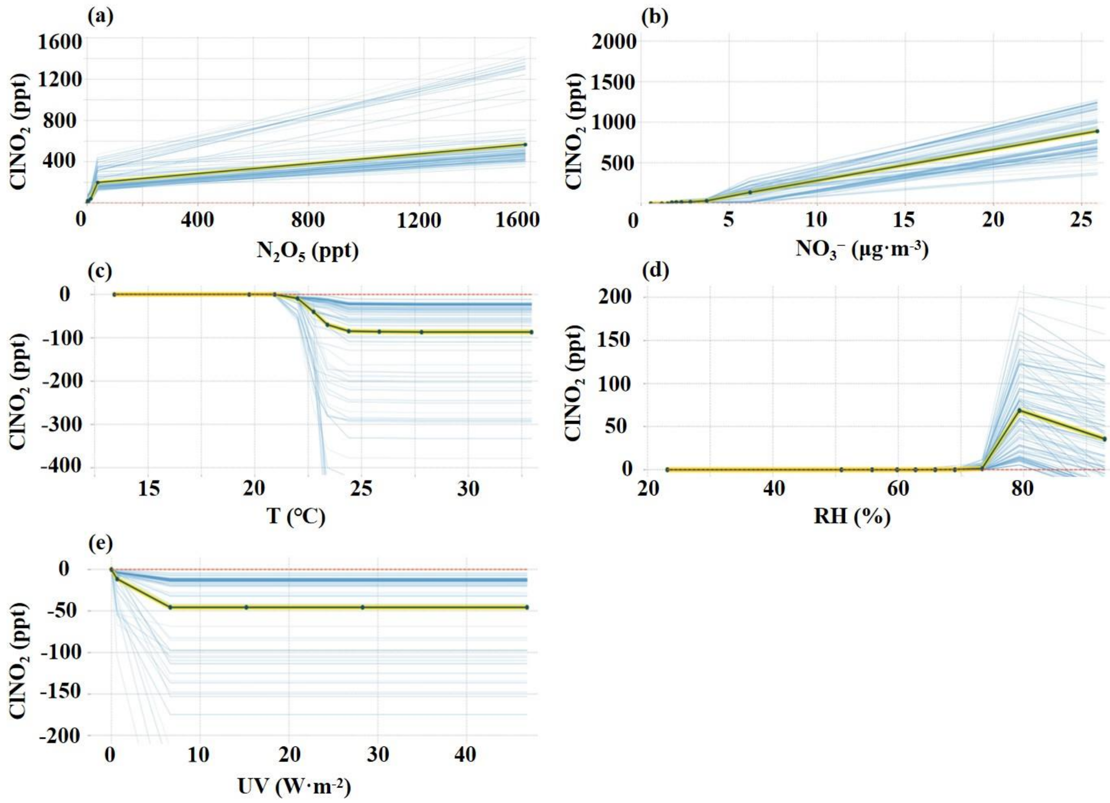

Figure 4. Isolation plots of PDP for N_2O_5 (a), NO_3^- (b), T (c), RH (d), and UV (e). The average variations of simulated $ClNO_2$ with factors' changes spline are indicated by the yellow and black curve, and blue curves presents all situations during the whole observation period.

References

Dalton, E. Z., Hoffmann, E. H., Schaefer, T., Tilgner, A., Herrmann, H., and Raff, J. D.: Daytime Atmospheric Halogen Cycling through Aqueous-Phase Oxygen Atom Chemistry, J. Am. Chem. Soc., 145, 15652-15657, https://doi.org/10.1021/jacs.3c03112, 2023.

Niu, Y.-B., Zhu, B., He, L.-Y., Wang, Z., Lin, X.-Y., Tang, M.-X., and Huang, X.-F.: Fast Nocturnal Heterogeneous Chemistry in a Coastal Background Atmosphere and Its Implications for Daytime Photochemistry, J. Geophys. Res. Atmos., 127, e2022JD036716, https://doi.org/10.1029/2022JD036716, 2022.

Tan, Z., Fuchs, H., Hofzumahaus, A., Bloss, W. J., Bohn, B., Cho, C., Hohaus, T., Holland, F., Lakshmisha, C., Liu, L., Monks, P. S., Novelli, A., Niether, D., Rohrer, F., Tillmann, R., Valkenburg, T. S. E., Vardhan, V., Kiendler-Scharr, A., Wahner, A., and Sommariva, R.: Seasonal variation in nitryl chloride and its relation to gas-phase precursors during the JULIAC campaign in Germany, Atmos. Chem. Phys., 22, 13137-13152, https://doi.org/10.5194/acp-22-13137-2022, 2022.