

EGUSPHERE-2025-5285 – Author comments to RC1

<https://doi.org/10.5194/egusphere-2025-5285>

Legend: Referee comments in *blue and italic*, author comments in **black**, changes in the manuscript text in **green**

Review #1

The paper presents a thorough multi-site analysis of the glyoxal-to-formaldehyde ratio (RGF) derived from ground-based MAX-DOAS measurements across four environments (tropical forest, temperate forest, and two urban sites). The authors attempt to clarify why the RGF values reported in different studies often diverge, emphasizing observational geometry, aerosol and NO_x regimes, and temperature dependence. The topic is timely and valuable for improving the interpretation of satellite and ground-based VOC proxies.

I should note that I am not a specialist in atmospheric chemistry or VOC oxidation modeling, so my comments focus mainly on the scientific reasoning, completeness of the analysis, and clarity of interpretation. Overall, the study is carefully executed and well documented, but it still lacks several essential elements that would make the findings more conclusive. Below I summarize major and minor scientific questions and suggestions.

We would like to thank the Referee #1 for the detailed feedback, helpful comments, and advice on the manuscript. The suggestions improved the manuscript message and extended the analysis in a great way.

Detailed responses to the reviewer's comments can be found below. We hope that we have incorporated all suggestions and comments in a satisfactory way. At the end of this document, you will find a general section on additional changes we made during the review process

- 1. The paper interprets RGF differences as signals of VOC source composition, yet it does not provide a quantitative link from precursor classes to CHOCHO and HCHO yields. Please add a concise mechanism-based discussion that relates RGF to isoprene, aromatics, alkenes, and oxygenates, including the role of NO_x regime and OH reactivity.*

We are working on a detailed box model study to investigate the sensitivities of R_{GF} to VOC source composition, NO_x regime, and OH reactivity.

Further, we have expanded the discussion in the RGF-NO₂ relationship section to include a mechanism-based discussion covering the influence of NO_x on the VOC chemistry and OH concentrations. Furthermore, we added the link VOC precursor classes to CHOCHO and HCHO formation. Specifically, we now relate R_{GF} to differences in product yields from larger VOCs like aromatics, alkenes, smaller VOCs like methoxy. The added text draws on

established chemical mechanisms summarized in Seinfeld and Pandis from 2006 (Lines 557-572).

2. *Differences in photolysis, OH loss, heterogeneous uptake, and wet removal for CHOCHO versus HCHO can alter R_{GF} independently of emissions. Please analyze, at least qualitatively, whether the observed diurnal and seasonal patterns can be explained by lifetime contrasts. A short sensitivity test or a conceptual lifetime budget would help.*

We added a paragraph mentioning the effects that can change R_{GF} independently of emissions in the diurnal cycle chapter (Lines 365-369). Furthermore, we added a conceptual figure in the supplement (Fig. A11) visualizing the effect different photochemical lifetimes can have on R_{GF} . The figure demonstrates that, even when neglecting other processes, differences in the effective lifetimes of CHOCHO and HCHO alone yield systematic variations in R_{GF} . This qualitative analysis supports the interpretation that lifetime contrasts may contribute to the observed diurnal and seasonal variability, although a quantitative attribution would require model-based investigation.

Lines 365-369: *Furthermore, other effects independent from emissions could have an influence, like differences in photolysis, OH loss, heterogeneous uptake, and wet removal, but our dataset does not allow to separate such effects. Resulting different photochemical lifetimes of CHOCHO and HCHO might also contribute to the shape of the diurnal cycle of R^*_{GF} . Under simplified conditions, a longer lifetime of CHOCHO compared to HCHO, results in an increase of R^*_{GF} and a decrease otherwise (see Fig. A11).*

3. *The manuscript reports a robust negative dependence of R_{GF} on temperature but does not present temperature-normalized R_{GF}. Please include a regression-based removal of temperature effects and show the residual R_{GF} variability that might be attributable to emission composition or chemistry. A supplemental figure would suffice.*

We thank the reviewer for this good suggestion. We account for the temperature dependence of R_{GF} using an outlier-robust orthogonal regression approach. Data points with residuals exceeding two standard deviations are down-weighted to reduce the influence of extreme values.

As the results were interesting, we added a subsection (3.2.2) discussing the temperature-normalised R_{GF} . The new figure (Fig. 10) shows that the seasonal variability disappears and the diurnal variability remains. Further, the response of R_{GF} due to other meteorological variables is significantly reduced after removing the temperature effect (new Fig. A2).

4. *Because photochemistry drives both CHOCHO and HCHO, actinic flux and cloudiness matter. Please document any clear-sky filtering and examine whether shortwave radiation or AOD anomalies co-vary with R_{GF}. Summarize this in Methods and expand the Discussion with a brief correlation or stratification analysis.*

Following the comment above, we added a sentence describing the filtering methodology, and clarifying that no explicit clear-sky filtering was applied. Instead, poor viewing conditions are indirectly screened using an RMS threshold (Line 116-119).

To incorporate the additional variables, we generalized the former temperature dependence section to “Link to meteorology” (3.2), where we discuss the response of R_{GF} to dew point temperature, relative humidity, boundary layer height, short wave radiation, and wind speed (Fig. A5 and A6). However, as the variables are strongly intercorrelated the variability drops significantly after accounting for temperature effects (see response to reviewer 1 point 3).

ERA5 data does not provide cloudiness or AOD, therefore, no analysis on aerosols and clouds was performed.

5. *Opposite R_{GF} -NO₂ relationships across urban sites are described but not explained. Please provide a mechanistic discussion of how NO_x modifies RO₂ and HO₂ pathways and shifts product yields, and reconcile the contrasting behaviors at the two city sites with local emission structure.*

We have now expanded the discussion of the R_{GF} -NO₂ relationship section (see response to point 1, reviewer 1).

Our working hypothesis is that differences in industrial sectors contributing to the NMVOC emissions lead to differences in precursor VOC composition, which, in turn, results in site-dependent yields driving the R_{GF} ratio (Lines 573-581).

6. *Quality control thresholds are described, but a quantitative error budget is missing. Please provide random and systematic uncertainties for CHOCHO and HCHO dSCDs, account for correlation between bands, and propagate to R_{GF} . Include an uncertainty table and show how uncertainties vary by season and site.*

We added a new subsection (2.2.1) in the methodology titled “Uncertainties of R_{GF} ” that draws on Pinaridi et al. (2013) to discuss random and systematic uncertainties for HCHO MAX-DOAS observations. Furthermore, we now discuss systematic differences between the sites (different processing at ATTO, higher viewing elevation at Incheon, high instrument altitude at Athens).

In addition, we added an uncertainty propagation for R_{GF} , using the uncertainties from the DOAS fit. The resulting annual and seasonal uncertainties are summarized in an additional table in the supplementary material (Fig. A9).

7. *The O₄ correction assumes similar vertical sensitivity for CHOCHO, HCHO, and O₄. Please discuss the validity range and potential failure modes, for example under lofted layers or*

strong stratification. A short diagnostic test or a reference to AMF differences would strengthen the argument.

We refined the paragraph discussing the O4 correction extending the examples of failure modes with the strong stratification. Further we added a reference (Sinreich et al., 2013) to guide readers to a more elaborated technical discussion of the usage of O4 dSCDs for light path estimations (Lines 133-138).

- 8. The paper notes that average-then-ratio versus ratio-then-average can produce different results but does not quantify the bias. Please add a numerical example using the actual time series to demonstrate the magnitude under realistic variability.*

The bias is illustrated in Fig. 13 for both the diurnal and monthly aggregation. As the bias depends on the distribution of datapoints within each bin, it varies and it cannot be easily quantified in a generalized manner. The strongest bias we observed is around 1%pt for the 10:00 local solar time bin of the diurnal cycle in Orléans. Lines 662-669

We added a figure in the supplement showing a time series of daily means of R_{GF} along with the instantaneous and global R_{GF} for comparison (Fig. A10).

- 9. As RGF is often used in satellite validation, please include a brief comparison to coincident satellite products if available, or at least set expectations by summarizing typical satellite RGF ranges for similar environments. A limited collocation analysis or a literature-based context paragraph would be valuable.*

As we use corrected dSCDs to compute R_{GF}^* , a direct comparison to satellite products would not be straightforward. Instead, we added a paragraph extracting R_{GF} from TROPOMI observations reported by Chen et al. (2023). The ranking of stations based on their annual means is consistent with our findings (excluding Athens as it could not be extracted due to its vicinity to the coastline). The wet-dry seasonal contrast at ATTO is also reproduced in the TROPOMI data.

Lines 452-457: Chen et al. (2023) published global R_{GF} maps based on the TROPOMI observations for the year 2019. Although our R_{GF}^* is derived from dSCDs and therefore does not correspond to the exact same measurement volume (Sect. 3.4.1), a comparison of the magnitude of annual means is still meaningful. Extracting R_{GF} values at our measurement sites from their maps for 2019 suggests the following ranking: Incheon > ATTO > Orléans. Athens could not be identified in their maps due to its vicinity to the coastline. Furthermore, Chen et al. (2023) maps show enhanced R_{GF} values during the wet season compared to the dry season at the ATTO site, which is consistent with our observations.

10. For the tropical site, discuss the role of biomass burning using fire counts or CO as a tracer. Show whether peaks in R_{GF} align with burning indicators. Add a short subsection or figure in Results for seasonality at the tropical site.

We thank the reviewer for this interesting suggestion. We agree it is a worthwhile avenue, but given the noise in our R_{GF} time series and the difficulty in robustly identifying individual events during the dry season, a rigorous analysis would require considerable effort and likely deserves a dedicated study. We therefore consider this beyond the scope of the present manuscript and leave it to future work.

11. R_{GF} here is derived from dSCDs that emphasize near-surface sensitivity. Please clarify how conclusions might change for VCD-based ratios or for in-situ concentrations. If possible, provide an AMF-based conversion for at least a sensitivity check.

We examined how our results change when using VCDs derived from 30° elevation dSCDs (geometric approximation). The number of available data points drops substantially, as the higher elevation angle yields smaller columns and therefore larger relative errors, causing more data to be removed by the quality filters. Relaxing these filters increases the number of valid observations, but increases the scatter in R_{GF} lowering the interpretability.

We added two figures in the supplementary material: The diurnal (Fig. A14) and seasonal cycles (Fig. A13) using the 30° dSCD data. The variability of the CHOCHO columns is strongly reduced, whereas HCHO columns behave similarly to the figures using low viewing elevations (Fig. 5 and 6). As a result, the seasonal cycle (driven primarily by the HCHO variability) is similar, but the diurnal cycle is (driven by CHOCHO variability) becomes very noisy.

To ensure clarity, we have added explanatory text in the methodology to justify our choice of dSCDs (Lines 106–108): *Measurements at 30° viewing elevation, representing a geometric approximation of the vertical column density (VCD), are shown in the supplement (Figs. A13 and A14). However, the limited number of data points remaining after filtering, together with the reduced variability of CHOCHO, renders these data unsuitable for the present analysis.*

12. Please define the representative footprint for each station under the low-elevation MAX-DOAS geometry and discuss how it affects inter-site comparability. A map and a paragraph would help readers interpret differences among sites.

In addition to the existing Fig. 2 that shows the site environment and the viewing orientation, we added a section (2.5.5) discussing coverage and representativeness of the four sites. Further, we ran backward sensitivity simulations using the FLEXPART model to quantify the sites footprint (Fig. 4).

13. Conclusions state that RGF is not universal. Please add two or three sentences on how chemical transport models or box models could assimilate site-specific RGF constraints to evaluate VOC mechanisms.

Box models or chemistry transport models are well-suited for investigating how different drivers influence R_{GF} . Rather than assimilating R_{GF} , it may be more appropriate to use it as a proxy to evaluate whether a model properly captures the VOC chemistry and other influencing factors. Assimilating CHOCHO and HCHO fields directly would likely be more straightforward than assimilating their ratio.

EGUSPHERE-2025-5285 – Author comments to RC2

<https://doi.org/10.5194/egusphere-2025-5285>

Legend: Referee comments in *blue and italic*, author comments in **black**, changes in the manuscript text in **green**

Review #2

This study investigated the diurnal, seasonal, and weekly characteristics of glyoxal-to-formaldehyde ratio RGF over four ground-based MAX-DOAS measurement sites, with additional investigations about the dependence of elevation, temperature, and NO₂ concentrations. The study also discussed the effects of different measurement techniques in the literature of ground-based versus column-based, vertical sensitivity, temporal sampling biases, and calculation order for RGF on the differences across various reports. The analyses are comprehensive and is helpful for the generalization on the usage of RGF on the understanding of VOC emission characteristics.

We would like to thank Referee #2 for the insightful comments and questions that helped to improve our manuscript.

Detailed responses to the reviewer's comments can be found below. We hope that we have incorporated all suggestions and comments in a satisfactory way. At the end of this document, you will find a general section on additional changes we made during the review process

General comments

- 1. The study argued that there are various calculation methods adopted in the literature. When comparing the results in this study against prior studies, the differences are dominated by the calculation methods or the real variation of RGF temporally and spatially?*

Quantifying these effects remains challenging. Differences in the order of averaging and ratioing are expected to introduce a systematic offset, while variations in the sampled air volume imply averaging over different spatial and temporal scales, leading to differing degrees of smoothing. However, the observed discrepancies across studies are more complex than either of these factors alone can explain. In particular, the contrasting reported R_{GF} values between biogenic and anthropogenic environments, with systematically lower values reported for biogenic sources in some studies and higher values in others, cannot be attributed to offsets or smoothing effects. We therefore conclude that, while methodological differences likely contribute to inter-study variability, real spatial and temporal variation in R_{GF} should still be the dominant factor.

- 2. How spatially representative are these ground-based measurement sites?*

We added an extra section “coverage and representativeness” (2.5.5) where we discuss the spatial representativeness (see reviewer 1 point 12)

Line 305-309: The horizontal orientation of the light paths introduces an additional spatial averaging that is inherent to MAX-DOAS measurements and is illustrated in Fig 2. The retrieved dSCDs represent the concentration along the effective light path, whose length within the boundary layer depends on atmospheric visibility. Under clear conditions, photons scattered at distances of up to approximately 15 km from the instrument can contribute to the signal (Seyler et al., 2017).

3. *As the reference was taken at the high elevation angles, then naturally the analyses about the elevation effects would be close to 1 with high elevation angles. Would the uncertainties increase with elevation angles as well?*

We agree that the elevation effects should be close to 1 for high elevation angles. However, we believe that the higher inter quartile range (IQR) of R_{GF} at higher viewing elevation is due to the lower signals of both trace gases, which lead to a high variability at higher elevation, i.e. dividing small values by small values.

Having the length of the manuscript in mind (reviewer 4) and given the complexity of interpreting the elevation effects, with and without the applied O_4 correction, we decided to remove the elevation section.

4. *Why were the slant column sensitivities chosen instead of vertical column sensitivities?*

See our response to reviewer 1 point 11.

Specific comments

1. *Line 474-475: what would be the emission ratio of NO_x and VOCs over these two measurement sites? Would different emitted VOCs from different sector contribute to the different values of RGF?*

Based on this important point raised by the reviewer, and to provide the reader with a clearer context for the sites, we computed the NO_x to NMVOC ratio and the anthropogenic-to-biogenic NMVOC ratio (AVOC/BVOC) using the CAMS - anthropogenic (CAMS-GLOB-ANT) and biogenic emissions (CAMS-GLOB-BIO) over the area indicated in the supplement (Fig. A3). For NMVOC we used the total (anthropogenic + biogenic) amount. The NO_x/VOCs ratio is 0%, 26%, 50%, 66% for ATTO, Orléans, Athens, and Incheon, respectively, and the AVOCs/BVOCs ratio is 0%, 174%, 2500%, 5800% for ATTO, Orléans, Athens, and Incheon, respectively. These ratios support our site characterization. The high AVOC/BVOC ratio in Athens and Incheon reflects their strong anthropogenic influence, while ATTO is dominated by biogenic emissions. They additionally highlight the intermediate (transitional) nature of Orléans. The ratios have been added to the new Table 1.

Coming to the second part of the question: This is indeed our hypothesis to explain the different R_{GF} behaviour between the two anthropogenic sites. Assuming a consistent VOC composition per sector, regardless of the location, the higher CHOCHO emissions in Athens imply that refineries and fugitive emissions would produce more CHOCHO relative to HCHO than industrial emissions. We revised the relevant paragraphs to improve our argumentation.

Lines 573-579: Checking the top five sectors contributing to the total non-methane VOCs emissions in both cities from CAMS-GLOB-ANT (Fig. 12) shows similar contributions from solvents and road transport but differences in other sectors. Industrial processes dominate in Incheon, whereas refineries and fugitive emissions are more prominent in Athens. Assuming a consistent VOC composition per sector, regardless of the location, the higher CHOCHO emissions in Athens imply that emissions from refineries and fugitive emissions would produce more CHOCHO relative to HCHO than industrial emissions. Possible species with high CHOCHO yields include aromatics (Chan Miller et al., 2016) or acetylene/ethylene (Fu et al., 2008). However, for aromatics, Nishino et al. (2010) found that CHOCHO yield decreases with increasing NO_2 levels.

2. *Line 510-512: how would the RGF values compare between MAX-DOAS measurement sites and that coincidentally sampled from TROPOMI observations?*

See our response to reviewer 1 point 9.

EGUSPHERE-2025-5285 – Author comments to RC3

<https://doi.org/10.5194/egusphere-2025-5285>

Legend: Referee comments in *blue and italic*, author comments in **black**, changes in the manuscript text in **green**

Review #3

The article presents an observational analysis of CHOCHO, HCHO, and their ratio (RGF) across contrasting environments using MAX-DOAS measurements. The study investigates temporal variability and relationships with meteorological parameters and NO₂, aiming to improve understanding of the applicability of RGF as an indicator for VOC source attribution. It also discusses the factors contributing to discrepancies among previous studies using different methodological approaches, thereby highlighting both the potential and limitations of employing RGF as a proxy for VOC source identification. The manuscript addresses a timely, underexplored topic and is generally well-structured and informative. Before publication, several issues should be clarified or strengthened as noted below.

We thank the Referee #3 for the constructive feedback to the manuscript. The raised issues helped us to further improve the manuscript.

Detailed responses to the reviewer's comments can be found below. We hope that we have incorporated all suggestions and comments in a satisfactory way. At the end of this document, you will find a general section on additional changes we made during the review process

Specific comments

- 1. It would be recommended that the authors further elaborate on how the substantial differences in instrument placement (10-500 m above ground) and non-overlapping measurement periods (2017-2024, varying by site) may influence the retrieved diurnal, weekly, and seasonal patterns. MAX-DOAS sensitivity to boundary-layer gases depends strongly on viewing geometry and instrument height, and differences in measurement years imply variations in meteorology and emissions among the sites.*

We added two new sections: “Uncertainties of R_{GF} ” (2.2.1) and “Coverage and representativeness” (2.5.5). The instrument placement is discussed as part of the systematic uncertainties. Lines 170-173: **At Athens, the instrument is located at 500 m above sea level, while the city centre lies near sea level. Under shallow boundary layer conditions, such as in winter, the effective light path may therefore only partially sample the polluted boundary layer, resulting in lower measured columns. However, since CHOCHO and HCHO concentrations peak in summer, when the boundary layer is typically well developed, this effect is expected to be small.**

The temporal patterns are discussed in the coverage and representativeness section. Lines 301-304: **The datasets were collected during non-overlapping periods, as the station**

locations originate from long-term measurement activities. While this limits strict temporal comparability, the analysis focuses on characteristic relationships rather than direct year-to-year contrasts.

2. *MAX-DOAS data were aggregated into 30-minute bins, whereas accompanying reanalysis meteorology is hourly and subsequently interpolated. Please justify this choice. Using hourly averages for MAX-DOAS would reduce noise and ensure temporal consistency with the meteorology without interpolation.*

We thank the reviewer for the comment. The aggregation into 30-minute bins was a remaining artefact, used for merging for datasets in extended analyses that were ultimately not included in the final manuscript. Based on the discussion about averaging-ratioing order in section 3.4.4, we decided to remove the pre-binning completely. We now interpolate the instantaneous ERA5 data to the exact timestamps of the MAX-DOAS measurements (<30 minutes) (Line 190).

3. *The diurnal cycles results show higher RGF-related values in ATTO than in Orléans, which seems counterintuitive given the expected dominance of HCHO in the Amazon. Figure 4 shows unscaled ATTO values as the lowest. Could the scaling applied introduce a bias in the RGF representation at ATTO?*

The O₄ scaling is applied to correct for reduced quality in the O₄ vis retrieval in ATTO. As the detector is missing the spectral coverage of the absorption band of O₄ at 477 nm in the visible, we believe that the unrealistically high difference between the O₄ dSCDs in visible and UV (Fig. 1a) originates from the lower retrieval quality in the visible. Since the O₄ ratio has a large impact on the average R_{GF}, we decided to scale the O₄ vis dSCD to align it with the realistic behaviour of other stations. While the scaling is not perfect as it depends on the determined slopes (Figs. 1a to 1c), it prevents the misinterpretation of a systematic retrieval artefact as a real signal. As a result, the average R_{GF} value in ATTO has a larger systematic uncertainty, but the interpretation is more robust with the applied scaling.

4. *In Figure 11, it is difficult to visually confirm a systematic increase in RGF with viewing elevation. Since O4-based light-path corrections have reduced influence at higher elevation angles, applying a uniform correction across sites with substantially different instrument heights (e.g., Athens at ~500 m a.g.l.) may introduce biases. Excluding the highest elevation angles appears to yield more consistent trends, and the median pattern remains relatively stable with only a slight U-shaped variation. Please clarify how these factors may influence the interpretation of elevation-dependent RGF behavior.*

See our response to reviewer 2 general point 3. The instrument altitude of Athens is discussed in response to reviewer 3 point 1.

5. *Including relative humidity datasets in the analysis or discussion may be valuable, as it strongly influences glyoxal uptake and therefore RGF variability in different environments.*

See our response to reviewer 1 point 4.

6. *The "seasonal" analysis currently reflects monthly averages. Because hemispheric seasons differ and the sites span both hemispheres, grouping by season or consistently referring to "monthly variability" would improve clarity.*

We thank the reviewer for raising this point. We have chosen to keep the term "seasonal cycle" as it is a standard label in atmospheric sciences describing the recurring annual pattern of variability, independent of hemisphere-specific seasons. However, we made sure to indicate the different seasons for ATTO compared to the other stations (Fig. 4, 7, A2).

7. *Differences in the correlation between RGF and NO₂ at the urban sites may reflect differing emission trends, such as the recent decline in NO_x levels in Seoul (Incheon). A brief note on this possibility could enhance interpretation.*

We agree with this comment and have added a paragraph mentioning that possibility.

Lines 579-581: *There is also the possibility, that the declining NO₂ levels in Incheon (Seo et al., 2021) during the measurement period lead to a more stable R_{GF}—NO₂ relationship. But as the Incheon dataset only covers one year, the effect should be minimal.*

8. *Line 113 indicates 1–3° viewing angles, but figures include data beyond 30°. Please clarify the full range of elevation angles used in the analysis.*

Based on reviewer 2 point 3, we decided to remove the elevation dependency. All other sections use only low elevation data, except new figures for reviewer 1 point 11. These figures use dSCDs from 30° elevation angle, but are only included in the supplement.

Line 102: *In this study, we use off-axis measurements at low elevation angles from 1–3°.*

9. *Line 213: It would be helpful to include the latitude and longitude of each measurement site to clearly define their locations.*

We have added a new table (Table 1) that contains an overview over stations (including latitude and longitude) and combines multiple reviewer comments (reviewer 2 specific point 1, reviewer 4 minor point 2).

10. *Line 348: While temperature variability in ATTO is limited, a slight increase during September-October coincides with peak HCHO values; this may be worth noting.*

We have added this sentence to the relevant paragraph. Lines 423-424: *While temperature variability in ATTO is limited, a slight increase during September—October coincides with peak HCHO values.*

11. *Line 496: Please clarify the spatial representativeness of the MAX-DOAS measurements. Over what approximate horizontal extension can these observations be considered representative, and how does this compare to a point measurement assumption?*

See our response to reviewer 2 general point 2.

12. *The description of how temperature dependence was evaluated is unclear. Were data grouped into temperature bins? How does this analysis add beyond what is shown in Fig. 7 where temperature influences are already visible?*

That is correct, R_{GF} was binned to temperature intervals and then the statistics per bin have been visualized in Fig. 9. As R_{GF} showed a strong dependence on temperature, we wanted to discuss it separately. Based on reviewer 1 point 3, we extended the discussion by investigating the temperature-normalised R_{GF} (3.2.2).

13. *Figure 9 indicates a possible weekend effect in RGF. It would be helpful to assess whether the differences between weekdays and weekends are statistically significant.*

Although this relative difference of about 10% is comparable to our calculated uncertainties (2.2.1), the uncertainties are expected to affect all days uniformly and should therefore not systematically bias the reduction during the weekend.

We added section 2.4 describing the statistical approach and Table A8 in the supplementary material to show the detailed results of the test. A paragraph discussing the significance was added. Lines 472-476: For both stations, the mean weekday R_{GF}^* exceeds the mean weekend value by 0.5 %pt., corresponding to a reduction of approximately 10 % on weekends. Although this relative difference is comparable to our systematic uncertainties, these uncertainties are expected to affect all days uniformly and should therefore not be relevant for the weekend effect. As described in Sect. 2.4, the differences between weekday and weekend are significant for Athens ($t = 4.4$, $p = 2 \cdot 10^{-5}$) and Incheon ($t = 2.7$, $p = 8 \cdot 10^{-3}$).

EGUSPHERE-2025-5285 – Author comments to RC4

<https://doi.org/10.5194/egusphere-2025-5285>

Legend: Referee comments in *blue and italic*, author comments in **black**, changes in the manuscript text in **green**

Review #4

Bittner et al. present a valuable multi-year analysis of RGF using ground-based MAX-DOAS measurements at four sites: two biogenic and two anthropogenic. The author found that the patterns of the glyoxal-to-formaldehyde ratio (RGF) varied across the four study sites, including diurnal, weekly, and seasonal cycles. They also investigate how the temperature and NO₂ drive the changes in RGF, and provide a valuable discussion of the four factors that reduce the comparability between different RGF values.

When the language is well used and data visualizations are very good, I feel the manuscript is very long, so I find it challenging to extract the take-home messages. Efforts should be made to improve the readability of the manuscript by combining certain sections and figures. More discussions should be included for other factors that can affect RGF, e.g., boundary layer height and air mass history. Machine learning could be a useful approach for this discussion. Overall, major revisions are required before the work can be published.

We would like to thank the Referee #4 for the valuable feedback that improved the manuscript significantly. We decided to adapt the proposed combined structure of the paper to make it easier to grasp. The results are now structured in (1) Temporal cycles (2) Link to meteorology (3) R_{GF}-NO₂ relationship (4) Comparability between R_{GF}. We further decided to remove the elevation section to make space for new content, as the elevation dependency was not easy to interpret (with and without O₄ correction) and R_{GF} did not show a clear signal. Furthermore, we computed the air mass history for each site (Fig. 4). Using backwards sensitivity runs with FLEXPART driven by ERA5 allows to visualize the footprints/sensitivities for all stations.

Detailed responses to the reviewer's comments can be found below. We hope that we have incorporated all suggestions and comments in a satisfactory way. At the end of this document, you will find a general section on additional changes we made during the review process

Major comments

- 1. Section 3.2 and Section 3.5: The seasonal patterns are mainly attributed to temperature-driven HCHO changes (Figure 7), but other factors like boundary layer height, cloud coverage, and photolysis rates also need to be discussed.*

See our response to reviewer 1 point 4.

2. *Section 3.6: The authors provide valuable empirical analysis of RGF-NO₂ relationships across diverse environments. But there is a lack of mechanistic interpretation of how varying NO_x regimes fundamentally alter CHOCHO and HCHO yields from different VOC precursors. In addition, it is unclear how the industrial processes in Incheon and Athens drive the observed difference in the CHOCHO-NO₂*

See our response to reviewer 1 point 1 for mechanistic interpretation and our response to reviewer 2 specific point 1 for an explanation regarding the emission sectors.

Minor comments

1. *Lines 58-86: The summary of the latest research on RGF is descriptive. It lacks clear motivation. I recommend that the authors make it more concise and compact for better readability.*

We have rewritten the literature review to be more conceptual and concise, focusing on the scientific motivation rather than a descriptive list of studies. Lines 057-082: The ratio of glyoxal-to-formaldehyde (R_{GF}) was proposed by Wittrock et al. (2006) and Vrekoussis et al. (2010) as a potential proxy for differentiating VOC source types. Because CHOCHO and HCHO have similar sources and loss processes, subtle differences in VOC mixtures or source-specific yields are expected to be reflected in R_{GF} . The interpretation of R_{GF} as a diagnostic for VOC sources has remained inconsistent since its introduction. Vrekoussis et al. (2010) analysed two years of GOME-2 satellite data and found a strong spatial correlation between R_{GF} and VOC source categories, proposing a threshold of 4 % to distinguish anthropogenic sources (below) from biogenic or pyrogenic origins (above). They further observed decreasing R_{GF} with higher NO₂ levels and increasing values with greater vegetation density, quantified by the Enhanced Vegetation Index (EVI).

Subsequent studies, however, produced mixed and sometimes contradictory results (Irie et al., 2011; DiGangi et al., 2012; MacDonald et al., 2012; Li et al., 2014; Chan Miller et al., 2014). Based on airborne in-situ data, Kaiser et al. (2015) shifted the focus toward VOC precursor speciation, finding that monoterpenes yield high R_{GF} values while isoprene yields low values. DiGangi et al. (2012) went further, proposing an interpretation opposite to that of Vrekoussis et al. (2010), with lower R_{GF} associated with biogenic sources and higher values with anthropogenic or pyrogenic origins. More recently, Chen et al. (2023) reported a positive correlation of R_{GF} with both EVI and NO₂ using TROPOMI data, and proposed that anthropogenic VOC emissions can be identified where $R_{GF} > 4\%$ with additional constraints on EVI and HCHO columns. Hong et al. (2024) further argued that primary HCHO emissions bias R_{GF} , and proposed the ratio of CHOCHO to secondary HCHO as a more reliable metric.

Further complexity was added by MAX-DOAS observations at rural and semi-urban sites in Southeast Asia. Hoque et al. (2018a, b) and Rawat et al. (2024) revealed pronounced seasonal and diurnal variability, while Xing et al. (2020) reported altitude-dependent changes in the diurnal cycle using vertical profile retrievals in China. Together, these studies found various influencing factors that contribute to the inconsistent results and highlight that the interpretation of R_{GF} remains challenging.

This study aims to systematically investigate the drivers and limitations of R_{GF} with the help of a multi-year, multi-site ground-based data set. MAX-DOAS observations from four sites in contrasting environments are analysed to investigate the overall magnitude of R_{GF} , temporal cycles (Sect. 3.1), link to meteorology (Sect. 3.2), and the R_{GF} -NO₂ relationship (Sect. 3.4). In addition, we identify and discuss four measurement-related effects in Sect. 3.4 that can hinder cross-study comparisons, with the aim of reassessing the suitability of R_{GF} as a proxy for VOC origin.

2. *Please include a table in Section 2.5 to summarize key characteristics: coordinates, type, instrument height, viewing direction, measurement period, and typical NO₂ level for each site. This would help readers quickly grasp information about the four study sites.*

In response to this excellent suggestion, we have added a new table (Table 1) that provides info on the four study sites.

3. *Figure 3: Please include the median and interquartile range for better statistics.*

Figure 3 was reworked to better match the observations. (1) The median and IQR were added. (2) Values during the night were excluded to better resemble the conditions during the measurements. (3) Precipitation is now summed per month/hour before computing the median, correcting a previous pre-processing error.

4. *Figure 4: The difference in mark sizes for data availability is not clear. The same applies to Figures 5-7 and other similar figures.*

The marker size shows the relative data availability in each panel. The bin with the maximum number of observations is defined as 100% per panel. If there are fewer data points in one bin (e.g. morning hours) the marker is smaller (line plot) or the alpha is lower (box plot).

To improve interpretability, the previous linear mapping of marker size to observation count has been replaced with a saturating function: marker size decreases rapidly below 30 % of the maximum number of observations, and more slowly above this threshold. The sentences introducing this convention have been revised accordingly. The updated text now reads (Lines 343--344): *Throughout this study, marker size is scaled to the number of observations per bin; smaller markers therefore indicate reduced bin size. The detailed mapping of bin sizes is provided in Fig. A8.*

5. *Figure 4: What are the uses of overpass times? There is no discussion in section 3.2.*

The overpass times are discussed in a later section (3.4.3). To avoid an extra figure, they are already included in Fig. 5 (previously Fig. 4).

6. *Lines 266-275: It is good that the authors mentioned findings in light of the literature. However, I feel that the comparison in R_{GF} between this study and the literature is lacking. The same applies to the Lines 324 -329 in Section 3.2 Seasonal cycle.*

We have reworked and expanded the literature comparisons in both sections.

Lines 375–383: A direct quantitative comparison with previous studies is complicated by methodological differences: whereas DiGangi et al. (2012) report in-situ point measurements and Hoque et al. (2018a, b) and Rawat et al. (2024) derive R_{GF} from VCDs, our R_{GF}^* is based on corrected dSCDs, which integrate over a slant light path and are therefore sensitive to a different effective measurement volume (Sect. 3.4.1). Despite this, the qualitative diurnal patterns are broadly consistent. The midday peak observed at Incheon is also reported for rural and semi-urban sites in Southeast Asia (Hoque et al., 2018a, b; Rawat et al., 2024). However, the occurrence of similar patterns across differently classified sites highlights a broader challenge in the literature: the lack of a uniform site categorisation complicates cross-study comparisons of R_{GF} . At our predominantly biogenic sites ATTO and Orléans, the diurnal cycle is comparatively flat, which is consistent with the weak diurnal variability reported by DiGangi et al. (2012) for high-altitude biogenic sites.

Lines 445–451: As with the diurnal cycle, a direct quantitative comparison is complicated by the fact that previous studies derive R_{GF} from VCDs, whereas our R_{GF}^* is based on corrected dSCDs at the lowest elevation angles, which correspond to a different effective measurement volume (Sect. 3.4.1). With this caveat in mind, the seasonal pattern at our anthropogenically influenced stations resembles most closely the winter enhancement reported by Xing et al. (2025) for Guangzhou. Our absolute R_{GF}^* values are lower than those reported by Xing et al. (2025), which may partly reflect the difference in measurement volume (dSCD vs VCD) rather than a true difference in R_{GF} . At our more remote stations, the magnitude of R_{GF}^* is comparable to that reported by Hoque et al. (2018b) for Pantnagar, even though no progressive annual increase is observed like at Phimai.

7. *Lines 280 - 282: The manuscript reports substantial differences in R_{GF} values between biogenic (2.2-3.1%pt) and anthropogenic sites (3.5-4.2%pt) (Line 280-282). However, no statistical test is performed to validate these differences. Given the considerable overlap in standard deviations (e.g., Athens 3.5±0.4%pt and Incheon 3.7±0.7%pt), statistical tests (e.g., ANOVA with post-hoc testing or non-parametric equivalents) should be applied to confirm whether the observed differences between sites are statistically significant beyond random variations. This is particularly important for supporting the central conclusion that*

RGF values systematically differ between environment types (Line 630). The same concerns on annual mean values of RGF (Line 331).

We decided to focus our testing on the annual values of RGF and therefore removed the mentioning of the daily averages from the diurnal cycle chapter. The new section 2.4 describes the applied methodology. Lines 210-220: To assess whether observed differences in mean values are caused by random variability, we apply statistical tests in Sect. 3.1.2 and 3.1.3. Since measurements are available approximately every 30 minutes, consecutive data points may sample the same atmospheric event. To increase statistical independence, the data are temporally aggregated prior to testing. Where appropriate, a logarithmic transformation is applied to approximate normality.

To compare biogenic and anthropogenic environments, represented by ATTO + Orléans and Athens + Incheon, the data are aggregated to monthly means (e.g., two years of data yield 24 values). Differences between groups are tested using Welch's t-test (Welch, 1947; Delacre et al., 2017) applied to the log-transformed data, which accounts for unequal variances and sample sizes. The same aggregation strategy is used for station-to-station comparisons. In this case, a Welch analysis of variance (ANOVA) (Welch, 1951; Delacre et al., 2019) is applied first to the log-transformed data to assess overall differences among stations. It is followed by a Games—Howell post hoc test (Games and Howell, 1976), which evaluates pairwise differences while accounting for unequal variances and sample sizes.

When discussing the average values between environments, the text now reads (Lines 409-416): Aggregating all data points by month and grouping them by dominant environment, i.e. Orléans and ATTO as biogenic and Athens and Incheon as anthropogenic, yields mean R_{GF}^* values of 3.2 ± 1.1 % in the biogenic environment and 4.2 ± 0.8 % in the anthropogenic environment. Looking at mean R_{GF}^* per station leads to 3.4 ± 0.9 %, 2.7 ± 1.3 %, 3.9 ± 0.8 %, 4.6 ± 0.7 % for ATTO, Orléans, Athens, and Incheon respectively. Applying statistical tests, as described in Sect. 2.4, leads to significant differences ($t = -5.8$, $p = 8 \cdot 10^{-8}$) between the biogenic and anthropogenic group. A Welch-ANOVA ($F = 19$, $p = 3 \cdot 10^{-8}$) combined with a Games-Howell post-hoc test resulted in significant differences for all station pairs except ATTO—Orléans and Athens—Incheon. More detailed results can be found in Table A6 + A7. It should be noted, that the aggregated data points maintain a significant autocorrelation due to the seasonal cycle.

8. *Figure 5 and Lines 296-297: It is not clear what information the authors want to convey with the O₄*

We use the O₄ ratio to correct for physical differences in the retrieval. Since R_{GF}^* depends on the O₄ ratio (Eq. 3), we include it in the plots to show its contribution.

9. *Figures 4 and 5 can be potentially combined into one plot. The same applies to Figures 6 and 7 and also other figures.*

We have combined the respective figures to one for every section.

10. *Lines 301-302: It is unclear how the dilution effects contribute to the observed shape.*

The original idea was that high HCHO and CHOCHO levels in the morning/evening and low levels over noon might be partly explained by sampling the planetary boundary layer (PBL) differently due to PBL growth (scattering point inside/above PBL). But we are not able to conceptualize it without multiple assumptions and geometric simplifications. Therefore, we decided to remove this statement to avoid speculation.

11. *Section 3.7: The discussion on the factors affecting RGF values between different measurement techniques is highly appreciated. It would also be good to include suggestions to guide future work, e.g., "Future work should quantify these effects through radiative transfer simulations or direct instrument intercomparisons."*

We have added this sentence to the conclusion paragraph discussing future work. Line 711: *Future work should quantify these effects through radiative transfer simulations or direct instrument intercomparisons.*

Technical comments

1. *Abstract: When mentioning HCHO and CHOCHO, it will be good to include their names.*

Replaced their structure formulas with their full names. Lines 7,8,11

2. *The use of NMVOC is unnecessary, as it is used only four times. Instead, using non-methane VOCs is fine.*

Replaced NMVOC with non-methane VOCs.

3. *Line 90: When "with meteorological variables" is mentioned, only temperature is analyzed throughout this study.*

See our response to reviewer 1 point 4.

4. *Line 275: "to" is repeated.*

Removed extra to.

5. *Do "%" and "%pt." mean the same thing?*

No, as R_{GF} is given in % it complicates talking about absolute and relative differences of R_{GF} . Relative changes use % and absolute changes use %pt. For example (Line 172): *For*

both stations, the mean weekday R_{GF} exceeds the mean weekend value by 0.5 %pt., corresponding to a reduction of approximately 10 % on weekends.

We have added an extra sentence to clarify this point (Line 146): Changes in R_{GF}^* are expressed in % for relative changes and in %pt. for absolute changes.

Other changes made to the manuscript during the review process:

- Every figure changed slightly except the map.
 - Removing the prebinning changed the numbers marginally (reviewer 3 specific point 2).
 - Changed to more outlier robust linear fit for Fig. 1, changed the O₄ scaling factor for O₄ vis dSCD in ATTO from 0.502 to 0.519.
- Added acknowledgements
 - Added missing persons and teams for supporting the stations
- Corrected wrong end year for Orleans: Figures already included two years of data (Jul 2023 until Jul 2025), whereas the text only stated until Jul 2024.
- Inclusion of backward sensitivity runs to get a better understanding of air mass history/station footprint with FLEXPART for reviewer comments (reviewer 4, reviewer 3 specific point 11, reviewer 2 general point 2, reviewer 1 point 12). Alexandros P. Poulidis who conducted the runs is now included in the co-authors.
- Restructured some sections for clarity. Removed introductory paragraphs and descriptive sentences to reduce bloat.
- Added figure showing diurnal cycles for different seasons (Fig A1).
- Removed ATTO O₄ scaling table and the other scatter plots used to determine ATTO O₄ vis scaling factor from the Appendix.
- Move most additional figures to the supplement material. Only two figures (Fig A1 & Fig A2) should stay in the appendix.