Manuscript: "Reactivity Over Abundance: Unveiling the True Kinetic Drivers of Urban Ozone Using Process-Informed Machine Learning" (egusphere-2026-2647)

General comments

This manuscript introduces a process-informed machine learning framework that uses net chemical consumption (ΔVOC), rather than ambient concentration, as the predictive feature for ozone formation sensitivity, and applies it to a one-month roadside dataset (PTR-ToF-MS and trace gases) collected in Seoul in April 2022. The central premise — that concentration-based ML attribution suffers from "survivor bias" in NOx-saturated, fresh-emission regimes and thereby undervalues highly reactive precursors — is well motivated and convincingly demonstrated. The authors substantiate the ML attribution through cross-validation against an MCM-based observation-based model (radical budgets), kinetic reactivity metrics (OFP, OHR), and a photochemical-clock analysis distinguishing chemical oxidation from physical dilution.

The recovery of the *k*OH dependence of ozone sensitivity from observations alone, together with mechanistically interpretable outliers (the radical-termination penalty for TMB and styrene; the ozonolysis-driven radical boost for monoterpenes), represents a meaningful step toward closing the gap between data-driven modeling and kinetic theory. The finding that high-ozone episodes reflect intensified turnover of specific rate-limiting precursors rather than a regime shift carries direct implications for VOC control prioritization in NOx-saturated urban environments and is well suited to the readership of ACP. The manuscript is clearly written, and the limitations (single site, single season, incomplete precursor speciation) are candidly acknowledged. I recommend publication after the following minor points are addressed.

Specific comments

1. **ΔVOC as a chemistry proxy.** The authors argue that the regularity of traffic emissions is captured by the 'hour of day' variable, so that the residual variance in ΔVOC represents chemical turnover (lines 117–127). However, Section 3.1 itself documents episodic, morning-confined solvent enhancements that do not follow a fixed diurnal schedule and could alias into ΔVOC. A weekday/weekend stratification, or a quantitative statement of how much ΔVOC variance is explained by the HoD variable, would strengthen this assumption.
2. **Choice of the 18-minute window.** According to Fig. S7, R^2 is higher at the 30-

minute window (0.809 vs. 0.792). The boundary-layer argument for preferring 18 minutes is reasonable, but a brief sensitivity check showing that the key conclusions (SHAP slopes of the principal species) are stable between 18 and 30 minutes would demonstrate that the results are not contingent on this choice.

3. **Interpretation of isoprene.** Given that strong daytime emissions can yield positive ΔVOC even during active oxidation, the explanation for the model's high isoprene sensitivity (detection of deviations from the expected accumulation rate) is plausible but remains qualitative. I encourage the authors to elaborate briefly in the main text, and to note there that isoprene was excluded from the linear fit in Fig. S8.

4. **NO_2 measurement and the HONO constraint.** The Thermo 42i employs a molybdenum converter, which is known to exhibit positive NO_2 interference from NOz species in roadside environments. Please briefly discuss the potential implications for the titration-sensitivity interpretation and for the OBM constraint of HONO = 2% of NOx, and, if available, report a sensitivity test of the OBM results to the HONO assumption.

5. **Unmeasured reactive alkenes.** The authors note the absence of light alkanes (ethane, propane), but light alkenes (e.g., ethene, propene) are likewise not quantified by PTR-MS and can contribute substantial OH reactivity at roadside sites. I suggest stating explicitly that the attribution is conditional on the measured precursor suite, with alkenes included in this caveat.

6. **Application of KORUS-AQ (2016) ratios.** The fixed ethylbenzene/xylene split (0.366) and the isomer-weighted kOH values are derived from the 2016 campaign and applied to 2022 observations. A sentence or two on the uncertainty introduced by possible changes in solvent regulations and fleet composition would be appropriate.

7. **Comparison with the WHO guideline.** The WHO peak-season guideline (31 ppb) is defined as the average of daily maximum 8-hour means over the peak season; comparing 83% of hourly measurements against it is not strictly consistent with the metric's definition. Please rephrase for accuracy.

8. **Numerical consistency.** The C8 aromatics concentration in the main text (1.40 ± 1.10 ppb, line 148) differs from the C_8H_{10} value in Table S1 (1.27 ± 1.03 ppb). Please clarify the scope of the calculation (e.g., whether styrene is included) or reconcile the values.

Technical corrections

- The axis and MAE units in Fig. S7 ("ppm") appear to be typographical errors for "ppb"; please verify.
- The toluene *k*OH is given as 5.7×10^{-12} in the main text but 5.63 in Table S2; please unify.
- Equation numbering is duplicated between the main text and the SI (both use "(1)"), and in SI Text S7 the SHAP equation is referred to as "Equation (1)" while labeled (2); please harmonize.
- Minor typographical issues throughout, e.g., missing space in "radicals(Tuazon" (line 275); subject–verb agreement in the Fig. 4 caption ("*k*OH for isomeric groups are").