Calibrating Interdependent Photochemistry, Nucleation, and Aerosol Microphysics in Chamber Experiments: Egusphere-2025-2412, Response to Reviewer Comments

We thank the reviewers for their constructive comments. Our responses to all three reviews follow, with one section per review. The reviewer comments are in bold and our responses are in a normal font.

1 Reviewer 1

The study by Donahue et al. presents an interesting approach to constraining H_2SO_4 production and loss based on particle number measurements, effectively through an optimization exercise across a multiparameter space coupled with modeling, ultimately claiming to minimize systematic and random calibration biases. Although lengthy, the paper is very well written, and the CLOUD team has made significant contributions in this area. I believe the paper could be publishable in AMT after the authors address some considerations, which I detail below.
Major comments:

1. **I would like to see some experimental validation of the model's input parameters. As it stands, the paper describes most relationships as if there were an ideal chamber. Even slight variability across this parameter space could significantly affect the particle number and other parameters. Considering that most chambers would likely experience even larger variability, I am not certain how useful this approach would be for the community.**

   As the reviewer notes immediately, this is a long and complicated paper. In truth, it is the most complicated paper I have ever attempted to write, and underwent multiple revisions over years to get to where it is. With help from the reviewers, we hope it will get to an even better level of polish. However, our responses to several comments will amount to "that would be lovely, but for the purposes of this paper, we have to draw a line somewhere". In spite of this, it is important to point out where we have drawn that line, and the potential consequences. So in the revised manuscript we will state clearly that we are taking the measurements we do not specifically adjust with a calibration factor to be perfect, in the sense of perfectly accurate and perfectly representative of a uniform well-mixed chamber. For example, our regularized method is to take measurements as provided as accurately calibrated and representative of the chamber, but then to adjust them with a factor $a_i \pm \sigma_i$. That is really the whole point of the paper. Regardless, a wide array of other boundary conditions could and should be represented as $b_j = 1 \pm \sigma_j$, with at least the variance in the parameter being an unknown (not to mention the covariance $\sigma_{i,j}$ ...). We do, however, need to draw the line somewhere.

   That being said, we *do* address the issue of uniformity in the manuscript. Specifically, we consider two different light sources with quite different degrees of uniformity. First, the so-called "UVH" Hg – Ar lamps, which are fed by fibers into the CLOUD chamber to provide a nearly uniform illumination, and second the so-called “UVX” 248 nm Ar – F excimer laser, which is fed through a single fiber into the chamber and so illuminates only a cone in the overall roughly cylindrical volume. This could matter here because the UV is the source of OH, which in turn oxidizes $SO_2 \longrightarrow H_2SO_4$. Nucleation, in turn, scales as $[H_2SO_4]^3$, and so any inhomogeneity will be amplified. Fortunately, the timescale for H_2SO_4 loss is relatively long (500 s to the walls, which can be roughly matched by the highest condensation sinks we observe), and so the chamber appears well mixed. Our observation that the same calibration factors for the UVH and UVX stages confirms that this holds water.

In terms of other chambers, if this is a problem, well then it *is* a problem, and we as a community need to know about it. However, it is hard to quantify without some form of measurement of variability (i.e. $\overline{[H_2SO_4]} + [H_2SO_4]'$). One thing we can do at CLOUD is to reverse the direction of the mixing fans, and research is being carried on with HO_x measurements that could constrain the degree of heterogeneity in shorter lived constituents. However, that is beyond the scope of this already sprawling paper.

2. **Similarly, there are many assumptions regarding instrument operation and performance (e.g., unreliable calibrations, assumptions about particle losses in tubing, particle measurement efficiency, and charging). Since these factors are crucial to the development of the proposed approach, the associated potential errors should be explicitly propagated.**

That would be lovely... Here we have chosen a few parameters to regard as uncertain and for the most part taken others as known. That being said, we *do* test near boundary conditions (i.e. the flows of O_3 and SO_2, which drive the observed and modeled concentrations of each). In the revised manuscript we will make these points where we take measurements as certain more obvious, and we will also reiterate that this is overall a proof of concept to demonstrate the possibility that this approach is feasible.

3. **So, although some effort has been made to characterize uncertainties, I would like to see error propagation carried out across all general dynamic equations (GDEs) and compared with both the measurement and model results.**

We absolutely plan this, and work on a formal parameter estimation is ongoing. However, in our opinion this current manuscript serves as the baseline for that ongoing effort. Here our standard is a “fit by eye” and without a formal propagation of error, including coefficients and covariances, we are not convinced that an intermediate level of error analysis would do anything beyond providing a possibly false veneer of formality. Before we can accomplish this, we need to establish objective functions as well as reasonable a priori bounds on parameters (for example, here we use two moments of the size distribution, the number and area (really condensation sink), but that is not necessarily the correct thing to do, and the two moments are quite obviously covariant. In our opinion, endeavoring to conduct a rigorous parameter estimation with error in the same document where we present this whole method would be simply overwhelming for any but the most dedicated reader (or author).

We do not want to be dismissive of these important comments. Our feeling is that we do need to draw a line somewhere, but we will indeed make that line clearer in the revised manuscript, and discuss the implications of pushing to a more complete consideration of error in future work. That being said, by far the largest uncertainty that emerges from this current work is the binary variable of whether the $(NH_3)_2 \cdot SO_4$ particles are growing wet or dry, and in our opinion none of the other potential uncertainties come close to that. This leading uncertainty stands above the rest in the current discussion, and we feel that is appropriate – too pedantic a discussion of all the potential uncertainties would run the risk of burying the lede.

Instead, the main conclusion here is that an experiment to test this binary (wet-dry) variable is necessary. In fact, we have carried out that experiment, and the analysis is forthcoming.

Minor comments:

1. **Although the paper is long and I personally like the narrative style, I believe it would be beneficial for readers if the length were reduced. In addition, a more concise section outlining the conceptual framework and the planned analysis would help focus the reader's mind. I understand that Section 2 broadly attempts to do this; however, it is quite lengthy and includes additional details such as the, chamber operation and instrument uncertainties.**

2. **Some of the figures are not very clear, and certain elements are difficult to discern (e.g., I cannot clearly see the statement in L178–180 on Fig. 3). In line with my earlier comment, it would also be helpful if the number of figures were reduced, or if some were moved to the Supplementary Information.**

In the revised manuscript we will add more roadmapping and also consider whether some of the figures can be moved to the supplementary material.

2 Reviewer 2

This manuscript presents a comprehensive and highly rigorous framework for the calibration of complex chamber experiments, using the CLOUD facility as a case study. The authors tackle the formidable challenge of disentangling interdependent variables – ranging from gas-phase photochemistry and wall losses to aerosol microphysics and charge dynamics – by employing a novel strategy of "disaggregating loss terms." I find this work to be of exceptional quality. The theoretical depth, particularly the treatment of Van der Waals enhancements to explain condensational narrowing and the nuanced analysis of "wet vs. dry" growth pathways, provides significant physical insights beyond mere calibration. The manuscript is honest in its assessment of uncertainties (e.g., the potential inhomogeneity of mixing) and demonstrates a high level of fidelity between the modeled processes and the observational data. This study not only validates the data quality of the CLOUD experiment but also establishes a new methodological benchmark for the interpretation of atmospheric simulation chamber data. I recommend publication subject to Minor Revisions. The manuscript is well-written and logically structured. However, I have identified a few areas where the clarity of the methodological innovation and the rigorousness of the mathematical notation could be improved. Addressing the following specific points will, in my view, further strengthen the impact and readability of this excellent work.

1. **I recommend a structural change to highlight the methodological advancement: Please separate the text in lines 28-33 (from “Our objective is to scale...” to “...among all of the factors”) into a new, standalone paragraph. These lines define the study’s core innovation – the shift from seeking scalar calibration factors (F) to optimizing a joint probability density function () that explicitly accounts for covariance. This separation is necessary to ensure the reader immediately grasps the paradigm shift in your calibration methodology, rather than viewing it as a mere continuation of standard procedures.**

   We will follow this suggestion in the revised manuscript. We will also add emphasis that we do *not* meet this objective in this work, but we are setting up that as the next (currently ongoing) stage (see the discussion in the response to reviewer 1 above). We also hesitate to claim that accounting for covariance among many uncertain measurements and calibration factors is novel per se. Covariance is a necessary part of any formal parameter estimation, though it is not infrequently ignorred

2. **There is a notation conflict regarding alt. In line 64, the text states “the amplitude of each light appears as a_{lt}”. However, in line 119, alt is redefined as the “calibration factor” while s_{lt} represents the amplitude. This inconsistency must be resolved to avoid confusing the measured signal with the calibration coefficient. Furthermore, given the paper’s focus on calibration, it is recommended that the authors explicitly break down a_{lt} into its physical components (i.e. cross-sections, quantum yields, and geometric efficiency). Additionally, line 263 introduces a new variable. What’s the difference between $\mathcal{L}^{\mathrm{cal}}_{\mathrm{lt}}$ and a_{lt}?.**

   Oops. This careful reading is useful. We will correct the notational confusion in the revision. As with several of our responses here, there is an entire paper about the light calibration in preparation. These notational slips are relics of the long history of this present manuscript, which started as a fairly simple “calibrate the lights” exercise that grew much farther than we envisioned at first. The full details of breaking those amplitudes down into their constituents require a fair bit of text (really a full paper) and so our decision was to simply take the recovered amplitudes as measured values for the purposes of this manuscript. We feel that is justified, but we shall add a brief discussion of how photolysis rates relate to the amplitudes.

3. **The authors express the production rate as line 258. At first glance, explicitly including the measurement calibration factor () in the definition of the physical production rate (P) seems counter-intuitive, as the actual chemical production is physically independent of the CIMS detection efficiency. I presume this formulation is a mathematical strategy to decouple the absolute magnitude scaling from the relative non-linearities, thereby keeping close to unity during the optimization process after reading line 259–262. If this is the case, please clarify this rationale in the text. Without explanation, it appears as though the authors are conflating the measurement equation with the process equation. Readers need to understand that this is a parameterization choice for the inverse problem, not a statement of physical dependency.**

   This is correct – it is a consequence of this, formally, being an inverse model. However, our presentation leaves something to be desired and we shall clean up this presentation.

3 Reviewer 3

This manuscript presents a study that integrates observations and models in a novel way attempting to reduce parametric uncertainty in both the collection of those observations and the modeling of those observations - a version of a "digital twin" of the chamber. This is quite an ambitious attempt, even with the qualifications the authors have made regarding the reduced scope of the work (e.g., not attempting the full Bayesian approach to uncertainty quantification). There a number of key findings in the study that are novel, including:

- **how even the temporal response (as opposed to the absolute value) of a gas-phase species concentration (e.g., sulfuric acid) can be a powerful constraint on both aerosol chamber processes (e.g., first-order loss) AND inherent sampling uncertainties associated with chemical ionization**
- **the deliberate choice of an experimental system (e.g., sulfuric acid/ammonia/water ternary nucleation) that is "constrained enough" (e.g., most of the model uncertainty rests in parameter uncertainty rather than process/structural uncertainty) to permit a somewhat straightforward "tuning" of the system**
- **the somewhat serendipitous aspects of the experimental runs that observed a sufficient dynamic range in process contributions that allowed for a clearer reduction in measurement and process uncertainty**
- **the identification and use of a reference instrument/observation (e.g., CPC total number concentration), in combination with a "simple" process system, to perform a step-wise constraint going "back-ward" from an endpoint (e.g., aerosol number concentration) to a starting point (e.g., photolysis).**
- **a workflow that can be used to address uncertainty more broadly in other systems.**

This work should be published, but I do have some concerns that I hope can be addressed regarding length and clarity as described below:

We thank the reviewer for a remarkably close reading of this manuscript, which will surely help to make a difficult presentation and easier read.

- **lines 29-30: Do these models also contain "parameters" with varying degrees of uncertainty (e.g., nucleation exponents) that should also be incorporated into the overall optimization? To what degree is model "structural" uncertainty (e.g., mis-represented/missing processes) a consideration in this workflow?**

  As with similar questions from other reviewers, the simple answer is "yes", but this is a much deeper question. Structural uncertainties are devilishly difficult to uncover, and all we can *ever* do is establish that a representation is consistent with observations, and ideally constrain parameters more tightly than their prior ranges. One example is the binary variable (in formal terms) that we introduce for "wet or dry". This only emerged after long consideration of the system, and there can be no guarantee that similar unrepresented processes are important. Perhaps the most obvious is inhomogeneity, which could be represented for all observables as $\overline{y} + y'$, but which would also require constraints. We do discuss this when the different light coverage of the UVH and UVX lights come up.

For what it is worth, the nucleation coefficients are not very important here, because we ultimately settled on selecting empirical coefficients within each broad light amplitude step. Determining J (nucleation rates) is *not* the goal of this work, and so we take those values to be whatever is needed to give the observed particle number (for a given stage – the exact nucleation rate still depends on measured $\mathrm{H_2SO_4}$). This is because we need to have the correct particle number in order for the size distribution to be even plausibly able to constrain the $\mathrm{H_2SO_4}$ condensation sink. Thus, while we use the number and surface area (condensation sink) moments of the distribution in our assessment, it is really the condensation sink moment that adds an extra constraint on the whole problem.

- **Figure 1:**
  - **Do the "arrows" represent process functional dependencies? (e.g., rates of production/loss as functions of concentrations and rate constants, etc.?) Can some explanation on what the "arrows" mean be included?**

    Yes these are loosely parametric dependences, which then reverse in the inverse problem (i.e. everything depends on the light intensities, etc). There is an element of production and loss as well, but it gets more involved when, e.g., $\mathrm{H_2SO_4}$ is the ultimate source of CS, but CS is the sink of $\mathrm{H_2SO_4}$.
  - **Would Figure 1 be a convenient place to introduce some of these process "reaction steps" that more explicitly include parameter definitions?**
- **lines 73-74: What does "semi" independently calibrated mean?**

  We will discuss more in the revised manuscript. The instruments are both calibrated with the same method, and the two groups jointly reviewed their data and calibrations before settling on these analyzed datasets. Thus, they are not blind calibrations, but there has been a degree of regularization.
- **line 76: Can you perhaps add some statement/reference on the observed robustness/stability of $\mathbf{O_3/SO_2}$ calibration seen in other studies, etc.?**

  We shall.
- **line 77: Should it be instead "All gas-phase instruments are corrected for vapor losses during sampling…"?**

  Yes.
- **line 82: Can an appropriate reference for the $\mathrm{CPC}_{2.5}$ be included? e.g., ideally point towards an experimentally-derived size-dependent counting efficiency during CLOUD?**

  We will do this.
- **line 89: Can we define these process symbols (e.g., $P(\mathrm{H_2SO_4})$, $j\mathrm{O(^1D)}$) either in text or in figure 1?**

  Yes.

- **line 93: How is the “amplitude” of the size distribution measurement being precise different from the “absolute magnitude” of the SMPS being less accurate described in line 95?**

  We are assuming that there is a size-independent counting (or flow measurement) error in the SMPS, as we have no evidence for (uncorrected) size-dependent errors. There are also other size distribution classifiers in the larger data ensemble that give us confidence in the overall assertion. This is yet another rabbit-hole we have elected not to go down in the interests of keeping this paper manageable as a proof of concept. However, a “real” optimal estimation would surely make use of all those data, with the assumption that the “real” size distribution had single value (was a single distribution), subject to (size-dependent) measurement error by numerous instruments.

- **line 95: Given the size range of the DMA-train, I would have thought that uncertainties in charging efficiencies (especially below 3 nm) would be an additional source of unknown uncertainty.**

  They are and we shall point that out – in the end this is why we use the combination of $\mathrm{CPC}_{2.5}$ to constrain the overall integrated particle number and the unified DMA train and SMPS data for the size distribution. Formally quantifying the uncertainty in the distribution within the constraint of the CPC is non-trivial.

- **line 96: Can a citation be included for the AIS that can point readers to “systematic uncertainties”?**

  Yes.

- **Figure 2:**

  - **panel (a) Can a short description/explanation of the vertical dashed gray lines referring to the individual “stages” be added to assist the reader?**

    Yes.

  - **panel (b) Are the time scales in panels a.) and b.) intended to be aligned? A consistent time-scale in those panels (run time vs. time of day) would be helpful to the reader.**

    The revised Figure 2a will have the common run time and not the absolute time.

  - **panel (b) Perhaps use “stage” instead of “sub-run” for reader consistency?**

    Oops. The term “sub-run” was from an earlier draft and should not be there. It however refers to the sequence of stages focused on a light source, e.g. the “UVH stages” are stages 01-07 and the “UVX stages” are stages 09 - 18. We will re-word the caption.

  - **panel (b) Revise to “follows the steps in light intensity as seen in panel (a)”?**

    Good idea.

  - **panel (c.) Can it be made more clear where the merge is between the nano-SMPS and SMPS data that contribute to the observed size distribution?**

    We will add a marker, but the change in size resolution (as well as the brief dropout of the DMA train near hour 16) makes this pretty obvious in our opinion.

- **lines 99-100: Can there be some consistency between the phrases “co-condensation of H_2SO_4 and NH_3” and “H_2SO_4-limited growth” that are frequently used throughout the manuscript?**

  We will search out these phrases and make them consistent.

- **line 104: Would it not be more accurate to state that the instruments provide direct measurements of particle concentration and number size distribution as opposed to “support constrained and accurate determination”?**

  Yes, thank you.

- **line 106: Would it not be more accurate to say that the size distribution and total number are directly measured (as opposed to well-constrained), and additionally provide constraints on other processes?**

  The instruments provide an experimental constraint. The term “directly measured” contains some assumptions as well, but we will consider this in the revision.

- **line 153: Can these parenthetical references (e.g., UVH, UVX, LED) be also added to the Figure 2a.) legend?**

  Yes.

- **line 162: Is the “H_2SO_4 depletion” referring to the more gradual periods of decreasing concentration mentioned just before (e.g., stages 5-6 and 14-15)? A clarification would be appreciated.**

  Yes. This is referring to a decrease compared with trivial expectation of no condensation (only wall) loss. In the revised manuscript we shall include a supplemental figure to complement Figure 8 showing the model predictions with no condensation loss.

- **Figure 3:**

  - **What is meant by “Raw” charged particle observations? Are these data prior to any correction or inversion?**

    We mean “uncorrected”. Specifically, the is some formation of small (“primary”) ions within the instruments, causing a “ghost” signal of very small primary ions even during “neutral” stages during which the high-voltage clearing field is on (and we are confident the ion concentrations in the chamber are below the detection limit. This instrumental background is subsequently removed later in the analysis.

  - **line 193: A more general comment that the term “first-order” is used extensively throughout the manuscript, and some clarification on what “first-order” refers to might be helpful, e.g., first-order with respect to number concentration.**

    We will do this.

  - **line 193: For the phrase “requiring only precise H_2SO_4 measurements”, would it not require also that the calibration is constant (albeit unknown) over the time period of the first-order decay, not just precise measurements?**

Perhaps one could define a power spectrum of instrument response with respect to a steady input signal (a calibration standard), and any variability at a higher frequency than some threshold could be defined as "noise" affecting precision, and any variability at a lower frequency than that threshold could be defined as "drift" affecting accuracy. There might even be a notable minimum in the power spectrum, making this separation obvious. In our opinion we have bigger fish to fry.

- **Figure 4:**
  - **What is meant by "raw" chemical drivers for H_2SO_4 formation? Is this raw data that has not been corrected or calibrated?**

    Not the best wording. These are the unfiltered (but calibrated) signals from instruments measuring the chemical drivers. Later we apply a median filter to smooth them.

- **line 204: "All the data in Figs 2 and 4 require some amount of refinement." Would the term "data conditioning" be more accurate than "refinement"?**

  Yes.

- **line 205: What is meant by "sharp edges"? Is it more like "maintain the temporal dynamics", or "sudden changes in concentration"?**

  When e.g. lights turn on, the amplitudes change suddenly. So it is more like "sudden changes".

- **line 211: How does the particle growth time compare to the chamber mixing time?**

  The chamber mixing time is of order 1 minute, so it is much faster than other timescales in the experiment. This is why we can treat the chamber as a continuous stirred tank reactor.

- **line 215: How does the number concentration uncertainty change as this "step-function" assumption is relaxed to the more experimentally-observed "sloping" size-dependence in the cut-off size range?**

  Very little. The point is that for the vast majority of the dataset, the bulk of the size distribution is well larger than the cutoff size, and so any uncertainty in the CPC size dependence is immaterial. This is another "rabbit-hole" that we are stepping over in this manuscript, but we feel this (skipping over it) is well justified as well.

- **line 296: Can it be re-mentioned that the multiple instruments contributing to the "merged" number size distribution include the DMA-train, nano-SMPS, and regular SMPS?**

  Yes it can – we shall.

- **line 305: Is this derivation adding substantively to the discussion/interpretation? Could it be possibly moved to a supplement? A more general comment: perhaps just emphasize the contributions of the various constraining parameters?**

That Van der Waals term is very important (the difference between the dashed and solid red curves in Fig 6d) and this is where it appears. More generally as well, our preference is to derive microphysics starting from the collision (small particle, or large Knudsen number) limit, and then to introduce terms such as diffusion limitations as particles grow larger. Because this differs from the more common continuum limit derivation that then grows "Cunningham slip corrections" and "transition regime corrections", we have elected to recapitulate the basic derivation in several publications (our experience is that readers are sometimes confused without this).

- **line 321: Is the term "$\alpha_{i,p}\, E^{\mu}_{i,p}\, \epsilon_{i,p}\, e_{i,p}\, B_{i,p}$" the "overall collision adjustment"? If so, having the term "$s_i/4$" right after it is somewhat confusing. In other words, it might be easier to read if the term "$\alpha_{i,p}\, E^{\mu}_{i,p}\, \epsilon_{i,p}\, e_{i,p}\, B_{i,p}$" follows right after the phrase "overall collision adjustment" so that term "$s_i/4$" can follow right after "nominal H_2SO_4 condensation speed".**

  Yes that construction is poor. We will clean this up.

- **Figure 5:**

  - **In the y-axis label in panel (a), where is "vdep/(c/4)" defined?**

    We will change that label. This should be $s^{\perp}_{i,p}/(\overline{s}_i/4)$.

- **line 332: How does this use of "data conditioning"/interpolants improve precision? The precision of retrieved parameters?**

  We will change this to "reduces noise".

- **line 333: Why not experimentally characterize (as well) the detection efficiencies of the combined nano-SMPS and regular-SMPS system so that quantitative closure could have been achieved between total number concentration ($\mathrm{CPC}_{2.5}$) and integrated total number concentration (SMPS)? Additionally, as an aside, how were charging and counting efficiency corrections at sizes below 7.2 nm determined and applied?**

  To some extent that is what we are attempting here. Fundamentally, we assume there is a single "true" size distribution $n^{\circ}(d_p) = dN/d\log_{10} d_p$ (or we could even include $n^{\circ}(d_p) + n'(d_p)$ to account for inhomogeneities). We are then sampling this in various ways, with various size dependent losses during sampling and then various size dependent transmission, detection, and charging functions in various instruments, which collectively provide a constraint on that true distribution. The actual charge distribution (fortunately, almost entirely +1, 0, and -1) is also very important, as we discuss.

  It is worth regarding this current paper as an essay on what one can do with a large ensemble of datasets for a single experiment. We definitely *do* have in mind some grand optimal Bayesian estimation of the "true" state of the experiment over time, including *all* the chamber, transmission line, and instrument response models in some form of grand inversion, relating actual raw instrument signals back to the true state of interest. Maybe we will get there some day. In this instance, we do not even get so far as to provided a single quantitative posterior parameter estimate, beyond "I can tell the difference between 1.0, 1.15, and 1.3 for the H_2SO_4 calibration factor, and 1.15 is much better". We also do want to motivate "we are pretty confident in the $CPC_{2.5}$ but almost everything else gets a little iffy; however, taken together they are quite robust. Finally, the journey ends with "if the integer (wet/dry) growth variable is 'dry', the H_2SO_4 cal factor is 2.0, and if it is 'wet' the factor is 1.15; the balance of evidence favors wet". All the rest of this insanely convoluted mess does not really alter that binary result. And we can test it (by trying at different RH).

- **Figure 6:**

  - **"Particle measurements after calibration to $CPC_{2.5}$ values". Strictly speaking, this is not a calibration, but a "scaling".**

    Correct and we shall correct it.

  - **"the constraints are especially strong". What does it mean for a constraint to be "strong" Is a "strong constraint" equivalent to "low measurement uncertainty"?**

    The better word is "tight".

- **line 339: "(after correcting for the SMPS undercount)". Is this argument of CPC and SMPS agreement after scaling resting on the observation that the "calibration factor" is "constant" over size?**

  Yes.

- **line 364-365: How is "accurate constraint" vs. "precise constraint" distinguished here?**

  OK. For me, "accurate" is "I know what the y-axis actually is, while "precise" is "I could put 'AU' on the y-axis but the relative values are good."

- **line 368: Should "run 1844" be "run 1833)?**

  Oh for the love of suffering Jesus do not tell me there is another run. Yes.

- **Figure 7:**

  - **How sensitive is the optimization/calibration workflow to the contribution of this "small source of unknown origin"?**

    Not at all. It is several orders of magnitude below where anything matters. That being said, it is really cool that the signal dips low while the condensation sink is high, then relaxes to a steady value – this very strongly suggests it is a real H_2SO_4 source and not just an instrumental background (common to both instruments).

- **line 376: "these (precise) level steps". What is "precise" about these level steps?**

  The H_2SO_4 signal is a nearly perfect square wave when the CS is low.

- **line 381-382: Is the "small unknown source of H_2SO_4" really considered "negligible" if it is also included in the optimization workflow?**

  It is just an interesting signal that is almost certainly real and could either be swept under the rug by selecting a higher y minimum or acknowledged and discussed.

- **line 403: The parenthetical "horizontal" can be removed here.**

  OK.

- **line 409-411: If condensational loss is contributing around 50% to the total loss during this period, this seems to imply to me that particle measurements (e.g., those from which CS is determined) are important, as opposed to "regardless of any particle measurements".**

  This is actually important. The point is that the pseudo first-order decay of H_2SO_4 is a very *accurate* measurement of the actual condensation sink. No matter what the particle measurements say, no matter what the Van der Waals Hamaker constant actually is, we really *know* what CS is, because we *know* the H_2SO_4 lifetime really well. It so happens that the CS we wind up observing agrees with this quite remarkably well (it is the Hamaker constant in the end that is the most "free" adjustable parameter in the whole deal). That is the central aspect of this whole story. At first, this Run 1833 was supposed to be a simple lamp calibration with H_2SO_4 vs lamp intensity, but people (un-named co-authors on this paper in the control room at the time) elected to add NH_3 and do a little nucleation as well because "Ammonia makes the party". Our first reaction was that this impulsive decision more or less ruined the run, because the H_2SO_4 during e.g. stages 05-06 did not step up to where it should but instead sagged away due to condensation loss. After a bit of thought, however, it became clear that this impulsive act was an act of genius – it is the dual constraints of the H_2SO_4 lifetime and the observed particle size distribution that lock the whole puzzle together.

- **line 422-423: Are these two models run in a "coupled" fashion?**

  No. See the long answer to the previous question. The gas-phase model (with observed CS) reproduces the observed H_2SO_4 extremely well. Consequently, we can run the gas phase model with the observed CS, and the particle model with the observed H_2SO_4 and compare them offline. We *can* run them coupled, but there is no reason.

- **line 441: "This is crude, but because the fraction removed is small, errors caused by this crude treatment are negligible." Can this be verified quantitatively?**

  We think so, but that awaits the more formal Bayesian parameter estimation that does this whole thing properly.

- **line 448-449: Why not simply implement the (cited) ternary nucleation parameterizations (exponents and prefactors) instead of treating the nucleation coefficients as "adjustable" parameters in the optimization? My thinking is that either consistently: 1.) fully use the best information available, or 2.) if not using the experimentally-derived parameterizations, then treat the ternary nucleation parameterization itself as a fully adjustable parameterization.**

Because if we get the particle number wrong we are ruined. We cannot calculate the CS unless we have the number.

- **line 464-465: Why does apparent "agreement" between model nucleation rates influence choice selection?**

  Because this is about making modeling decisions – we need to inform the ion concentrations and ion-induced nucleation rates in some way, and the primary ion observations are quite noisy. In the ideal world, we would have an optimal estimator for the actual ion concentration constrained by measurements, but we need to make some choices about where to focus.

- **line 474: "and the figures are easier to read with this narrower time range". It's not clear to me how "readability" should be used as a criteria for which stages to analyze?**

  Oh we analyzed them all. Do you really want to see them?

- **line 482: "When species are not effectively non-volatile, that in turn depends on the volatility of the vapor species". This statement seems awkward and redundant. It equivalently reads "when species are volatile, that in turn depends on the volatility..." Can this be clarified?**

  That sentence definitely needs to be reconstructed.

- **line 485: Can the partition coefficient "$K_{i,p}$" be defined?**

  It sure can, especially because it is the Kelvin term, not a partitioning coefficient. The Kelvin term will make the activity of a species in the particle higher by an order of magnitude when $d_p = d_{K10} \simeq 5$ nm, and this is why the water content (and for very small particles, and quite critically here, NH_3) is lower for small particles.

- **line 529-530: Was there not sample drying prior to charging and mobility classification?**

  The state of the sheath flow is another deep rabbit hole. If the particles are growing "dry" without water, none of this matters. If they are "wet", however, a great deal matters. The sheath flow in the instruments is filtered air from the CLOUD chamber so in theory the RH in the instruments is the RH in CLOUD – so as a simple answer "no" they are not dried (and they should not be, ideally). Furthermore, as discussed above, we *know* the actual condensation sink because we know the first-order loss coefficient for H_2SO_4. So, in the end, the uncertainties here will be strongly covariant with the Van der Waals term (the Hamaker constant) but will not substantially influence other conclusions in the paper.

- **line 567-570: Why would the nucleation coefficient vary with stage? Is the nucleation mechanism different between stages?**

  As discussed above, rather than assuming a single "correction" factor for the nucleation coefficients, we decided to do this stage by stage because of the large stepwise changes in UV intensity and thus H_2SO_4. This is because so very much of the analysis depends on getting the total number correct. The good news is that when we turn things around and try to measure the nucleation rate (which is more or less the whole point of CLOUD), we are in fact very sensitive to J and so measure it well. However, for this analysis, J is the very last coefficient of interest, and so the best choice was to just ensure that each stage had as close to the observed J as possible. It does not make a huge difference – simply scaling the assumed parameterization does a reasonable job – but it is notably worse.

- **Figure 10:**
  - **"Essentially all the features of the observations are reproduced by the simulation with good fidelity, including the timing, intensity, growth rate, and final size of the nucleation bursts during each stage." Where is this model-measurement comparison of timing, intensity, growth rate, and final size described (and/or presented) quantitatively?**

    Nowhere. We perform no formal uncertainty or goodness of fit analysis. That being said, the agreement is obviously, to the eye, well within 10%. Based on that, we feel justified in asserting that this is a "quantitative" agreement, without actually getting to the point of quantifying it. That comes next.

- **line 579-580: Where is the growth rate comparison illustrated?**

  In Fig. 10. The observations (a) and model (b) agree almost perfectly.

- **line 590-592: What does "inhomogeneity" refer to? Spatial distribution?**

  Yes, spatial inhomogeneity in the CSTR.

- **line 600: Where is it shown/illustrated that the nucleation rate coefficients remain nearly constant as sulfuric and ion concentrations vary? (and what does "nearly" constant mean"?)**

  In the revised paper we will present the distribution of "calibration factors" applied to the nucleation parameterization. They are quite large (of order 75). We are not keen to belabor this because the parameterization covers a huge range of conditions (temperature, RH, and both H_2SO_4 and NH_3 concentrations), whereas this run 1833 covers a much smaller slice of that phase space. It is very interesting to wonder whether the (quite large) discrepancy between the parameterization and the specific nucleation constants (calibration factors) here reflect real errors in that parameterization or are some form of artifact of the large parameter space. One of the main goals of this workflow is to establish a procedure so we can ensure that, campaign to campaign, measurements of these basic parameters are sufficiently accurate that we can establish intercomparability. We are not quite there yet.

- **line 611-613: How is "clear signs of three modes" consistent with "hint of a maxima at 5 nm"? To me, "hint" is not equivalent to "clear". It might be worth including some cautionary statement on the potential for "over-fitting" modes.**

  There is no fitting here. The data show a clear mode. The model shows a hint. It is that simple.

- **Figure 11:**
  - **"(a) Measured size distribution, combining DMA train and scaled SMPS values". Were the DMA-train data also scaled to ensure good "overlap" with the scaled SMPS data?**

No. We only scaled the SMPS.

- **line 653: Can “light intensity” be slightly changed to “changes in light intensity”?**

  Yes.

- **line 657: “At a minimum it is important to interpolate these observed mixing ratios”. What does this mean? Is this interpolation over time?**

  We will clean this up. What we mean is that the O_3 and SO_2 are important boundary conditions for $P(H_2SO_4)$ and so we need to get the values right, but it is also interesting to try to understand why they vary.

- **line 659: Is the “wall timescale” referring to the “wall-loss timescale”?**

  Yes.

- **Figure 12:**

  - **“(c) Measured and modeled H_2SO_4 (with modeled HO_2 and OH)”. I’m assuming the red circles are observations? Can this be stated? Also, what is the red solid line referring to? And are the HO_x and HO_2 lines overlapping?**

    The circles are the (interpolated) observations and we will add this to the caption. The legend refers to the model, so the red line is the modeled H_2SO_4. Yes the HO_2 and HO_x are overlapping – this is almost always true, as OH is 2 orders of magnitude lower.

- **line 673: Should “major contributing” be revised to “major contributing processes”?**

  Sure.

- **Figure 13:**

  - **“Simulated (curve) and observed (points) gas-phase precursors with smoothed and interpolated measurements”. Just for confirmation, is it the observed points that are smoothed and interpolated? Does it not make more “sense” to interpolate model output to observed time scales?**

    No it does not. The model needs to be able to call for a measured value at any arbitrary time, because in some of the (very many) variants of the model we are constraining values to the measurements. Consequently, measured $H_2SO_4(t)$ needs to be valid for arbitrary t. Furthermore, all the instruments and measurements have independent time bases. It so happens that the median filtering and smoothing that goes into the interpolant is also effective at reducing noise.

- **line 690: The phrase “there is nowhere to hide” feels editorial and non-technical. Is it necessary to include?**

  This is such a long and taxing paper, occasionally showing that an actual human is writing the words does not seem like too much to ask.

- **line 736: Does nucleation at 2 nm need to be consistent with the sizes at which nucleated clusters form in the sectional model? (e.g., at 0.8 nm)?**

  This is a simple "quasi" modal scheme. We might ideally apply some form of survival probability to 2 nm, but so far we have not. That is in the sectional microphysical model, which is the one we use for the quantitative comparison between measurements and model.

- **Figure 14:**

  - **panel a: Are the points the observed ion signals? Can this be stated?**

    Yes and yes.

  - **panel c: What are symbols/line colors for the simulated and observed data?**

    We should do better there. The observations are the black dots, and the blue curve is the sum of the two magenta curves (the two modes larger than 2.5 nm).

  - **"Finally, the simulated N_5 trace reasonably reproduces the CPC$_{2.5}$ observations (shown in gray)". Are these referring to the filled circles? They don't look gray to me.**

    They are the filled circles, and they are gray, but there are rather a lot of them and they come across as very dark gray.

- **line 774: "this modal simulation reproduces the observations with good fidelity." What are the quantitative criteria for determining "good fidelity"?**

  The criterion is a visual assessment of "goodness of fit". This is a qualitative paper in the end. The formal, quantitative analysis is coming.

- **line 794: "quantitatively reproduces the concentration dips". Without well-defined metrics for evaluation, this reproduction seems more qualitative than quantitative.**

  We agree (see the immediately prior answer). However, we also do not – there is a clear difference in solutions with 10% different parameter values. There is no easy way to describe this beyond this discussion, but "qualitative" seems like a harsh descriptor for an overall agreement on numerous observables that is just plain obviously good to within 10% or so. We absolutely are working to make that formal, but in the mean time we are using the term "quantitatively" qualitatively.

- **line 817: instead of " are only meaningful", could you not use instead "are interpretable"? E.g., these experiments are meaningful only because they are interpretable.**

  Yes.

- **Figure 19:**

  - **"The system is often not in steady state (the dashed net rate is not always zero)". The net rate line does not appear to be "dashed" to my eyes.**

    Thank you - one of our co-authors hated on the dashed line and we turned it into a very thin solid line, but we missed the text.

- **line 941: "Uncertainties in this analysis rest on the accuracy of $CPC_{2.5}$ data and the precision of the H_2SO_4 measurements". Does analysis uncertainty not also critically depend on DMA-train uncertainties (e.g., size-dependent counting and charging efficiencies below 7 nm)?**

  No. The real meat of the analysis and the uncertainty arises when we have a CS similar to the wall loss, and that is always when the particles are larger.

- **line 944-945: Can further justification for "correcting" the SMPS data be provided?**

  This is empirical. In the end, something has to give, and we picked the CPC as the most reliable measurement. Again, at the end of the day the H_2SO_4 lifetime constrains this in a completely orthogonal way.

- **line 945-946: What does it mean for "precise H_2SO_4 behavior" to be "strong"?**

  It means nothing – we will re-word.

- **line 955-957: I find it somewhat concerning that there was no sample drying prior to charging and mobility classification in the SMPS. Additionally, would this also not impact CPC measurement in particular for sizes that have strong size-dependent activation?**

  This is a two-headed monster. We need to know the actual, not dry, size distribution in the chamber. The real world has humidity and water activity. If RH gets too high, many of the measurements are challenged. However, different instruments also have different protocols – in the revised paper and especially the followup work we shall endeavor to be clearer in our description.

- **line 964: "calibration factor is constrained to within approximately 0.1". What does this mean? That the calibration factor uncertainty is constrained to with 10%?**

  The calibration factor is 1.15 ± 0.1 to the best of our qualitative judgement.

- **line 997: "A formal optimal error estimation will require a rigorous Bayesian treatment with a microphysical model that fully couples both the gas-phase chemistry and the evolving particle charge distribution." I would argue that a rigorous treatment would also include formal instrument simulators (and their transfer functions) that capture "transfer function error/uncertainty"**

  We agree. This is massive undertaking. We are proposing to address this step by step, starting here, then undertaking a formal parameter estimation with covariance, and subsequently adding in the instrument models.