Olin et al. introduce an interesting study on how aerosol particle number emissions described with a previous inventory can be improved by implementing newer emission factor data. The authors derive emission factors for traffic in different size bins from their observations, apply these to improve the description of emissions in an air quality model, estimate the composition of the particles in different size ranges, compare the original and updated modelled particle size distributions and compositions with observations, and discuss the improvements of the update in terms of human health impacts. The topic is timely and the results are interesting and would deserve to be published in ACP, but in the current version the description of the methods and stepwise presentation of results is not adequate for fully understanding the results.

I list first my major comments related to the methods and their illustration, and below this, more detailed comments related to the text. If the authors can satisfactorily reply to these comments and modify the manuscript accordingly, I can recommend the publication in ACP.

Major comments:

The authors do not present any figures on the determination of EFPSD, which is one corner stone of this study. They refer to their earlier study in which a similar method was applied for sub-3 nm particles. However, I would assume that the determination of emission factors for larger particles is not as simple, due to their longer lifetimes which causes more varying back-ground concentrations. Where one can expect that sub-3 nm particles at the kerbside are fresh particles either from the traffic or from NPF, larger particles may originate from sources further away and their concentration can be expected to be less sensitive to nearby sources, especially with time resolution as low as 9 minutes. The authors also conclude that the derived EFPSD agrees with the one reported by Hietikko et al. (2018) and that this implies the method used in this article is acceptable. To evaluate the acceptability, the reviewers and the readers need to see how the data look like.

The authors use PMF on the diurnal patterns of the EUCAARI emission inventory to extract the contribution of traffic. Does the detailed category-level specification of emissions not
exist anymore? Have the producers of the emission data (Denier van der Gon and others at TNO) been contacted to inquire for such specification? If they have been contacted and the specification does not exist, this should be clearly stated and the personal communication to TNO could be used as a reference. Otherwise, I would strongly recommend the authors to reconsider making this contact. The study would seem much more exact or accurate if the original traffic emission output data could be applied.

The authors use CFD modelling for determining the composition of European wide aerosol emissions from traffic. If I understood correct, this CFD model result is based on one diesel bus. Since the CFD model is described only in (non-peer reviewed) MSc thesis (in Finnish), the description of the model and main results should be given in the article. Part of this could be included in the supplementary material. The composition results are not compared to any previous article and the composition is not discussed in the introduction. In the current form, the part on composition should not be published in ACP.

Minor comments:

Lines 33-34: Traffic emissions in Paasonen et al (2016) are not based on EUCAARI inventory, but on EU FP7 project TRANSPHORM.

Lines 80-85: is the composition of diesel bus exhaust assumed for the whole fleet? Some words about how this assumption may bias the results.

Line 92: Some words about how the interpolation may bias the results. Why not using simply 9 min averages of CO2 as well?

Lines 100-103: some references to the observation sites should be included.

Line 106: Could refer to recent Okuljar et al. article.

Lines 145-155: Why do you not investigate further factors 7 and 11, which both seem to be related to rush hours? How is their size distribution and is there a good reason to exclude them? At least factor 7 would contribute significantly to overall result.

Line 148: Since the Denier van der Gon -report is not available without request from the project office, the count mean diameter and/or other features of their PSD should be listed.

Line 176: Luoma et al seem to have the trend calculated mainly for periods 2015-2019. They also mention that the applied trend, -7.1 %/a for PM2.5, is the only trend they could not determine a statistically significant trend. Can you somehow justify extending similar trend to 2008?

Lines 216-218: While this is possibly the case, this sentence should be reconsidered when the Paasonen et al. (2016) paper is notified to be updated in terms of traffic emissions from EUCAARI to TRANSPHORM. One way to discuss the representativeness of Paasonen et al. emissions may be to reflect their comparison to emission size distribution calculated from long-term observations in Kontkanen et al. (2020, https://doi.org/10.5194/acp-20-11329-2020).

Lines 220-227: An equation including the three different modes, or at least better explanation, should be given. I am not entirely sure if I understand what the trimodal fit here means. The authors often refer to previous studies in a way that even basic understanding of what is done in this study is not possible to get without reading the other
lines (same holds for the determination of emission factor, see my first major comment).

Lines 228-232: Would be good to mention here also that PMF 6 is presumably related to diesel particles only. And that NCA emissions are not described for other sources in EUCAARI inventory.

Line 275: Some references required for the underestimation of PSD measurements in sub-10 nm size range.

Line 281-282. It would be interesting to see separate figures for the different sites, especially to those where one would expect the traffic emissions to play a big role in <10 nm concentrations (Kumpula and possibly Melpitz).

Lines 290-293: Increasing condensation sink is a plausible reason for decreasing modelled concentrations, but do you have any evidence of that being the reason? N>100 does not seem to change (Table 2), but how much does N>50 change? Or could you draw a map of change in coagulation sink as well? Also later, in Fig. 7, it is difficult to understand the difference in sink: Fig 7c shows very similar size distribution for original and updated model run in sizes >30nm, actually with higher concentration in original run in 40-50 nm and 500-600 nm size ranges. It looks like the sink in the original run in 7c is higher than in the updated run, whereas the interpretation of the differences in lines 341-343 suggests the opposite. Additionally, related to Fig 7c, I wonder why the modelled PSD jumps so much up and down between different size ranges: the difference in concentrations in neighbouring size bins can be up to two orders of magnitude and the nucleation, Aitken and accumulation mode do not show any modal distribution. Also in Fig 6., the modelled size distributions are surprisingly unsmooth for monthly means.

Lines 355-356: Is the line for observations in London missing, or are there no observations? Why wouldn’t the authors use Kumpula observations instead or additionally to London as an example of site with traffic in vicinity? London is not even mentioned in Section 2.2. If there is no data from London, adding Kumpula (or Melpitz) data becomes even more crucial.

Language overall: There are many quite long and difficult sentences, due to which I suggest the authors to doublecheck the language in general. Additionally, the use of the word “unity” instead of “one” or “one-to-one” in “ratio over or below unity” does not sound good to me. In my understanding unity is something that cannot be exceeded.