Interactive comment on "Methane emission estimates using chamber and tracer release experiments for a municipal waste water treatment plant" by C. E. Yver-Kwok et al.
Anonymous Referee #2

This paper describes an experiment where various measurement techniques were applied to the same source but with different spatial scales. As such the paper is within the scope of the journal. Overall I find that the paper could still be more compact and that the authors should consider if some of the material could be moved to a supplement (e.g. fig 6, 7 and some of the descriptions there of). On the other hand I realize that the journal is on measurement techniques and I readily admit I am more interested in the results and got a little impatient with the time it took me to get there. I have some comments on section 1 to 5, mostly minor.
I have more serious concerns about section 6 and 7. Here I think the authors make a fundamental error in the comparison with the EDGAR emissions data base. The main point being that they should compare emission factors, not inventories. I will address all points in more detail below and in order of appearance in the paper.

We would like to thank the reviewer for the helpful comments and suggestions. Our answers and corrections are detailed following the reviewer's comments.

Abstract: The last sentence is true but not very relevant, also the statement relies heavily on the other CH4 sources of Valence that are not investigated (If these sources are over or underestimated the 1.5% would change). It should be removed or replaced by a comment on the derived emission factors (see later comments on section 6 and 7)
This sentence will be revised according to the reviewer's suggestion about emission factors.

Section 3 p 2964, l 6 – Why use a conic shape? If you have boxes and you know the waterlevel, the volume can always easily be calculated.
We would like to thank you for this suggestion. We agree that using a box would help us minimize the error associated with the water surface area.
Having used a conic shape, we were still able to calculate both the chamber volume and water surface area. What we discussed on top of p 2964 was how variation in the water level might have affected our total estimate uncertainty. Indeed, with conic shape, the water level affects both the uncertainties associated with volume and surface area, while a box design would help us eliminate the surface area uncertainty. With your permission, we would like to include this insight in the discussion, where we contemplate further improvement of our methods.

P 2978, Line 22:
In the case of the closed-chamber, the water area enclosed by the chamber and the air volume in the chamber are the parameters associated with the strongest uncertainties. They eventually depend on the uncertainty of the water level. Consequently, a more accurate measurement of the water level in the chamber and a minimization of its variation should be aimed at if lowering of the total uncertainty is desired . The error associated with the water surface area can be fully eliminated by choosing a box over a conic chamber. With a box, the variation of the water level would not affect the surface area across which the exchange takes place.

Section 5.1 p2973, l9 "supports the hypothesis of very local emissions" This is a bit odd as you base this conclusion on a comparison with a background station located about 500 km away. There could be lots of reasons why the local concentrations around Lyon could be different than Paris?

We compare measurements both localized in suburban areas with a relative common environment (highways, fields, buildings). However at Gif-sur-Yvette, there are no local sources at the exact location where the measurements are performed. Hence we observed a relatively elevated background but no peaks, while in Valence, measurements were collocated with the WWTP. Such high signals have to be local, high concentrations from Lyon, for example, would be highly diluted before reaching the instrument.

Section 5.3 p2974 l 25 Although I understand the need to recalculate everything to daily fluxes this is not correct when looking at erratic events. You measure for 10 min. and then calculate a flux per day. Strictly for such events you can only calculate an average flux per minute (based on a 10 min measurement) but you have no idea if it would lead to such an average flux per day.
We agree with your objection. Overall, we wanted to express all fluxes in the same unit. However, erratic events are very difficult to quantify. That is why we presented this "thought experiment", in which such erratic fluxes measured over 10 min are assumed to persist the whole day (P2975, 18-19). We know that this is hypothetical and we tried to express this. In the revised version of the paper, we changed the units of the erratic event in g/min in this section and in Table 2. On the next page, we still formulated this "thought experiment", but we rephrased so that it is clear that we are unable to derive a daily flux estimate.

P2974 l.7-25
Only 4 out of 8 floating chamber measurements on the clarification basin exhibited an approximately linear increase (chamber runs 2 (from minute 7 on), 3, 4, 7, see Fig. 6, panels b, c, d and f). The emissions calculated from these measurements averaged 3.8 mg min-1 (for the individual values see Table 2). The SD, calculated to assess the spread of the individual measurements, was 2.6 mg min-1. It is reasonable that upscaling to the whole basin introduces uncertainty when not all locations on the basin were covered by our measurements. The uncertainty in volume and area contributed to the squared total error by 52 and 48 %, respectively, for all four diffusive flux measurements. The uncertainties associated with CV, pressure and temperature were negligible. Based on our four measurements, we consider the obtained average of 3.8 mg min-1 or 5.4 g d-1 to give the order of magnitude of the diffusive exchange flux, which represents the lower limit of the total emissions from the clarification basin.

For the other four measurements (see Fig. 6, panels a, e, g and h), the increase cannot be linearly approximated. Due to the very sudden increase of the methane concentration in the chamber, we think that erratic methane emissions caused this non-linearity, i.e., ebullition. Since such events might occur more frequently close to the rotating arm and the number of measurements is too small for estimating the frequency of such events, it is difficult to estimate the methane flux from the basin generated by erratic events. However, we can state that the highest average flux for these measurements over a 10 min period was 169 mg min-1 (chamber run 5, Fig. 6 panel e).

P 2975 l 18-19. This is repetition, you say exactly the same on previous page l25.
We will try to make the distinction clearer. On page 2974, l.25, we state what we actually measured. On this page, we state what the maximum flux from this basin would be if it were only for such erratic events.

P 2975 l 18-19:
Within the short time of measurements on the basin (1 day), it was not possible to do a systematic study of the methane emissions due to these erratic events. Therefore, here, we can only provide an approximate estimate for erratic fluxes from the basin. We choose this approximate estimate in a way that it expresses the maximum erratic flux that we can consider possible according to our measurements. For this, we took the highest of the four erratic fluxes we measured and assumed that this flux, measured over 10 minutes, occurred for 24 h over the entire area of the basin. In that case, the emissions would sum up to 243 g d-1.

P 2977 l20-25 Something is wrong here in the numbers. Degassing basin 1.13 kg/day and clarification basin 0.8 kg/day. That would never lead to the conclusion that the latter could be neglected. Please recalculate, Fig 9 clearly shows the conclusion that clarification basins are not important – so one of the two figures given here is wrong.
Please check also if clarification basin is the correct English term – it sounds a bit strange to me, but might be correct.
The number 0.8 kd/day come from assuming that the erratic event would happen continuously on the whole basin which is most probably not the case. Using only the values from the diffusive emissions, the number is lowered down to 0.005 kd/day which is then indeed negligible compared to 1.13 kg/day. This will be rephrased in the revised text.

P 2977 l. 22:
Given the wind direction, this emission number could include emissions from one or more of the clarification basins. However, whilst the floating chamber measurements showed that maximum emissions from the clarification basins were comparable (0.8 kg d−1), on average they were about one quarter of this amount. As this figure includes erratic fluxes, for which a conservative upper limit was given, the true fluxes are likely to be much lower. For example, if only diffusive emissions were included, then the flux would be smaller than 0.01 kg/d per basin. This compares to 1.13 kg/d from the degassing basin. Conclusively, the emissions from the clarification basin contribute only very little to the emissions from aquatic surfaces in the WWTP.

Section 6.1, last sentence. The biggest uncertainty for more robust WWTP emission estimates is probably that you measure only 4 days and try to derive an representative flux from this short period. Especially when you aim for more robust estimates, some uncertainties matter little. The results show that the chamber methods can show which basins are important but they cannot provide a good overall WWTP estimate as they miss some of the most important sources (as shown by the tracer method). The uncertainty coming from the short campaign period with little information if this is also representative episode for winter, spring etc. should be mentioned here.
This point is valid and we will mention it in the revised text.

P2979 l.23
Moreover, longer measurement campaigns over different times of the year would also allow to catch the variability of the emissions of the site. Finally, if aiming for a general estimate, several WWTP have to be investigated.

Section 6.2 p 2980 l3 onwards – Here we come to a fundamentally wrong way of comparison. To compare the results of this study a comparison should be made to emission factors (EFs) not inventories like EDGAR. EDGAR is a combination of statistical data and EFs on a national scale and then subsequently spatially distributed using certain maps, e.g with location of WWTPs all over the country. The use of EDGAR for checking an individual grid cell is irrelevant. This also applies to the

comparison of line 25 onwards. The reader and /or other scientists learn nothing from a conclusion that for a particular pixel a global emission inventory presents a too high figure. In l20 it says “18 higher” – I assume this is about being a factor 18 higher, so the word factor should be added.
What is relevant than? The comparison of emission factors used in inventories like CITEPA, EDGAR, IPCC. This can be done using your data. First in section 2 you provide key data for Valence: “The station is managed by Veolia France and treats the water for 150 000 inhabitant-equivalents, which represents about 2800m3 h-1 with an exiting BOD of 35 kgm-3 . This can be compared with the Table 6.4 of http://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/5_Volume5/V5_6_Ch6_Wastewater.pdf Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories. From this you can derive whether the Valence station is exceptional or average. If you want to make a comparison with EDGAR (although this reviewer thinks it is pointless) than you would have to make a comparison at the activity data (statistics) level and the EF level. Next the IPCC guidelines say “For domestic wastewater, inventory compilers can compare country-specific values for Bo with the IPCC default value (0.25 kg CH4/kg COD or 0.6 kg CH4/kg BOD). When you take your data you can compare with this IPCC default value and comment on a possibly lower Emission factor being derived for a French WWTP [However, I did not do the comparisons because I am uncertain if the units in Section 2 are all correct.] EDGAR most likely uses the IPCC default factor but again I did not check this – You should be able to find this in the EDGAR documentation. It can also be compared with the EFs that CITEPA uses in their national inventory. This is the relevant discussion and may possibly lead to an argument that an representative EF for France should be derived from new measurements. Again the EDGAR pixel being too high or low will be related to choice of emission factors (+ possibly some redistribution errors of calculated emissions on the national scale). Moreover, the Valence WWTP or Valence city emissions of CH4 are irrelevant at the regional, national or larger scale but the possible implication for revising emission factors are relevant at all these scales. Hence, this is what the wider public and community should be informed about.
So, to conclude this reviewer would like to see a comparison with default emission factors as used by IPCC and possibly CITEPA (EDGAR being optional) and some indicative conclusion from this comparison. Of course the short episode of measuring will prohibit directly replacing EFs in guidebooks but at least it could be a recommendation to repeat such measurements, especially at the plant scale (tracer methods) if the gap appears wide between measured EF and default EF values. Although this change to comparing EFs instead of inventory grid cells is, in my opinion, critical, I have advised “accepted with minor revisions” because I think the authors can easily make this adjustment. I trust this will be done accordingly.

Thank you for taking the time to detail your suggestion. We will revise as suggested in the text using emission factors.

P2980

We can also compare this estimate with inventories estimates. The European Database for Global Atmospheric Research (EDGAR, OLivier et al., 1996) ,which provides gridded maps, as well as the CITEPA , which is responsible for the French inventory, use the IPCC methodology to estimate CH4 emission factors from WWTPs.

EF=BOD.365.Bo.Σ(WSx.MCFx)
with Bo the maximum CH4 production capacity, WSx, the percentage of a certain process used in the WWTP and MCFx the conversion rate of this process.

By hypothesis, the conversion rate for a WWTP like Valence (aerobic treatment) should be at the maximum 0.1 with a maximum of 0.4 if not well managed (IPCC chapter 6 table 6.3). The Bo is usually estimated to be 0.6kgCH4/kgBOD. CITEPA estimate for France an average emission factor of 74g/yr/inhabitant which is very close to the Valence estimate.
Using the data from the WWTP, we can recalculate the conversion rate and compare it to the expected value. Here, we find a conversion rate of 0.07 which is in the expected range.
From these first measurements, it seems then that the Valence WWTP is an average French WWTP in term of CH4 emissions.