Comments to “Measurements of carbonyl compounds around the Arabian Peninsula indicate large missing sources of acetaldehyde” by Wang et al.

Overview:

The manuscript presented ship-borne measurements of carbonyl compounds (carbonyls) around the Arabian Peninsula during the AQABA campaign (end of June to end of August 2017) and interpreted the measurement with a chemical transport model (EMAC). Analyses were focused on 3 classified carbonyl groups—aliphatic ($C_nH_{2n}O$), aromatic ($C_nH_{2n-8}O$), and unsaturated ($C_nH_{2n-2}O$)—in terms of their regional distribution, source characterization and partitioning, model-measurement comparison in 8 geographical regions, and finally an in-depth discussion on the missing sources of acetaldehyde (C2 aliphatic carbonyl).

The manuscript is well organized with clear methods, thorough analyses, and coherent arguments. The topic is of great importance given that carbonyls can provide key clues on reactive carbon source and chemistry as an oxygenated and stable group of VOCs, and acetaldehyde has drawn increased attention recently with growing evidence of its large missing sources. This manuscript therefore provided valuable information in advancing the knowledge of atmospheric reactive carbon in the Middle East. I support the publication of the manuscript once the following issues being addressed.

Scientific comments:

1) The title is interesting but over-weights the topic towards acetaldehyde and meanwhile lacks a reflection of the large portion of work done on the overview analysis of different carbonyls in this region as presented in the manuscript.
2) L22. “A smaller but still significant discrepancy …”, are you trying to say that the model-measurement discrepancy of other carbonyls is smaller comparing to that of acetaldehyde? This sentence is not clear and needs to be reorganized.
3) L130. How did you deal with below LOD data?
4) L134-140. There are other measurement techniques that can differentiate ketone and aldehyde at a shared mass, better to give an average or a range of in-situ reported ketone:aldehyde ratio (e.g., acetone:propanal ratio) to prove the dominance of ketone.
5) Figure2. The description of air mass characteristics is not clear. For example, what’s the difference between “source nearby” vs “source nearby, aged”, “remote, aged” vs “remote, clean”?
6) L246-247. “Interestingly, the mean acetaldehyde mixing ratio (0.62 ± 0.59 ppb) over Suez was twice the level found over the Mediterranean Sea, whilst the acetone level was only slightly higher”, why is this interesting? What are you trying to imply?
7) L247-248. “Besides the local-scale emission and photochemical production contribution to the acetone over Suez, the longer lived acetone could be also transported from the Mediterranean Sea (where acetone was high)”. Acetone is quite similar in Mediterranean and Suez. “where acetone was high” seems to indicate that acetone source in Suez is dominated by transport rather than local production. We do not know the contribution of transport vs. local source to acetone in Suez region, so need to provide evidence or reorganize the sentence.

8) L249-250. “Although the mean mixing ratios of aliphatic carbonyls over Suez were lower than those over the Arabian Gulf, larger variations were observed.” I’m not seeing the higher variation over Suez than Arabian Gulf.
9) L254-258. The arguments here are based on back trajectory. How about the proposed deep sea local source of ethane and propane in Bourtsoukidis et al. 2020? This was found not sufficient to explain the model bias of acetaldehyde later in the paper, but it worth being brought up here as one potential explanation of high carbonyls in the Red Sea North and also briefly mention what you found later in the model …
10) Table 2. Are you summarizing only surface or tower measurement? Need to clarify.
    Too few examples of Oil & Gas region and Forest, add more.
11) L291, the title of Section 3.2 “Chemistry of aliphatic carbonyls”, does not reflect the source characterization discussed in the two case studies in Section 3.2.2.
12) L303, Eq. (2). Varying formats of this equation have been widely applied but I have one question regarding its application in secondary OVOCs from aged air. Eq. (2) assumes that both precursor hydrocarbon and aliphatic carbonyls experienced the same degree of OH exposure (or Δt). This assumption is reasonable for fresh air mass, but might not be valid for aged air, in which carbonyls kept being produced along the way (depending on its precursor’s lifetime) and thus experienced different degrees of OH exposure. So, for long-transported air masses, this could lead to an overestimated OH exposure, an underestimated carbonyl secondary production, and thus an underestimated fractional contribution from oxidation to the measured carbonyl mixing ratio.

    Would this uncertainty affect your conclusion (P12, L334-336; P22, L615-618) that hydrocarbon oxidation plays a more important role in polluted regions than other less-polluted regions, given the potential low bias in the estimated contribution for less-polluted/more-remote regions?

13) L308-312. Table S4. How is the yield being derived from MCM?
14) L318. Did you use 0-5AM data as in Yuan et al. 2012? Clarify that.

    L319. How do your calculated emission ratios compare to literature reported values from similar sources?

15) L329, “In general, the direct oxidation fraction varied from area to area for C1 to C3 carbonyls (formaldehyde, acetaldehyde and acetone).” Give the range of fractions for each species.
16) L414, Section 3.3 title “air chemistry of unsaturated carbonyls” does not reflect the discussion in this section, which is more focused on potential precursor/source of large unsaturated carbonyls.
17) L470-471. “A strong natural non methane hydrocarbon source from deep water in the Northern Red Sea was implemented in the model (Bourtsoukidis et al., 2020).” This needs to be added in the method section where you describe the model configuration. There you have stated that the configuration is following Bourtsoukidis et al. 2020, but also worth pointing out that this newly proposed source is applied.
18) L520. “the air-sea submodel (Pozzer et al., 2006) was augmented to include …” what is the air-sea submodel? Is that a model coupled into the EMAC model? Or it’s a separate model you used to get the ocean acetaldehyde concentration field? Need to clarify. And if it’s coupled to EMAC, need to add in the EMAC model configuration section.

19) P21, Section 3.5.3, “Anthropogenic primary sources”. Comparing to Section 3.5.1 and 3.5.2, this section is lack of analysis-based discussion.

    Also, if the revised version kept the same length or longer, better to separate into two paragraphs, with the first one discussing the underestimated anthropogenic primary emissions of acetaldehyde precursors (L550-558) and the second talking about underestimated anthropogenic primary emissions of acetaldehyde itself (L558-end).

    L551. Does the carbon number of unmeasured hydrocarbons start from C12? Figure 3 and 6 only considered HCs up to C8.

    L552-555. “Bourtsoukidis et al. (2020) compared measured hydrocarbons (ethane, propane, ethene etc.) with the results from model simulations (the same model used in this study) and periodically found significant model underestimation in both regions. This indicates that not all sources were present in the model’s emission inventory”. This seems to argue that the low bias of ethane, propane, ethene, etc. in the model inventory is one reason for the model underestimation of acetaldehyde, which contradicts the argument (P13, L340-349) based on Figure 3. The argument of non-measured and non-modeled HCs before this sentence is fine, but this sentence makes it unclear.

    L555-550. 1) need to say that the high ozone in case studies was observed during “nighttime”. 2) Does the modeled occasional underestimate of ethene persist in nighttime too?

    L565-567. “Therefore, the active petroleum industry located in the Arabian Gulf and intensive marine transportation in Suez are likely primary sources of acetaldehyde and other carbonyls which were not well constrained in the model.” This is not convincing.

20) L587, “Another acetaldehyde formation pathway …” suggest starting a new paragraph to discuss pyruvic acid.

    L590-592. “Zhou and Mopper (1997) pointed out that the net exchange direction for pyruvic acid is expected to be from the air to the sea due to its high partition coefficient (high solubility). Therefore, only low levels of pyruvic acid would be expected in the remote marine boundary layer”. High air-to-sea partition tendency does not guarantee low level of pyruvic acid in the air-side, if gas-liquid partitioning is the only limiting process, its concentration in the air also depends on its degree of saturation in the surface seawater. A review of surface seawater concentration of pyruvic acid and an air-sea equilibrium calculation might be helpful.

    L597-598. “As the air-sea exchange of pyruvic acid is limited, low levels of pyruvic acid were expected” Not clear what this means. How did you know that the air-sea exchange of pyruvic acid is limited? And why would that result in low levels of pyruvic acid?

L598-599. “Even if we fully assign the m/z 89.024 to pyruvic acid, the contribution to acetaldehyde via photolysis of pyruvic acid is negligible compared other sources” Clarify why it is negligible, using 58 ppt and a 100% yield? Is this contribution detectable by intrument?

Minor/technical comments:

L17: typo “3,8”

L89: section numbering error, should start with 2.1

L131 and Table S1: 3σ in text but 1σ in Table S1

L135: add “of” between “number” and “three”

L156: “interaction with ocean/land” includes “dry deposition”

Table 1. L193, L207. “Aliphatic CCs, Aromatic CCs, Unsaturated CCS”, “CC” hasn’t been defined.

Table 1. L193, L207. Extend the name of “S” to “Suez”

L348. “section” to “sections”

Figures 4, 5, and 7. Change x-axis label from “Dateandtime” to “Dateandtime (UTC)”, or clarity in the figure caption.

Figures 4 and 5. missing units for several y-axis, “J_NO2”, “[OH]t”, “Wind_D”

Figure 7a left panel. “S” to “Suez”

Figure S6. Only 6 regions shown in the figure.