Review of “Understanding the vertical temperature structure of recent record-shattering heatwaves” by B. Hotz, L. Papritz and M. Röthlisberger submitted to Weather and Climate Dynamics

General comment:

In this paper, the authors analyse the vertical temperature structure of three record-breaking heatwaves. In a first step, they use the diagnostics developed by Zhang and Boos (2023) to asses to what extent convective stability/instability played a role in determining the magnitude and duration of the heatwaves. In a second step, the authors then perform a detailed Lagrangian analysis based on the diagnostics developed in Röthlisberger and Papritz (2023). They quantify to what extend horizontal advection, vertical advection, and diabatic heating contributed to the specific events. By doing so, they go beyond many other Lagrangian-based studies in that they look at the entire vertical structure and not just the near ground. They find that the contribution from the individual processes varies significantly across the troposphere, with horizontal advection generally being a key process for establishing positive temperature anomalies in the mid to upper troposphere and adiabatic and diabatic warming in the lower troposphere; whereby large differences between the events have also occurred. Many of the results are consistent with the existing literature; in some cases the authors find contradictions. For sure, the analysis will stimulate future work in understanding the formation mechanisms of heatwaves.

The manuscript is well-thought, well-written, and really worth reading. I especially enjoyed those parts that dealt with the Lagrangian analysis, which corresponds to the better part of the manuscript. Generally, I‘m a fan of the Röthlisberger-Papritz-diagnostics and I‘m convinced that its application brings us a good deal closer to understanding the underlying mechanisms in the development of temperature anomalies. Unfortunately, I had difficulties to follow the reasoning in Section 3.1, which deals with the role of convective instability, and how it relates to the (results of the) Lagrangian analysis.

Below I have compiled a list of questions and comments, and I am eager to hear the authors' responses. If properly revised, I find the manuscript well suited for publication.

Major comments:

Section 3.1: I have some difficulty following your reasoning in this section. My main problem is that I thought Zhang and Boos (2023) were arguing that a heat wave persists as long as the atmosphere is stable for moist convection, and that once convective instability is reached, precipitation sets in, ending the heat wave. However, for the PNW and the RU heatwaves you do see convective instability, but this does not end the heatwaves. It might help my understanding if you could describe what the plots in Figure 2 should look like to conclude that convective instability did not play a significant role in the heatwaves. Also, addressing my questions about L173-175 and L176/176 will certainly be helpful as well.

L173-175: Here you write “Based on MSE*500-MSEs we find that during the PNW and RU heatwaves, the atmospheric vertical structure reached neutrality to moist convection during their peek phases. Hereby, negative values of MSE*500-MSEs are somewhat surprising.” But I thought you were inferring from the negative values of the MSE*500 MSEs that the vertical structure achieved neutrality. So why are the negative values surprising if the atmospheric vertical structure has reached neutrality? My understanding is that it would be better if you wrote, e.g., "During the PNW and RU heatwaves, we find negative values of MSE*500 MSEs during their peak phases, indicating that the atmospheric vertical structure has achieved neutrality to moist convection at that time. This is somewhat surprising."

L176/176: You write that one should “note that […] the near-surface temperatures peaked when negative […] values started to appear”. But I cannot see any line in Fig. 2 indicating temperature. I think this is also the reason why it‘s pretty hard for me to follow your reasoning in the whole section.

L266-270: Here, it‘s not clear to me why the pronounced diurnal cycle of the PBL height in combination with the constant vertical extent of the lower tropospehereic diabatic T‘ shows that there is heat accumulation in the PBL. Could you explain this a bit more? E. g. What would the PBL height and diabatic T' look like if there had been no heat accumulation?

L279-281: As mentioned earlier, I‘m not entirely convinced that you see evidence for a top-down control of T‘ via convective stability/instability, nor for the multi-day heat accumulation. But this probably changes if you clarify this a bit more in the earlier paragraphs.

Minor comments:

Title: My feeling is that “understanding” is a very strong word that means to know the causes/reasons why things are the way they are. With your diagnostics, you can attribute a certain process to a certain T‘ and this is really great. But, is it enough to really understand the vertical temperature structure, i. e. to give a cause for it? In light of that (and your comment in L455458), you might consider rephrasing the title a bit, e.g. into “Towards an understanding of the vertical temperature structure of ...”. Or: “Disentangling the vertical temperature structure of ...”.

L23/24: I think what you do is that you analyse the large-scale dynamics (advection, subsidence, WCB), the moist convection (stability/instability), and a bit the boundary layer processes (PBL height), but at the moment I do not see that you really discuss their “interplay”. If you think so too, consider rephrasing this sentence.

L51: Which events do you mean with “events mentioned above”? The PNW, RU, and UK heatwaves or the “benign” heat waves from the previous paragraph?

L91/125: Here, I’m interested in a technical detail. If you compute the mean T on modellevels, what do you then take as Δp in the computation of the vertical gradient? E. g. do you compute a climatological mean pressure for each modellevel at each grid point, from which you then infer the Δp between each modellevel at each grid point? Or do you take the instantaneous values of p?

L93: I do not fully understand what you mean by “9-year windows centered on the time step of interest”. Especially I was wondering how you handle dates at the edges of the ERA5 timeseries when there are no 4 years left, e. g., dates in the year 2021 or 2022. Could you clarify this?

L121: Can the LAGRANTO trajectories account for turbulent mixing in the PBL?

L125: I was just wondering why you trace specific humidity q. I suppose just in case it‘s needed at some point? If so, you could drop q in this listing here to prevent confusion.

L160/161: Do the regions and periods of the other two heatwaves also correspond to the event definitions in other studies?

L166/167: This goes back to my previous comment. You write that your results are “insensitive to horizontal shifts of the heatwave regions by a few degrees, enlarging or shrinking the regions by a few degrees, and shifts of the heatwave periods by 1-2 days.” For the PNW and the UK heatwaves, I can well imagine that this ist true. However, for the RU heatwave, which was quite extensive spatially and lasted over a month in total, your results may not be representative of the entire heat wave period/region. So it could well be that a shift of a few days doesn't matter, but a shift of, say, 10 days might actually matter and your results would have looked different if, e. g., you had chosen a period closer to the onset phase of the heatwave. I think you should mention this a bit more prominently (see also my comment to L347).

L206: In Fig. 3b, I can't see that the air parcels first ascended and then descended. I think this is because the trajectories overlap and cover each other and only a small part of the colorbar is used. I think the figure would improve if you adjusted the colorbar accordingly and made the trajectories a little transparent. Maybe it would also be a good idea to take trajectories that end up more apart. The same suggestions apply to Fig. 6 and 9.

L254: Determining the Lagrangian age and formation distances would be possible without doing the full T‘ decomposition. Thus, reformulating this sentence into e.g. “By determining the Lagrangian age and formation distances, the temporal and spatial scales over which the temperature anomalies form can be quantified.” might be more precise.

L283 (and L21 and L453): Here you argue that the aging of T' suggests the concept of a "heat dome" in which "air recirculates and accumulates heat". I agree with you that a heat dome would have exactly these properties, i.e., aging of T', and that it is plausible to observe this during the PNW heatwave. However, the aging of T' is not conclusive evidence that the air masses are actually recirculating. What can "only" be seen is that the air masses that happen to arrive in the heat wave region tend to be older (in terms of temperature anomaly lifetime) than before. But you can't be sure if it's always the same air masses you observe, or if it's new air masses that happen to have an older lifetime. If you agree, I would appreciate a comment on this when you make statements about the "heat dome".

Fig. 5/8/11c: Never mentioned/discussed in the text.

L347: The phrase "clearly ageing" confuses me a bit, since T' only ages in the period you defined as the RU heatwave. In fact, if you also took into account the days before July 31 (where it was already very hot), you could actually see deageing between July 30 and July 31.

L455-463: I think that is a very wise comment and it provides, in my opinion, a major reason why the research regarding heatwaves cannot be finished at this point!

Technical corrections:

L6: Replace “Western Russia” by “western Russian” as in L26 (or in L26 “western Russian” by “Western Russia”)?

L60: Not totally sure, but I think you should replace “the stability of the atmospheric profile to moist convection” by “the stability of the atmosphere to moist convection” (since it‘s not the profile, which is stable/ unstable to convection, but the atmosphere itself). Same in L98 and 102.

L86: “the” in front of “Lagrangian analyses”?

L91: I find the word “transient” a bit confusing here. I would just drop this part of the sentence, since the following part of the sentence makes clear what you mean.

Fig 2/5/8/11: It‘s difficult to identify which of the date label corresponds to which tick on the x-axis. I suggest to rotate the date labels by 45°, such that they are aligned vertically.

L181: “19 July” instead of “20 July”?

L198: I suggest putting “(visible in the top left of Fig. 3a)” right after “was an upstream cyclone” (since it‘s the cyclone that is visible in the figure and not how it deepened rapidly).

Fig 3/6/9 (left column): The black and purple lines are really hard to see. Maybe you show less, but thicker lines to improve the visibility? Furthermore, I think the labeling with numbers is not needed at this point and dropping them may improves the visibility of the figure as well.

Fig 3 (caption): L1:Change “(i) the T‘ and its contribution” into “(i) the near-surface T‘ and its decomposition”?; L4: Drop “dashed” (or make the rectangle dashed in the figure); L7/8: What do you mean with “maximum 5-day daily T‘” (same in L295)? I thought the gray lile denoted the period that you identified as your heatwave period. And I suggest putting the grey line all the way to the front (and not hiding it behind the light blue).

L225: “Causes” is a quite strong word. Maybe change “physical causes” into e.g. “physical (or underlying) processes”?

L232: “700” instead of “600”?

L241: Comma in front of Fig. S2?