

Overall:

This manuscript introduces a data set associated with a dense network of weather stations making observations during the June – August 2020 FESST@HH field experiment. The goal of the field experiment was to shed light on sub-mesoscale features of convective cold pools, and the data also seem to produce interesting insights about urban climate effects (although this is not my area of expertise). Despite the excessive length of my review, I should emphasize that the manuscript and data set are generally well constructed. However, the data set and associated manuscript contain several limitations that I would like to see addressed before I can recommend acceptance.

My review is organized into a section on the manuscript and a section on the data set. Each section is organized primarily chronologically. In the manuscript section, I have separated the substantive questions and comments from the minor technical corrections. The data set section is organized roughly into requests for additional data and comments on existing data. I have also uploaded some Matlab code and results associated with sample analyses that I completed.

Manuscript

Substantive Comments on the Manuscript:

Line 18: Please explain the full field experiment acronym. As I read the manuscript, I found myself wondering why *H* appeared in the acronym twice. I did not realize until I saw the 2 May 2022 version of the data set that “@HH” meant “at Home, Hamburg” (or something similar, I think).

Line 33: All of these references examine oceanic convection exclusively. Drager et al. (2020) discuss these triggering mechanisms in a continental environment, and I recommend citing this article here.

Drager, A. J., Grant, L. D., & van den Heever, S. C. (2020). Cold pool responses to changes in soil moisture. *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001922.

<https://doi.org/10.1029/2019MS001922>

Line 34: There are many earlier references regarding cold pool influences on the organization of convective clouds that should be cited here. These include:

Purdom, J. F. W. (1976). Some uses of high-resolution GOES imagery in the mesoscale forecasting of convection and its behavior, *Mon. Weather Rev.*, 104(12), 1474–1483, [https://doi.org/10.1175/1520-0493\(1976\)104%3C1474:SUOHRG%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(1976)104%3C1474:SUOHRG%3E2.0.CO;2)

Purdom, J. F. W. (1982). Subjective interpretation of geostationary satellite data for nowcasting. In K. Browning (Ed.), *Nowcasting* (pp. 149–166). London: Academic Press.

Rotunno, R., Klemp, J. B., & Weisman, M. L. (1988). A theory for strong, long-lived squall lines. *Journal of the Atmospheric Sciences*, 45(3), 463–485. [https://doi.org/10.1175/1520-0469\(1988\)045%3C0463:ATFSL%3E2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045%3C0463:ATFSL%3E2.0.CO;2)

Wilson, J. W., & Schreiber, W. E. (1986). Initiation of convective storms at radar-observed boundary-layer convergence lines. *Monthly Weather Review*, 114(12), 2516–2536. [https://doi.org/10.1175/1520-0493\(1986\)114<2516:IOCSAR>2.0.CO;2](https://doi.org/10.1175/1520-0493(1986)114<2516:IOCSAR>2.0.CO;2)

Line 38: I’m not sure what “Daw” refers to (there is no year provided, and there is no corresponding entry in the References list), and in any case, some earlier references should be provided here since Li et al. (2015) was certainly not the first to discuss this topic. A couple of suggested references are below.

Deardorff, J. W. (1980). Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-Layer Meteorology*, 18(4), 495–527. <https://doi.org/10.1007/BF00119502>

Smagorinsky, J. (1963), General circulation experiments with the primitive equations. I. The basic experiment, *Mon. Weather Rev.*, 91(3), 99–164, [https://doi.org/10.1175/1520-0493\(1963\)091<0099:GCEWTP>2.3.CO;2](https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2)

Lines 45-48: As someone who participated in the C³LOUD-Ex field campaign, I take issue with some aspects of the way it is characterized here. Firstly, “convective updrafts and downdrafts” is not quite correct. We were interested in characterizing convective updrafts and cold pools (the “O” in C³LOUD-Ex stands for “Outflows,” i.e., cold pools). Moreover, although there were indeed only three surface stations, the surface stations were portable so as to allow for targeted observations of cold pools. The line of surface stations (with quasi-stationary UAS hovering above each surface station, along with radiosonde launches, arranged in a coplanar “flying curtain”) was always oriented parallel to the advancing gust front, such that a time-to-space conversion would be able to characterize the three-dimensional structure of the cold pool as the gust front passed through the flying curtain. There certainly exist limitations to this approach, and I do not doubt that more surface stations would have been useful. However, the comparison of the targeted observations of C³LOUD-Ex to the stationary surface station network of FESST@HH is not quite an apples-to-apples comparison. I would request the following rewording in order to clarify more precisely the contrast between C³LOUD-Ex’s scope and FESST@HH’s scope:

In a recent study, van den Heever et al. (2021) described the use of uncrewed aerial systems, regional radars, radiosondes, and surface stations during the C³LOUD-Ex campaign to characterize convective updrafts and cold pools on scales between $O(100)$ m and $O(1)$ km. Their portable “Flying Curtain” sensing network collected in-situ observations of cold pools over three closely-spaced geographical locations and revealed that cold pool temperatures can

exhibit variability on spatial scales of $O(100)$ m as well as $O(1)$ km. However, given the small geographical extent of the in-situ network, C³LOUD-Ex was unable to characterize the complete surface temperature structure of any individual cold pool.

Line 58: In line 246, you mention that there were 37 cold pool events during the three-month measurement period, which translates to ~ 12 cold pool events per month. This seems to be significantly higher than the expected amount (7 cold pool events per month). Was 2020 a lucky year for cold pools, or is the discrepancy due to the greater area covered by the FESST@HH domain, as compared to the single location examined in Kirsch et al., 2021a?

Section 2.1: A major limitation of this data set is the APOLLO stations' lack of humidity and wind measurements. Wind is an essential metric for identifying cold pool boundaries (e.g., Fournier and Haerter, 2019, citation below). Additionally, as is discussed by Kirsch et al., 2021a, there is still much about the near-surface moisture structure of cold pools (e.g., moisture rings) that is poorly understood. Could you discuss why the decision was made not to include these measurements?

Fournier, M. B., & Haerter, J. O. (2019). Tracking the gust fronts of convective cold pools. *Journal of Geophysical Research: Atmospheres*, 124, 11,103 – 11,117.
<https://doi.org/10.1029/2019JD030980>

Sections 2.1 and 2.2: I am wondering what impact the lack of sensor aspiration may have had on the results obtained during the field experiment.

Lines 79-80: A few questions here:

- Is this an e -folding time constant [i.e., time to complete $100 \times (1 - 1/e) \approx 63.2\%$ of the response to a step-function increase or decrease], or is it some other metric?
- What procedures were used in obtaining this estimated time constant? “Wind tunnel laboratory tests” is not a sufficient description.
- How does the response time vary as a function of wind speed?

Line 94: Do you have plans to draft instructions for acquiring the required parts and building the stations? Is the €300 figure still accurate? What are the labor costs for assembling the stations?

Lines 95-97: A few questions here:

- Could you provide more details regarding the comparison with the “inertia-free ultrasonic sensor”? For example, what type of ultrasonic sensor was used?
- What does “almost instantaneously” mean – can you quantify this?
- What does “high-frequency” mean – can you quantify this? I’m not sure whether I should be focusing on the oscillations in Figure 3 that last for a few seconds or those that last for about a minute.

Lines 98-99: A few questions here:

- What does “considerable” mean in this context – can you quantify the impact?
- How was this assessed? I am interested to see the “not shown” data and learn more about how they were collected.
- Was this assessed under a variety of cold pool conditions, including clear versus precipitating conditions?

Section 2.2: What impacts do obstacles and surface properties have on the cold pool signal in the WXT data? This is not discussed.

Line 136: What is the significance of the number following each \pm symbol? Are these standard deviations? The full ranges? Please make this explicit.

Line 148: What is meant by “near” and “large”? Can you either quantify these or provide a map showing these particular bodies of water?

Lines 150-151: Can you provide exact dates? I do not see any data from early September in the data set.

Line 182: Does “midnight” refer to local time or UTC?

Lines 182-183: What is a measurement “telegram”? (Note: This word was also used in Line 88.) I am unfamiliar with the use of this word outside of the context of 19th-century technology (telegraphs, Morse code, etc.). Can this wording be made more straightforward?

Sections 4.1 and 4.2: Can the level 0 and level 1 data be added to the public data set?

Lines 205-207: Two things:

- I am not convinced that 15 K is an appropriate threshold here, given that a cold pool may affect only one station at a given time as it begins to spread across the domain. Cold pool temperature drops of 15 K do not seem implausible.
- What does it mean that these three stations had erroneous temperature sensors? How is this defined, and how was it determined that these three sensors were erroneous?

Lines 213-214: These biases could have changed over the course of the measurement campaign, as it is possible for individual sensors to drift by different amounts over time. When did the week-long calibration occur, relative to the time period of the measurement campaign? Ideally I think this calibration should have been done both before and after the measurement campaign, at the very least. How were the WXT sensors calibrated? What about the pressure sensors in the APOLLO stations? The manuscript does not seem to provide any information in this regard.

Section 5.2: How is a “day” defined/delineated? Is this from midnight to midnight UTC, or is a different definition used (e.g., midnight to midnight local time)?

Lines 247-248: How is “associated with rainfall” determined/defined?

Lines 250-251: How was this information determined? Were radar retrievals used, and if so, can they be shown here?

Lines 254-255: The cold pool appears to extend beyond the eastern edge of the network, so it seems difficult to draw any conclusions regarding the actual size of the cold pool.

Line 269: Is this gradient based on the absolute temperatures or the perturbations? Please clarify.

Lines 273-275: Can you expand upon how the expected range was determined?

Lines 280-282: Other studies have discussed non-hydrostatic pressure effects associated with downdrafts and cold pools, and they should be cited here. Here’s one example, though there are probably some earlier studies as well:

Wu, F., & K. Lombardo (2021). Precipitation enhancements in squall lines moving over mountainous coastal regions. *Journal of the Atmospheric Sciences*, 78(10), 3089–3113.
<https://doi.org/10.1175/JAS-D-20-0222.1>

Lines 289-290: What does it mean for an environment to be “weakly sealed”? How is “sealing” quantified? I am not familiar with this concept, and I expect that many (if not most) readers—at least those whose primary interest is cold pools—will not be familiar with it either. Could you provide a brief definition, or at least cite and refer readers to some sources that explain this concept?

Lines 290-291 and Figure 10: According to <https://data.giss.nasa.gov/modelE/ar5plots/srlocat.html>, the day length in Hamburg (53.55°N, 10.10°E) changes by 3.32 hours over the course of the interval from 15 June to 31 August (maximum of 17.05 hours around 20 June, minimum of 13.73 hours on 31 August). Daily mean insolation is also changing, with the 24-hour average maximizing at 482.27 W/m² on 20 June and minimizing at 334.70 W/m² on 31 August.

With the caveat that I am not an expert on quantifying the urban heat island effect:

Given that the sunset and sunrise timing are changing over the averaging period, I am wondering if it would make more sense to shift and rescale the time axis prior to averaging so that the sunset and sunrise times are aligned. This would preserve the abruptness of any temperature shifts associated with sunset and sunrise, and would minimize “smearing.” Alternatively, since the daily sunset and sunrise times do not change very much near the solstice, limiting the averaging period to 15-30 June, and excluding the months of July and August, could have a similar effect.

Lines 320-321: Can you expand on how these additional experiments were conducted, and how this conclusion was drawn?

Figure 4: The four lowermost white-text labels (“3-m mast” and below) are displaced far to the left of the items they describe. Please either add arrows or move the labels so that it is obvious what they are referring to.

Figure 8: Please plot the WXT stations using a different shape (perhaps a square) so that they can be distinguished more easily from the APOLLO stations.

Table 1: What is the relative accuracy of the pressure sensor for temperatures below 25°C? Please provide this information.

Table 2: Can you provide additional wind speed and direction accuracy estimates for weaker winds, given that winds as strong as 10 m/s seem not seem to occur very often here (per Figure 8)?

Minor Technical Comments:

Spelling, grammar, rewording (including minor alterations in the sentence meaning)

Line 3: sub-mesocale → sub-mesoscale

Lines 9-10: supported the station maintenance → supported station maintenance

Line 10: add a comma after “their measurement characteristics”

Line 12: add a comma after “inside a cold pool”

Line 13: its size → a cold pool’s size

Lines 14-15: as an expression of → associated with

Line 18: Conventional meteorological → Conventional mesoscale meteorological

Line 19: essential to obtain → essential in order to obtain

Line 20: techniques help → techniques can help

Line 21: precipitation or wind speed → precipitation and wind speed

Line 23: The information → Information

Line 25: This fact qualifies them as the → Therefore, cold pools are the

Line 29: several tens of km → hundreds of km (note: squall lines can extend across hundreds of km)

Line 30: temperature falling by more than 10 K → temperature decreases that can exceed 10 K

Line 30-31: propagates away from → propagates horizontally away from

Line 36: Grant and Heever → Grant and van den Heever

Line 42: aircrafts → aircraft

Line 49: the the → the

Line 52: aim → idea

Line 56: cheap → inexpensive (Note: “cheap” can carry connotations of poor quality)

Line 60: allows to study → allows for the assessment of

Line 62: in response to → due to

Line 68: cold pool fronts → cold pool gust fronts
Line 69: had to operate independent of → had to be able to operate independently of
Lines 69-70: to facilitate the search for appropriate site locations → given site location constraints
Line 70: According to → Based on
Lines 71-72: WXT weather stations based → WXT weather stations (see Section 2.2) based
Line 73: APOLLO → APOLLO stations
Line 76: source and → source, and
Line 82: whereas → and
Line 84: written → recorded
Line 87: to correct → correction of
Line 87: drifting → drift
Line 100: WXT weather station → WXT stations
Line 104: direction and precipitation → direction, and precipitation
Line 112: rains drops → rain drops
Line 118: The units of the station are powered by → Each station is powered by
Line 120: sun light → sunlight
Line 126: in response to → associated with
Line 128: home office → pandemic-related
Line 133: randomly → non-uniformly
Line 133: the tendency for → generally
Line 134: The random arrangement results → The arrangement of stations results
Line 139: altitude of all measurement site lies between → altitudes of all measurement sites lie between
Line 141: impact → impacts
Line 148: 10 → Ten (Note: this may depend on ESSD's editorial style)
Line 152: permissions at short notice for using → permission at short notice to use
Line 165: home office not only affected → pandemic-related restrictions affected not only
Line 169: the data and → the data, and
Line 169: public ground → public grounds
Line 175: In contrast → By contrast
Line 190: regular → standardized
Line 206: sensor → sensors
Line 207: applied and → applied, and
Line 210: near by → nearby
Line 212: sensor; → sensor; and
Line 214: (114UHW); → (114UHW).
Line 218: meta data → metadata
Line 219: above ground and → above ground; and
Line 229: higher stability of the power supply → greater stability of the WXT power supply
Line 229: generally higher availability → generally greater availability
Line 235: more than → maximum daily temperatures exceeding
Line 236: by a phase of relatively cold temperatures with → by relatively cold temperatures, with

Line 237: Rainfall → Measurable rainfall

Line 238: on half of → on approximately half of **OR** on exactly half of (Note: exactly half would be 46 days)

Lines 239-240: June when a strong convective event occurred → June in association with a strong convective event

Line 244: further → additional

Lines 251-252: surface-layer air masses → surface-layer air

Lines 254-255: cold pool deepened to → cold pool temperature perturbation strengthened to

Line 258: early stages of its life → early stages of the cold pool's life

Line 259: cold air masses during → cold-air region during

Line 262: cold pool origin experienced → cold pool center experienced

Line 263: reached its maximum → reached a maximum

Line 266-267: the maximum cooling decreases to about one-third away from the center → the maximum amount of cooling decreases by about two-thirds away from the center

Line 270: (2021) who → (2021), who

Line 270: cold pool properties on scales between $O(100)$ m and $O(1)$ km → cold pool temperatures of order 1–2 K on scales between $O(100)$ m and $O(1)$ km

Line 284: effects as → effects, as

Line 285: surrounding → surroundings

Line 305: by the local → by local

Line 306: condition like → conditions like

Line 312: effectively reduced → effectively reduces

Line 334: their ground as → their grounds as

Lines 335-336: they were never used → they had never been used

Line 340: urban meteorology as the → urban meteorology, as the

Line 347: glass for sub-mesoscale → glass for revealing sub-mesoscale

Line 354: prove of concept → proof of concept

Line 362: analyses helped to → analyses. FA also helped to

Line 363: conflict of interests → conflicts of interest

Line 388: this entry is missing the author(s) and year

Line 416: and Heever, S. C. v. d. → and van den Heever, S. C.

Line 416: Geohphys. → Geophys.

Line 453: Szoeki, S. P. D. → de Szoeki, S. P.

Figure 3 caption: by co-located → by a co-located

Figure 5 caption: sub-class is considered → sub-class is depicted

Figure 7 caption: 5%-quantile and 95%-quantile → 5th percentile and 95th percentile

Figure 8 caption: perturbation smaller than → perturbation stronger than

Figure 8 caption: Length of reference arrow refers to → The length of the reference arrow corresponds to

Table 3 caption: Meta data of → Information about

Data Set

Requested Additions to the Data Set:

1. The data set would benefit greatly from the addition of appropriate radar and/or geostationary satellite data. This would enable users to contextualize the cold pool observations relative to the locations of rain shafts and clouds, without having to track down these data elsewhere. If there are licensing or other restrictions preventing the inclusion of such data, then it would be helpful if an instructional guide for locating and downloading such data could be included.
2. Please add a list of the 37 cold pool events, either as a table in the manuscript or as part of the data set.
3. Please add some derived variables to the data set, such as:
 - a. u- and v-components of the horizontal wind at WXT stations
 - b. water vapor mixing ratio and/or specific humidity at WXT stations
 - c. potential temperature at APOLLO and WXT stations
 - d. equivalent potential temperature at WXT stations
 - e. virtual potential temperature at WXT stations

Specific Comments on the Data Set

Note: Except where stated otherwise, I analyzed the following version of the data set:

<https://www.fdr.uni-hamburg.de/record/8973>

4. When I unzipped the main data files using the Finder app on my MacBook (as opposed to, e.g., using a command-line interface), the default names of the directories were not helpful. For instance, for the version of the data set uploaded 2 May 2022, the files “fessthh_uhh_apollo_l2_202006.zip” and “fessthh_uhh_wxt_l2_202006.zip” both unzip to create directories named “fesst-at-home”. I realize that these can be altered quite easily; it’s just a minor inconvenience.
5. ~~readme_fessthh_data.txt: Typo (“humuidity”)~~ → This seems to have been fixed in the version of the data set uploaded 2 May 2022. (However, the “Last updated: 17 March 2021” line at the bottom of the document was not updated!)
6. ~~Naming convention for station images: SO should be SE (Southeast instead of Süden Øst)~~ → This seems to have been fixed in the version of the data set uploaded 2 May 2022.
7. The number of seconds since 1 January 1970 at 00:00 UTC does not appear take leap seconds into consideration. This causes an offset of 27 seconds when I analyze the data using Matlab’s datetime capabilities.

8. The long name and comment for zsl seem to be contradictory. Why not just say “altitude of ground level above mean sea level” or similar?

Details:

From ncinfo (in Matlab) for fessthh_uhh_apollo_12_ta_v00_20200810000000.nc:

```
zsl
  Size:      82x1
  Dimensions: station
  Datatype:  single
  Attributes:
    standard_name = 'altitude'
    long_name     = 'altitude of instrument location above mean sea level'
    comment       = 'altitude of ground level'
    units         = 'm'
```

9. Wind direction: How are the degree values handled when wind is from the north, oscillating between 0° and 360°, given that these are averages of the 4-Hz data? Are there any cases where combinations of slightly east of north and slightly west of north average out to 180° (for example)? Or is this done by averaging the u- and v-components over time?
10. Wind gust: I am concerned because of the large gust length (3 seconds) relative to the time interval (10 seconds). What if there’s a stronger gust that straddles two time intervals? I think it might make more sense to calculate such wind gusts on a rolling basis or something like that.
11. Number of characters allocated for station ID, LCZ, etc., as well as rows vs. columns:

Why are 40 characters allocated for each station ID, when each station ID is only 6 characters long? Also, the names are arranged in columns, rather than in rows, which makes them difficult to read in their raw format. This is what it looks like when I read the data into Matlab:

From ncinfo for fessthh_uhh_apollo_12_pa_v00_20200810000000.nc:

```
station_id
  Size:      40x82
  Dimensions: character,station
  Datatype:  char
  Attributes:
    long_name = 'station identifier code'
```


I obtained a nearly identical list for temperature data as for pressure data, so I did not include both lists here.

I was curious to see whether anything weird happened when data were not missing, so I plotted 100-second time series of the temperature and pressure centered at 31-Aug-2020 13:00:02, 14:00:02, 15:00:02, 17:00:02, and 18:00:02, and I did not see any anomalous signal in pressure or temperature at the XX:00:02 time stamp.

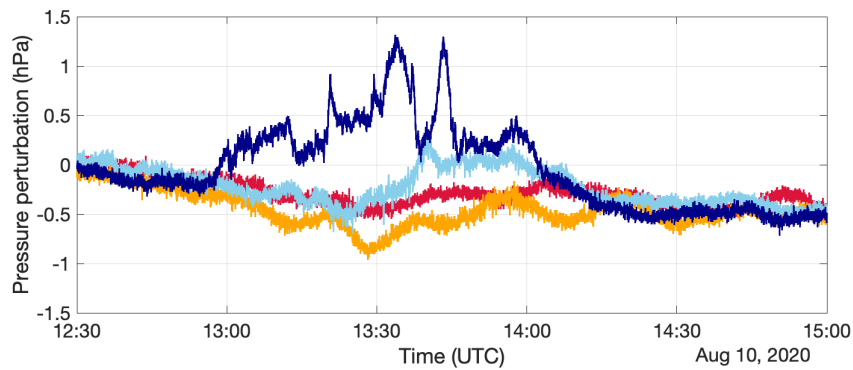
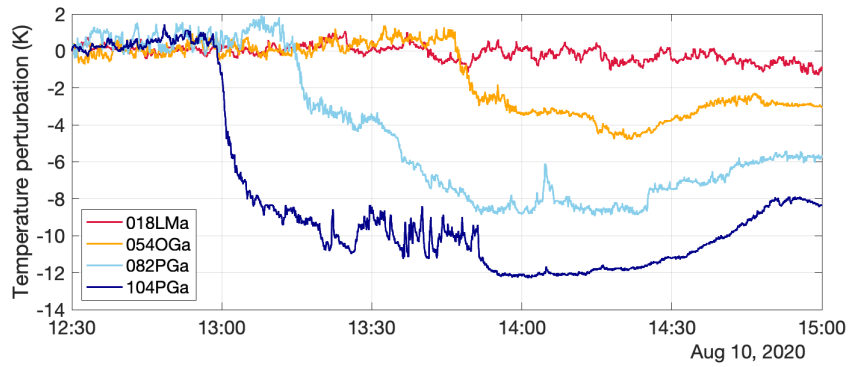
This issue may exist at other stations, and I will ask the authors to investigate how pervasive the issue is, as well as the potential cause of the missing data. Since the GPS time is logged once per hour, perhaps there is sometimes sufficient clock drift to result in an immediate jump from XX:00:00 to XX:00:02 upon GPS synchronization, resulting in missing data at XX:00:01.

Although this issue seems to result in a very small proportion of missing data, it is inconvenient that the data are missing very close to the XX:00:00 time stamps, as users are more likely to sample the data at/near the “clean” XX:00:00 time stamp than at other, more “random-looking” time stamps like (perhaps) XX:52:23.

I might suggest creating (and documenting) an alternative version of this data set with any one-second gaps filled in with the average of the surrounding values, if this is considered permissible and ethical. Either way, this issue should be documented more fully so that users are aware.

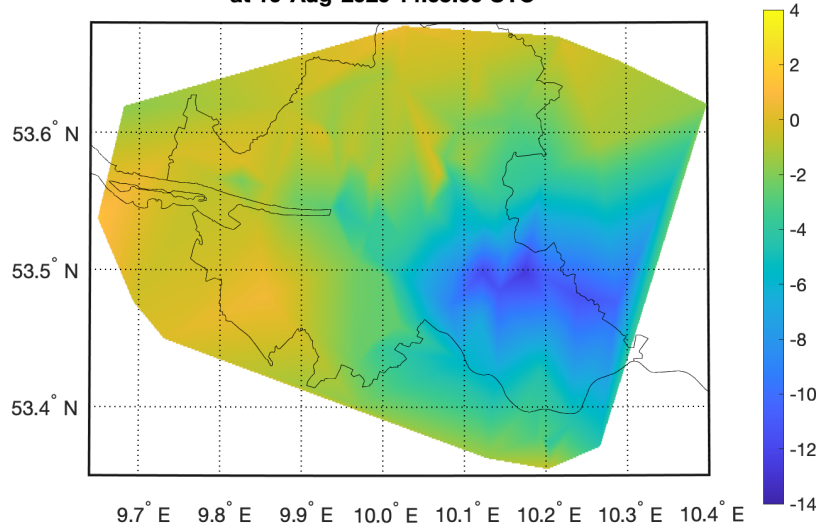
My own exploration of the data using Matlab R2021a:

- I was able to reproduce Figure 9 using the data set:



- I interpolated data onto a Cartesian grid and plotted the corresponding maps.

**Temperature perturbation (K) relative to 10-Aug-2020 12:30:00 UTC
at 10-Aug-2020 14:05:00 UTC**



The above sample map, which corresponds to Figure 8d in the manuscript, contains some strange spatial oscillations near 53.5°N, 10.15°E. It is unclear to me what is causing these oscillations. It is possible that I have made a mistake somewhere. I have uploaded my code as part of this review, along with the images/animations that were produced for this 10 August 2020 cold pool case-study.