Details of the revisions and responses to **Reviewer 2** comments on the manuscript entitled '*GSDM-WBT: Global station-based daily maximum wet-bulb temperature data for 1981–2020*' (essd-2022-309)

We would like to thank the reviewer for your insightful and constructive comments that help to enhance the overall quality of our manuscript. Our responses (on a comment-by-comment basis) are indicated in blue text, and all updates to the original submission will be highlighted in the revised manuscript.

Comments: I really like what the authors have set out to do, to make a homogenised station-based dataset of T_w with complete temporal coverage. I know that this is a complex task, and they should be commended for doing so. They have validated their product against a reanalysis (which is as independent as one can get for these kinds of observations) and a similar station-based product which uses alternative methods.

My major comment relates to the use of the NCEP DoE reanalysis, and how sensitive this data product is to the choice of this (older) reanalysis. Most other comments suggest improvements for the figures, the readability or other minor clarifications.

Response: Thank you for recognizing our new dataset and providing the comments. All responses to your major and minor comments are as follows.

Major Comment:

Line 95 - The NCEP DOE reanalysis is a relatively old product, and a number of studies have found issues with it. There are more recent reanalyses available (albeit being usually larger, these are more complex to use). How do you think your results depend on the choice of reanalysis?

Response: In our study, the NCEP-DOE reanalysis dataset was introduced to complement series when the values in each time step of all candidate stations were missing. This process was essential but with relative low effect of reanalysis series on the whole results. As shown in the Methods of manuscript, we selected the reanalysis series which had the top 10% correlation coefficients ($p<0.05$) with station-based series to improve the reliability of reference series. We could also investigate the number/percentage of void time steps in all series in different station zones (Table R1), and it was found that the percentages (0.04%-2.59%) relative to 14610 total time steps were low. Therefore, the selection of NCEP-DOE and its complementary series affects the eventual results slightly. We will demonstrate the effect of reanalysis series in the Results in the next version of manuscript.

Table R1. The effect of complementary series in different station zones.

Station zone	Number of complementary series	Number of all stations	Number of void time steps in all stations	Percentage of void time steps (%) in all stations
Z1	1	5	27	0.18
Z2	1	8	8	0.05
Z3	1	9	14	0.10
Z4	1	5	36	0.25
Z5	5	54	12	0.08
Z6	1	9	10	0.07
Z7	9	87	5	0.03
Z8	1	11	8	0.05
Z9	1	12	378	2.59
Z10	45	451	6	0.04
Z11	2	20	8	0.05
Z12	1	12	6	0.04
Z14	3	31	15	0.10
Z15	4	35	12	0.08
Z16	1	8	24	0.16
Z17	1	9	27	0.18
Z18	4	41	7	0.05
Z20	1	5	169	1.16
Z21	2	18	9	0.06
Z22	3	34	13	0.09
Z23	19	187	6	0.04
Z24	4	41	7	0.05
Z27	6	64	12	0.08
Z28	1	5	67	0.46
Z30	6	56	7	0.05
Z31	4	38	35	0.24
Z32	1	12	46	0.31
Z33	1	8	41	0.28
Z34	4	36	5	0.03
Z35	1	7	145	0.99
Z36	1	7	123	0.84
Z37	1	9	21	0.14
Z38	1	8	21	0.14
Z39	2	24	41	0.28
Z40	1	9	32	0.22
Z41	1	11	17	0.12

Besides, there are several reasons why the NCEP-DOE could be as a selection in this study, especially after considering the availability and data volume. Firstly, its temporal resolution is consistent with our criteria of data quality control (i.e., at least one TW every six hours per day).

In addition, although other reanalysis (e.g., ERA5-land hourly data) have higher temporal resolutions, users might be required to compute the input variables of calculating TW by themselves (e.g., calculate the humidity based on the dewpoint temperature). The NCEP-DOE directly provides the 2m air temperature (K), 2m specific humidity (kg/kg) and surface pressure (Pa) to calculate TW, which is in concert with the use of HadISD. In previous research on heat-related events (Mora et al., 2017; Wang et al., 2021), NCEP-DOE was also widely applied.

Minor Comments:

Line 12 - the phrase "and the response on human health" doesnt quite read correctly for me. I see what youre trying to say, but that sentence doesnt flow well.

Response: Thank you for the comment. We rewritten this sentence as "*The daily maximum TW can be effectively used in monitoring humid heatwaves and their effects on health*".

Line 16 - insert "stations" after "These"

Response: Thank you for the suggestion. We have inserted the "stations" after "These".

Line 18 - its not clear at this point what the offsetting mentioned does. Is it possible to clarify a little (I appreciate it is the abstract, so space is limited).

Response: Thank you for the comment. "Offset" here means that using the GSDM-WBT could avoid the underestimation of TW calculated from reanalysis dataset. We changed the sentence as "*GSDM-WBT handles stations with many missing values and possible inhomogeneities, and also avoid the underestimation of the TW calculated from reanalysis data*".

Line 24 - "..resulting in the increase of the frequency...

Response: Thank you for the suggestion. We have changed the phrase.

line 29 - "response to"

Response: Thank you for the suggestion. We have corrected the word.

line 31 - I think it would be important to indicate here that the "relatively low air temperature" refers to heat wave conditions, I initially read this as relatively low absolute air temperatures (which for me is 0-5C).

Response: Thank you for the suggestion. We have changed this sentence as "*For example, extreme humid-heat combining with low air temperature but a high humidity might still cause*

lethal and even deadly events".

line 35 - Im not sure this is the representation of the WBT on its own. However my understanding is that, combined with biophysical metrics, it can be used to determine whether sweating would still be effective.

Response: Thank you for the comment. Based on the thermodynamic definition, wet-bulb temperature is the temperature of an air parcel cooled to saturation by the evaporation of water into it. In the physiological research, sweating is considered as the main cooling mechanism under extreme heat conditions. So, the TW could also be used as the lowest temperature reached by evaporative cooling of sweating. For better understanding, we changed this sentence according to your suggestion as "*...the higher TW could dampen the evaporative cooling of sweating*".

Figure 1 - include in the caption a statement saying that the numbers in parentheses are the station counts left at each stage. Im not sure that "Correlationship" is a word (though I get what youre trying to say). The arrow from the Complemented series back to Climatol is missing a head.

Response: Thank you for the suggestions. We have added the statement of the numbers in the parentheses. We changed the "Correlationship" to "Correlations". We added the arrow from the "Complemented series" to "Climatol 3.1.2".

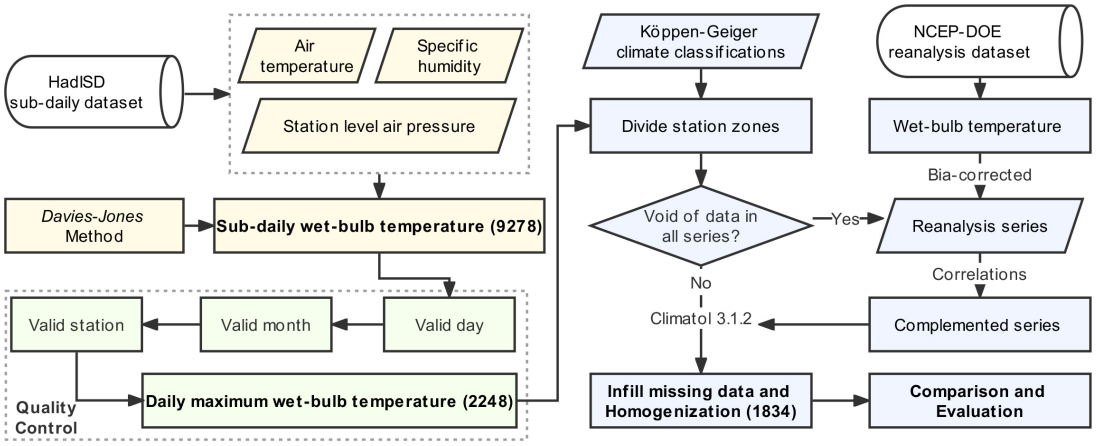

Figure 1. Procedures of producing global daily maximum TW (GSDM-WBT) dataset. The numbers in the parentheses indicate the counts of stations remained after each procedure.

Line 82 - add suitable reference for the ISD (e.g. Smith et al, 2011)

Response: Thank you for the suggestion. We added the citation of ISD here, as Smith et al., 2011.

Line 88 - how did you account for the shift to/from summer time when using the local time for the stations?

Response: Thank you for the comment. In this study, the local time was adjusted based on the time zone calculated from the longitude of each station. We did not consider the shift of the summer time and winter time. One main reason is that this shift is not applied all over the world, for example, there is no daylight saving time in China. Furthermore, the impact of one-hour transition between summer time and winter time could be generally low because we controlled our data as at least one TW every six hours per day. But the time zones (±12h) are quite important to change the distributions of sub-daily TW in UTC time to the real diurnal variations.

Line 121 - there is no section 3.1.1

Response: Thank you for the comment. We changed "section 3.1.1" to "section 3.1".

Line 127 - restate that the 6-hourly intervals are in local time.

Response: Thank you for the suggestion. We added the "in local time" in this sentence.

line 135 - "(of a total of..."

Response: Thank you for the suggestion. We corrected the phrase to "of a total of…".

line 137 - "contain" (delete "s")

Response: Thank you for the suggestion. We corrected the "contains" to "contain".

line 139 - insert a line break before "According"

Response: Thank you for the suggestion. We inserted a line break here.

line 156 - replace "for" with "so that there were"

Response: Thank you for the suggestion. We changed "for" to "so that there were".

line 160 - perhaps rephrase as "...depends on how many of the surrounding stations had missing data at this step"

Response: Thank you for the suggestion. We have rewritten the phrase as "*...depends on how many missing data of surrounding stations at this step*".

line 165 - "needed"

Response: Thank you for the suggestion. We corrected “need” to “needed”.

Figure 2 - can you adjust the colourmap so that it diverges when centred around 0. I appreciate there are few stations which have a bias of <0 but using a symmetrical scale will make that clearer than currently where blue-colours cover negative values and those 0-0.15C.

Response: Thank you for the suggestion. We redrew the figure by adjusting the colourmap.

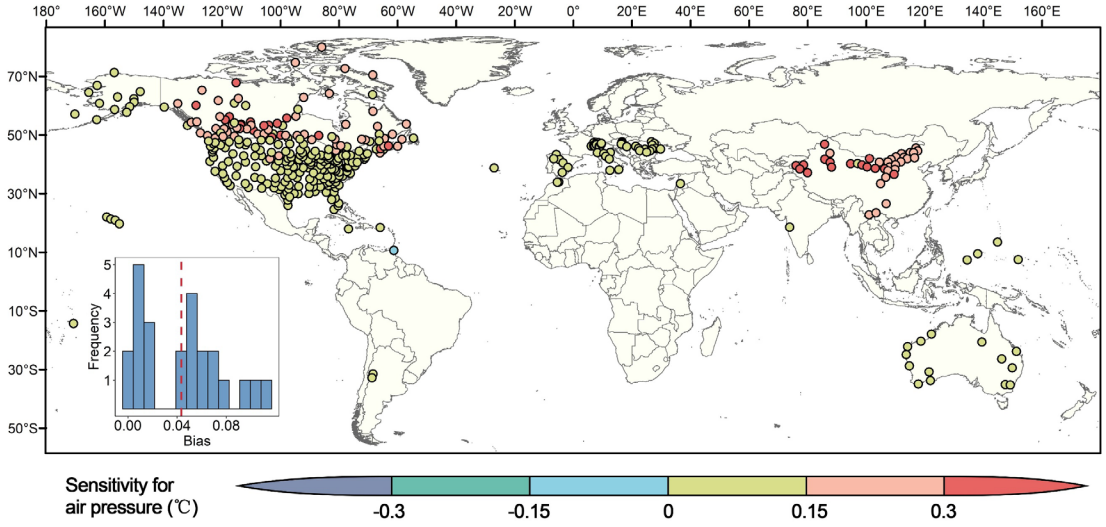

Figure 2. Sensitivity of air pressure on TW. Sensitivity, or average bias, was calculated by subtracting the daily maximum TW based on long-term average pressure by daily maximum TW calculated from sub-daily pressure. Sub-plot showed the histogram of average bias when average daily maximum TW was more than 20°C, where the red dashed line indicated the mean (0.04°C).

line 241 - move "in theory" to the front of that clause, to before "the corrected series"

Response: Thank you for the suggestion. We moved the “in theory” to the front of “the corrected series”.

line 262 - please check this sentence as there cannot be more than 30/31 missing days in a month.

Response: Thank you for the comment. Here the missing days were counted in each month during all forty years, so the total of days in each month for each station is about 1200 days. For better understanding, we have revised this sentence to “*The median number of missing days in each month over past forty years in the Northern Hemisphere is less than 100 days…*”.

line 272 - please add a suitable reference for the PHA algorithm

Response: Thank you for the suggestion. We added the related citation as “Menne and Williams,

2009".

Figure 3 - similar to Figure 2, please can you adjust the colourmap so that it diverges when centred around zero. Currently blue and pink colours are both negative, which could be confusing.

Response: Thank you for the suggestion. We thought you pointed out the problem of Figure 6, and we redrew this figure by adjusting the colourmap.

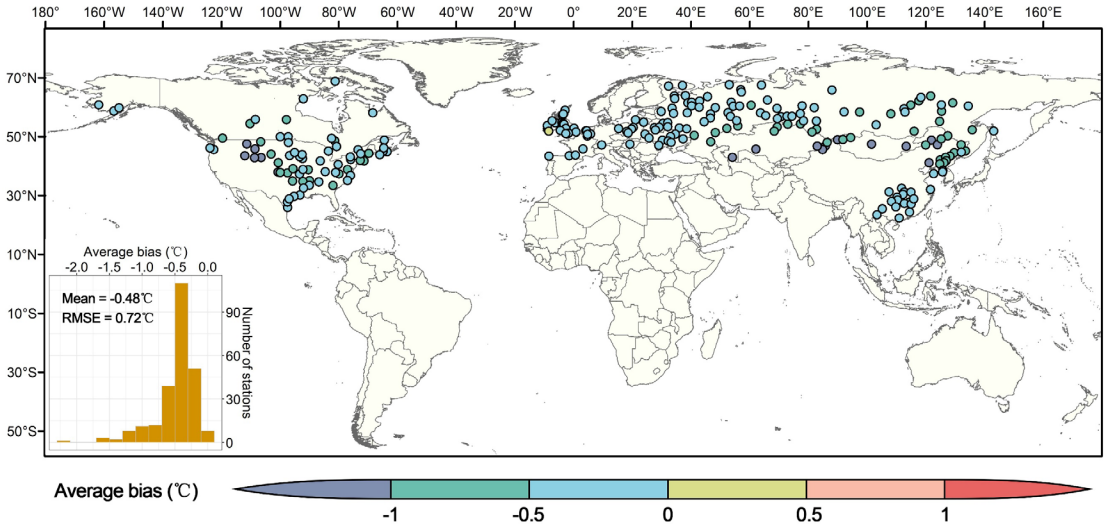

Figure 6. Average bias between daily maximum TW of GSDM-WBT and HadISD-Humidity.

line 282 - please add a suitable reference for ERA5

Response: Thank you for the suggestion. We added the citation as "Hersbach et al., 2020".

line 283 - add parentheses around "2021" in the Yan et al reference.

Response: Thank you for the suggestion. We have added the parentheses.

Figure 7 - similar to Figure 2, please adjust the colourmap to center the divergence around 0. Further guidance on this can be found at https://colorbrewer2.org. Currently it I found it difficult to identify the high and low bias regions.

Response: Thank you for the suggestion. We redrew the figure by adjusting the colourmap.

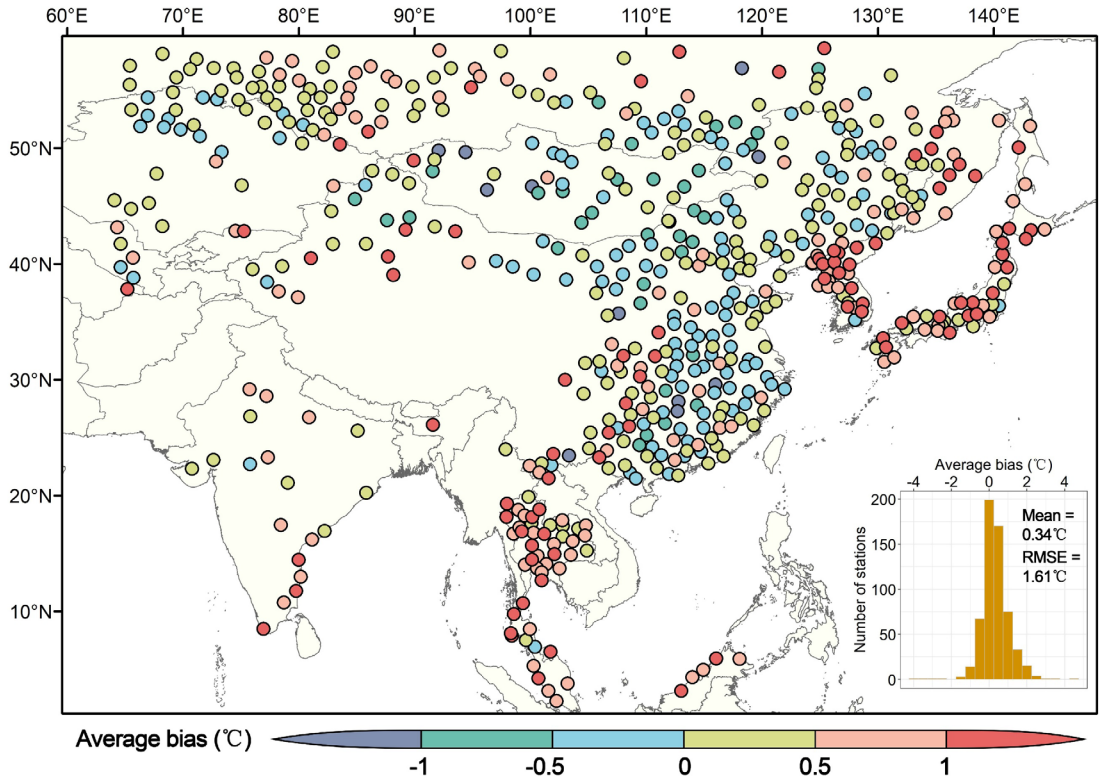

Figure 7. Average bias between station-based daily maximum TW of GSDM-WBT and that of the nearest grid points in HiTiSEA.

line 314 - replace "index" with "indices"

Response: Thank you for the suggestion. We have replaced “index” with “indices”.

line 351 - perhaps replace "cognizing" with "characterizing"

Response: Thank you for the suggestion. We have replaced “cognizing” with “characterizing”.

Figure S1 - add the number of valid and invalid stations to the caption. The grey crosses are hard to see at first glance.

Response: Thank you for the suggestion. We added the number of valid and invalid stations to the caption and also redrew the figure.

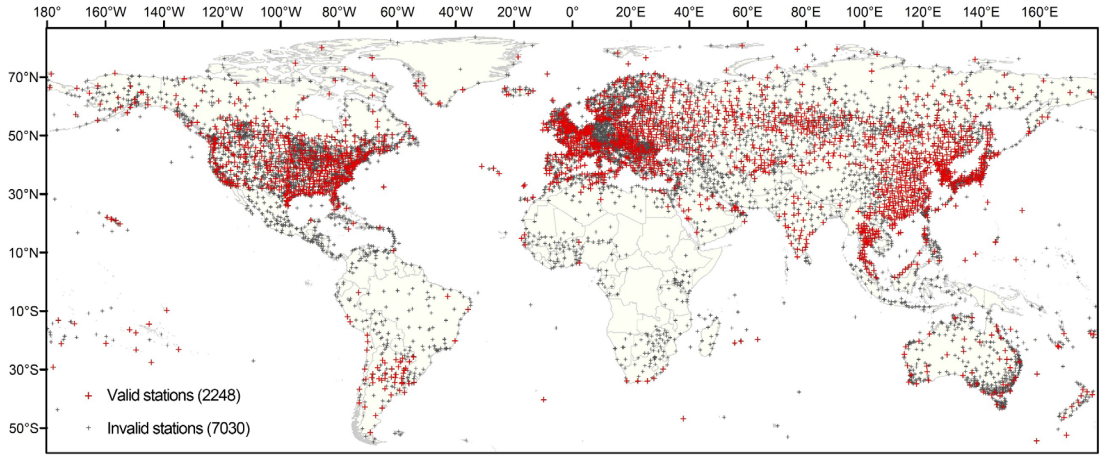

Figure S1. Spatial patterns of valid stations selected by data quality control.

References:

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q J R Meteorolog Soc, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Menne, M. J. and Williams, C. N.: Homogenization of temperature series via pairwise comparisons, J Climate, 22, 1700–1717, https://doi.org/10.1175/2008JCLI2263.1, 2009.

Mora, C., Dousset, B., Caldwell, I. R., Powell, F. E., Geronimo, R. C., Bielecki, C. R., Counsell, C. W. W., Dietrich, B. S., Johnston, E. T., Louis, L. V., Lucas, M. P., McKenzie, M. M., Shea, A. G., Tseng, H., Giambelluca, T. W., Leon, L. R., Hawkins, E., and Trauernicht, C.: Global risk of deadly heat, Nature Clim Change, 7, 501–506, https://doi.org/10.1038/nclimate3322, 2017.

Smith, A., Lott, N., and Vose, R.: The Integrated Surface Database: Recent Developments and Partnerships, Bull Am Meteorol Soc, 92, 704–708, https://doi.org/10.1175/2011BAMS3015.1, 2011.

Wang, P., Yang, Y., Tang, J., Leung, L. R., and Liao, H.: Intensified humid heat events under global warming, Geophys Res Lett, 48, e2020GL091462, https://doi.org/10.1029/2020GL091462, 2021.