A DETAILED LIST OF RESPONSES

TO THE EDITOR

We greatly appreciate your careful reading of the manuscript, insightful comments, and valuable suggestions. Your thoughtful review has enhanced our paper considerably. The manuscript has been revised thoroughly according to your comments and those of the individual reviewers, with our point-by-point responses detailed below.

(1) Given the importance of uncertainty quantification for the drought dataset, I think that Figure S4 should be added to the main body of the manuscript AND that there should be a suitable discussion around its interpretation. Such a discussion would ALSO include general limitations of uncertainty quantification in the presence of sparse observations.

Response: Many thanks for your comments. We have put the uncertainty graph (form Figure S4 to Figure 7) in the manuscript and discussed it around three main aspects: interpolation method uncertainty, interpolation station density, and consideration of covariates. Specifically, in Section 4.2 (Lines 371-390) we added the following: "<u>In addition, we quantified the uncertainties of SPI, SPEI and EDDI at different time scales (Figure 7). We used standard deviation to quantify the results, which were similar to those in Figures S2–S4. The regions with higher standard deviation, such as the arid northwest, highlight the spatial variability in uncertainty across different datasets. This suggests that the drought indices calculated from these datasets may show obvious discrepancies in regions with sparse observational coverage. These uncertainties may have the following reasons: (1) The variability in interpolation techniques across datasets is a critical factor contributing to uncertainty. For instance, the CHM dataset employs advanced interpolation techniques based on high-density observational stations, while the CRU and CN05.1 datasets utilize thin plate smooth spline (TPSS) and inverse distance weighting (IDW) methods, respectively (Harris et al., 2020; Xu et al., 2009). These methodological differences become particularly pronounced in areas with complex topography, such as the arid northwest. Xu et al. (2022) demonstrated that TPSS performs well in capturing broad climate gradients, it may overly smooth the</u>

results in data-sparse regions, leading to underestimation of extremes. Conversely, IDW might overemphasize local station values, causing biases in interpolated fields (Shen et al., 2023). (2) Sparse observational coverage is another significant source of uncertainty. Liu et al. (2009) highlighted that the density of interpolation sites is the key factor influencing interpolation accuracy. They found that the performance of interpolation methods, such as kriging or IDW, deteriorates significantly as the number of sites decreases. (3) Differences in the inclusion of auxiliary covariates, such as topography, land cover, or climate zones, further contribute to dataset discrepancies. The CHM dataset incorporates high-resolution digital elevation models (DEM) as covariates, while the CRU dataset primarily relies on planar spatial gradients without explicitly considering terrain effects (Harris et al., 2020). This leads to substantial differences in regions with complex orography."

(2) To increase transparency of the methods: Figures R2-R7 should be added to the Supplement. To do so they should be modified to (1) include measures of the fit of the exponential decay (e.g. variance explained) and (2) ideally changed to contour/ density plots to better visualize the distribution.

Response: Many thanks for your comments. We have added the contents of the CDD section to the supplementary document, including the CDD concept and method description, as well as the CDD density map with these meteorological variables added in Figure S1. Specifically, in Supplementary document (Lines 26-39) we added the following: "The ADW interpolation method used for this study was the modified Shepard's algorithm, which introduces the concept of correlation decay distance (CDD), also called correlation length scale or decorrelation length (Shepard, 1984; Dunn et al., 2020). The CDD is defined as the distance at which the correlation between one station and all other stations decays below 1/e, approximately corresponding to the significance level of 0.05 for the correlation within large samples (Jones et al., 1997; Harris et al., 2020). The number of stations for interpolating the target grid cell is well constrained by the CDD, thus improving the interpolation precision (New et al., 2000; Mitchell and Jones, 2005; Hofstra and New, 2009). For every station, correlations (r) and distances

(x) for each variable are shown in Figure S1, and the ordinary least-squares method was used to fit an exponential decay function: $r = e^{-x/CDD}$, take the meteorological variable Wind (Figure S1a), for example, the estimated CDD is 361 km (95 % confidence interval: 361 km) at the 0.05 significance level."

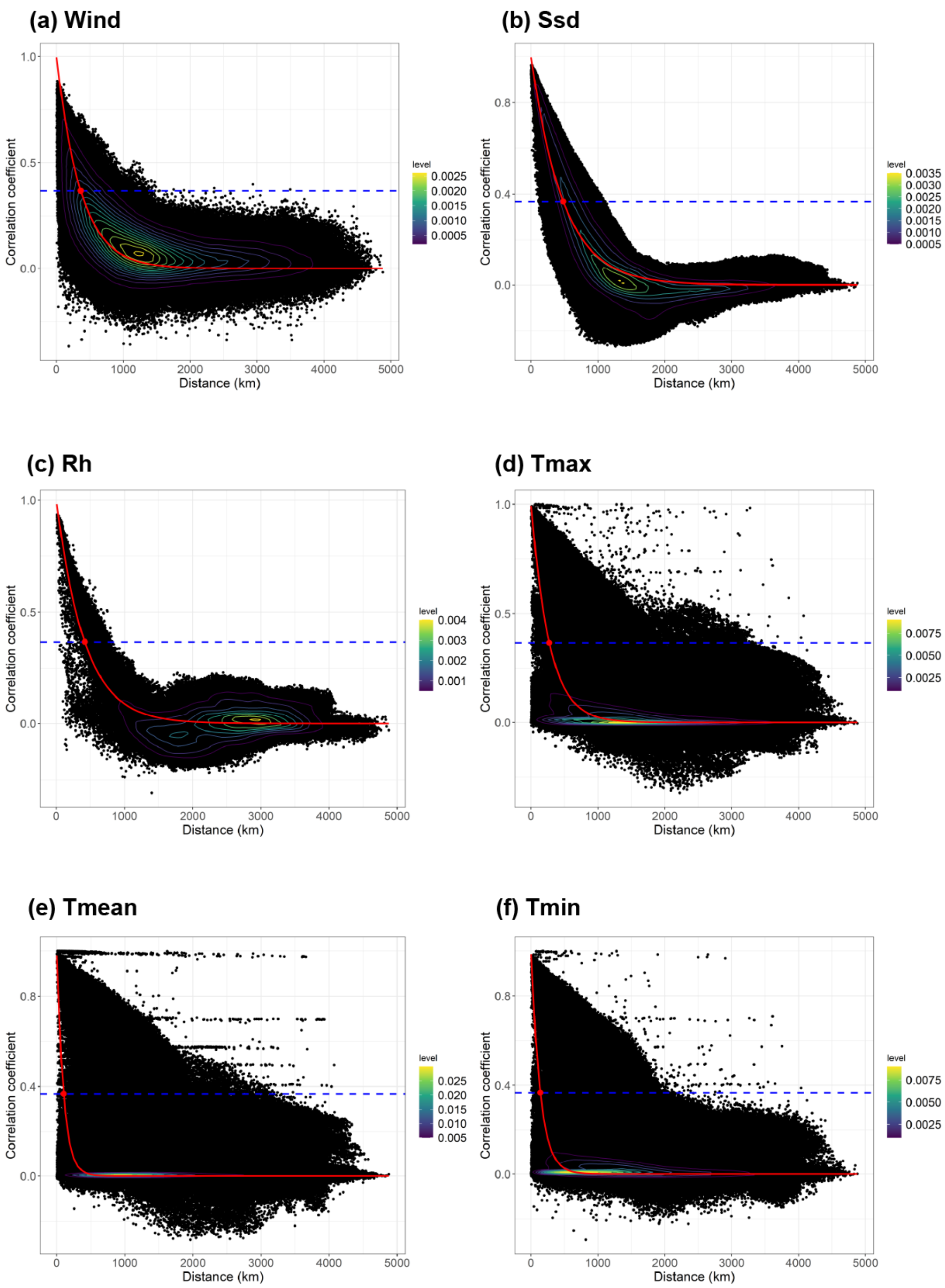

Figure S1: Kernel density visualization of the Correlation Decay Distance (CDD) and the distribution for meteorological variables (Wind≈361, Ssd≈480, Rh≈420, Tmax≈ 272, Tmean≈ 99, Tmin≈ 136) for all stations within the interpolated domain. Black points show the distance–correlation pair for each station. The blue line is the exponential curve fitted to the data by ordinary least squares. The red dashed line marks where correlation equals 1/e.

(3) Technical:L125: PrecIpitation, L351 > no comma after THAT

Response: Many thanks for your comments. We have modified.

References:

Han, J., Miao, C., Gou, J., Zheng, H., Zhang, Q., and Guo, X.: A new daily gridded precipitation dataset for the Chinese mainland based on gauge observations, Earth Syst Sci Data, 15, https://doi.org/10.5194/essd-15-3147-2023, 2023.

Dunn, R. J. H., Alexander, L. V., Donat, M. G., Zhang, X., Bador, M., Herold, N., Lippmann, T., Allan, R., Aguilar, E., Barry, A. A., Brunet, M., Caesar, J., Chagnaud, G., Cheng, V., Cinco, T., Durre, I., de Guzman, R., Htay, T. M., Wan Ibadullah, W. M., Bin Ibrahim, M. K. I., Khoshkam, M., Kruger, A., Kubota, H., Leng, T. W., Lim, G., Li-Sha, L., Marengo, J., Mbatha, S., McGree, S., Menne, M., de los Milagros Skansi, M., Ngwenya, S., Nkrumah, F., Oonariya, C., Pabon-Caicedo, J. D., Panthou, G., Pham, C., Rahimzadeh, F., Ramos, A., Salgado, E., Salinger, J., Sané, Y., Sopaheluwakan, A., Srivastava, A., Sun, Y., Timbal, B., Trachow, N., Trewin, B., van der Schrier, G., VazquezAguirre, J., Vasquez, R., Villarroel, C., Vincent, L., Vischel, T., Vose, R., and Bin Hj Yussof, M. N. A.: Development of an updated global land in situ-based data set of temperature and precipitation extremes: HadEX3, J. Geophys. Res.-Atmos., 125, e2019JD032263, https://doi.org/10.1029/2019JD032263, 2020.

Jones, P. D., Osborn, T. J., and Briffa, K. R.: Estimating sampling errors in large-scale temperature averages, J. Climate, 10, 2548–2568, https://doi.org/10.1175/15200442(1997)010<2548:ESEILS>2.0.CO;2, 1997.

Harris, I., Osborn, T. J., Jones, P., and Lister, D.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, Sci. Data, 7, 109, https://doi.org/10.1038/s41597-0200453-3, 2020.

Hofstra, N. and New, M.: Spatial variability in correlation decay distance and influence on angular-distance weighting interpolation of daily precipitation over Europe, Int. J. Climatol., 29, 18721880, https://doi.org/10.1002/joc.1819, 2009.

Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol., 25, 693–712, https://doi.org/10.1002/joc.1181, 2005.

New, M., Hulme, M., and Jones, P.: Representing twentiethcentury space–time climate variability. part II: development of 1901–96 monthly grids of terrestrial surface climate, J. Climate, 13, 2217–2238, https://doi.org/10.1175/15200442(2000)013<2217:RTCSTC>2.0.CO;2, 2000.

Shepard, D. S.: Computer Mapping: The SYMAP Interpolation Algorithm, in: Spatial Statistics and Models, edited by: Gaile, G. L., and Willmott, C. J., Springer Netherlands, Dordrecht, 133–145, https://doi.org/10.1007/978-94-017-3048-8_7, 1984.

-- end line ---

For your convenience, to make the review of our revisions easier, we have marked all responses and related revisions in light blue.