Supplement of manuscript on

A global dataset of daily near-surface air temperature at 1-km resolution (2003-2020)

Tao Zhang[1], Yuyu Zhou[1], Kaiguang Zhao[2], Zhengyuan Zhu[3], Gang Chen[4], Jia Hu[1], Li Wang[5]

[1] Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA, 50011, USA

[2] School of Environment and Natural Resources, Ohio Agricultural Research and Development Center, The Ohio State University, Wooster, OH, 44691, USA

[3] Department of Statistics, Iowa State University, Ames, IA, 50011, USA

[4] Laboratory for Remote Sensing and Environmental Change (LRSEC), Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte, NC, 28223, USA

5 Department of Statistics, George Mason University, Fairfax, VA, 22030, USA

Correspondence to: Yuyu Zhou (yuyuzhou@iastate.edu)

S1 Estimating missing values at weather stations

We found that there were only a few weather stations with valid Ta measurements for some days of the study period in Africa, and Latin America (Fig. S1). When the number of weather stations with valid measurements was too small for a specific day, the performance of the SVCM-SP algorithm might be reduced. Therefore, we implemented a 5-day moving average algorithm (Kemp et al., 1983) to fill the missing Ta measurements at weather stations in Africa and Latin America. As shown in the example in Fig. S2, for a specific weather station, we checked each 5-day temporal window where the central day has a missing Ta measurement in the original time series data. When there are at least two days have valid measurements in the 5-day local window, the missing value in the central day was filled by using the average value of the local window. After filling the missing values, the number of weather stations with valid measurements was largely increased (Fig. S3). Taking the day 316 of 2016 in Africa as an example, after filling the missing values, the number of weather stations with valid measurements increased from 9 to 248 stations, and the spatial distribution of stations was improved significantly (Fig. S4). We also implemented leave-one-out cross-validation by randomly removing valid measurements from the original time series of Ta measurements at weather stations. Results of cross-validation indicated that the performance of the 5-day moving window algorithm performed well with RMSE values ranging from 1.6 to 2.6 °C in Africa and Latin America (Fig. S5).

Figures

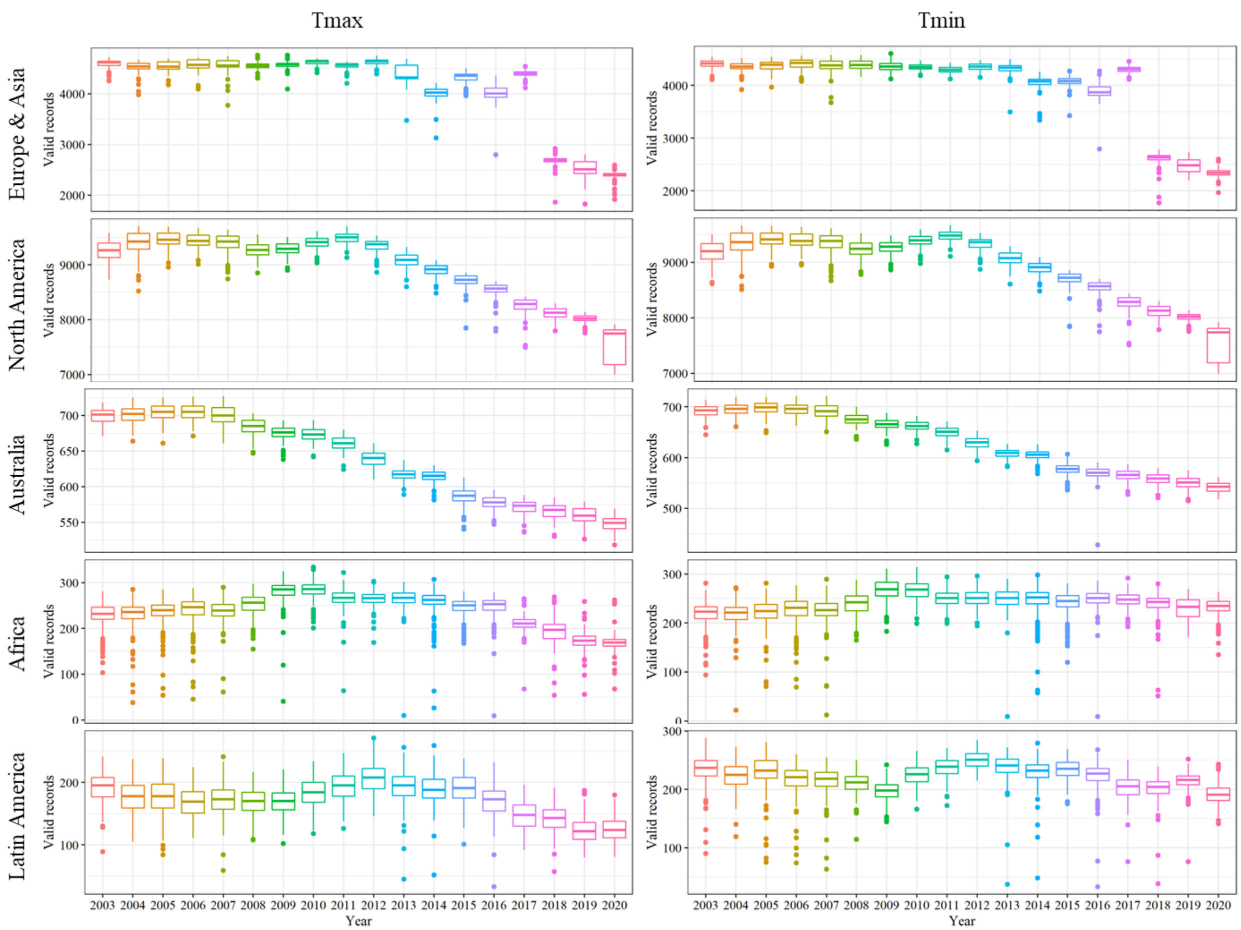

Figure S1: Number of stations with valid observations (Y-axis) in five regions.

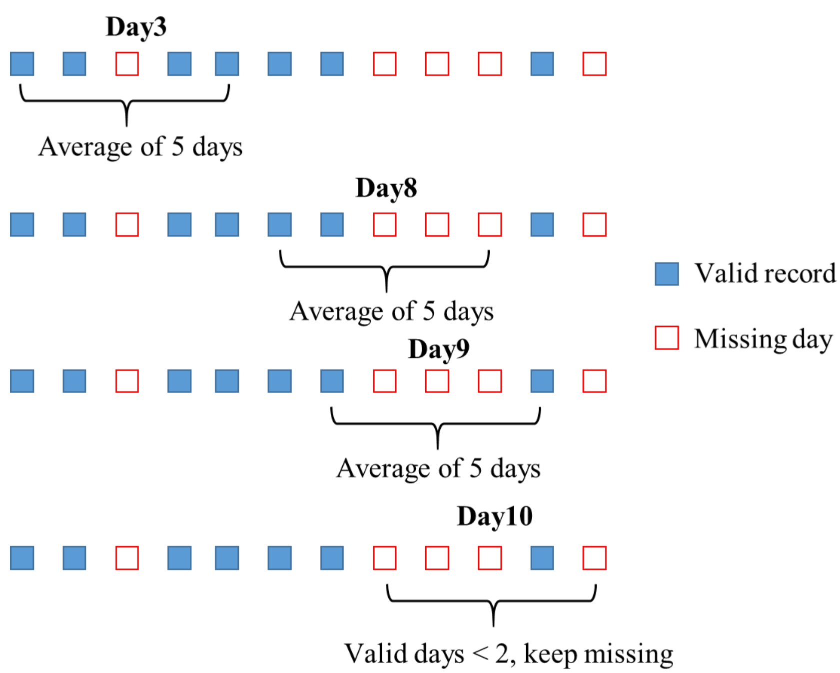

Figure S2: An example of filling missing Ta values at weather stations.

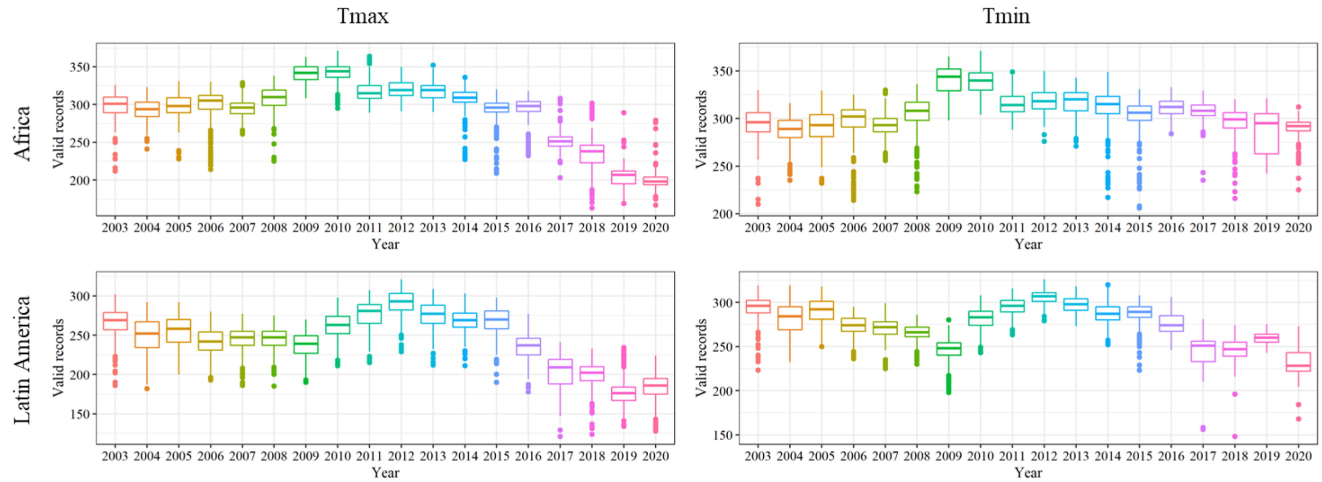

Figure S3: Number of stations with valid observations (Y-axis) after filling missing values in Africa and Latin America.

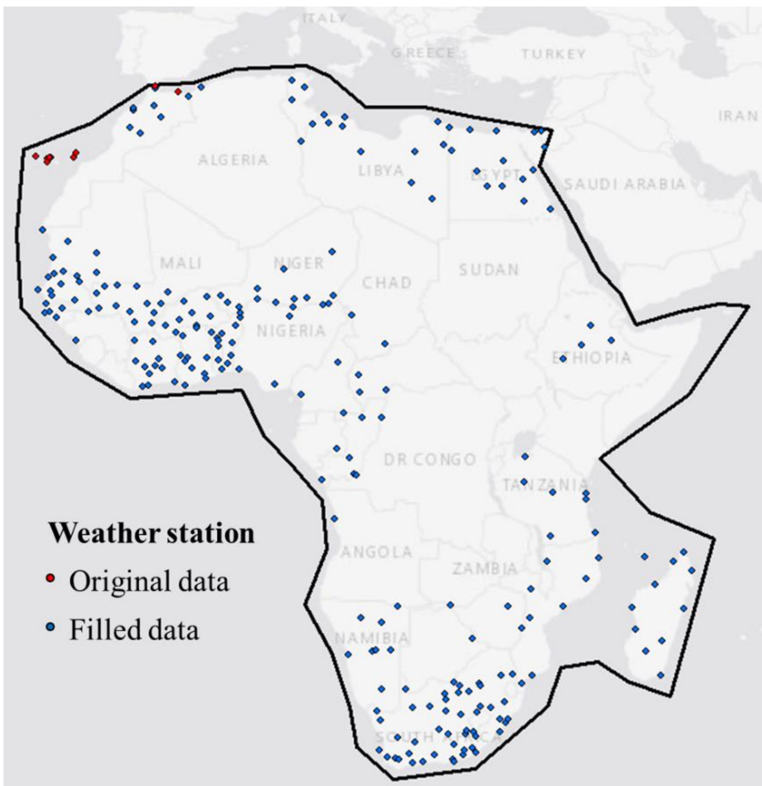

Figure S4: An example of valid observations in Africa (Tmax on day of 316 in 2016).

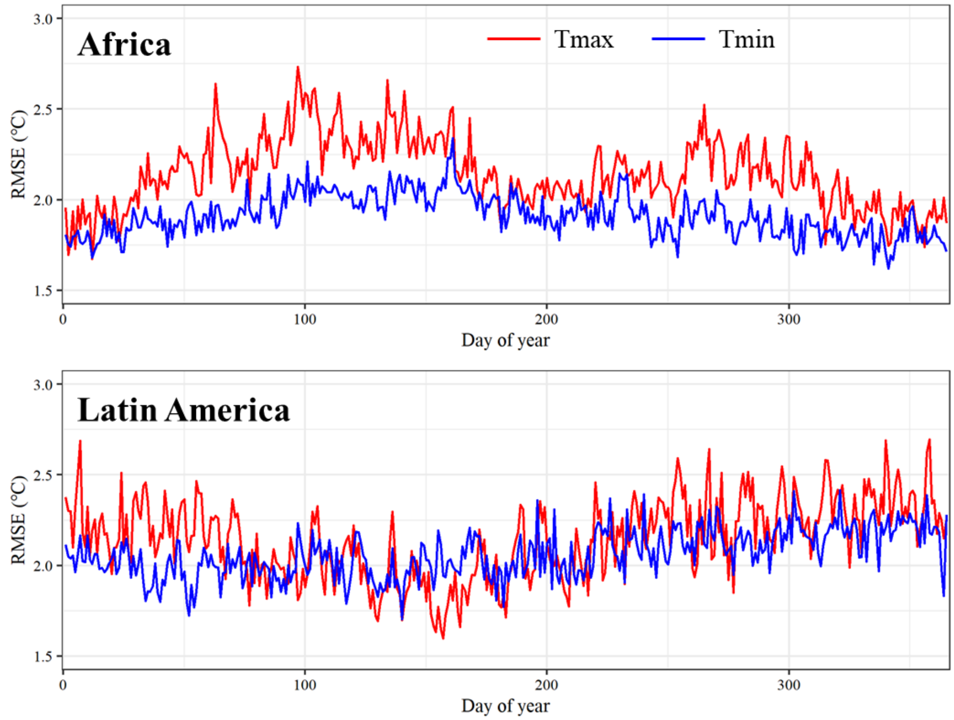

Figure S5: RMSE of filled missing values at stations in Africa and America based on the cross-validation.

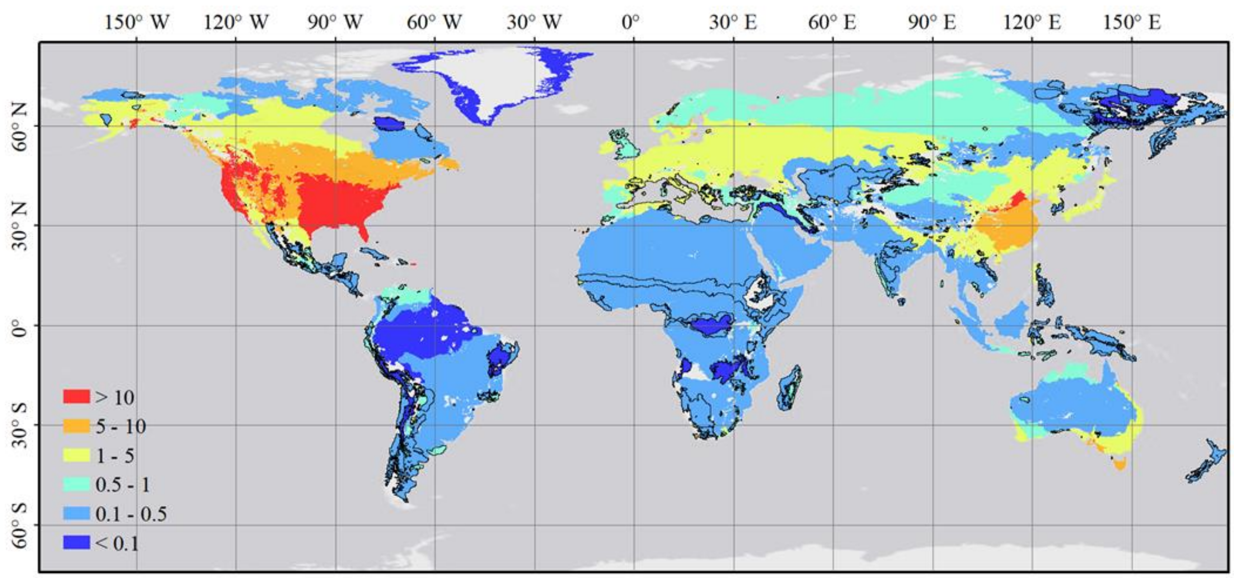

Figure S6: Spatial pattern of station density in different climate zones in 2003-2020 based on all available stations. Climate zones with black boundaries do not contain available weather stations within 50 km of training stations.

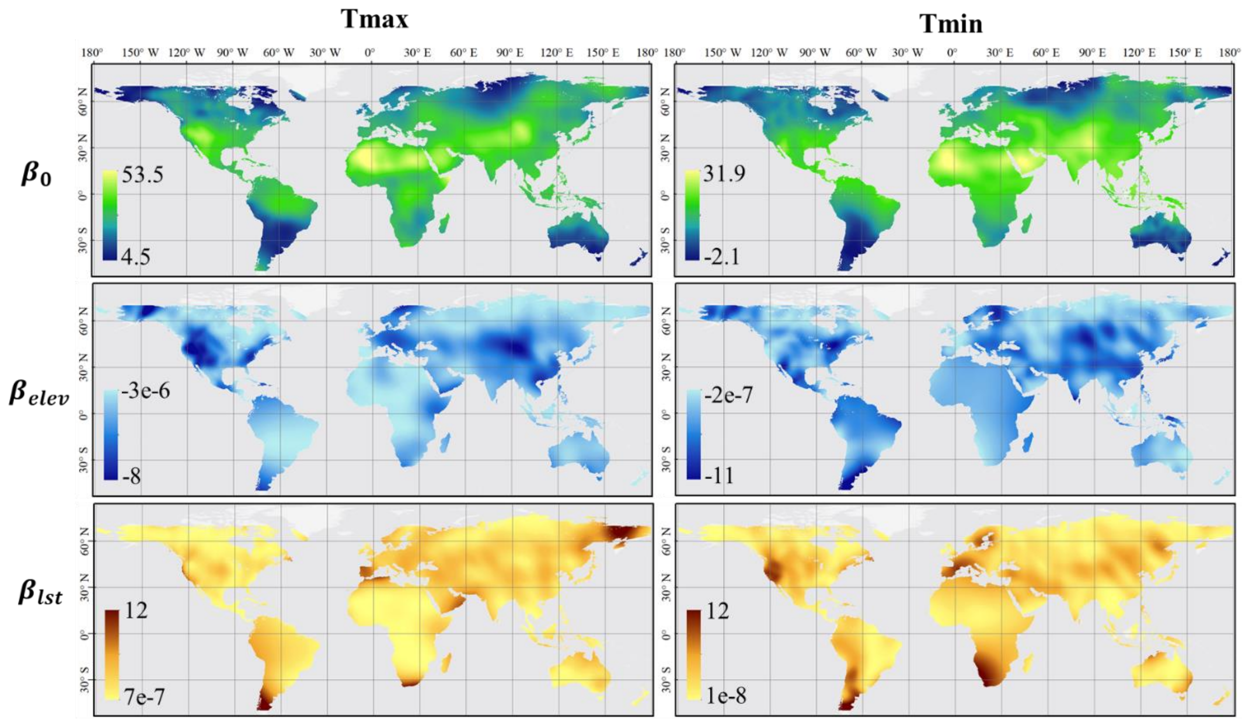

Figure S7: Spatial patterns of intercept (β_0), standardized regression coefficients of elevation (β_{elev}) and LST (β_{lst}) in day 200 of the year 2010.

Tables

Table S1. Average RMSEs (±standard deviation) for Tmax and Tmin in five regions from 2003 to 2020 (Unit: °C)

Year	North America		Latin America		European&Asia		Africa		Australia	
	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin
2003	2.49±0.39	2.42±0.36	2.09±0.41	1.75±0.26	1.79±0.19	1.74±0.27	2.09±0.56	2.10±0.37	1.20±0.27	1.63±0.23
2004	2.46±0.39	2.38±0.39	1.92±0.42	1.81±0.32	1.74±0.17	1.74±0.27	2.19±0.69	2.20±0.39	1.17±0.25	1.65±0.24
2005	2.44±0.39	2.36±0.38	2.01±0.47	1.90±0.30	1.79±0.20	1.73±0.26	2.20±0.63	2.28±0.43	1.15±0.22	1.63±0.23
2006	2.41±0.36	2.40±0.35	1.91±0.41	1.87±0.31	1.79±0.18	1.73±0.25	2.35±0.70	2.44±0.51	1.20±0.27	1.74±0.27
2007	2.45±0.36	2.42±0.36	1.93±0.43	1.83±0.29	1.74±0.18	1.72±0.23	2.26±0.58	2.60±0.53	1.18±0.25	1.64±0.25
2008	2.50±0.45	2.48±0.44	1.97±0.50	1.84±0.36	1.74±0.17	1.71±0.23	2.22±0.54	2.33±0.48	1.16±0.22	1.66±0.26
2009	2.48±0.41	2.40±0.43	1.94±0.47	2.01±0.33	1.81±0.18	1.75±0.25	2.07±0.51	2.19±0.37	1.19±0.27	1.67±0.22
2010	2.36±0.32	2.32±0.36	1.97±0.42	1.95±0.34	1.84±0.22	1.76±0.34	2.00±0.48	2.23±0.39	1.14±0.23	1.57±0.25
2011	2.45±0.37	2.39±0.43	1.86±0.42	1.90±0.32	1.81±0.18	1.71±0.26	1.98±0.49	2.12±0.34	1.14±0.21	1.62±0.29
2012	2.41±0.36	2.39±0.35	1.85±0.40	1.92±0.31	1.81±0.19	1.71±0.27	2.03±0.53	2.21±0.40	1.16±0.23	1.71±0.28
2013	2.44±0.42	2.40±0.42	1.98±0.45	2.09±0.29	1.81±0.21	1.76±0.27	2.10±0.43	2.29±0.39	1.20±0.26	1.75±0.23
2014	2.50±0.49	2.41±0.44	2.04±0.48	2.22±0.27	1.84±0.20	1.64±0.23	2.19±0.38	2.17±0.41	1.23±0.27	1.73±0.26
2015	2.39±0.40	2.35±0.40	1.93±0.38	2.23±0.27	1.83±0.20	1.66±0.22	2.23±0.42	2.15±0.37	1.22±0.26	1.71±0.27
2016	2.37±0.38	2.34±0.38	1.91±0.41	2.00±0.28	1.84±0.18	1.63±0.25	2.18±0.46	2.13±0.37	1.21±0.27	1.64±0.25
2017	2.45±0.41	2.39±0.40	1.92±0.44	1.96±0.30	1.81±0.19	1.69±0.25	2.44±0.58	2.15±0.35	1.24±0.27	1.74±0.28
2018	2.43±0.41	2.38±0.39	1.97±0.34	2.04±0.30	1.83±0.18	1.98±0.34	2.40±0.60	2.04±0.37	1.27±0.28	1.76±0.29
2019	2.43±0.42	2.38±0.42	1.88±0.40	2.09±0.25	1.79±0.19	1.94±0.32	2.48±0.66	2.08±0.40	1.30±0.32	1.81±0.27
2020	2.41±0.42	2.42±0.36	1.77±0.35	2.03±0.32	1.72±0.16	1.84±0.23	2.52±0.67	2.04±0.38	1.25±0.28	1.69±0.28

Note: The selected testing stations were within 50 km^2 surrounding the training stations.

Table S2. Average MAEs (±standard deviation) for Tmax and Tmin in five regions from 2003 to 2020 (Unit: °C)

Year	North America		Latin America		Europe&Asia		Africa		Australia	
	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin
2003	1.86±0.30	1.79±0.27	1.57±0.29	1.31±0.19	1.29±0.15	1.27±0.20	1.55±0.40	1.62±0.28	0.89±0.18	1.24±0.19
2004	1.83±0.29	1.77±0.29	1.43±0.29	1.34±0.22	1.25±0.14	1.27±0.20	1.60±0.49	1.71±0.29	0.87±0.18	1.26±0.20
2005	1.82±0.30	1.75±0.28	1.50±0.32	1.38±0.22	1.29±0.16	1.26±0.19	1.59±0.40	1.76±0.32	0.86±0.15	1.25±0.19
2006	1.81±0.27	1.79±0.27	1.42±0.27	1.40±0.22	1.30±0.15	1.27±0.18	1.70±0.49	1.84±0.34	0.89±0.19	1.33±0.23
2007	1.82±0.27	1.80±0.27	1.42±0.29	1.37±0.21	1.26±0.15	1.25±0.18	1.59±0.38	1.94±0.35	0.88±0.17	1.25±0.21
2008	1.87±0.34	1.84±0.33	1.49±0.33	1.39±0.26	1.25±0.14	1.25±0.18	1.56±0.36	1.76±0.34	0.86±0.16	1.26±0.22
2009	1.85±0.31	1.78±0.32	1.48±0.33	1.51±0.24	1.31±0.15	1.28±0.19	1.49±0.35	1.68±0.27	0.89±0.19	1.27±0.18
2010	1.76±0.24	1.71±0.27	1.50±0.31	1.48±0.26	1.34±0.18	1.29±0.25	1.45±0.31	1.70±0.28	0.86±0.16	1.19±0.21
2011	1.83±0.28	1.77±0.32	1.42±0.28	1.45±0.24	1.30±0.15	1.25±0.19	1.44±0.32	1.62±0.24	0.85±0.16	1.22±0.24

2012	1.79±0.27	1.78±0.27	1.39±0.26	1.46±0.23	1.3±0.15	1.25±0.20	1.51±0.38	1.68±0.29	0.86±0.17	1.30±0.23
2013	1.82±0.31	1.78±0.32	1.48±0.30	1.55±0.21	1.31±0.18	1.29±0.20	1.57±0.31	1.72±0.29	0.89±0.18	1.32±0.19
2014	1.87±0.36	1.79±0.33	1.52±0.31	1.67±0.21	1.32±0.16	1.21±0.18	1.62±0.28	1.63±0.27	0.92±0.20	1.32±0.21
2015	1.79±0.31	1.75±0.30	1.45±0.27	1.63±0.21	1.32±0.17	1.22±0.17	1.63±0.28	1.64±0.27	0.91±0.18	1.30±0.22
2016	1.78±0.29	1.74±0.28	1.45±0.28	1.47±0.21	1.33±0.16	1.19±0.18	1.61±0.30	1.64±0.27	0.90±0.20	1.25±0.21
2017	1.83±0.31	1.77±0.29	1.45±0.31	1.48±0.23	1.30±0.16	1.24±0.19	1.77±0.36	1.66±0.26	0.92±0.20	1.33±0.24
2018	1.82±0.30	1.77±0.30	1.47±0.24	1.52±0.23	1.28±0.14	1.46±0.24	1.76±0.38	1.56±0.26	0.94±0.21	1.34±0.24
2019	1.82±0.31	1.77±0.31	1.39±0.24	1.53±0.18	1.26±0.15	1.43±0.23	1.84±0.43	1.58±0.27	0.96±0.23	1.38±0.22
2020	1.80±0.31	1.80±0.27	1.34±0.25	1.52±0.23	1.20±0.13	1.35±0.17	1.85±0.42	1.54±0.26	0.93±0.20	1.28±0.23

Note: The selected testing stations were within 50 km^2 surrounding the training stations.

Table S3. Summary of publicly available gridded near-surface air temperature datasets

Literature	Spatial extent	Spatial resolution	Temporal frequency	Coverage time	Accuracy (default: RMSE)
Kalnay et al., 1996	Global	2.5 º	4-times daily	1948 – present	--
Hersbach et al., 2018	Global	0.25 º	Hourly	1979 – present	--
Thornton et al., 2021	North America, Hawaii, and Puerto Rico	1 km	Daily (Tmax and Tmin)	1950 – present calendar year	MAE of 1.52 ºC (Tmax) and 1.78 ºC (Tmin)
Oyler et al., 2015	CONUS	30-arcsec (~800m)	Daily (Tmax and Tmin)	1948 – 2016	MAE of 1.03 ºC (Tmax) and 1.06 ºC (Tmin)
Hooker et al., 2018	Global	0.05 º	Monthly	2003 – 2016	1.14–1.55 ºC
Meyer et al., 2019	South Africa	3.5 km	Hourly	2010 – 2014	2.61 ºC
Werner et al., 2019	Northwestern North American	1/16 º	Daily	1971 – 2000	~2.2–3.5 ºC
Nashwan et al., 2019	Central North region of Egypt	0.05 º	Daily (Tmax and Tmin)	1981 – 2017	1.41 ºC (Tmax) 1.31 ºC (Tmin)
MacDonald et al., 2020	North American	60-arcsec (~2 km)	Monthly (Tmax and Tmin)	1901 – 2016	MAE of 0.7 ºC (Tmax) and 1.02 ºC (Tmin)
Crespi et al., 2021	North-eastern Italy	250 m	Daily (Tmean)	1980 – 2018	MAE is 1.5 ºC
Chen et al., 2021	Mainland China	1 km	Daily (Tmean)	2003 – 2019	1.615–1.957 K (leave-location-out cross-validation)
Fang et al., 2021	China	0.1 º	Daily (Tmax, Tmin, and Tmean)	1979 – 2018	0.86–1.78 ºC (Tmax), 0.78–2.09 ºC (Tmin), 0.35–1.00 ºC (Tmean)
This study (Zhang and Zhou, 2022)	Global	1 km	Daily (Tmax and Tmin)	2003 – 2020	1.20–2.44 ºC (Tmax), 1.69–2.39 ºC (Tmin)

References:

Chen, Y., Liang, S., Ma, H., Li, B., He, T. and Wang, Q.: An all-sky 1km daily land surface air temperature product over mainland China for 2003-2019 from MODIS and ancillary data, Earth Syst. Sci. Data, 13(8), 4241–4261, doi:10.5194/essd-13-4241-2021, 2021.

Crespi, A., Matiu, M., Bertoldi, G., Petitta, M. and Zebisch, M.: A high-resolution gridded dataset of daily temperature and precipitation records (1980-2018) for Trentino-South Tyrol (north-eastern Italian Alps), Earth Syst. Sci. Data, 13(6), 2801–2818, doi:10.5194/essd-13-2801-2021, 2021.

Fang, S., Mao, K., Xia, X., Wang, P., Shi, J., Bateni, S., Xu, T., Cao, M. and Heggy, E.: A dataset of daily near-surface air temperature in China from 1979 to 2018, Earth Syst. Sci. Data Discuss., (October), 1–37, doi:10.5194/essd-2021-309, 2021.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R. and Rozum, I.: ERA5 hourly data on single levels from 1979 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), ECMWF, 147, 5–6, 2018.

Hooker, J., Duveiller, G. and Cescatti, A.: Data descriptor: A global dataset of air temperature derived from satellite remote sensing and weather stations, Sci. Data, 5, 1–11, doi:10.1038/sdata.2018.246, 2018.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, Bull. Am. Meteorol. Soc., 77(3), 437–472, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Kemp, W. P., Burnell, D. G., Everson, D. O. and Thomson, A. J.: Estimating missing daily maximum and minimum temperatures., J. Clim. Appl. Meteorol., 22(9), 1587–1593, doi:10.1175/1520-0450(1983)022<1587:EMDMAM>2.0.CO;2, 1983.

MacDonald, H., McKenney, D. W., Papadopol, P., Lawrence, K., Pedlar, J. and Hutchinson, M. F.: North American historical monthly spatial climate dataset, 1901–2016, Sci. Data, 7(1), 1–11, doi:10.1038/s41597-020-00737-2, 2020.

Meyer, H., Schmidt, J., Detsch, F. and Nauss, T.: Hourly gridded air temperatures of South Africa derived from MSG SEVIRI, Int. J. Appl. Earth Obs. Geoinf., 78(February), 261–267, doi:10.1016/j.jag.2019.02.006, 2019.

Nashwan, M. S., Shahid, S. and Chung, E. S.: Development of high-resolution daily gridded temperature datasets for the central north region of Egypt, Sci. Data, 6(1), 1–13, doi:10.1038/s41597-019-0144-0, 2019.

Oyler, J. W., Ballantyne, A., Jencso, K., Sweet, M. and Running, S. W.: Creating a topoclimatic daily air temperature dataset for the conterminous United States using homogenized station data and remotely sensed land skin temperature, Int. J. Climatol., 35(9), 2258–2279, doi:10.1002/joc.4127, 2015.

Thornton, P. E., Shrestha, R., Thornton, M., Kao, S. C., Wei, Y. and Wilson, B. E.: Gridded daily weather data for North America with comprehensive uncertainty quantification, Sci. Data, 8(1), 1–17, doi:10.1038/s41597-021-00973-0, 2021.

Werner, A. T., Schnorbus, M. A., Shrestha, R. R., Cannon, A. J., Zwiers, F. W., Dayon, G. and Anslow, F.: A long-term, temporally consistent, gridded daily meteorological dataset for northwestern North America, Sci. Data, 6, 1–16, doi:10.1038/sdata.2018.299, 2019.

Zhang, T. and Zhou, Y.: A global 1 km resolution daily near-surface air temperature dataset (2003 – 2020), Iowa State Univ., doi:10.25380/iastate.c.6005185, 2022.