Supplementary Materials for

Characterizing heatwaves based on land surface energy budget

Yinglin Tian[1,2,3], Axel Kleidon[4], Corey Lesk[5,6], Sha Zhou[7], Xiangzhong Luo[8], Sarosh Alam Ghausi[4], Guangqian Wang[1], Deyu Zhong[1*], Jakob Zscheischler[2,9*]

[1]State Key Laboratory of Hydroscience and Engineering, Key Laboratory of Hydrosphere Sciences of the Ministry of Water Resources, Department of Hydraulic Engineering, Tsinghua University, Beijing, China

[2]Department of Compound Environmental Risks, Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany

[3]Earth System Analysis, Potsdam Institute for Climate Impact Research (PIK) – Member of the Leibniz Association, Potsdam, Germany.

[4]Biospheric Theory and Modelling, Max Planck Institute for Biogeochemistry, Jena, Germany

[5]Lamont-Doherty Earth Observatory, Columbia University, New York, USA

[6]Department of Geography and Neukom Institute, Dartmouth College, Hanover, USA

[7]State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing, China

[8]Department of Geography, National University of Singapore, Singapore

[9]Technische Universität Dresden, Dresden, Germany

Correspondence to: Prof. Dr. Deyu Zhong (zhongdy@tsinghua.edu.cn) and Prof. Dr. Jakob Zscheischler (jakob.zscheischler@ufz.de).

Supplementary Text

Text S1. Composite decomposition of the temperature anomalies during all summer heatwave days

As depicted in Fig. S1a, the temperature anomalies exhibit greater magnitude in high latitudes. Adiabatic heating predominates in regions with higher elevations, such as the Rocky Mountains, Andes, Nordic, Alps, Greater Caucasus, Kunlun, Himalayas, South Siberian, and East Siberian mountains (Fig. S1b). Meanwhile, advective heating is primarily observed in coastal areas at high latitudes and hyper-arid or arid regions, including the Sahara, Arabian Peninsula, Kalahari, Patagonian Desert, Gobi Deserts, Central Asia, and the Australian Outback (Fig. S1c). In general, diabatic terms make a positive contribution to temperature anomalies across global land areas (Fig. S1d), with adiabatic and advective terms partially offsetting each other. Moreover, in diabatic heating, solar radiation has a more substantial heating effect in high latitudes compared to mid-latitudes and the tropics, except for East Asia, while it has a cooling effect over Greenland Island and the Sahara Desert (Fig. S1e). Downward longwave radiation contributes positively to temperature anomalies in mid- and high-latitude regions but has a negative impact in the tropics (Fig. S1f). Regarding surface turbulent flux, latent and sensible heat flux anomalies generally counterbalance each other (Figs. S1g and S1h), with the former contributing to heat extremes in semi-arid regions. The ground heat flux typically induces a cooling effect (Fig. S1i).

Text S2. Sensitivity test of clustering

We test our results with different settings of the ratio between diabatic heating and skin temperature variation (Figs. S5 and S6), moving window of threshold (Fig. S7), selection of the proxy of suraface adiabatic and advective heating (Fig. S8), and clustering method (Fig. S9), and found our clustering remained robust to a variety of specifications (Figs. S4-S9). In general, the optimal cluster number remains 4 according to the Silhouette score (Fig. S4), and the relative distribution and the percentage of the different types of heat extremes remained largely consistent when comparing Figs. 1, S6- S9.

To assess the validity of our assumption that temperature changes resulting from diabatic heating can be accurately represented by variations in skin temperature (f=0.91), we conduct a cluster analysis using f=0.91 as well, which is the ratio between $\Delta T_{diabatic} = \Delta T_{2m} - (\Delta T_{adiabatic} + \Delta T_{advection})$ and ΔT_s considering all the summer heatwave-days (Fig. S5), and the results are shown in Fig. S6. We find that the outcomes are not significantly affected the f values.

Fig. S7 shows the clustering results with 7-d moving window instead of 15-day-moving-average.

To test the robustness of the clustering results to our definition of surface advective and adiabatic heating term, we use the advective (and adiabatic) heating anomalies on 850 hPa instead of on 950 hPa as the proxy of the surface advective (and adiabatic) heating for heatwave days with surface pressure ranging 900~1100 hPa, i.e.,

$$\Delta T_{advective,test} = \begin{cases} \Delta T_{advective,850\ hPa}\ (p_s \geq 800\ hPa) \\ \Delta T_{advective,750\ hPa}\ (800\ hPa > p_s \geq 700\ hPa) \\ \Delta T_{advective,650\ hPa}\ (p_s < 700\ hPa) \end{cases} \quad \text{Eq. S1}$$

$$\Delta T_{adiabatic,test} = \begin{cases} \Delta T_{adiabatic,850\ hPa}\ (p_s \geq 800\ hPa) \\ \Delta T_{adiabatic,750\ hPa}\ (800\ hPa > p_s \geq 700\ hPa) \\ \Delta T_{adiabatic,650\ hPa}\ (p_s < 700\ hPa) \end{cases} \quad \text{Eq. S2}$$

As Fig. S8 shows, with $\Delta T_{advective,test}$ and $\Delta T_{adiabatic,test}$ (Eqs. S1-S2) instead of $\Delta T_{advective}$ and $\Delta T_{adiabatic}$ (Eqs. 4-5), the proportion of advective and adiabatic heatwave days is slightly higher compared with Figure 1, which is consistent with higher upper tail of adiabatic and advective heating at higher pressure. But the main results stay consistent.

Finally, we repeat the analysis by using K-Medoids instead of K-Means as clustering method and the results are shown in Fig. S9, which identifies the same four types of heatwave days as K-Means.

Text S3. Normalization by the Manhattan norm

Normalization by the Manhattan norm (L1 norm normalization) involves scaling the components of a vector such that the sum of the absolute values of the components equals 1.
Given an array $x = (x_1, x_2, x_3, \dots, x_n)$, the Manhattan norm is:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \quad \text{Eq. S3}$$

To normalize the array by its Manhattan norm, each component of the vector is divided by the Manhattan norm:

$$x_{normalized} = \left(\frac{x_1}{\|x\|_1}, \frac{x_2}{\|x\|_1}, \frac{x_3}{\|x\|_1}, \dots, \frac{x_n}{\|x\|_1}\right) \quad \text{Eq. S4}$$

This ensures that the sum of the absolute values of the components in the normalized array is 1:

$$\sum_{i=1}^{n} \left|\frac{x_i}{\|x\|_1}\right| = 1 \quad \text{Eq. S5}$$

For example, for an array $x = (3, -4, 5)$, the Manhattan norm would be $\|x\|_1 = |3| + |-4| + |5| = 3 + 4 + 5 = 12$, and normalized array would be $x_{normalized} = \left(\frac{3}{12}, \frac{-4}{12}, \frac{5}{12}\right) = (0.25, -0.333, 0.4167)$.

This process is useful for scaling data, particularly in machine learning and clustering tasks, to ensure uniformity in feature magnitudes.

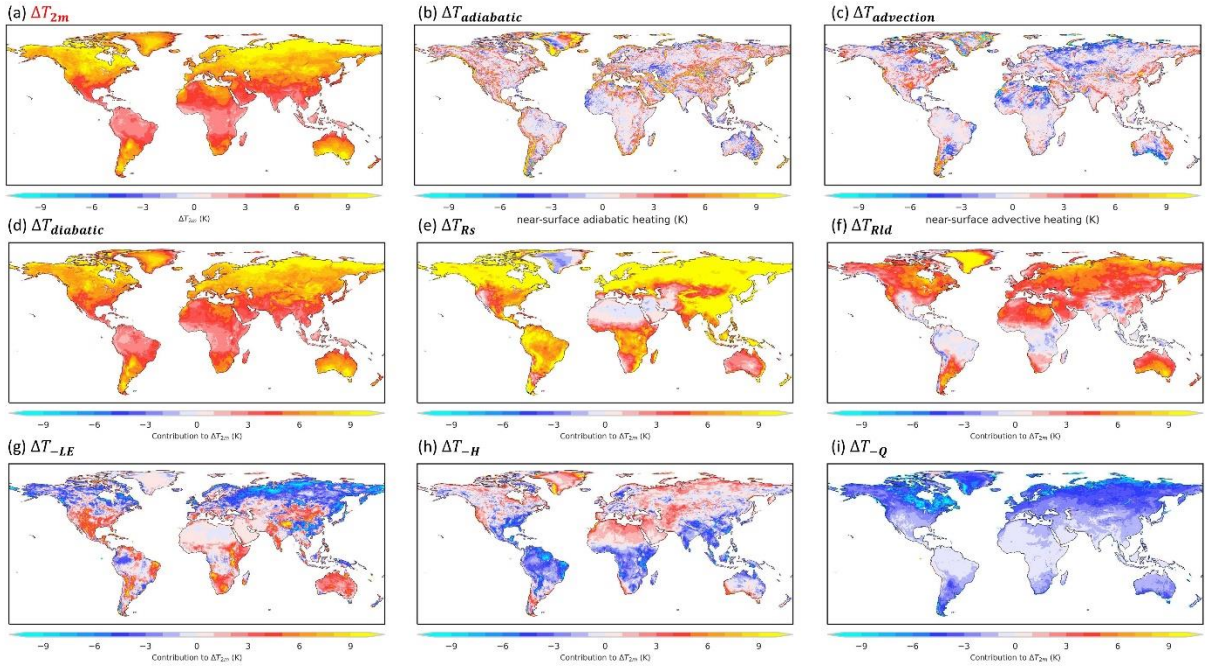

Fig. S1. Composite decomposition of the temperature anomalies during heatwave-days. (**A**) Temperature anomalies during all kinds of heatwave-days (8463177 days in total), and its decomposition into the contribution of (**B**) adiabatic heating, (**C**) advective heating, (**D**) diabatic heating, with the diabatic heating term further decomposed into contributions of (**E**) net solar radiation, (**F**) downward longwave radiation, (**G**) latent heat flux, (**H**) sensible heat flux, and (**I**) ground heat flux.

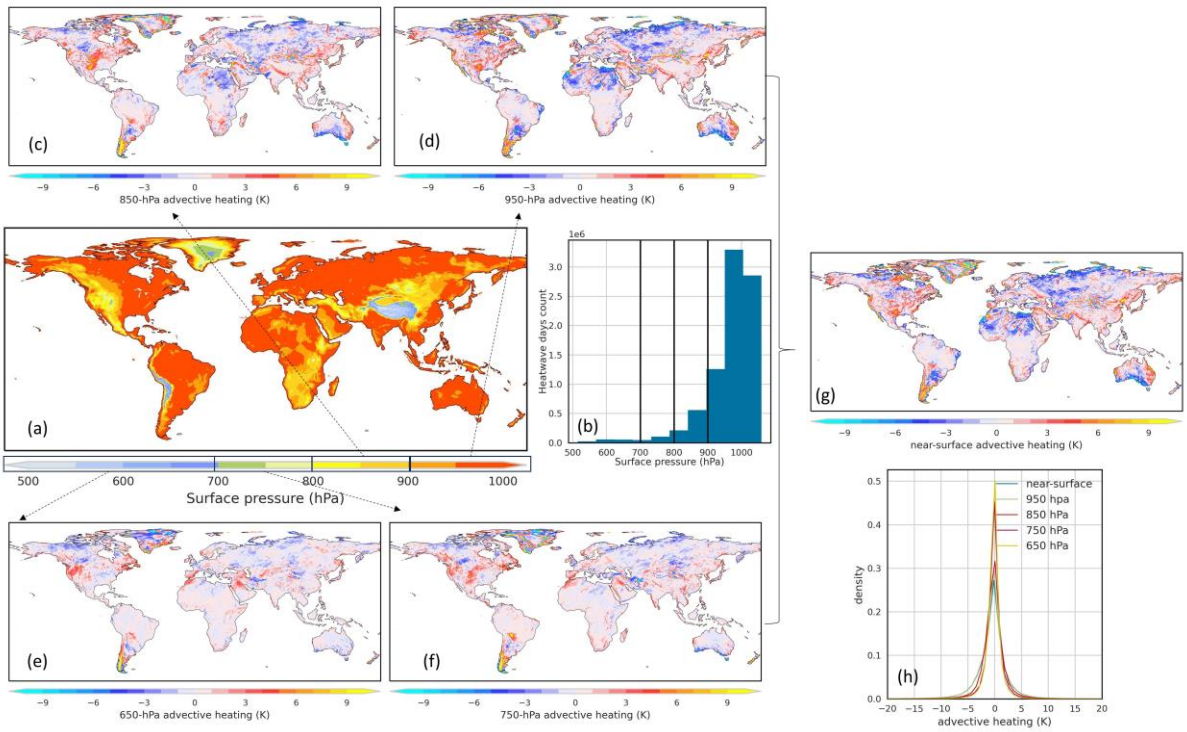

Figure. S2 Calculation of near surface advective heating. (A) Surface pressure on heatwave days. **(B**) Histogram of the surface pressure on heatwave days. **(C-F)** Anomalies in advective heating term during heatwave days on 850 hPa, 950 hPa, 650 hPa, and 750 hPa. **(G)** Anomalies in near-surface advective heating term during heatwave days. **(H)** Histogram of anomalies in advective heating term during heatwave days on pressure levels and near surface.

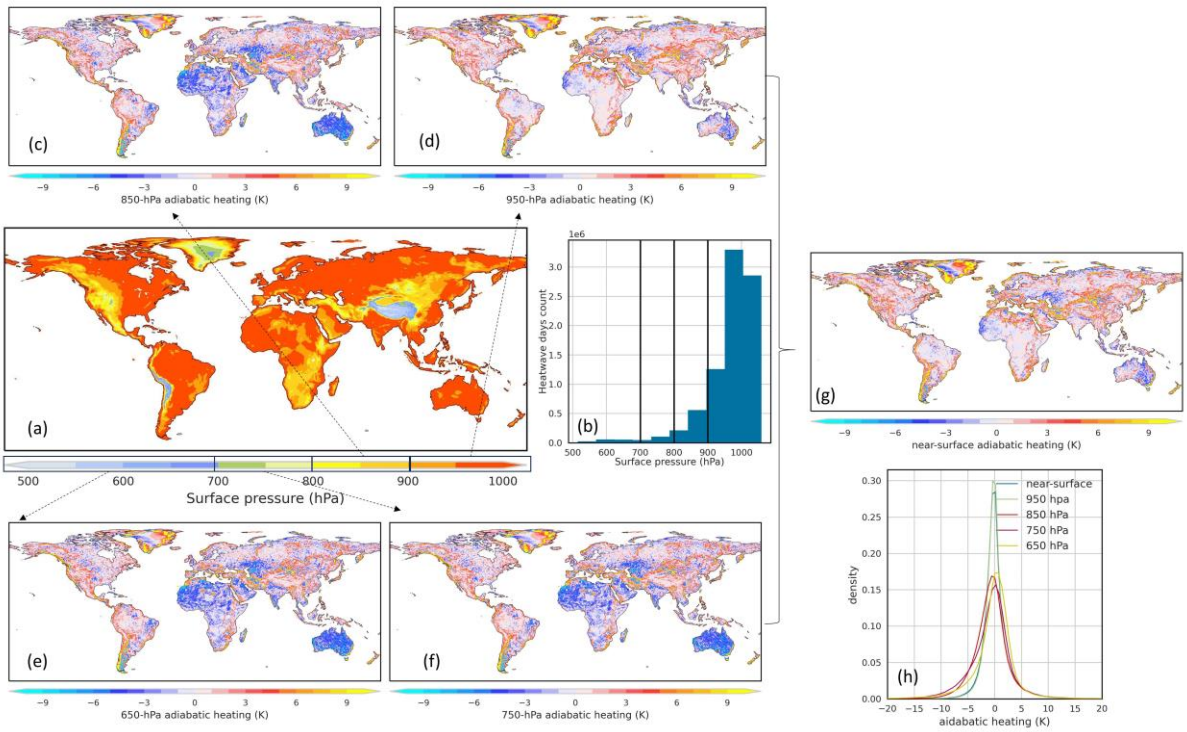

Figure. S3 Calculation of near surface adiabatic heating. (A) Surface pressure on heatwave days. (B) Histogram of the surface pressure on heatwave days. **(C-F)** Anomalies in adiabatic heating term during heatwave days on 850 hPa, 950 hPa, 650 hPa, and 750 hPa. **(G)** Anomalies in near-surface adiabatic heating term during heatwave days. **(H)** Histogram of anomalies in adiabatic heating term during heatwave days on pressure levels and near-surface.

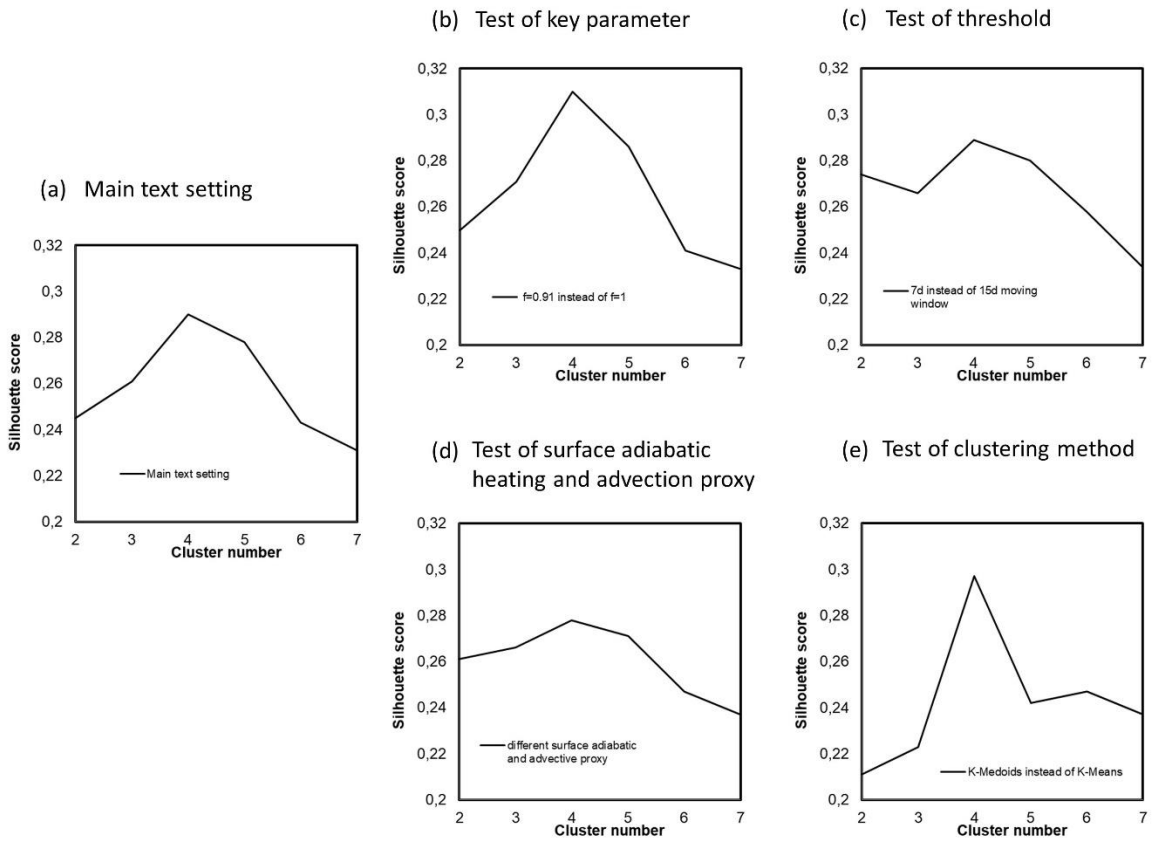

Fig. S4. Silhouette score of clustering based on different settings.

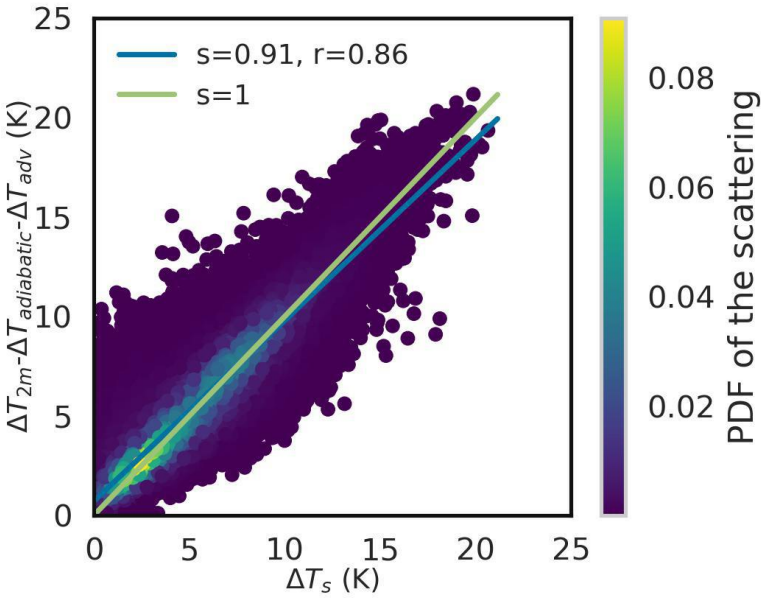

Fig. S5. The correlation between $\Delta T_{diabatic} = \Delta T_{2m} - (\Delta T_{adiabatic} + \Delta T_{advection})$ and ΔT_s during all the summer heatwave-days (8,463,177 days in total).

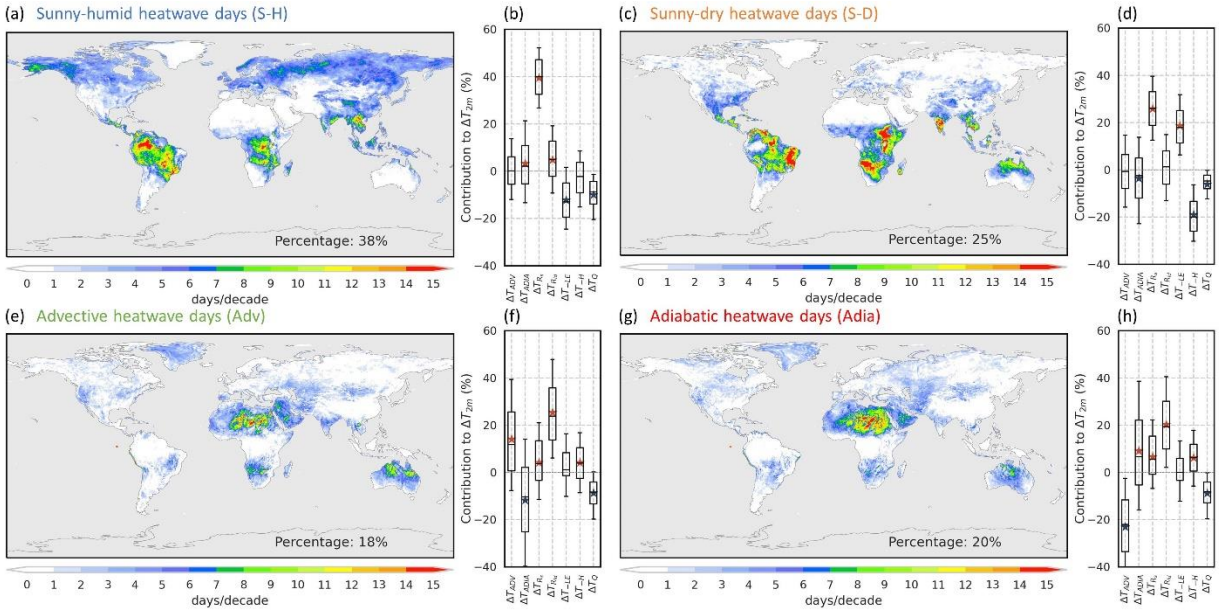

Fig. S6. The same as Figure 1 (panel a-h) but with *f*=0.91 instead of *f*=1.

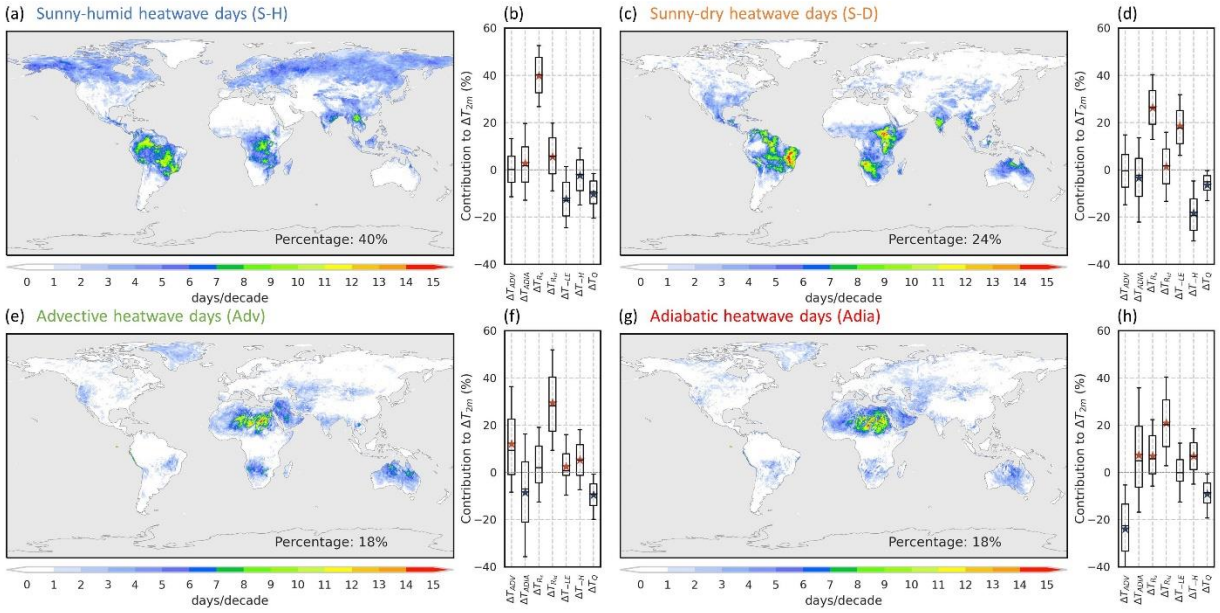

Fig. S7. The same as Figure 1 (panel a-h) but with a threshold of hot days defined based on a 7d-moving window instead of 15d-moving window.

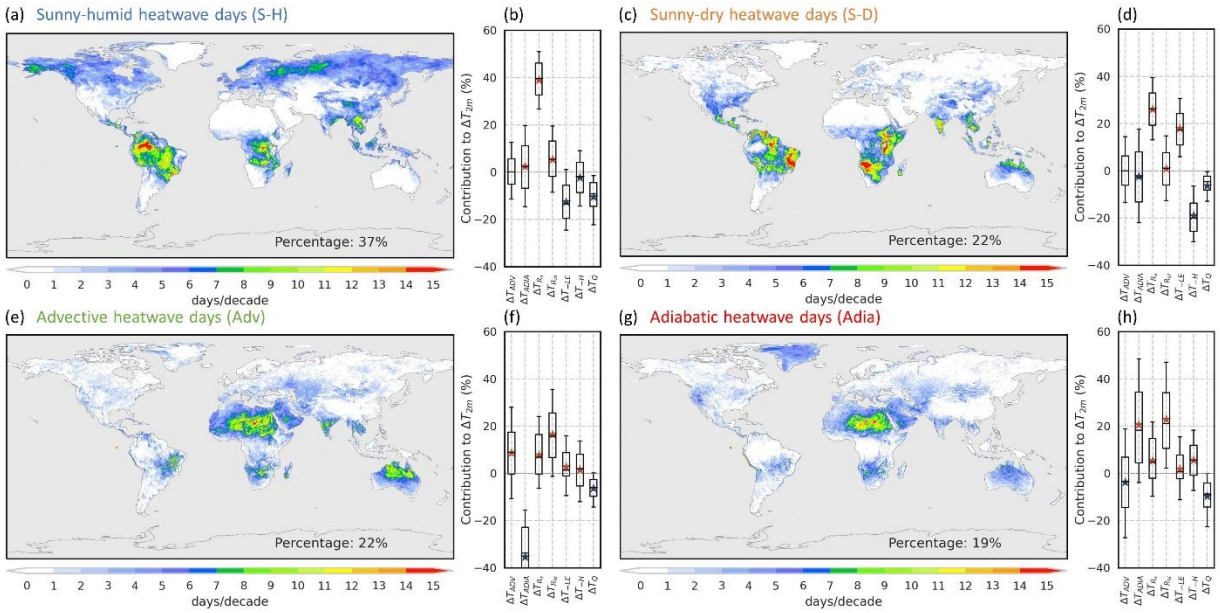

Fig. S8. The same as Figure 1 (panel a-h) but with 850 hPa instead of 950 hPa used to calculate advective and adiabatic heating term for locations with surface pressure ranging from 900 hPa to 1100 hPa, i.e., based on $\Delta T_{advective,test}$ and $\Delta T_{adiabatic,test}$ instead of $\Delta T_{advective}$ and $\Delta T_{adiabatic}$.

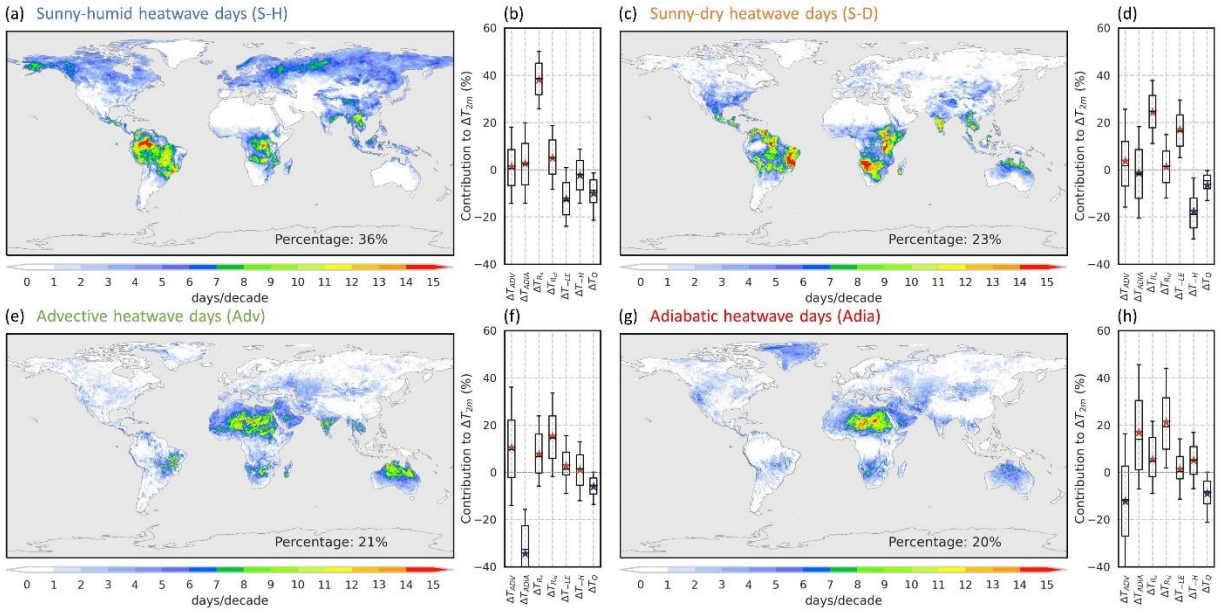

Fig. S9. The same as Figure 1 (panel a-h) but with K-Medoids instead of K-Means as clustering method

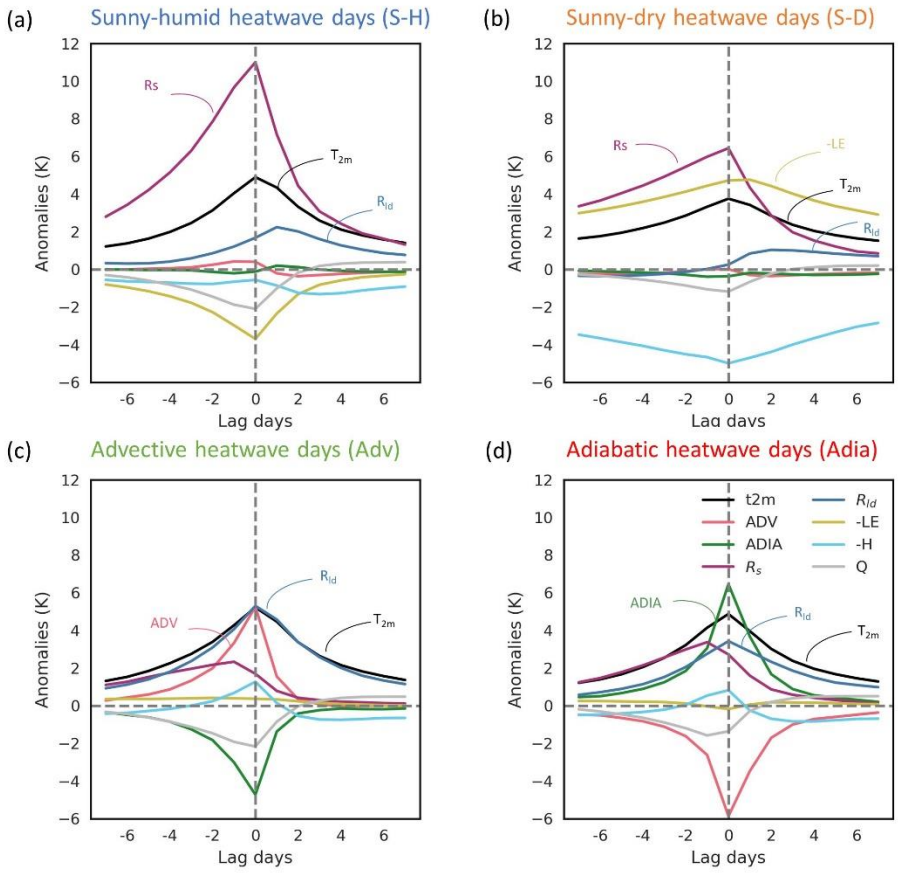

Fig. S10. **Evolution of the contribution of seven heating terms before and after the four kinds of heatwave-days.** Composite evolution of the contribution of 7 heating terms during 7 days before and 7 days after the days of (**A**) sunny-humid, (**B**) sunny-dry, (**C**) advective, and (**D**) adiabatic heatwave days.

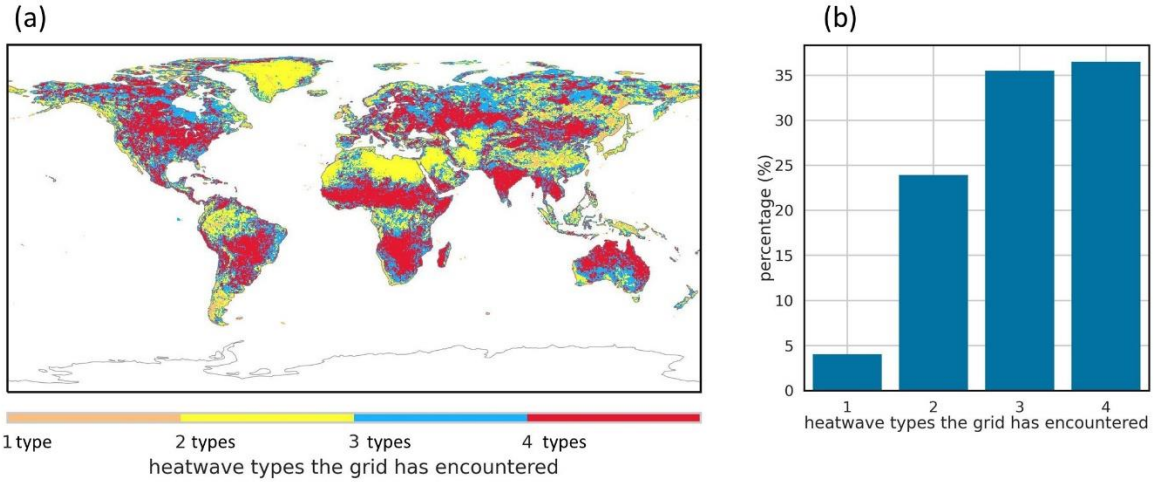

Fig. S11. Number of the heatwave types that the grids has experienced during 1979-2020. (A) The distribution. **(B)** The percentage of the grids that has encountered 1-4 types of the heatwave days.

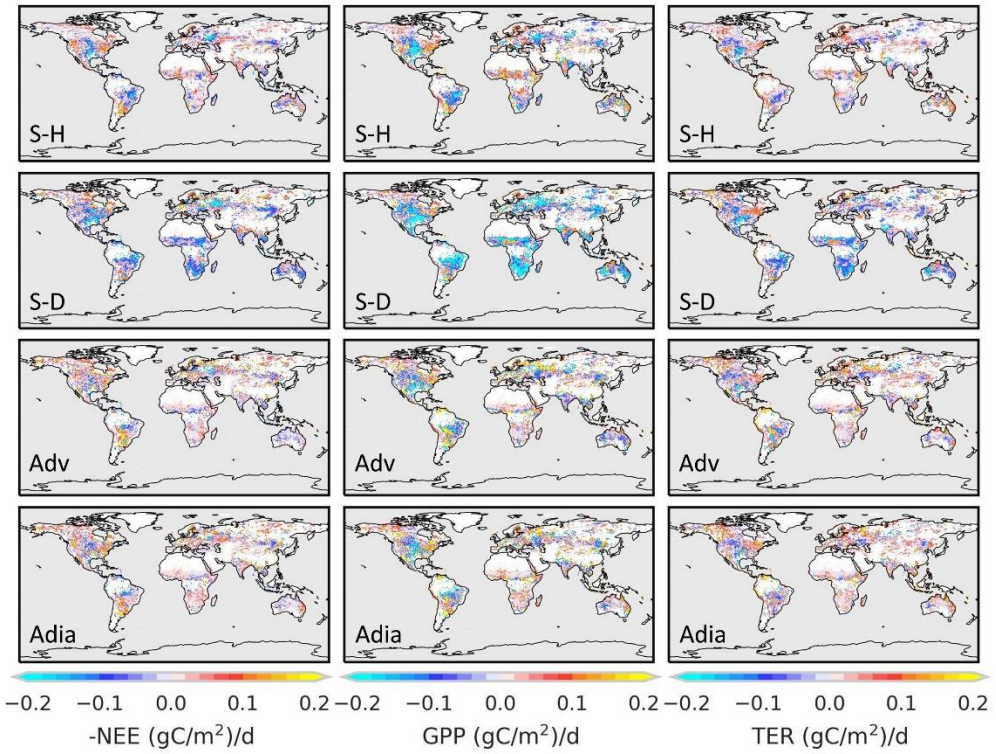

Fig. S12. The distribution of the impacts of four kinds of heatwave-days on ecology system. The maps of anomalies in -NEE (the first column), GPP (the second column), and TER (the third column), during sunny-humid (S-H, the first row), sunny-dry (S-D, the second row), advective (Adv, the third row), and adiabatic (Adia, the forth row) heatwave days. Anomaly refers to the deviation from the multi-year average of the corresponding calendar day. All the variables shown here are detrended. Values are shown where the grids have experienced all four kinds of heatwave days during 1979-2020. Contours are only shown over the harvest area (crop area fraction > 0).

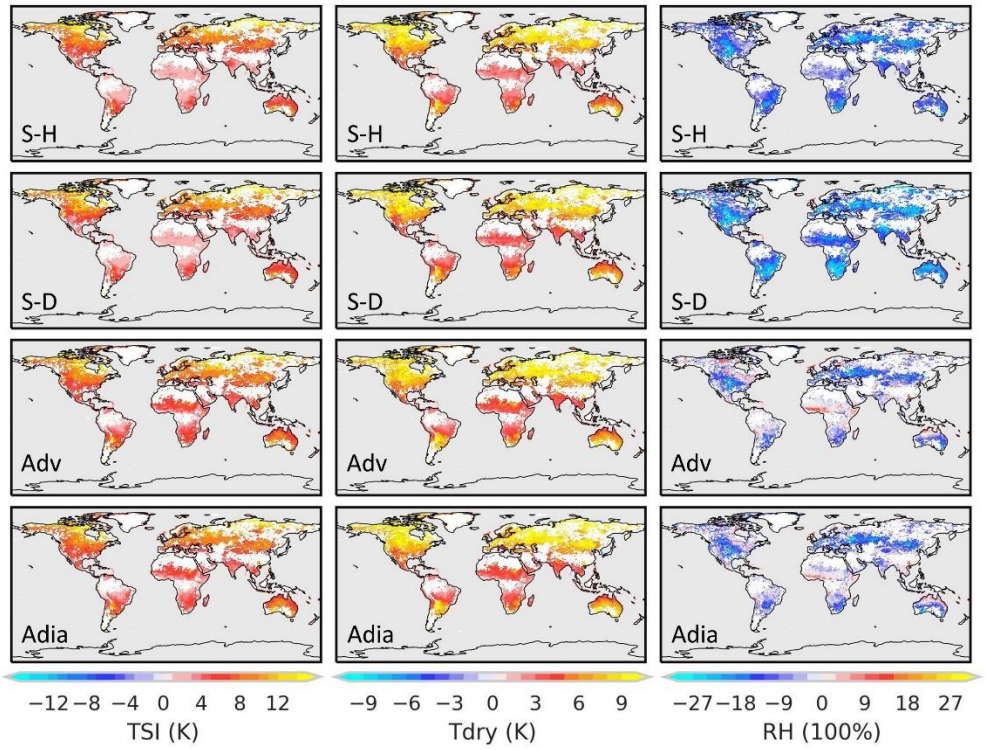

Fig. S13. The distribution of the impacts of four kinds of heatwave-days on human system. The maps of anomalies in thermal stress index (TSI, the first column), dry-bulb temperature (Tdry, the second column), and relative humidity (RH, the third column), during sunny-humid (S-H, the first row), sunny-dry (S-D, the second row), advective (Adv, the third row), and adiabatic (Adia, the forth row) heatwave days. Anomaly refers to the deviation from the multi-year average of the corresponding calendar day. All the variables shown here are detrended. Values are shown where the grids have experienced all four kinds of heatwave days during 1979-2020. Contours are only shown over populated regions (population density ≥1 person per km^2).

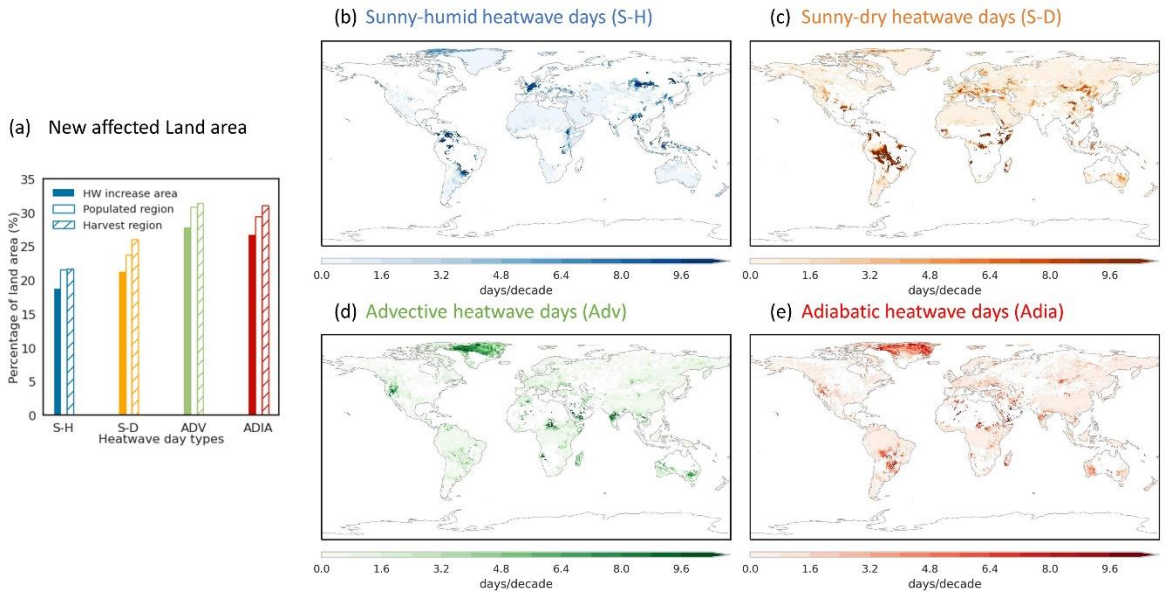

Fig. S14. New-affected area of the four types of heatwave days. (A) Percentage of the new-affected land area, i.e., the regions that experienced the corresponding type of heatwave days only after 2000. (**B**-**E**) Distribution of the difference in frequency of the four types of heatwave days between 2000-2020 and 1979-1999 over new-affected land area.