Discontinuous daily temperatures in the WATCH forcing data sets

Henning W. Rust *, Tim Kruscke, Andreas Dobler,

Madlen Fischer, Uwe Ulbrich

Institut für Meteorologie, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10, D-12165 Berlin, Germany

* *Corresponding author address:* Henning W. Rust, Institut für Meteorologie, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10, D-12165 Berlin, Germany.

E-mail: henning.rust@met.fu-berlin.de

Abstract

The WATCH forcing data sets have been created to support the use of hydrological and land surface models for the assessment of the water cycle within climate change studies. They are based on ECMWF reanalysis products (ERA-40 or ERA-Interim) with temperature (among other variables) adjusted such that their monthly means match the monthly temperature data set from the Climatic Research Unit. To this end, daily minimum, maximum and mean temperatures within one calendar month have been subjected to a correction involving monthly means of the respective month. As these corrections can be largely different for adjacent months this procedure is potentially leading to unplausible differences in daily temperatures across the boundaries of calendar months. We analyze day-to-day temperature fluctuations within and across months and find that across months differences are significantly larger, mostly in the tropics and frigid zones. Average across-months differences in daily mean temperature are typically between 10% to 40% larger than their corresponding average within-months temperature differences. However, regions with differences up to 200% can be found in the tropical Africa. Particularly in regions where snow-melt is a relevant player for hydrology, a few degrees difference can be decisive for triggering this process. Daily maximum and minimum temperatures are affected in the same regions but in a less severe way.

1. Introduction

An assessment of the global water cycle requires reliable data sets with global coverage of the meteorological variables driving the water cycle. The EU WATCH project has created global data sets meant to meet these needs: e.g., the WATCH Forcing Data 20th Century (Weedon et al. 2010, 2011, WFD,). With the same methodology the WATCH-Forcing-Data-ERA-Interim (WFDEI, Weedon et al. 2014) was created. This data set has been frequently used in the context of hydrological modeling (e.g., Gudmundsson et al. 2011; Koch et al. 2013; Prudhomme et al. 2014).

The WFD is based on the ECMWF 40-yr reanalysis (ERA-40, Uppala et al. 2005) with variables adjusted by observational products, while the WFDEI uses ERA-Interim (ERA-Interim, Dee et al. 2011) as a basis. Sub-daily temperatures are adjusted such that their monthly means match the corresponding Climatic Research Unit (CRU) dataset's monthly temperatures running from January 1958 to December 2001 (New et al. 1999, 2000; Mitchell and Jones 2005; Weedon et al. 2010, 2011). This implies potentially very different adjustments for adjacent days belonging to different calendar month. Hence, the difference in daily temperatures between the first day of month and the last day of the previous month might be very different from the day-to-day temperature difference within the same month. Ideally this day-to-day differences at the beginning or end of calendar months are of the typical size of day-to-day temperature variations and hence should neither be detectable nor pose serious problems for further application of the WATCH forcing data sets in a hydrological context. But what if the typical across-months differences are notably larger than typical within-months differences? Implausible temperature differences between adjacent days occur. To quantify this problem, we compare the distribution of day-to-day differences in daily temperatures within calendar months to differences across months for both the WFD and the WFDEI. Specifically, we count events with across-months daily temperature differences being larger in magnitude than extreme within-months daily temperature differences and we compare the magnitude of across-months temperature differences to the corresponding within-months differences in total and on a monthly resolved basis.

The problem of discontinuities in adjusted (bias-corrected) reanalyses has been discussed, e.g., in (Hempel et al. 2013). Hagemann et al. (2011) and Piani et al. (2010) pointed already to potential jumps between months and suggested a continuous correction.

The WATCH forcing data sets (WFD and WFDEI) as well as the methods used to detect and quantify untypical daily temperature differences are presented in Sec. 2. Section 3 shows the affected regions and quantifies the resulting discontinuities for the whole datasets and for transitions from specific month to the subsequent one; conclusions are given in Sec. 4.

2. Data and Methods

We consider mean (tas), minimum (tasmin) and maximum (tasmax) near-surface temperature (2m) from the WFD and WFDEI. The data can be obtained from the International Institute for Applied Systems Analysis (`ftp://rfdata@ftp.iiasa.ac.at/`).

Plotting interquartile ranges (IQR, difference between third and first quartile) of daily mean temperature across the year for a grid-box from Ethiopia in the WFD data set provides a first visual impression of the problem (Fig. 1): The larger than usual differences between daily temperature at the boundaries of months become particularly evident between March and April, May and June, September and October, as well as November and December; these are of the size of the inter-quartile range of within-month temperature differences.

In the following, we compare day-to-day temperature fluctuations between days i and $i-1$

$$\Delta T_i = T_i - T_{i-1} \tag{1}$$

within months ($\Delta T_{i,\mathsf{in}}$) to those across months ($\Delta T_{i,\mathsf{across}}$). Figure 2 shows the histograms of absolute temperature differences $|\Delta T_{i,\mathsf{in}}|$ ($N_{\mathsf{in}} = 45 \cdot (365 - 12)$, gray bars) and $|\Delta T_{i,\mathsf{across}}|$ ($N_{\mathsf{across}} = 45 \cdot 12$, orange bars) as well as the differences in their sample means (dashed lines) for the same grid-box in Ethiopia used for Fig. 1. The 0.95-quantile of inner-monthly daily fluctuations is marked as a solid vertical line and we classify values below this line as *normal* day-to-day variations, which is expected to be exceeded in only 5% of day-to-day variations. For the given example grid-box, we find 24.5% of across moths day-to-day variations exceeding this line and thus more than the expected 5%. To assess whether this number of exceedances is significantly different from the expected 5% inner-monthly exceedances, we construct the following hypothesis test: under the null hypothesis H_0 we assume that $|\Delta T_{i,\mathsf{across}}|$ are realizations of independent and identically distributed random variables (iid) with the same distribution as $|\Delta T_{i,\mathsf{in}}|$. We further assume the sample estimate of the 0.95-quantile being an adequate estimate of the true quantile and, hence, a probability of $p = 0.05$ for $|\Delta T_{i,\mathsf{across}}|$ exceeding it. For 45 years with 12 months, we have $N = 45 \times 12 = 540$ beginnings of months, i.e. N trials of a Bernoulli experiment with probability $p = 0.05$ for exceeding this quantile. We thus expect 27 exceedances and deduce from the binomial distribution that the number of exceedances is smaller than 36 for 95% of all trials. These

36 exceedances correspond to about 7% of $N = 540$ trials and hence 7% marks a critical value which we consider as not being consistent with H_0 at a 5%-level of significance.

Additionally, we analyze the direction of the discontinuities, that is whether the across-month temperature differences are positive or negative on average. These average across-month temperature differences $\overline{\Delta T_{m,\mathsf{across}}}$ are compared separately for every month m to *normal* variations, estimated from the temperature differences before and after the transitions across months

$$\overline{\Delta T_{m,\mathsf{in}}} = \frac{1}{2n} \sum_{y=1}^{n} (T_{m,f-1,y} - T_{m,f-2,y} + T_{m,f+1,y} - T_{m,f,y}) \tag{2}$$

with the indices (m, f, y) and $(m, f+1, y)$ denoting the first (f) and second $(f+1)$ day of calendar month m in year y, respectively; $f-1$ and $f-2$ are thus the last and second last day of the previous month. Additionally, we consider the normalized difference in mean values of within-months and across-months fluctuations

$$t_m = \frac{\overline{\Delta T_{m,\mathsf{across}}} - \overline{\Delta T_{m,\mathsf{in}}}}{\sqrt{\frac{s^2_{\Delta T_{m,\mathsf{across}}}}{n} + \frac{s^2_{\Delta T_{m,\mathsf{in}}}}{2n}}} \tag{3}$$

with $s^2_{\cdot}$ being the associated sample variances for month m and n the number of years available.

3. Results

For every grid-box in the WFD and WFDEI data set, we obtain fractions of days with absolute across-month temperature fluctuations $\Delta T_{i,\mathsf{across}}$ above the estimated 0.95-quantile of absolute within-months daily fluctuations $\Delta T_{i,\mathsf{in}}$. These fractions (in percent) are depicted in Fig. 3 for minimum (tasmin), mean (tas) and maximum (tasmax) daily temperature. For both data sets, regions with very pronounced differences in the mean temperature (tas, Fig. 3, middle row), i.e. fractions of 20% or larger, are the horn of Africa, the south of the Arabian Peninsula and Angola. Large parts of Africa, South America, Greenland, Siberia, India and South-East Asia are also affected but in a less severe way. Minimum (tasmin, Fig. 3, top row) and maximum (tasmax, Fig. 3, bottom row) daily temperature show the same regional patterns but in a less pronounced way. Additionally, Fig. 4 gives ratios of mean absolute values $\overline{|\Delta T_{\mathsf{across}}|}/\overline{|\Delta T_{\mathsf{in}}|}$ for the daily minimum, mean and maximum temperatures. The above mentioned regions show ratios of 2 and more, indicating that the average across-months fluctuation is of twice the magnitude than the average within-months fluctuation. Corresponding results for minimum (tasmin, Fig. 4, top row) and maximum (tasmax, Fig. 4, bottom row) daily temperatures again show basically the same spatial patterns. The differences in magnitude between across-months and within-months is, however, not as pronounced as for daily mean temperatures. Day-to-day fluctuations of extreme temperatures are in general larger (more variable) than mean temperatures and thus the adjustment does not lead to outstanding across-months fluctuations that easily. Regions which are little or not affected (i.e. a fraction of less than 7%) roughly correspond to areas where the CRU temperature data set profits from a particularly good coverage of observational data, see Mitchell and Jones (2005).

To resolve seasonal effects, Fig. 5 shows normalized differences (Eq. (3)) of across-month and surrounding within-month temperature fluctuations for every transition from one to the subsequent month for WFD. The color bar is chosen such that color starts for $|t_m| > 2$ which roughly corresponds to a two-sided t-test on a 95% level of significance (Welsh's t-test with modified degrees of freedom, von Storch and Zwiers 1999). A positive (negative) t_m indicates a positive (negative) deviation of the climatological annual cycle caused by the adjustment scheme. The affected regions are identical to those mentioned before. However, implausible positive and negative across-month fluctuations are distributed differently across the year for different regions. For a given region, there are typically transitions with both, large positive and large negative fluctuations. This is an indication that the seasonal cycle of the underlying reanalyses do not match the cycle of the CRU temperature series. Discontinuities are most severe in the transition seasons for the tropical regions, as well as the Northern Hemisphere sub-polar to polar latitudes. While the direction of discontinuities is predominantly downward in spring and upward in summer to autumn over Greenland, Siberia exhibits an opposite behavior. The Arabian Peninsula shows downward (upward) jumps from late winter through early summer (late summer and autumn). Northern and Southern Africa exhibit similar signals, equatorial Africa an opposite behavior. These patterns extend more or less zonally with India and Southeast Asia behaving like equatorial Africa and opposite signs for South America and Northern Australia. In order to allow a first assessment whether these statistically significant implausible across-months fluctuations are relevant for further analyses to be build upon, Fig. 6 shows non-normalized differences $\overline{\Delta T_{m,\mathsf{across}}} - \overline{\Delta T_{m,\mathsf{in}}}$ for mean temperature which are probably more intuitive than Fig. 5. Discontinuities in the tropical regions which were found to be statistically highly significant in Fig. 5, are comparably small in absolute numbers. They hardly exceed a magnitude of ± 2K and, hence, might be neglectable in terms of potential impacts for applications such as hydrological modelling. However, results for the extra-tropics are completely different, magnitudes of across-months fluctuations are much larger here. Especially those in (boreal) autumn, winter, and spring are striking, representing widespread sudden jumps of mean temperature in the order of ±5K and beyond. In North America, Greenland and the frigid parts of Asia discontinuities of a few degrees exists (e.g., up to 7K in Greenland for December to January). A few degrees difference in these regions are responsible for triggering snow-melt and are thus important for hydrological modeling. The corresponding figure for WFDEI (not shown) depicts very similar spatial structures and values but differs in detail.

4. Conclusion

The availability of consistent and homogeneous sets of global forcing data is of great importance for hydrological modeling and climate research. Particularly, hydrological impact studies need a set of reference data with consistent temperature and precipitation. The WFD and WFDEI have been designed to meet these demands for the global land areas. However, in certain regions implausible day-to-day differences in temperature across the boundaries of calendar months arise. These result most likely from a combination of the adjustment scheme used in combination with a mismatch of the seasonal cycle of the two data sets under consideration: The adjustment scheme involves CRU-TS monthly mean values and thus adjustments change abruptly at the boundaries of calendar months if the seasonal cycles of CRU-TS and the ERA reanalyses differ in phase or amplitude. Hence, the introduction of discontinuities it is not an effect of the adjustment scheme alone, nor of the CRU temperature data set; problems occur when the seasonal cycles in the reanalysis and reference data sets differ in amplitude or phase. While this spurious effect of the adjustment scheme may not become evident in seasons and locations where "‘normal"’ inner-monthly daily temperature fluctuation are comparably large, we found noticeable and potentially problematic differences in the tropics and frigid regions.

Daily mean temperature exhibits fluctuations between two consecutive months of about 150% or more of the size of the "‘normal"’ inner-monthly daily temperature fluctuations for regions in South America and central Africa; for the Arabian peninsula and Ethiopia twice the size is found. As extreme temperatures (tasmin, tasmax) exhibit larger day-to-day temperature fluctuations than daily mean temperatures (tas), the tolerance for inhomogeneous adjustments is larger for minimum and maximum daily temperatures. The few degrees discontinuities in the frigid regions can be decisive for triggering snow-melt and are thus relevant for hydrological modeling.

According to Weedon et al. (2011, Tab. 1), the adjustment scheme based on monthly values was applied to 2-m temperature, downward shortwave radiation, rainfall rate, snowfall rate. Hence these variables are potentially affected by discontinuities in the same way if similar mismatches between their seasonal cycles in the ERA reanalyses and the reference data set exist. However, for variables as precipitation potential discontinuities are more difficult to detect as the day-to-day variability is a larger than for temperature.

Analysis of the individual months’ transitions reveals most severe discontinuities in the transition seasons, organized in approximately zonally symmetric patterns. Africa and the Arabian Peninsula exhibit the largest magnitudes of these discontinuities with downward jumps in boreal late winter to early summer and upward jumps in autumn.

A detailed assessment of related impacts on hydrological applications is beyond the scope of this study. However, we assume that such impacts exist and do imply significant consequences for hydrological applications. This is particularly reasonable when temperatures are close to 0°C and triggering of snow-melt come into play. In conclusion, any application of the WFD or WFDEI in the mentioned regions should be aware of implausible temperature fluctuations across boundaries of calendar months.

Acknowledgments.

We thank the Water and Global Change program (WATCH) for providing the WFD data set and the EU FP7 EMBRACE project for supporting the development of the WFDEI data set.

References

Dee, D. P., et al., 2011: The era-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137 (656)**, 553–597, doi:10.1002/qj.828, URL http://dx.doi.org/10.1002/qj.828.

Gudmundsson, L., L. M. Tallaksen, K. Stahl, and A. K. Fleig, 2011: Low-frequency variability of european runoff. *Hydrology & Earth System Sciences*, **15 (9)**.

Hagemann, S., C. Cui, J. O. Haerter, J. Heinke, D. Gerten, and C. Piani, 2011: Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. *J. Hydrometeor*, **12 (556–578)**.

Hempel, S., K. Frieler, L. Warszawski, J. Schewe, and F. Piontek, 2013: A trend-preserving bias correction – the isi-mip approach. *Earth Syst. Dynam.*, **4 (2)**, 219–236, doi:10.5194/esd-4-219-2013, URL http://www.earth-syst-dynam.net/4/219/2013/.

Koch, H., S. Liersch, and F. F. Hattermann, 2013: Integrating water resources management in eco-hydrological modelling. *Water Science & Technology*, **67 (7)**.

Mitchell, T. D. and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, **25 (6)**, 693–712, doi:10.1002/joc.1181, URL http://dx.doi.org/10.1002/joc.1181.

New, M., M. Hulme, and P. Jones, 1999: Representing twentieth-century space–time climate variability. part i: Development of a 1961–90 mean monthly terrestrial climatology. *Journal of Climate*, **12 (3)**.

New, M., M. Hulme, and P. Jones, 2000: Representing twentieth-century space-time climate variability. part ii: Development of 1901-96 monthly grids of terrestrial surface climate. *Journal of Climate*, **13 (13)**.

Piani, C., G. Weedon, M. Best, S. Gomes, P. Viterbo, S. Hagemann, and J. Haerter, 2010: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *J. Hydrol.*, **395 (3-4)**, 199–215.

Prudhomme, C., et al., 2014: Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences*, **111 (9)**, 3262–3267, doi:10.1073/pnas.1222473110, URL `http://www.pnas.org/content/111/9/3262.abstract`, `http://www.pnas.org/content/111/9/3262.full.pdf+html`.

Uppala, S. M., et al., 2005: The ERA-40 re-analysis. *Quart. J. R. Meteorol. Soc.*, **131**, 2961–3012, doi:10.1256/qj.04.176.

von Storch, H. and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge ; New York, X, 483 S. : Abb. ; 28 cm pp.

Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo, 2014: The wfdei meteorological forcing data set: Watch forcing data methodology applied to era-interim reanalysis data. *Water Res. Research*, **50**, doi:10.1002/2014WR015638.

Weedon, G. P., S. Gomes, P. Viterbo, H. Österle, J. C. Adam, N. Bellouin, O. Boucher, and M. Best, 2010: The watch forcing data 1958-2001: A meteorological forcing dataset for land surface-and hydrological-models. *WATCH Technical Report*, **22**.

Weedon, G. P., et al., 2011: Creation of the watch forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *J. Hydrometeor.*, **12**, 823–848, doi:http://dx.doi.org/10.1175/2011JHM1369.1.

List of Figures

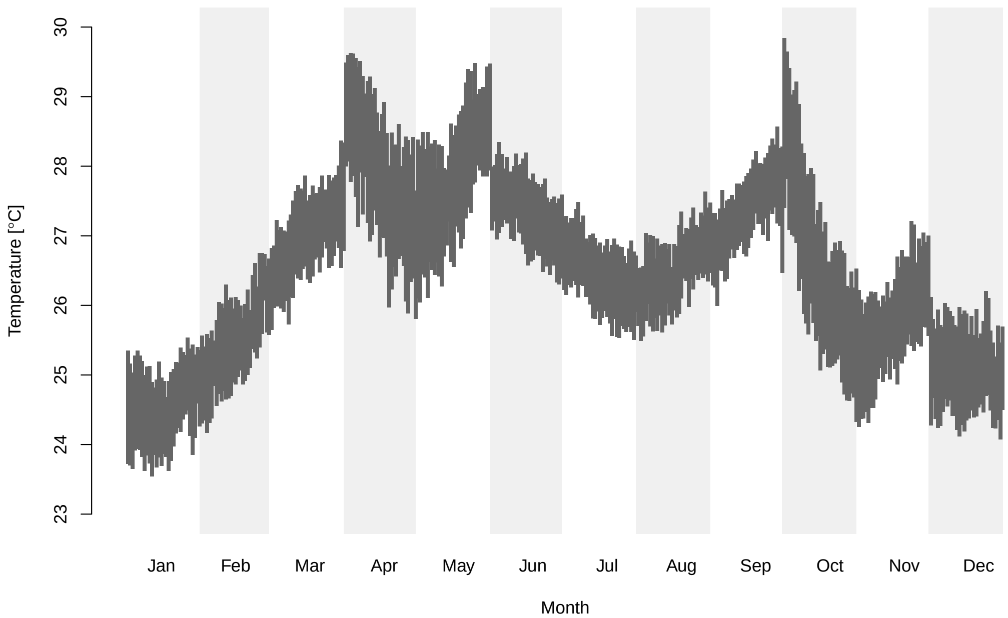

Fig. 1. Interquartile ranges (IQRs) of daily mean temperature from WDF for a grid-box in Ethiopia (40.75°E,11.25°N. The dark gray bars mark the IQR of the 45-year temperature sample of a given day in the year. Alternating gray and white shadings separate different calendar months.

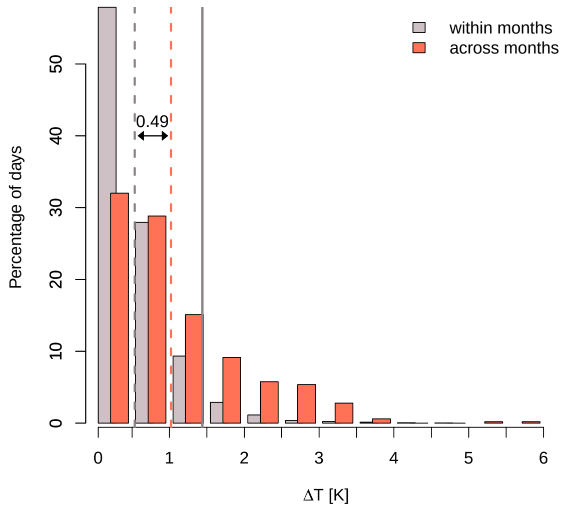

FIG. 2. Histogram of absolute daily temperature fluctuations from WDF within-months $|\Delta T_{i,\mathsf{in}}|$ (45 years with 365-12 days, gray bars) and across-months $|\Delta T_{i,\mathsf{across}}|$ (45 years with 12 days, orange bars). The dashed vertical lines mark the mean of the corresponding distributions, the solid gray line marks the 0.95 quantile of $|\Delta T_{i,\mathsf{in}}|$.

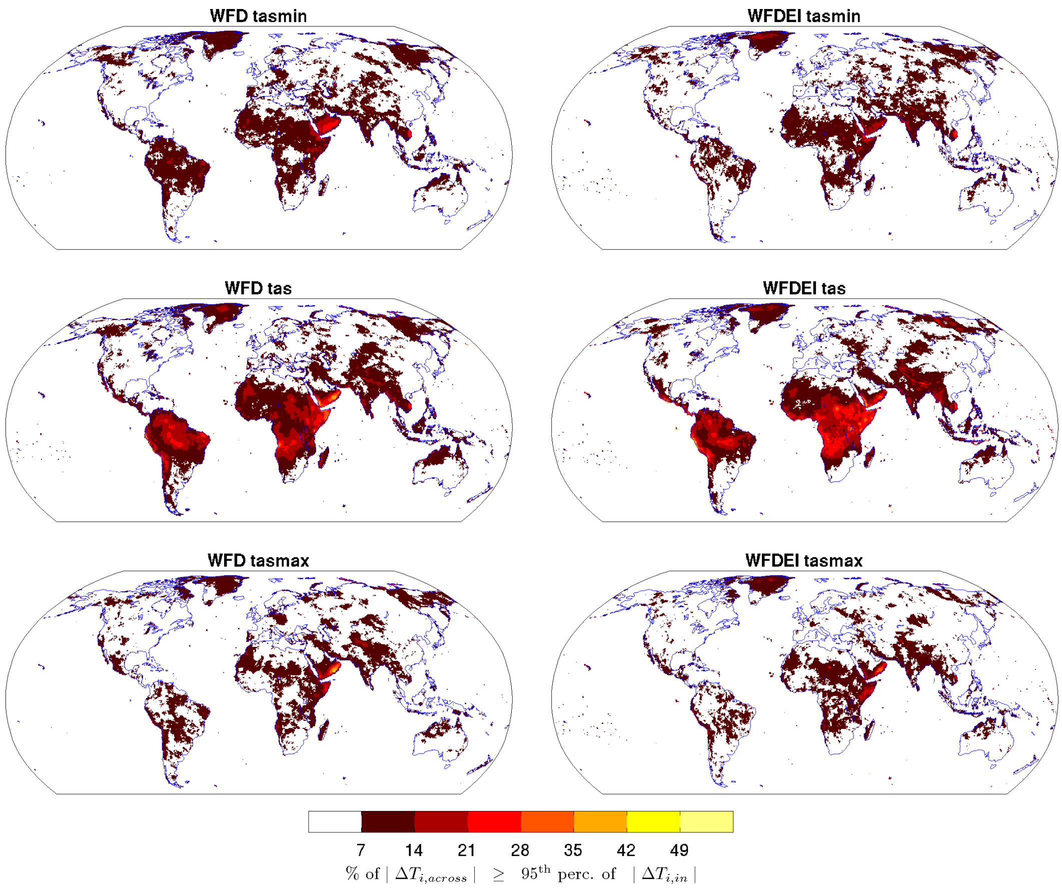

FIG. 3. Fraction of days with absolute across-months temperature fluctuations greater than the 0.95-quantile of absolute within months fluctuations for minimum daily temperature at surface (tasmin, top row), mean (tas, middle row), max (tasmax, bottom row) for WFD (left column) and WFDEI (right column).

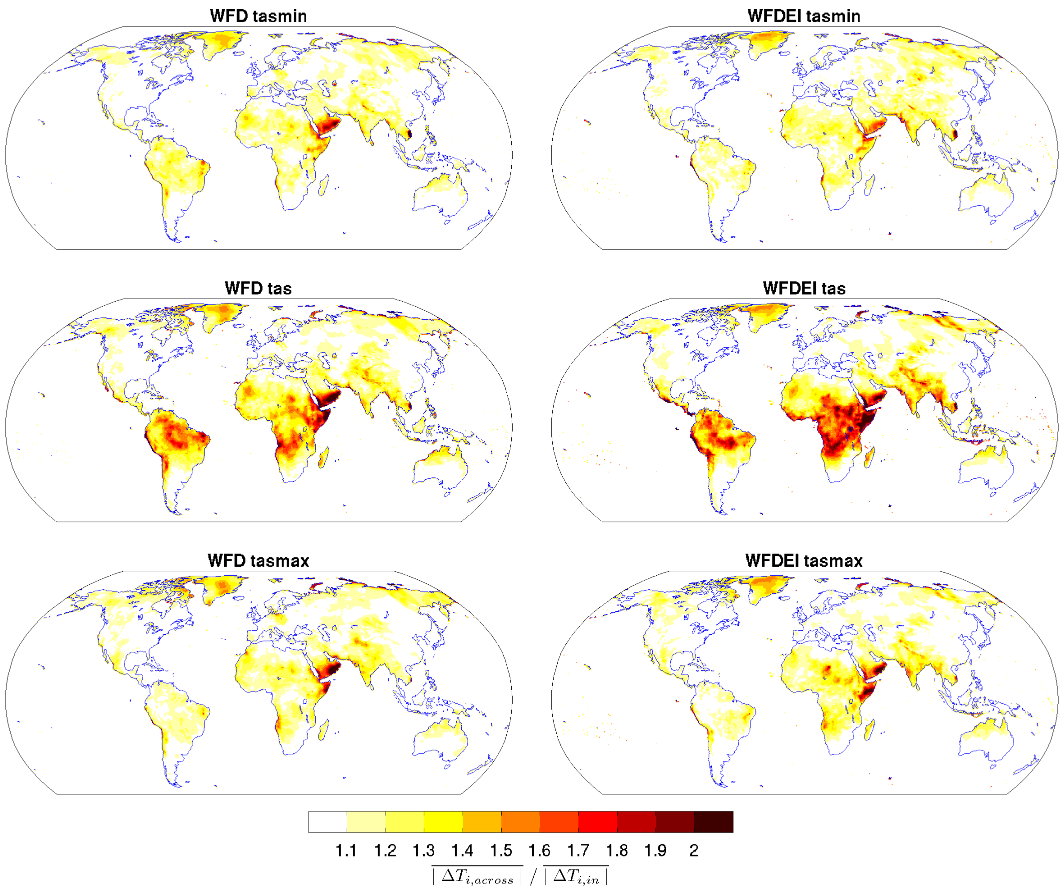

FIG. 4. Ratio of absolute across-moths and within months temperature fluctuation for minimum daily temperature at surface (tasmin, top row), mean (tas, middle row), max (tasmax, bottom row) for WFD (left column) and WFDEI (right column). Relative differences larger than 2 exists but are not depicted in separate color.

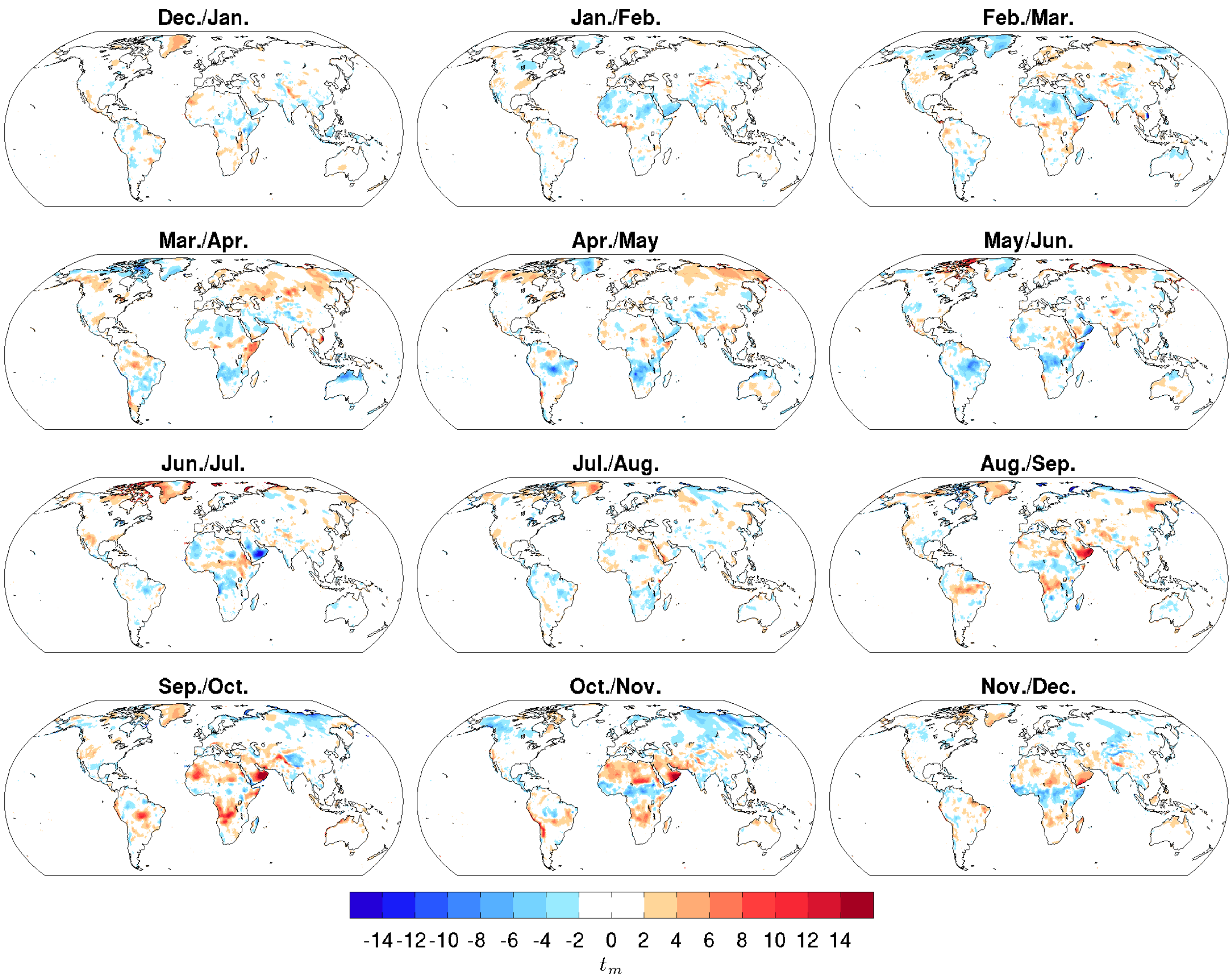

FIG. 5. Normalized mean differences between across-months and surrounding within months daily mean temperature fluctuations (Eq. (3)) for WFD, e.g. mean difference 31 Dec./1 Jan. related to mean difference 30 Dec./31 Dec. and 1 Jan./2 Jan. (top left).

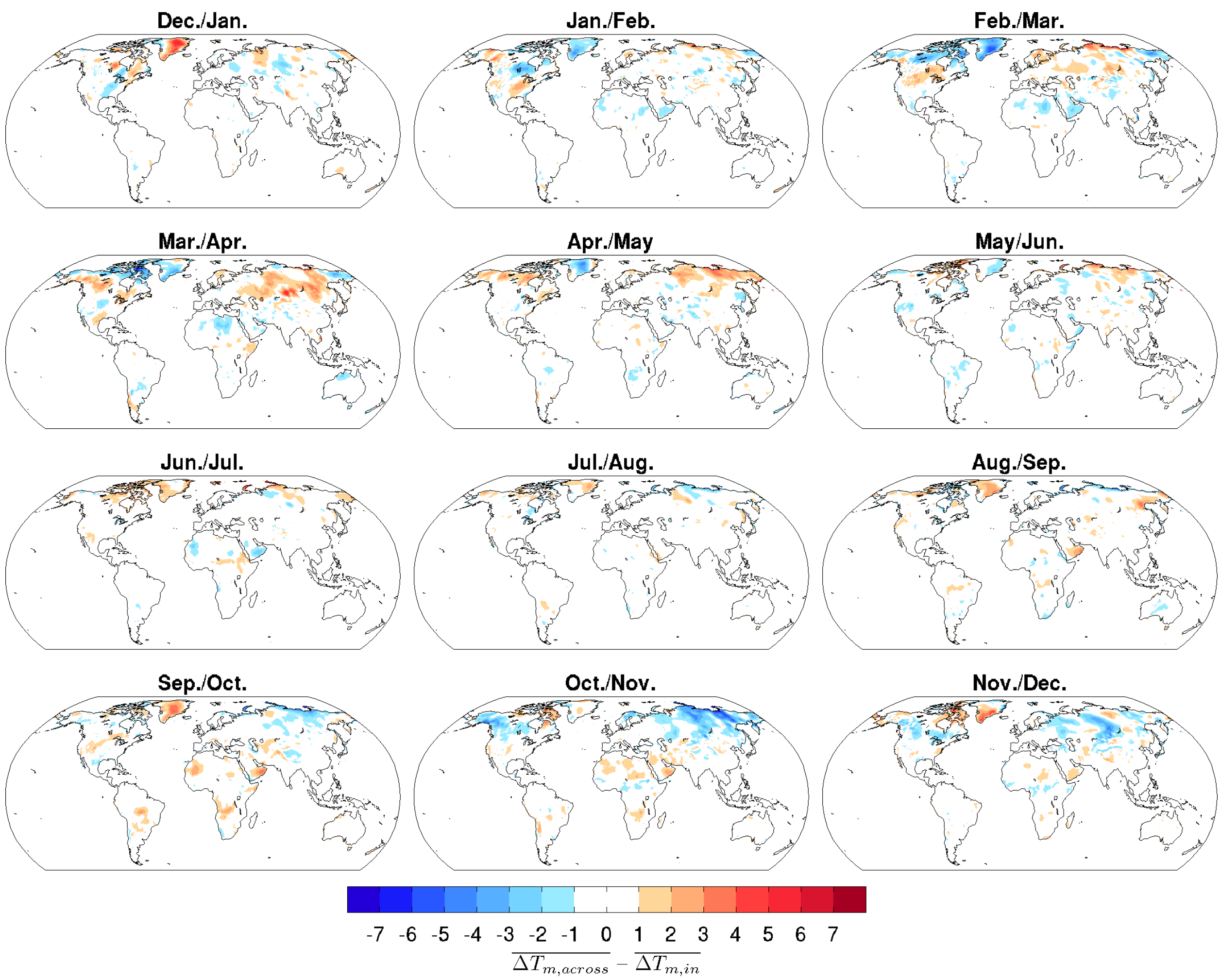

FIG. 6. Non-normalized mean differences between across-months and surrounding within months daily mean temperature fluctuations (numerator of Eq. (3)) for WFD, e.g. mean difference 31 Dec./1 Jan. related to mean difference 30 Dec./31 Dec. and 1 Jan./2 Jan. (top left).