



OPEN

DATA DESCRIPTOR

Development and validation of the CHIRTS-daily quasi-global high-resolution daily temperature data set

Andrew Verdin¹✉, Chris Funk^{2,3}✉, Pete Peterson³, Martin Landsfeld³, Cascade Tuholske^{3,4} & Kathryn Grace⁵

We present a high-resolution daily temperature data set, CHIRTS-daily, which is derived by merging the monthly Climate Hazards center InfraRed Temperature with Stations climate record with daily temperatures from version 5 of the European Centre for Medium-Range Weather Forecasts Re-Analysis. We demonstrate that remotely sensed temperature estimates may more closely represent true conditions than those that rely on interpolation, especially in regions with sparse *in situ* data. By leveraging remotely sensed infrared temperature observations, CHIRTS-daily provides estimates of 2-meter air temperature for 1983–2016 with a footprint covering 60°S–70°N. We describe this data set and perform a series of validations using station observations from two prominent climate data sources. The validations indicate high levels of accuracy, with CHIRTS-daily correlations with observations ranging from 0.7 to 0.9, and very good representation of heat wave trends.

Background & Summary

This manuscript is focused on the development and validation of CHIRTS-daily: a high-resolution ($0.05^\circ \times 0.05^\circ$) daily maximum and minimum temperature data series spanning 60°S–70°N. The CHIRTS-daily data set builds on the Climate Hazards center InfraRed Temperature with Stations T_{\max} data set (CHIRTS $_{\max}^1$) (Fig. 1a), which leverages remotely sensed infrared land surface emission temperatures and a dense global network of approximately 15,000 Berkeley Earth *in situ* station observations to provide robust high-resolution ($0.05^\circ \times 0.05^\circ$) estimates of monthly mean maximum 2-meter air temperature (T_{\max}). The CHIRTS $_{\max}$ product focuses only on maximum air temperature estimates because minimum temperatures are more difficult to distinguish from the cool cloud-top temperatures observed by satellites. For T_{\max} values, cloud screening can be used to isolate the thermal infrared land surface temperature signal, which correlates well with the monthly maximum 2-meter temperatures observed at weather stations. To support scientific modeling, as well as health, agriculture, and ecological applications that often demand finer temporal resolution, we have created a daily CHIRTS product (CHIRTS-daily) by using daily T_{\max} and T_{\min} fields from version 5 of the European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA5; <https://cds.climate.copernicus.eu/cdsapp#!/home>) to disaggregate the monthly CHIRTS $_{\max}$ product. The ERA5 data set is an advanced reanalysis that uses weather and land models, forced with satellite and *in situ* observations, to derive a complete suite of physically consistent variables describing many aspects of the Earth. The daily ERA5 diurnal temperature range (DTR; defined as maximum temperature minus minimum temperature) is then used to create CHIRTS-daily minimum temperatures. The ERA5 DTR values are based on the hourly reanalysis simulations, which incorporate physically based land and atmospheric modeling components, and hence are not affected by potential cloud contamination. Figure 1b provides a graphical representation of the CHIRTS-daily development process. The monthly CHIRTS $_{\max}$ is built around an accurate very high resolution climatology, thereby providing important spatial detail. The month-to-month variations in CHIRTS $_{\max}$

¹Minnesota Population Center, University of Minnesota Twin Cities, Minnesota, USA. ²U.S. Geological Survey, Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, USA. ³Climate Hazards Center, University of California, Santa Barbara, USA. ⁴Department of Geography, University of California, Santa Barbara, USA. ⁵Department of Geography, Environment and Society, University of Minnesota Twin Cities, Minnesota, USA. ✉e-mail: averdin@umn.edu; chris.funk@ucsb.edu

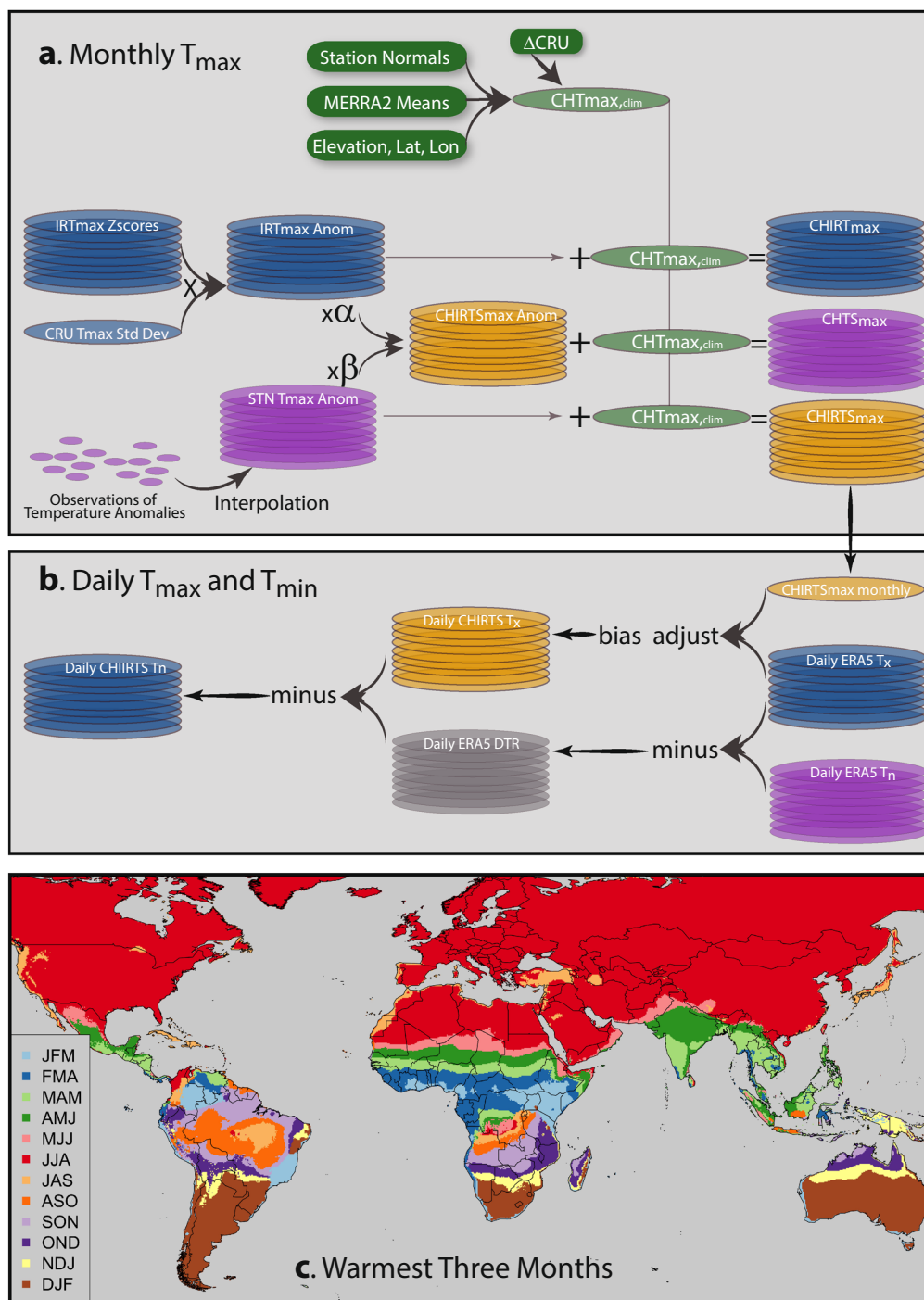


Fig. 1 Description of the CHIRTS-daily development and validation processes. **(a)** Description of the CHIRTS_{max} monthly T_{\max} development and validation processes. **(b)** Schema of CHIRTS-daily T_{\max} (T_x) and T_{\min} (T_n) development process. **(c)** Hottest three consecutive months derived from CHT_{clim} (CHIRTS_{max} climatology).

leverage the power of satellite observations to provide skillful estimates in data-sparse regions. Within any given month, the dynamically consistent ERA5 provides accurate assessments of day-to-day temperature anomalies. This approach uses the monthly CHIRTS_{max} to provide a reliable and consistent representation of the climate, and the daily ERA5 to represent weather.

Validation of the CHIRTS-daily maximum (CHIRTS_x) and minimum (CHIRTS_n) temperature products is carried out using daily climate station observations collected from the Global Historical Climatology Network (GHCN²) and the Global Summary of the Day (GSOD; <https://data.noaa.gov/dataset/global-surface-summary-of-the-day-gsod/>). We focus on validating in anomaly space to reduce the influence of the seasonal cycle on validation statistics. However, a complementary comparison of mean bias is also presented.

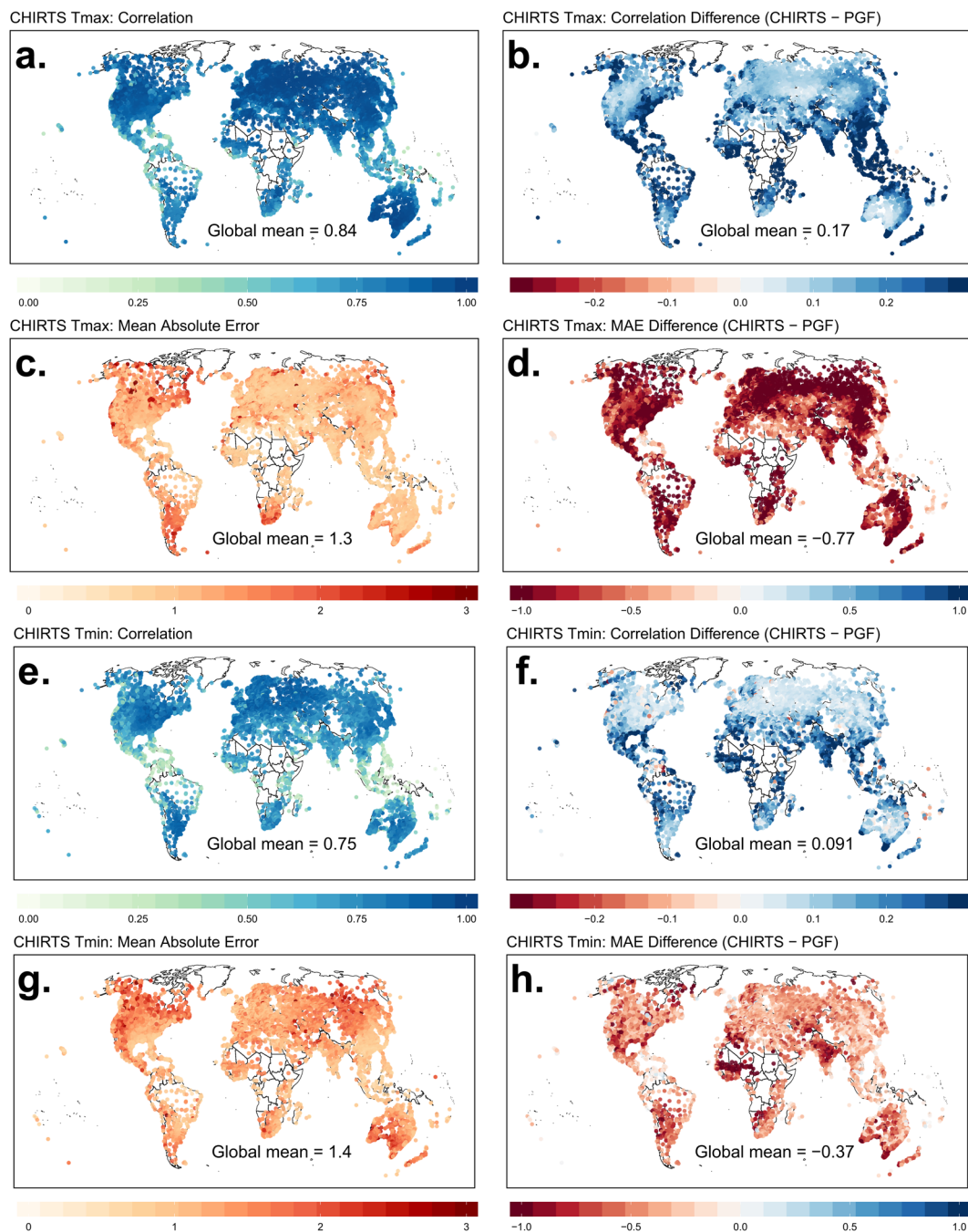


Fig. 2 Validation statistics for maximum and minimum temperatures and a comparison between CHIRTS and PGF. **(a)** Correlation between CHIRTS_X and validation data sets. **(b)** Difference between correlations for CHIRTS_X and PGF_X . **(c)** Mean absolute error of CHIRTS_X with respect to validation data sets. **(d)** Difference between mean absolute errors for CHIRTS_X and PGF_X . **(e)** Correlation between CHIRTS_N and validation data sets. **(f)** Difference between correlations for CHIRTS_N and PGF_N . **(g)** Mean absolute error of CHIRTS_N with respect to validation data sets. **(h)** Difference between mean absolute errors for CHIRTS_N and PGF_N .

This series of validations is limited to the three hottest consecutive months (Fig. 1c), identified using long-term monthly $\text{CHIRTS}_{\text{max}}$ averages. We focus on the three hottest consecutive months under the assumption that many users will be interested in performance during the hottest times of the year. For each GHCN and GSOD station, CHIRTS_X and CHIRTS_N estimates for the nearest pixel are extracted and validation measures computed.

The performance of CHIRTS_X and CHIRTS_N is compared to minimum and maximum temperatures from the Princeton Global Forcing (PGF^3) data set, which is produced by the Terrestrial Hydrology Group at Princeton University. We find that CHIRTS-daily exhibits greater skill in capturing temperature anomalies, most notably in data-sparse regions (Fig. 2). As discussed at length in the $\text{CHIRTS}_{\text{max}}$ manuscript¹, the number of global *in situ* station observations is low and in decline—in fact, much of Earth is a data-sparse region (Fig. 3).

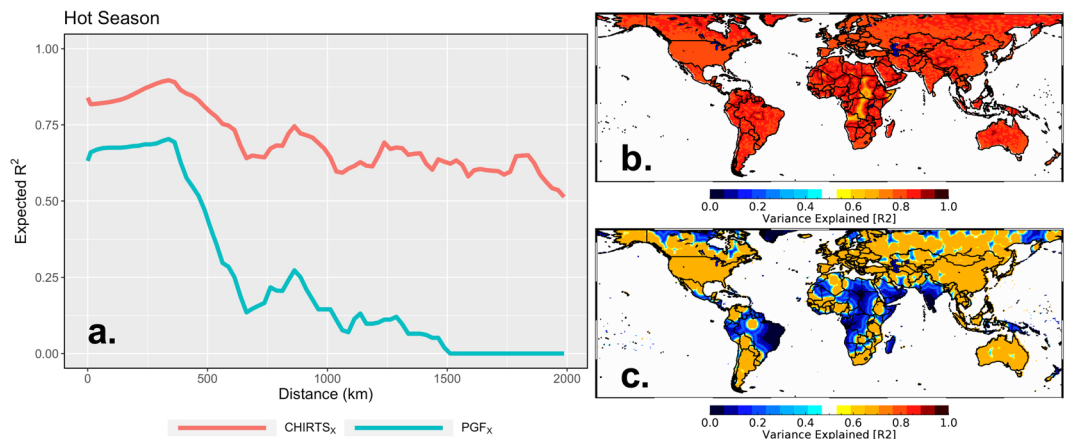


Fig. 3 Performance of CHIRTS_x and PGF_x, expressed as variance explained (R^2) as a function of distance from the nearest CRU station, for the hottest three months. **(a)** Expected R^2 as a function of distance from the grid cell to the nearest station. **(b)** Expected R^2 of CHIRTS_x based on distribution of Berkeley Earth stations for January 2016. **(c)** Expected R^2 of PGF_x based on distribution of CRU stations for January 2016.

Spatial covariograms (Fig. 3a) quantify the performance of both products as a function of distance from the nearest station used in the gridded data set from the Climatic Research Unit at the University of East Anglia (CRU⁴). The x-axis of this graph shows binned distance values, and the vertical axis displays variance explained (R^2) values. Performance (R^2) for both products decays as the distance from the nearest CRU station increases, but this decay is much less in the CHIRTS-daily product. We use the CRU station archive because it is generally considered to be the gold standard, and because PGF is bias corrected using the monthly CRU. Subsequently, we produce maps of expected performance based on each product's station density for January 2016 (Fig. 3b,c). These maps are based on the empirical covariograms shown in Fig. 3a. CHIRTS-daily is bolstered by not only the ~15,000 monthly Berkeley Earth station observations, but also the thermal infrared backbone from the monthly CHIRTS_{max}, which better captures temperature anomalies in data-sparse regions. Areas where CHIRTS-daily begins to have relatively poor performance include a relatively large swath across central Africa, the Horn of Africa, and a small bubble in northern Mali. However, relatively poor performance for CHIRTS-daily indicates an expected R^2 of about 0.60, whereas the expected performance of PGF in these same regions is closer to an expected R^2 of 0.20.

Methods

Motivation. Assessing weather-related hazards in a changing climate requires data sets that have accurate high-resolution spatial mean fields, good performance in data-sparse regions, and limited sources of non-stationary errors. Many scientists are interested in the impacts associated with a degree or two of warming. But non-stationary errors—errors that create time-varying biases in data sets—can easily be as large or larger than this climate change signal. Spatial fidelity also matters. High-resolution mean fields are important because impacts to health, agriculture, and other sectors are always local and typically non-linear. Impacts on humans or crops will be related to extremes in specific locations, and these impacts will often be strongly related to the variations in the absolute value of the weather variable under consideration. A 1- or 2-degree change in mean temperature, for example, can dramatically alter the number of days exceeding some specific temperature threshold. This spatial accuracy is important for monitoring extreme events on a year-to-year basis, and for assessing the impacts of climate change. As demonstrated below, data sets that are too cool may also underestimate changes in the frequency of heat waves.

Another important consideration is performance in areas with low densities of publicly available weather stations. Climate hazards typically convolve weather-related shocks, human exposure, and human vulnerability. The most vulnerable populations, and those with the most rapidly expanding populations and exposure, are often in areas (like Africa, Central America, and parts of Asia) with few available *in situ* weather observations.

A third consideration is consistency and stationarity, both of which can be important in data-sparse areas, where the spatial location and density of available *in situ* observations changes over time, and typically declines. When weather station observations are blended with spatially explicit background fields, non-stationary systematic errors can arise, either through changes in the observational network and/or discontinuities in the spatially explicit background fields. Accurate high-resolution mean fields can reduce homogeneities arising from shifts in observational networks. When the spatially explicit background fields track closely with the *in situ* observations, disruptions associated with changing networks are minimized. Spurious systematic non-stationary errors in the background fields, however, can create large incorrect changes.

In general, there are two main approaches to overcoming these limitations: the creation of tailored data sets that combine satellite proxy information with weather station observations, or alternately, the use of ready-made modern reanalysis systems.

Drought early warning systems, especially those focused on data-sparse regions, must grapple with these issues. While the accurate and early identification of drought conditions can trigger mitigation activities that save

lives and livelihoods, consistent, accurate, high resolution data sets are required to make such assessments. For 20 years, scientists at the University of California, Santa Barbara's (UCSB) Climate Hazards Center (CHC) have focused on developing high-quality precipitation estimates suitable for supporting famine early warning, and crop and hydrologic modeling in data-sparse regions. The result of this work, the Climate Hazards center InfraRed Precipitation with Stations (CHIRPS⁵) is now one of the most widely used products for global drought monitoring. CHIRPS has been adopted by the World Food Program, the US Agency for International Development's (USAID) Famine Early Warning Systems Network (FEWS NET, www.fews.net), the European Union, the Food and Agricultural Organization (FAO), and a host of regional and national agencies. The spatial resolution, accuracy, and consistency of CHIRPS also make it widely useful for applications such as crop insurance and climate change studies. In a typical month, more than 700 unique users download more than 100 gigabytes of CHIRPS data. The CHIRPS data set is hosted at the CHC, whose large computational capacity arises from ongoing support for and by FEWS NET. Every month, CHIRPS, along with many other valuable environmental data sets, helps FEWS NET guide billions of humanitarian assistance dollars to millions of extremely food-insecure people.

At present, there is a dearth of accurate information supporting the monitoring and evaluation of extreme temperatures in many food-insecure regions. Such extremes can wilt crops or decimate livestock herds, setting the stage for famine. Yet our ability to track these extremes in countries without weather station observations remains limited.

To address this limitation, the CHC has developed a modeling philosophy based on a geostatistical framework that decomposes environmental variables into static mean fields and time-varying anomaly fields. There are two stages to this process: a monthly T_{\max} algorithm, described in the CHIRTS_{max} manuscript¹ and illustrated in Fig. 1a, and a daily disaggregation procedure, described in Fig. 1b and in the following sections. In the CHC's approach, great attention is given to building the high-resolution ($0.05^\circ \times 0.05^\circ$) mean fields. The CHC's method for this (Moving Window Regression, or MWR) builds localized regression models using large sets of *in situ* climate normals. Complicated statistical modeling is supported by the fact that there are typically many more available stations to estimate the long-term average conditions than there are to represent variations on a given day or month. The CHC's MWR process also makes unique use of high-resolution satellite mean fields as predictors. The MWR approach and satellite-means allow the CHC climatologies to perform well in data-sparse regions, and even in regions with complex topography.

Within the CHC's approach, temporal variations are represented by combinations of *in situ* observations and geostationary satellite-based thermal infrared (TIR) observations. The monthly CHIRTS_{max} product uses a unique cloud-screening process to produce accurate global 2-meter T_{\max} anomalies. This accuracy arises through a maximum-compositing process similar to that used in developing gridded vegetation index data sets, such as the Normalized Difference Vegetation Index. Atmospheric water vapor can cause spurious declines in greenness indices. In the case of surface temperatures, partial cloud cover can reduce the temperatures observed by a satellite. In both cases the signature of the contamination is known to suppress the signal. Taking maximum composites over a period of time, therefore, can be used to minimize contamination. For the CHIRTS_{max}, this process provides a robust global and very high resolution (0.05°) set of monthly TIR-based temperature anomalies. These anomalies, when combined with the CHC's high-resolution climatology, provide an accurate and consistent source of estimates, even when there are no nearby weather stations.

Unfortunately, the CHC's maximum compositing approach cannot work on daily data, because at daily time steps it is difficult to distinguish TIR signal contributions from the land surface and clouds with measurements from only the 11 μm band provided by the GridSat⁶ data set—hence the need for a different approach to disaggregation (Fig. 1b). For this, the CHC uses modern reanalyses. Modern reanalyses use atmospheric models and assimilation schemes to merge vast quantities of information to produce physically based syntheses that provide a complete description of the land and atmosphere. For example, ERA5 uses a four-dimensional assimilation scheme to assimilate satellite radiances from 25 infrared and microwave sources and satellite scatterometer data from four sources. This rich set of data sources provides valuable information about land surface temperatures, soil moisture, atmospheric water vapor, atmospheric air temperatures, precipitation, clouds, and atmospheric circulation anomalies.

This rich set of information, and all the benefits accruing from physically modeling the Earth's systems, provides an excellent source of information about diurnal temperature variations. At the same time, it should be recognized that the inclusion of these multiple data sources also creates a stream of input data that is heterogeneous in time. For example, many infrared and microwave-based sounders and profilers only appear late in the data record, typically arriving in the late 1990s or early 2000s. Even relatively consistent imagery coming from geostationary satellites can be substantially influenced by inter-satellite calibration issues or orbital changes in any given satellite. Reanalyses that ingest station data, furthermore, face threats related to large shifts in the station data that go into these reanalyses. Both changes in the satellite systems and observation networks can alter the local energy and water budget, potentially introducing spurious random errors.

The four sources of information used in the CHIRTS_{max} contribute in different ways (Fig. 1a,b). The spatial mean fields provide local context. The carefully validated and curated monthly station and TIR temperature anomalies provide a consistent source of climate information, carefully constructed to reduce potential non-stationary systematic errors. The Berkeley Earth organization (www.berkeleyearth.org) was founded in 2012 to collect, quality control, and analyze an integrated set of global air temperature observations. Details on this data set and methods can be found at <http://berkeleyearth.org/methodology>. Finally, ERA5 reanalysis information is used to disaggregate within a specific month, and greatly reduces any possible issues associated with changes in reanalysis inputs.

CHIRTS_{max}. The monthly CHIRTS_{max} product is the foundational data set from which the CHIRTS-daily products are developed. CHIRTS_{max} combines three components: a high-resolution ($0.05^\circ \times 0.05^\circ$) climatology,

interpolated *in situ* temperature anomaly fields, and remotely sensed infrared land surface emissions anomalies based on GridSat⁶ B1 Thermal Infrared geostationary weather satellite observations. Complete details can be found in the CHIRTS_{max} manuscript⁴, though a brief description is provided in this subsection and Fig. 1a for completeness.

There are three components that are combined to create the CHIRTS_{max}:

- (1) CHT_{clim}, a high-resolution (0.05° × 0.05°) monthly maximum temperature (T_{max}) climatology developed using Moving Window Regression⁵ with FAO station normals, ERA5 long-term average 2-meter temperatures, latitude, longitude, and elevation as predictors.
- (2) CHIRT_{max}, a high-resolution (0.05° × 0.05°) monthly time series of satellite-based T_{max} anomalies.
- (3) CHTS_{max}, a high-resolution (0.05° × 0.05°) monthly time series of interpolated monthly T_{max} anomalies based on Berkeley Earth (<http://www.berkeleyearth.org>) and Global Telecommunication System T_{max} air observations.

Let \bar{C} denote the long-term average (CHT_{clim}). Let I' and S' denote, respectively, the CHIRT_{max} and CHTS_{max} anomalies from their individual long-term means. Then, the final CHIRTS_{max} estimate T is a weighted linear combination of these three components, as follows:

$$T = \bar{C} + \alpha I' + \beta S'$$

where α and β are weights that sum to 1 and are derived using the expected variance explained by the CHTS_{max} and CHIRT_{max} estimates. The variance explained by the CHTS_{max} component is based on an empirical covariogram and the distance to the closest station. The variance explained by the CHIRT_{max} component is assumed to be 0.25. The weights α and β are proportional to these variance values. The final CHIRTS_{max} estimate, therefore, is an adjusted version of the climatology (CHT_{clim}). The adjustment is based upon a weighted combination of satellite-derived estimates of T_{max} anomalies (CHIRT_{max}) and interpolation-based estimates of station-observed T_{max} anomalies (CHTS_{max}). In data-sparse regions, the satellite-derived anomalies will receive greater weight than their interpolation-based counterparts. Conversely, in regions with high station density, the interpolation-based anomalies will receive the greater weight, effectively leveraging the strengths of both data sources.

Downscaling ERA5. Daily temperatures from the ERA5 are critical to developing the CHIRTS-daily products, as they define the relative evolution of daily temperatures within a given month. The most apparent limitation to using the ERA5 simulations in tandem with CHIRTS_{max} is the difference in spatial scales between the data products. The spatial scale of CHIRTS_{max} is approximately 5 km by 5 km (0.05° × 0.05°), while that of ERA5 is approximately 25 km by 25 km (0.25° × 0.25°). To bridge this gap and facilitate the merging of the two data products, the ERA5 simulations are downscaled using bilinear interpolation in the Interactive Data Language (IDL⁷) using the CONGRID command. Maximum and minimum temperatures for each day are treated independently in the downscaling procedure. Additionally, there is no explicit treatment of day-to-day temporal dependence. We assume that the inherent temporal autocorrelation is captured by the ERA5 simulations and is preserved in the downscaling routine. We also assume the dependence between maximum and minimum temperatures on a given day is preserved in the downscaling process. Ultimately, the decision to use the ERA5 simulations to disaggregate the monthly CHIRTS_{max} to daily scale is motivated by the latency of the product. Our goal is to provide updates to the CHIRTS-daily product with minimal delay. Collaboration with partners at the National Oceanic and Atmospheric Administration (NOAA) should ensure timely updates to the monthly CHIRTS_{max} product. These updated CHIRTS_{max} products will then be disaggregated with ERA5, providing a much-needed source of information that can be used to monitor extreme temperature conditions. These conditions can have dire impacts on human and livestock health and crops, while also setting the stage for potentially extensive wildfires.

CHIRTS-daily. To produce the CHIRTS-daily T_{max} (CHIRTS_x) values, the downscaled ERA5 T_{max} are first translated into anomalies from the monthly ERA5 T_{max} average. These daily T_{max} anomalies are then added to each month's CHIRTS_{max} value. The resulting CHIRTS_x product thus varies on monthly timescales with the CHIRTS_{max} while tracking the day-to-day variations of the ERA5 reanalysis. The ERA5 T_{max} and T_{min} are then used to determine the daily diurnal temperature range (DTR) at each 0.05° pixel. DTR is then used to produce CHIRTS-daily T_{min} (CHIRTS_y) by subtracting the DTR from CHIRTS_x.

The steps taken to produce the CHIRTS-daily temperature fields are summarized as follows.

1. Compute the DTR using the downscaled ERA5 fields:

- a. $DTR_t = ERA5_{X_t} - ERA5_{N_t}$ for $t = 1, \dots, T$

2. Convert the downscaled ERA5 T_{max} to anomalies:

- a. $ERA5_{X_t}^{m,anom} = ERA5_{X_t}^m - ERA5_{X_t}^m$ for $t = 1, \dots, T$ and $m = 1, \dots, M$

3. Apply the results of Step 2 to the monthly CHIRTS_{max} values to produce CHIRTS_x:

- a. $CHIRTS_{X_t}^m = CHIRTS_{max}^m + ERA5_{X_t}^{m,anom}$ for $t = 1, \dots, T$ and $m = 1, \dots, M$

4. Apply the results of Step 1 to $CHIRTS_x$ to produce $CHIRTS_N$:

$$a. CHIRTS_{N_t} = CHIRTS_{x_t} - DTR_t \text{ for } t = 1, \dots, T$$

In the above equations, T represents the total number of days in the CHIRTS-daily data record, where (1, 2, 3, ..., $T-2$, $T-1$, T) represents (Jan 1 1983, Jan 2 1983, Jan 3 1983, ..., Dec 29 2016, Dec 30 2016, Dec 31 2016); M represents the total number of months in the $CHIRTS_{max}$ data record, where (1, 2, 3, ..., $M-2$, $M-1$, M) represents (Jan 1983, Feb 1983, Mar 1983, ..., Oct 2016, Nov 2016, Dec 2016).

Ancillary data fields and Heat Index calculation. As a convenience to end users, several ancillary daily variables have been derived from the ERA5 archive and provided at a downscaled 0.05° resolution matching that of CHIRTS-daily. The downscaling procedure used to produce these ancillary data fields is the same used to downscale the ERA5 temperature fields (i.e., CONGRID command in IDL). These data have not received additional validation but are provided to facilitate research for the end users. The ERA5 data set, on which these fields are based, has been widely used and validated. Hourly temperature and dew point temperature from the ERA5 data set are used to estimate relative humidity (RH). These derived RH values, along with CHIRTS-daily, are used to calculate the Heat Index (HI) using the series of equations and rules provided by the National Weather Service (NWS).

Relative humidity (RH) is estimated from temperature and dew point temperature using the Magnus equation⁸, as follows:

$$RH \cong 100 * \exp\left(c * b * \frac{(TD - T)}{(c + T) * (c + TD)}\right)$$

where the constants are defined as $b = 17.625$ and $c = 243.04$ ⁸.

The HI is calculated using the equations provided by the NWS. The main HI equation is a refinement of the multiple regression analysis from a 1990 NWS Technical Attachment (SR 90-23). This regression modeled a set of estimated “apparent temperatures” based on a model of human biophysical thermal temperature equilibria⁹. The following set of equations, referred to as the Rothfusz regression, is abstracted from the NWS website (https://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml):

$$\begin{aligned} HI = & -42.379 + 2.04901523 * T + 10.14333127 * RH - 0.22475541 * T * RH \\ & - 0.00683783 * T^2 - 0.05481717 * RH^2 + 0.00122847 * T^2 * RH \\ & + 0.00085282 * T * RH^2 - 0.00000199 * T^2 * RH^2 \end{aligned}$$

where T is temperature in degrees Fahrenheit (F) and RH is relative humidity in percent. HI is the heat index expressed as an apparent temperature in degrees F. If the RH is less than 13% and the temperature is between 80 and 112 degrees F, then the following adjustment is subtracted from HI :

$$ADJ_1 = \frac{13 - RH}{4} * \sqrt{\frac{17 - \text{abs}(T - 95)}{17}}$$

If the RH is greater than 85% and the temperature is between 80 and 87 degrees F, then the following adjustment is added to HI :

$$ADJ_2 = \frac{RH - 85}{10} * \frac{87 - T}{5}$$

The Rothfusz regression is not appropriate when conditions of temperature and humidity warrant a heat index value below 80 degrees F. In those cases, we mask and flag the cells where these conditions occurred.

Data Records

The CHIRTS-daily products are available on the Climate Hazards Center website¹⁰. To the extent possible under the law, we have waived all copyright and related or neighboring rights to CHIRTS. The data have been assigned to the public domain using the Creative Commons CC0 1.0 Universal waiver. This work is published from the United States. The Climate Hazards Center website also provides access to the validation data used in this paper (<http://data.chc.ucsb.edu/products/CHIRTSdaily/v1.0/ValidationData.zip>).

At the moment, CHIRTS-daily are available as a climate data record that spans 1983–2016. The 1983 starting point corresponds with the starting point of a reasonably complete global thermal infrared geostationary satellite archive. Future efforts will update the $CHIRTS_{max}$ and CHIRTS-daily on a routine basis. We anticipate that there will be an early release CHIRTS-daily product, available with a delay of approximately 1 month. The final CHIRTS-daily products will be released with an anticipated latency of 2 months.

CHIRTS-daily data are provided in GeoTiff format, though alternate formats may be available upon request. Units are degrees Celsius.

Technical Validation

Station screening. To construct the validation data set, we leverage daily station observations from the GHCN and GSOD archives provided by NOAA’s Climate Prediction Center. A stringent and robust station screening process ensures that the validation station data are consistently reported over the $CHIRTS_{max}$ era (1983–2016). Four independent screenings are carried out: one for each source (GHCN and GSOD) and each

variable (T_{\max} and T_{\min}). Below, we describe the screening process for GHCN T_{\max} stations. This screening process is applied to provide more reliable data for comparison with CHIRTS-daily. The other three data sources (GHCN T_{\min} , GSOD T_{\max} , GSOD T_{\min}) are screened in an identical fashion.

For the GHCN T_{\max} screening process, we extract all the daily GHCN maximum temperature station data for the period 1 January 1983 to 31 December 2016, which amounts to ~115 million observations across 17,355 stations. The first step in the screening process is to exclude all stations that have fewer than 2,920 observations (an approximation of 80% reporting for at least 10 years), which leaves ~100 million observations across 12,830 stations. The remaining steps are quantitative in nature, and described in the following rules:

1. Calculate the station median and standard deviation across all days and years for each month.
2. Calculate the z-score for each day, using the following equation:

$$z_t = x_t - m_i/s_i, \quad i = 1, \dots, 12,$$

where z_t is the computed z-score for day t , x_t is the observation for day t , and m_i and s_i are the station median and standard deviation for the i^{th} month (January = 1, February = 2, etc.).

3. Remove **observations** that satisfy at least one of the following conditions:

- a. $z_t < 4.0$
- b. $z_t > 4.5$
- c. $x_t - m_i > 20^\circ\text{C}$

* These z-score checks (3a and 3b) were intended to screen out potentially mis-coded data. It is common, for example, that data can be recorded unintentionally scaled by a factor of 10. Z-scores beyond ± 4 are extremely uncommon. A larger positive z-score (4.5) was used because we expect climate change-related non-stationarity.

4. Perform a median check for false zero contamination. There are times when the median across all days and years is exactly zero and/or the number of observations that are exactly zero is excessive (~15% of the time). Remove **stations** that meet either of these criteria. Repeat two more times, computing new station medians and standard deviations each time.
5. Finally, remove **stations** that satisfy at least one of the following conditions. Note that CHT_{clim_i} denotes the monthly T_{\max} climatology for the i^{th} month, and m_i again denotes the station median for the i^{th} month.

- a. $\text{abs}(m_i - CHT_{\text{clim}_i}) > 5^\circ\text{C}$
- b. $\text{abs}(m_i - CHT_{\text{clim}_i})/s_i > 3$
- c. The number of remaining observations is less than <2,920

This leaves 58.3 million GHCN T_{\max} observations across 8,587 stations. We then repeat this screening process for GHCN T_{\min} , and again for GSOD T_{\max} and T_{\min} . What remains are 15,713 T_{\max} stations from GHCN (8,587) and GSOD (7,126), with an estimated overlap of 1,612 co-registered stations, and 13,997 T_{\min} stations from GHCN (7,747) and GSOD (6,250), with an estimated overlap of 1,390 co-registered stations. We identified a station as potentially co-registered between sources if its coordinates in one source are similar enough to be within 1 km of any set of coordinates in the other source. This method is required, as station identification codes are inconsistent between sources. Note that because the station screening processes are independent of source and variable, some stations will not necessarily be included in both (T_{\max} and T_{\min}) validation data sets.

CHIRTS_x validation results. The CHIRTS_x product is validated alongside the PGF³ T_{\max} product (hereafter PGF_x for convenience) using the GHCN and GSOD stations that passed the screening process. To remove any inherent correlation due to the seasonal cycle, all observations in the validation data sets are converted to anomalies at the monthly scale. That is, the monthly mean T_{\max} for January is subtracted from all daily January observations, and so on, for all months. This ensures that the mean of the observations at any station will be zero, and that the seasonal harmonics are removed. The same normalization is applied to the CHIRTS_x and PGF_x products. Each station in the validation data set is then paired with the nearest CHIRTS_x and PGF_x pixel, using the Euclidean distance from the station coordinates to the pixel centroid. Validation statistics are computed for the hottest three-month period, which is defined using the CHT_{clim} (see Fig. 1c).

Panels A-D in Fig. 2 show the validation statistics for CHIRTS_x in the left column and the difference between CHIRTS_x and PGF_x in the right column (computed as $CHIRTS_x - PGF_x$). CHIRTS_x consistently exhibits a stronger agreement with the validation data set than PGF_x. The prevalence of blue in panel B indicates that CHIRTS_x has consistently higher correlations with the stations than PGF_x—the global average is 0.17. Similarly, the abundance of red in panel D indicates CHIRTS_x has consistently lower mean absolute error (MAE) than PGF_x—the global average is -0.77°C . Table 1 summarizes these validation statistics by region/continent, which further illustrates the skill of CHIRTS_x.

CHIRTS_N validation results. We next compare the performance of CHIRTS_N to that of the PGF T_{\min} product (PGF_N for convenience) for the same hottest 3-month period as the CHIRTS_x validation (Fig. 1c). We did this in order to maintain consistency with the CHIRTS_x validation, and to focus our validation on performance during the hottest times of the year.

We convert the T_{\min} observations from GHCN and GSOD to monthly anomalies to remove inherent correlation. Each station is paired with the nearest CHIRTS_N and PGF_N pixels, and validation statistics are computed.

T_{\max}	Correlation		Mean Absolute Error	
	CHIRTS _X	PGF _X	CHIRTS _X	PGF _X
Global	0.84	0.66	1.3	2.2
N. America	0.80	0.63	1.6	2.4
S. America	0.72	0.45	1.3	2.2
Africa	0.81	0.57	1.1	2.0
Eurasia	0.91	0.74	1.0	1.9
Australia	0.92	0.75	1.0	2.1

Table 1. Validation statistics for maximum temperature products.

T_{\min}	Correlation		Mean Absolute Error	
	CHIRTS _N	PGF _N	CHIRTS _N	PGF _N
Global	0.75	0.65	1.4	1.8
N. America	0.78	0.71	1.4	1.8
S. America	0.64	0.48	1.2	1.6
Africa	0.67	0.46	1.2	1.8
Eurasia	0.78	0.67	1.3	1.6
Australia	0.60	0.51	1.9	2.3

Table 2. Validation statistics for minimum temperature products.

T_{\max}	Mean Bias Error (°C)	
	CHIRTS _X	PGF _X
Global	0.12	-0.24
N. America	0.01	-0.06
S. America	-0.16	-0.30
Africa	0.11	-0.23
Eurasia	0.21	-0.43
Australia	0.39	0.00

Table 3. Mean bias error in degrees Celsius for the maximum temperature validation data set. Positive (negative) values indicate the data product is consistently warmer (cooler) than observations.

Panels E-H in Fig. 2 show the validation statistics for CHIRTS_N in the left column and the difference between CHIRTS_N and PGF_N in the right column (computed as CHIRTS_N-PGF_N). Consistent with the CHIRTS_X validation, the CHIRTS_N correlations are consistently higher and the MAE are consistently lower than that of PGF_N. The global average difference in correlations between CHIRTS_N and PGF_N is 0.09; the global average difference in MAE is -0.37°. Table 2 summarizes these validation statistics by region/continent.

Mean bias error. It is important to acknowledge the fact that we are validating the CHIRTS-daily data series in anomaly space. It is therefore crucial to quantify potential mean bias error in these data sets. To this end, for every GHCN and GSOD station and every month, we compute the station climatologies and compare these values to the mean of the CHIRTS_X and PGF_X for the same collection of days. We compute these differences as CHIRTS_X-station and PGF_X-station, thus a positive (negative) value indicates a warm (cool) bias in the data product. This produces 12 difference values at every validation station, which we then collapse (average) into a mean bias error statistic. Table 3 summarizes the mean bias error globally and by region. These results indicate that CHIRTS_X is much less biased from a global perspective, and consistently better for many regions because of its low bias and good performance in data-sparse regions. However, CHIRTS_X exhibits a significant positive bias for Australia that is not present in PGF_X.

Performance in data-sparse regions. What distinguishes the CHIRTS_{max} from other gridded products is its use of remotely sensed thermal infrared temperatures. Where large gaps in station networks exist, the resulting interpolation-based gridded products are highly uncertain. Furthermore, station-based gridded products tend to revert to climatology as the distance to the nearest *in situ* observation increases. This likely underestimates the variance in data-sparse regions. It follows that the performance of CHIRTS_X in data-sparse regions is bolstered by remotely sensed observations, which inherently makes it more reliable in data-sparse regions.

To verify this assumption, we analyze the performance of CHIRTS_X and PGF_X as a function of distance from the nearest station included in the monthly CRU data product. We focus on CRU stations because the CRU product is widely considered the gold standard of gridded climate products. Additionally, the CRU gridded products are used to bias correct the PGF in its development workflow. Empirical covariograms are estimated for each

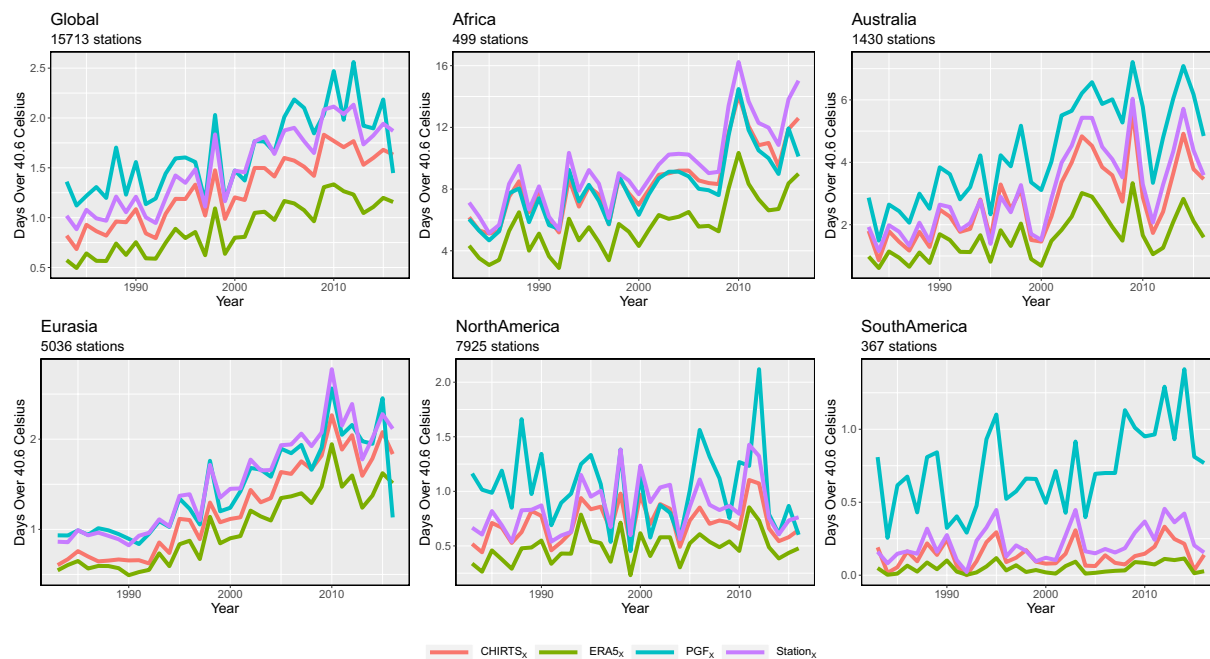


Fig. 4 Time series of the average number of days over 40.6 °C for stations in various regions based on the observations (Stations), CHIRTS_x, ERA5_x, and PGF_x data sets.

product at a set of sequential distances bins, calculated as $R^2 = 1 - \text{MSE}/\text{SS}$, where MSE is the Mean Squared Error, and SS is the Sum of Squares of the station anomalies.

Figure 3a shows the expected performance for the warmest 3 months, expressed as R^2 values for CHIRTS_x and PGF_x. When both products are assessed at a station location (i.e., at a distance of zero), CHIRTS_x has an R^2 that is about 0.20 greater than that of PGF_x. At about 300 km, the performance of both products begins to decline at the same rate. However, at about 650 km, the performance of PGF_x begins a slow, oscillating decline. The performance of CHIRTS_x hovers around $R^2 = 0.63$, while that of PGF_x is closer to $R^2 = 0.15$. At 1500 km, the performance of PGF_x drops to $R^2 = 0$, while CHIRTS_x maintains a value of $R^2 = 0.50$. It can be clearly stated that leveraging the thermal infrared temperatures in the CHIRTS_{max} product effectively enhances CHIRTS_x in data-sparse regions.

Figure 3b shows a map of the expected variance explained by CHIRTS_x based on the distribution of stations in the Berkeley Earth archive for January 2016. Figure 3c shows the same for PGF_x based on the distribution of stations in the CRU archive for January 2016. The performance of CHIRTS_x is bolstered by not only the thermal infrared background of the monthly CHIRTS_{max}—a satellite-only component of the monthly CHIRTS_{max}, which was shown to have a correlation of about 0.80 with *in situ* observations globally¹—but also a much denser distribution of stations than PGF_x, which results in a considerably more reliable data set.

Usage Notes

This section provides two examples of usage case studies, representing typical applications for the CHIRTS-daily maximum temperatures. The first application uses a 40.6 °C temperature threshold. This threshold, often used in human health studies, represents a temperature level at which the human body often has difficulty maintaining adequate cooling of internal organs. The second application uses a threshold of 30 °C, a common threshold used to identify agricultural heat stress. The first case study is global. The second case study focuses on Ethiopia, expanding on supplemental material provided in the CHIRTS_{max} manuscript¹.

Figure 4 presents a comparison of the daily station data, CHIRTS-daily, ERA5, and PGF archives, for the warmest 3 months. The CHIRTS-daily, ERA5, and PGF values have been extracted at the station locations, supporting a one-to-one comparison. What is quite striking in all the regions examined is that ERA5 dramatically underestimates the number of hot days in all regions. The PGF archive, on the other hand, substantially overestimates in Australia and South America. The CHIRTS-daily, in all cases, tracks the station-based estimates very closely. Note that the daily GHCN and GSOD data likely contributed to the monthly Berkeley Earth data used in the monthly CHIRTS_{max}, so Fig. 4 should not be interpreted as an independent validation study, which is beyond the scope of this data descriptor. Nevertheless, CHIRTS-daily does appear to be fit-for-purpose for analyzing trends in temperature extremes, especially in data-sparse regions. It should be noted that Africa stands out as a region with large increases in the number of very hot days (Table 4). While all the gridded data sets underestimate the station data change estimate (+5.7 days), the CHIRTS-daily estimate was the closest (+4.7 days). More analysis of the spatial pattern and health hazards associated with these large increases appears warranted.

We next turn to a typical agro-climatic risk assessment. Building on results presented in the supplemental material of the CHIRTS_{max} manuscript¹, this case study focuses on July T_{max} temperatures in the Amhara province

	CHIRTS	ERA5	PGF	Station
Global	+0.77	+0.55	+0.74	+0.88
N. America	+0.13	+0.14	-0.01	+0.20
S. America	+0.04	+0.03	+0.45	+0.13
Africa	+4.70	+3.40	+4.30	+5.70
Eurasia	+1.20	+0.92	+1.00	+1.20
Australia	+1.80	+0.88	+2.90	+2.20

Table 4. Change in the mean number of hot days (over 40.6 °C) between 2007–2016 and 1983–1992.

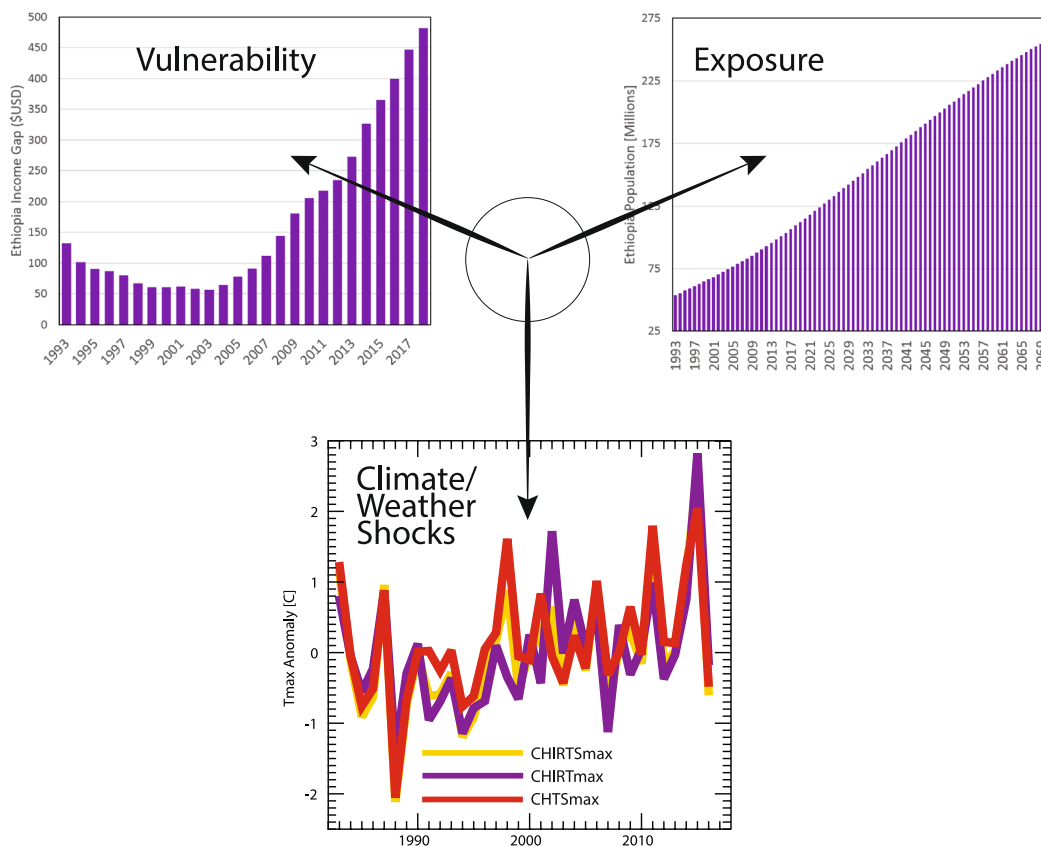


Fig. 5 A schematic representation of climate hazard dimensions for Ethiopia. Climate/weather shocks are represented by monthly July T_{\max} anomalies for the satellite-only CHIRT_{\max} data set, the station-only CHTS_{\max} archive, and the blended satellite-station CHIRTS_{\max} product. The vulnerability to price and crop production shocks is represented by the income gap between the poorest 20% of Ethiopians and the median income. Exposure is represented by national population estimates.

of northern Ethiopia. El Niño-related rainfall deficits in this region¹¹ contributed to “the worst drought in 50 years” and led to widespread crop failures that helped push approximately 11 million people into crisis levels of food insecurity.

These low rainfall levels were also accompanied by exceptionally warm July air temperatures (Fig. 5). The unique nature of the monthly CHIRTS_{\max} archive provides independent satellite-only temperature estimates (CHIRT_{\max}) and station-only temperature estimates (CHTS_{\max}), as well as the blended “best estimate” CHIRTS_{\max} product. Humanitarian relief agencies often use a convergence-of-evidence approach to guide drought assessments. The fact that the independent CHIRT_{\max} and CHTS_{\max} archives both indicated historically extreme air temperatures provides convergent evidence of severe potential crop stress.

Climate hazards typically involve this type of climate shock with underlying vulnerability and exposure—Fig. 5 represents this schematically. Vulnerability is represented by the gap between the median and 20th percentile World Bank per capita incomes. Despite increases in agricultural productivity, the number of extremely food insecure Ethiopians has increased. This increase may be due to a series of recent climate shocks, combined with an increased price-spike vulnerability, for poorer Ethiopians (<http://blog.chc.ucsb.edu/?p=634>). Figure 5 represents the third dimension of climate hazards—exposure—using United Nation estimates of the Ethiopian

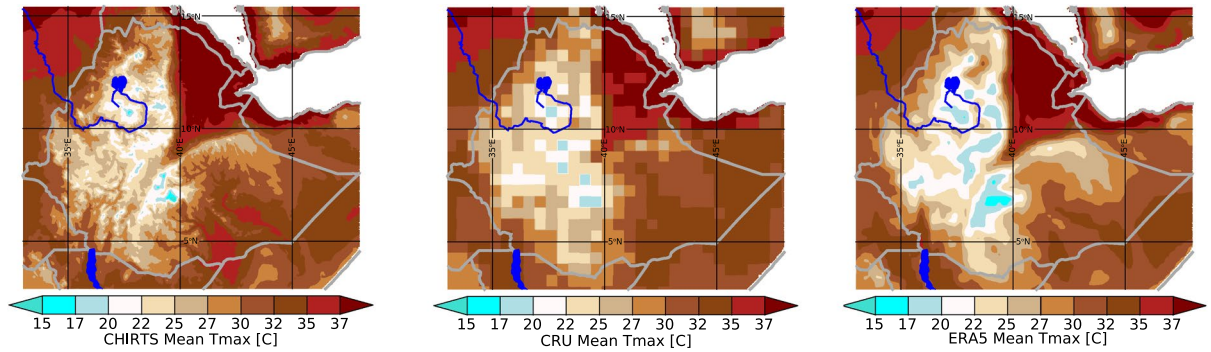


Fig. 6 A comparison of July T_{\max} 1983–2016 mean fields for Ethiopia. From left to right, the panels show means derived from CHIRTS_{max}, CRU, and ERA5 data sets.

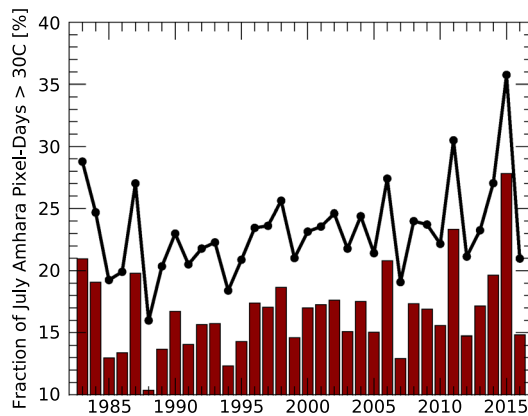


Fig. 7 Time series of observed and projected frequency of agricultural heat stress for Ethiopia's Amhara province in July. The barplot shows the frequency based on CHIRTS-daily maximum temperatures. A value of 25 means that 25% of the day-pixels had maximum temperatures exceeding 30°C. The black line with filled circles represents a plausible climate change scenario in which the temperatures have been increased by +1.3°C.

population. Between 1993 and 2070, Ethiopia may experience a five-fold increase in population, with the number of Ethiopians increasing from about 50 to 250 million people.

While the $\sim +2.5^\circ\text{C}$ temperature anomaly shown in Fig. 5 appears concerning, it is difficult to interpret from an agronomic perspective. Plants typically respond to the actual temperature values. In colder areas, crops may actually benefit from warmer temperatures. In hot areas, temperature increases can result in increased wilting and moisture loss. This dependence means that an accurate background climatology can improve the utility of climate hazard assessments. Figure 6 shows long-term July T_{\max} values averaged over 1983–2016 for the CHIRTS-daily, CRU, and ERA5 data sets. The high-resolution CHIRTS-daily product clearly represents Ethiopia's complex orographic influences. We find some of the steepest temperature gradients on Earth, with highland areas having mean maximum temperatures ranging from 15 to 20°C, while nearby lowland areas may have mean values of greater than 37°C. The Moving Window Regression⁵ modeling process used to construct the CHIRTS background climatology uses satellite TIR mean fields as a predictor, and these fields help capture these complex gradients. These patterns are not captured well by the coarse CRU product. The physically based ERA5 reanalysis captures the overall pattern with reasonable fidelity, but many fine details are missed. Furthermore, as noted above in the global case study, there appears to be a consistent tendency to underestimate the magnitude of the mean maximum temperatures, making this an inappropriate product for estimating the hazards associated with temperature extremes.

For agricultural impact assessments, days over a 30°C threshold is a common metric of heat stress. The bar plot shown in Fig. 7 shows a time series of this metric for consecutive Julys in the Amhara province of Ethiopia—a critical crop-growing region strongly impacted in 2015. This time series gives us a meaningful basis for assessing agricultural temperature-related impacts. First, note that this time series is quite variable. In some years, only 10% of the pixel-days exceeded 30°C. In 2015, on the other hand, almost 30% of the pixel-days exceeded 30°C. This suggests widespread temperature impacts that negatively influenced crop production—capturing such impacts from space may be quite valuable. Current crop water accounting approaches, such as the Water Requirement Satisfaction Index metric (<https://earlywarning.usgs.gov/fews/product/126>), did not adequately capture the magnitude of 2015 agricultural yield anomalies in Amhara.

We conclude with a brief 2070 climate change “projection” for Amhara, based on a perturbed version of the observed CHIRTS-daily data set. Note that our goal here is to simply provide a representative usage case, not to perform a detailed climate change assessment. In general, it is not unreasonable to assume that future air temperatures may be represented by current air temperatures (CUR) perturbed by a smoothly varying set of increases (DELTA). Furthermore, let CUR be a data set with high spatial and temporal resolution, like CHIRTS-daily. DELTA may be derived from climate models such as those represented in the Phase 5 or 6 Coupled Model Intercomparison Projects. Here, we use a single value of +1.3 °C, which is the change in the Phase 5 multi-model ensemble T_{\max} average, for Amhara in July, between 2070 and 2020, based on the 6 Wm⁻² Representative Concentration Pathway. Perturbing our data set with this value and recalculating the fraction of pixel-days warmer than 30 °C results in the black line and filled circles in Fig. 7. By 2070, Ethiopia is likely to have a population of ~250 million people (Fig. 4), and the average heat stress fraction may be around 23%, a level exceeded in only the most extreme months in the observational record.

Code availability

The CHIRTS-daily data products are derived using approximately 500 lines of code written in the Interactive Data Language. While not written as a portable library or toolset, access to the code is not restricted, and it is available upon request.

Received: 5 February 2020; Accepted: 14 August 2020;

Published online: 14 September 2020

References

1. Funk, C. *et al.* A high-resolution 1983–2016 Tmax climate data record based on infrared temperatures and stations by the Climate Hazard Center. *J. Clim.* **32**, 5639–5658 (2019).
2. Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E. & Houston, T. G. An overview of the Global Historical Climatology Network-Daily Database. *J. Atmos. Oceanic Tech.* **29**, 897–910 (2012).
3. Sheffield, J., Goteti, G. & Wood, E. F. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Clim.* **19**, 3088–3111 (2006).
4. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 Dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
5. Funk, C. *et al.* The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci. Data.* **2**, 1–21 (2015).
6. Knapp, K. R. *et al.* Globally gridded satellite observations for climate studies. *Bull. Amer. Meteor. Soc.* **92**, 893–907 (2011).
7. IDL v.4 (*Harris Geospatial Solutions, Inc.*, 1995).
8. Magnus, G. Versuche iiber die Spannkraft des Wasserdampfs. *Ann. Phys. Chem.* **61**, 225–247 (1844).
9. Steadman, R. G. The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. *J. Appl. Meteorol.* **18**, 861–873 (1979).
10. Climate Hazards Center. CHIRTSdaily. *Climate Hazards Center* <https://doi.org/10.15780/G2008H> (2020).
11. Funk, C. *et al.* Assessing the contributions of local and east Pacific warming to the 2015 droughts in Ethiopia and Southern Africa. *Bull. Amer. Meteor. Soc.* **97**, S27–S29 (2016).

Acknowledgements

This work was supported by the National Science Foundation Award Abstract #1639214 INFIEWS/T1: Understanding multi-scale resilience options for vulnerable regions, US Geological Survey cooperative agreement #G09AC000001, USAID cooperative agreement: Early identification and forecasts of reduced-yield agricultural seasons in the Developing World, and the USGS Drivers of Drought project. The Climate Hazards Center acknowledges support from the Defense Advanced Research Projects Agency (DARPA) World Modelers Program under Army Research Office (ARO) prime contract no. W911NF-18-1-0018. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the position or the policy of the Government or the Prime Contract (DARPA and ARO), and no such official endorsement by either should be inferred. The views and conclusions presented in this article do, however, represent the views of the U.S. Geological Survey. The authors thank Climate Hazards Center’s technical writer, Juliet Way-Henthorne, for providing professional editing.

Author contributions

A.V. designed and performed the validations, wrote the initial manuscript draft, contributed to the usage notes section, and contributed to the final manuscript draft. C.F. designed the CHIRTS-daily development process, contributed to the design of the validation process, developed the usage notes analyses and figures, and contributed to the final manuscript draft. P.P. extracted and screened the validation stations, wrote and ran the CHIRTS-daily development process, and contributed to the methods section of the manuscript. C.T. contributed to validation and application of data and contributed to the final manuscript draft. K.G. contributed to discussion of interpretation and application of data and contributed to the final manuscript draft.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.V. or C.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2020