ABSTRACT: As human exposure to hydroclimatic extremes increase and the number of in situ precipitation observations declines, precipitation estimates, such as those provided by the Integrated Multisatellite Retrievals for Global Precipitation Measurement (GPM) (IMERG) mission, provide a critical source of information. Here, we present a new gauge-enhanced dataset [the Climate Hazards Center IMERG with Stations (CHIMES)] designed to support global crop and hydrologic modeling and monitoring. CHIMES enhances the IMERG Late Run product using an updated Climate Hazards Center (CHC) high-resolution climatology (CHP_clim) and low-latency rain gauge observations. CHP_clim differs from other products because it incorporates long-term averages of satellite precipitation, which increases CHP_clim's fidelity in data-sparse areas with complex terrain. This fidelity translates into performance increases in unbiased IMERG_late data, which we refer to as CHIME. This is augmented with gauge observations to produce CHIMES. The CHC's curated rain gauge archive contains valuable contributions from many countries. There are two versions of CHIMES: preliminary and final. The final product has more copious and better-curated station data. Every pentad and month, bias-adjusted IMERG_late fields are combined with gauge observations to create pentadal and monthly CHIMES_prelim and CHIMES_final. Comparisons with pentadal, high-quality gridded station data show that IMERG_late performs well ($r = 0.75$), but has some systematic biases which can be reduced. Monthly cross-validation results indicate that unbiasing increases the variance explained from 50% to 63% and decreases the mean absolute error from 48 to 39 mm month$^{-1}$. Gauge enhancement then increases the variance explained to 75%, reducing the mean absolute error to 27 mm month$^{-1}$.

KEYWORDS: Precipitation; Hydrometeorology; Climate records; Microwave observations; Remote sensing; Climate services
Accurate, timely, unbiased, consistent gauge-enhanced satellite precipitation can help understand and manage agricultural and hydrologic risks. Unfortunately, the number of global gauge observations is very low, covering an area of less than a single football field or soccer pitch (Kidd et al. 2017). Globally, the number of gauges\(^1\) has declined between 1981 and 2021, from about 32,000 to 14,000, while in Africa, the corresponding figures are just 3,400 and 600. These declines are caused by a combination of data policy and decaying observation and transmission systems. Precipitation thus remains a leading source of uncertainty for agroclimatic risk assessments and related agricultural modeling applications (Ruane et al. 2021). Gridded satellite rainfall datasets, therefore, provide a key resource for understanding weather, providing effective early warning, and assessing changes in extremes in a warming world (Kirschbaum et al. 2017). There are three primary categories of satellites used to study and quantify precipitation from space. The oldest category dates back to the Television and Infrared Observation Satellite (TIROS), which launched in 1960 and hosted imagers capable of capturing visible and infrared wavelength Earth observations (EOs). Modern descendants of TIROS occupy geostationary orbits over the equator. Circling in step with Earth’s rotation, they provide continuous information about cloud-top temperatures (Janowiak et al. 2001; Knapp et al. 2011). These geostationary infrared (GEO-IR) observations, however, do not directly observe precipitation, but rather, just the highest layer of clouds and/or the land surface. Passive microwave observations (PMW), on the other hand, can observe hydrometeors, supporting the retrieval of rainfall intensities using techniques like the Goddard profiling algorithm(s) (GPROF) (Kummerow et al. 1996, 2001, 2015). Seminal combined estimation approaches merge PMW and GEO-IR estimates using weights determined by standard errors (Huffman et al. 1997; Xie and Arkin 1997), use sophisticated artificial neural nets to translate GEO-IR observations into precipitation (Hsu et al. 1997), or use collocated PMW and GEO-IR observations to tune local GEO-IR estimates based upon PMW observations (Huffman et al. 2007; Turk et al. 2003, 2010). The neural net approach can use cloud classifications, as in the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks Cloud Classification System (PERSIANN-CCS) (Hong et al. 2004).

A third category of satellite—characterized by the 1997–2015 Tropical Rainfall Measuring Mission (TRMM) (Kummerow et al. 1998, 2000) and the 2014–present Global Precipitation Measurement mission Core Observatory (GPM-CO) (Hou et al. 2014; Skofronick-Jackson et al. 2017)—provided single platforms that contained multiple sensors explicitly designed to monitor precipitation. TRMM had a Ku-band radar and a passive microwave multifrequency imaging radiometer; these instruments were capable of producing three-dimensional views of precipitation events (Kozu 2001). The GPM-CO is even more sophisticated, with additional channels on the microwave imager and a dual-frequency precipitation radar (Hou et al. 2014; Skofronick-Jackson et al. 2017).

Building on techniques used to develop the experimental multisatellite TMPA and TMPA-RT products (Huffman et al. 2018, 2020) (Huffman et al. 2007), the U.S. GPM team now combines

---

\(^1\) Time series available at http://data.chc.ucsb.edu/products/CHIRPS-2.0/diagnostics/stations-perMonth-byRegion/pngs/.
information from the GPM-CO, PMW constellation, and GEO-IR-based PERSIANN-CCS to produce the IMERG Early Run and IMERG Late Run (IMERG\textsubscript{late}) datasets with latencies of about four and fourteen hours (Huffman et al. 2018, 2020). The PMW constellation\textsuperscript{2} includes several types of radiometers: the Advanced Microwave Scanning Radiometer-2 (AMSR-2) on JAXA’s Global Change Observation Mission–Water (GCOM-W1), the Advanced Technology Microwave Sounder (ATMS) on the Suomi NPP and NOAA-20 polar orbiters, the multichannel Sondeur Atmospherique du Profil d’Humidite Intertropicale par Radiometrie (SAPHIR) on the Megha-Tropiques satellite, the Microwave Humidity Sounders (MHS) on the U.S. NOAA-19 and European MetOp platforms, and the Special Sensor Microwave Imager/Sounder (SSMIS) instruments on U.S. Defense Meteorological Satellite Program (DMSP) satellites. These PMW observations are assembled and converted to estimates of precipitation intensity using GPROF. These high-quality, half-hourly PMW estimates are then used to adjust the GEO-IR-based PERSIANN-CCS. Finally, a semi-Lagrangian time-interpolation scheme, adapted from the CMORPH–Kalman filter (Joyce and Xie 2011), is used to approximate the temporal evolution of the PMW-observed precipitation, based on motion vectors estimated from the GEO-IR. This process is repeated to produce the IMERG\textsubscript{early} and IMERG\textsubscript{late} products, with the latter being of higher quality due to it using more data and a time interpolation of data, rather than the forward-only propagation calculated in IMERG\textsubscript{early}. After a two month latency, an interpolated monthly station data analysis provided by the Global Precipitation Climatology Centre (GPCC) is combined with IMERG\textsubscript{late} to create an IMERG\textsubscript{final} product. This blending procedure is similar to that used in the TRMM Multisatellite Precipitation Analysis (Huffman et al. 2007) and was originally developed for the Global Precipitation Climatology Project (Huffman et al. 1997). The satellite-gauge blending process is achieved through 1) large-scale bias correction of the satellite estimates, followed by 2) blending with interpolated gauge fields. The blending process is guided by weights proportional to the inverse of the associated error variances—this process gives greater weight to the satellite estimates in data-sparse regions. This blending produces the IMERG\textsubscript{final}.

Here, we describe and evaluate a bias-adjusted version of IMERG\textsubscript{late} [Climate Hazards Center IMERG (CHIME)] and a station-enhanced version of CHIME [CHIME with Stations (CHIMES)]. CHIMES enhances the IMERG\textsubscript{late} product using a high-resolution climatology and low-latency rain gauge observations, resulting in a product that is similar to IMERG\textsubscript{final}, but with a lower latency. The value of this bias-correction procedure arises from the strong performance of the CHC’s approach to developing high-resolution climatologies. The core insight guiding this work is that long-term averages of satellite QPE provide excellent inputs into the production of high-resolution climatologies (Funk et al. 2015a). The spatial correlation between in situ precipitation normals and mean QPE fields is large and consistent (\textgreater;0.6), while the relationship with elevation is local, complicated, and typically weaker (Fig. 1 in Funk et al. 2015a). Furthermore, many gauge-based climatologies use complex mathematical fits, such as thin-plate splines, that can perform poorly in data-sparse areas, producing complicated orographic features that are not well supported by the available data. Other climatologies, like those underlying the Global Precipitation Climatology Centre’s interpolated gauge product, are based on interpolated climate normals—which may miss important details because of station limitations.

To support early warning efforts like the Famine Early Warning Systems Network (FEWS NET) (Funk et al. 2019) and hydrologic modeling efforts like the FEWS NET Land Data Assimilation System (FLDAS), we have developed an update to the Climate Hazards Center Precipitation Climatology. As shown in (Funk et al. 2015a), this approach performs substantially better than the GPCC climatology and other competing products. Focusing on Ethiopia (Figs. 10–12 in Funk et al. 2015a), we see that the GPCC captures the general shape of mean
Ethiopian rainfall, but still exhibits substantial biases. A sample west-to-east transect at 7°N shows that the GPCC captures the overall gradient reasonably well, but misses local maxima associated with the Rift Valley escarpments at 37° and 40°E. Other global assessments have confirmed the CHC climatology’s relatively strong performance (Beck et al. 2017). Just as Funk et al. (2015a) suggested improvements over the GPCC mean fields, the results presented here suggest that the IMERG\textsubscript{late}’s mean performance can be improved by use of an updated CHC climatology.

We conclude our introduction with a few concrete examples of how these data might be used in the context of food insecurity. Since 1999 (Verdin et al. 1999), FEWS NET and the CHC have focused on developing “climate science for early warning” (Verdin et al. 2005). Today, blended gauge-satellite QPE can be combined with forecasts to provide “staged” (Funk and Shukla 2020; Shukla et al. 2021) early warning systems that combine climate, weather, land surface models, and satellite observations to provide multiple opportunities to anticipate and monitor hydroclimatic extremes. These efforts contribute to FEWS NET’s sophisticated Food Security Outlook process (Magadzire et al. 2017), reports and alerts produced by the Group on Earth Observations Global Agricultural Monitoring Initiative (GEOGLAM) Crop Monitor project (http://www.cropmonitor.org/), the capacity-building and decision-support activities of NASA SERVIR (https://www.nasa.gov/mission_pages/servir/index.html), and NASA hydrologic modeling and forecasting activities, based on the FLDAS (McNally et al. 2017) and NASA Hydrological Forecast and Analysis System (NHyFAS; Arsenault et al. 2020).

While these efforts are already leveraging satellite observations to help guide billions of dollars of humanitarian assistance that help provide aid to millions of people every month, more needs to be done to address rapid increases in hydroclimatic risks. The human and economic costs of hydroclimatic shocks are massive and increasing. Between 2015 and 2019, the impacts and losses associated with droughts, floods, and hurricanes surged through the interaction of increasing exposure and more extreme weather and climate (Funk 2021). In 2020, the Aon-Benfield Reinsurance Company (2021) reported that cyclones, floods, and droughts resulted in $165 billion in losses. And between 2015 and 2021, the number of extremely food-insecure people has tripled; in 2021, more than 147 million people, one out of every 55 people, requires emergency food assistance, as the world faces a severe food security crisis (UN Food Security Information Network 2021). While conflict, food price inflation, and COVID-19 largely fuel this crisis, both droughts and extreme precipitation/flooding have played a substantial role. Improved QPE products, like CHIMES, will help us monitor, mitigate, and manage these increased hazards.

An overview of the CHIMES data products and process

While no dataset is perfect, authoritative data are trustworthy and capable of supporting important decisions that can save lives and livelihoods and trigger positive financial and societal responses to hazards. Authoritative data are actionable. Surveys of the user community (Kirschbaum 2020) indicate that continuity and gauge enhancement for IMERG\textsubscript{early} and IMERG\textsubscript{late} are key interests for the IMERG user community. The incorporation of timely gauge data can increase the utility and accuracy of gridded datasets. Instead of two often-quite different sources of information (gauge and satellite), users are provided a single harmonized data product. CHIMES builds on the techniques and resources used to produce the widely used CHC Infrared Precipitation with Stations (CHIRPS2) product (Funk et al. 2015b). The satellite-only component of CHIRPS2 (CHIRP2) is based solely on GEO-IR observations, while IMERG\textsubscript{late} also draws on the PMW and GEO-CO constellation. As we will demonstrate, IMERG\textsubscript{late} can be substantially improved by the bias adjustment and gauge enhancement provided by CHIMES. As our validation results indicate, this will provide a product that performs better than CHIRP2, but with a shorter period of record, beginning in 2001 rather than 1981. While
IMERG begins in mid-2000, CHIMES starts in 2001 to provide complete years. Performance will be evaluated via comparison with a completely independent set of high-gridded quality gauge observations (Contractor et al. 2020), and quantified via correlation, mean bias error, ratios of standard deviations, mean absolute errors, and dry spell hit ratios.

The basic objective of CHIMES is to enhance IMERG late using in situ station observations (Fig. 1a). This enhancement results in three products: the climatologically adjusted daily satellite-only adjusted IMERG late (CHIME), the moderate-quality station-adjusted submonthly (pentadal) CHIMES prelim, and the high-quality station-adjusted monthly CHIMES final. We use “S” to denote station enhancement.

Supporting CHIMES is an extensive station archive developed and curated by the CHC. Table 1 lists the data sources used in the CHC Station Climatology Database (CSCD), and appendix A describes the CHC’s extensive quality control and curation processes. Appendix B describes the CHIMES procedures mathematically. CHIMES benefits from three categories of station data. Used judiciously, information from these sources can “stack” in an additive manner (Fig. 1b). Numerous climate normals are used to generate high-resolution mean fields, which are used to produce unbiased daily CHIME fields. Then, every “pentad,”3 CHIMES prelim estimates blend CHIME pentads with relatively sparse, but low-latency, submonthly gauge data. The pentadal station observations are quality controlled using automated procedures. Because of the difficulty in identifying problematic daily station data, interpolation is not performed at the daily time scale, but daily CHIMES prelim fields are produced via disaggregation based on the satellite-only daily CHIME. In the last phase, after the careful manual curation (appendix A), the sum of the previous months’ CHIMES prelim pentads are blended with in situ observations to produce CHIMES final. Proportional disaggregation is then used to produce final pentad and daily CHIMES values.

While involved, this type of staged process can support effective early warning and early action, as rapid initial estimates (daily CHIME) are refined using available submonthly station observations (CHIME prelim pentads). Then, carefully screened monthly gauge totals are used to produce an authoritative but timely “final” product. Analysts and modelers have access to interoperable information sources at a range of latencies that extends from 2 to 18 days, with performance at the “final” stage similar to other high-quality monthly gauge-enhanced products (Becker et al. 2013).

The updated CHC Precipitation Climatology (CHP clim)
As periods of accumulation and latencies decrease, going from long-term means, to months, to pentads, the number of available stations declines. Hence, there is more information at the long-term mean time scale, and complex geospatial modeling techniques (Funk et al. 2015a) can be used to build high-resolution (0.05°) long-term mean fields (information source 1 in Fig. 1a). We refer to these climatologies as the “Climate Hazards Precipitation Climatology,” or CHP clim. As described in appendix B, monthly background mean fields are created using local regressions trained to the previous CHC climatology (Funk et al. 2015a). IMERG final and elevation data are used as predictors. If \( b_0, b_1, \) and \( b_2 \) represent the local regression coefficients, the estimate at a grid cell will be \( Y_{est} = b_0 + b_1 \text{IMERG}_{final} + b_2 \text{Elevation}. \) These \( Y_{est} \) fields are then adjusted using monthly gauge-undercatch-corrected precipitation gauge averages. In the next step, 1981–2010 CSCD averages are calculated and combined with GPCC observations (Schneider et al. 2013), with preference given to the CSCD values. As described in appendix B, a modified inverse distance interpolator is used to blend the station observations and \( Y_{est} \) fields. The 12 monthly climatologies are disaggregated to daily and pentadal means by 1) creating a padded 14-month 427 array of days, 2) assigning each day the associated monthly

3 Pentads break months into six accumulation periods. The first five pentads always have five days. The sixth pentad will contain the remaining days of every month.
value, and finally 3) running a smoothing filter many times to produce a smoothly varying annual cycle.

Ironically, the most rapidly produced product, the satellite-only daily CHIME, is adjusted using the slowest source of station information. Long-term climate normals are incorporated into $C_{\text{clim}}$, which is then used to bias adjust the $C_{\text{clim}}$ (information source 2, Fig. 1a). As we shall see, this substantially improves performance, as measured by mean bias errors, mean absolute errors, ratios of standard deviations, and hit ratios. This type of bias adjustment also improves the alignment with station observations, especially in data-sparse areas like East Africa (Dinku et al. 2018). Reducing systematic differences between gauge observations and satellite estimates can reduce inhomogeneities as station networks shift over time.

The CHC station archive and blending process
During the “preliminary” and “final” modeling steps, pentadal, and monthly station observations are blended with either the CHIME or CHIMES$_{\text{prelim}}$ (information source 3, Fig. 1a, appendix B). Numerous international and national station data sources (Table 1, described in more detail in appendix A) are combined and quality controlled to produce sets of in situ pentad and monthly totals.

Central to both the CHC quality control and pentadal/monthly interpolation procedures is the development of a set of well-organized “anchor stations”—a superset of all the possible gauge observation locations. Multiple data sources (Table 1) are then used to produce as-complete-as-possible time series, with higher-quality data sources given preference. These time series support the identification of improbable outliers. The pentadal quality control (QC) process is automatic. The final monthly station archive is also carefully checked, visually and analytically, by a team at the CHC. The pentadal station observations are obtained from

Fig. 1. (a) The CHIMES development process. (b) CHIME(S) products. (c) Map of the wettest three months used in the validation study; numerals correspond to the first month. (d) REGEN (Contractor et al. 2020) validation locations and regions. High-quality interpolated gauge data in these regions are compared with satellite-only estimates.
rapidly updated automated station archives, with a 2-day latency. At the monthly time step, many additional gauge observations are contributed by numerous partner agencies, including national meteorological agencies. These contributions often dramatically enhance national coverage.

Pentads, as opposed to days, were chosen as a preferred accumulation period for several reasons. One pragmatic constraint is uncertainty about the actual reporting time associated with many data sources. Another issue involves quality control. Statistical error detection typically becomes easier over longer time periods. The spatial coherence of precipitation fields increases with accumulation period, increasing the accuracy of spatial interpolation. And finally, many early warning practitioners, especially those interested in precipitation deficits (i.e., not floods) find that a 5–10-day submonthly accumulation period is an appropriate time scale (Ross et al. 2009).

This process uses IMERG\textsubscript{final} to generate spatial distance decay functions for every pixel and time period. The spatial coherence of IMERG\textsubscript{final} informs the spatial influence of station observations. For each pixel, decay functions, neighbors, and neighbor distances are precomputed, supporting rapid global interpolation speeds.

The CHIMES process (Fig. 1a) draws from four information sources: 1) dense sets of climate normals, 2) the GPM constellation/IMERG\textsubscript{late}, 3) in situ station archives, and 4) IMERG\textsubscript{final}’s spatial coherence. Appendix B provides a mathematical description of the CHIMES algorithms.

### Evaluating the performance of IMERG\textsubscript{late} and CHIME

We next present the first of two brief evaluation studies. The first study focuses on the performance of the satellite-only IMERG\textsubscript{late} and CHIME products. The second study uses cross validation to assess the contribution provided by station enhancement. While a detailed multiproduct assessment is beyond the scope of this brief paper, we also include three other representative satellite-only products: CHIRP2, PERSIANN-CCS (Hong et al. 2004), and the GOES precipitation index (GPI) (Arkin and Meisner 1987). The GPI and CHIRP2 are both based on cold cloud duration (CCD) values calculated from the Climate Prediction Center GEO-IR archive (Janowiak et al. 2001) using a fixed 235-K temperature threshold. The 235-K threshold represents the expected temperature of cloud-top temperatures, so CCD is related to deep convection, especially in tropical and subtropical monsoon regions. The GPI estimates are calculated by multiplying the CCD totals by 3 mm h\textsuperscript{-1}.

Validations are carried out using the high-quality daily 1° “Rainfall Estimates on a Gridded Network” (REGEN; or rain in German) (Contractor et al. 2020), which was obtained from the Frequent Rainfall Observations on Grids (FROGS) database (Roca et al. 2019). To facilitate global analyses, we define the wettest three months for each location (Fig. 1c), and focus our evaluations on the wettest three months and the 18 pentads within those months. Figures 2–5

### Table 1. CHC rain gauge archive contributions by source, along with approximate numbers of stations.

<table>
<thead>
<tr>
<th>Category</th>
<th>Source</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>GPCC</td>
<td>78,000</td>
</tr>
<tr>
<td>Mean</td>
<td>CHC merger</td>
<td>43,000</td>
</tr>
<tr>
<td>Monthly</td>
<td>GHCN Monthly</td>
<td>16,000</td>
</tr>
<tr>
<td>Monthly</td>
<td>Ethiopia (NMA)</td>
<td>140</td>
</tr>
<tr>
<td>Monthly</td>
<td>Somalia (SWALIM)</td>
<td>75</td>
</tr>
<tr>
<td>Monthly</td>
<td>El Salvador (MARN)</td>
<td>20</td>
</tr>
<tr>
<td>Monthly</td>
<td>Ecuador (INAMHI)</td>
<td>150</td>
</tr>
<tr>
<td>Monthly</td>
<td>Bolivia</td>
<td>20</td>
</tr>
<tr>
<td>Monthly</td>
<td>Mozambique</td>
<td>36</td>
</tr>
<tr>
<td>Monthly</td>
<td>Panama (ETESA)</td>
<td>15</td>
</tr>
<tr>
<td>Monthly</td>
<td>Southern Africa (SASSCAL)</td>
<td>74</td>
</tr>
<tr>
<td>Monthly</td>
<td>Honduras (COPECO)</td>
<td>61</td>
</tr>
<tr>
<td>Submonthly</td>
<td>Mexico (Conagua)</td>
<td>775</td>
</tr>
<tr>
<td>Submonthly</td>
<td>TAHMO (UG/KY/ET)</td>
<td>150</td>
</tr>
<tr>
<td>Submonthly</td>
<td>Guatemala (INSIVUMEH)</td>
<td>88</td>
</tr>
<tr>
<td>Submonthly</td>
<td>Brazil (Cemaden)</td>
<td>2,693</td>
</tr>
<tr>
<td>Submonthly</td>
<td>Chile (Chile-Met)</td>
<td>20</td>
</tr>
<tr>
<td>Submonthly</td>
<td>Costa Rica (IMN)</td>
<td>15</td>
</tr>
<tr>
<td>Submonthly</td>
<td>Colombia (IDEAM)</td>
<td>506</td>
</tr>
<tr>
<td>Submonthly</td>
<td>GTS</td>
<td>7,500</td>
</tr>
<tr>
<td>Submonthly</td>
<td>GSOD</td>
<td>7,300</td>
</tr>
<tr>
<td>Submonthly</td>
<td>GHCN Daily</td>
<td>10,000</td>
</tr>
</tbody>
</table>
show results based on seasonal accumulations. For each pixel, a fixed wettest 3-month period was identified (Fig. 1c), and then 16 years of 3-month total precipitation was calculated for the interpolated station (REGEN) and five satellite-only datasets. The five satellite-only datasets do not incorporate time-varying gauge observations. Table 2 shows results based on the 16 years × 18 pentadal accumulations for each pixel’s wet season. REGEN ends in 2016, hence this study uses a 2001–16 evaluation period, which is the longest whole-year period containing complete data for all sources.

We use the number of average daily observation stations in each 1° grid cell as a metric of data certainty. Cells with averages of at least seven daily observations are identified with blue circles in the bottom panel of Fig. 1c. The summary statistics presented in Table 2 are based on the average of results for these well-instrumented locations. To facilitate visual interpretations, these cells are also noted in Figs. 2–5. Poorly monitored locations, with less than an average of one station observation per day, are identified with gray circles.

**Correlation results.** Figure 2 shows 2001–16 wet season correlation maps focused on the four best-instrumented regions. The performance of the CHIME/IMERG estimates is compelling. Over large portions of North America, Europe, Iran, and Australia, seasonal correlations exceed 0.8, and most areas have correlations of at least 0.6. Information from GPM-CO and PMW instruments is clearly enhancing performance, since correlations for GEO-IR-based CHIRP2/
GPI and the PERSIANN-CCS are substantially lower and less consistent. The PERSIANN-CCS correlations are the lowest of the five products, with typical seasonal correlations of less than 0.6. The pentadal correlation results, shown in Table 2, present a similar tableau. Overall, the CHIME/IMERG pentadal correlations were high (0.78 on average) and consistently above the CHIRP/GPI and PERSIANN-CCS. The ability of IMERG late to track wet season precipitation from space with such a high degree of fidelity bodes well for applications utilizing IMERG late and CHIME.

**Ratio of standard deviations results.** Figure 3 shows maps displaying 2001–16 ratios of interseasonal standard deviations (satellite/REGEN). In North America, Europe, and Iran, the performance of CHIME is close to parity. In these regions, the IMERG-to-CHIME bias correction brings the variance structure into closer alignment with REGEN. In central Australia, CHIME may underestimate the variance, but this extremely dry region is also poorly instrumented.

In terms of variance, CHIRP2 performs poorly, substantially underestimating variability. The CHC team has tracked this problem back to the intercept terms used to calculate CHIRP2 pentads. PERSIANN-CCS and GPI perform better than CHIRP2, but both have areas where the variance of the satellite precipitation is much greater than the observations, particularly in the western United States and central and northern Iran.

**Ratio of means results.** Figure 4 shows maps displaying 2001–16 ratios of seasonal wet season means. The CHIRP2 and CHIME means are very close to REGEN averages. This
fidelity enhances their accuracy and utility, especially for applications sensitive to the absolute quantity of precipitation. For example, groups like FEWS NET rely on satellite rainfall grids to estimate the start of crop growing seasons, and estimate “crop water satisfaction” (Senay and Verdin 2003; Verdin and Klaver 2002). In semiarid areas, a small bias (say ±5 mm per pentad), can add up to a large bias when accumulated over a 3- or 4-month crop-growing season. These biases can also lead to inaccurate growing season start estimates. Hydrologic modeling applications like FLDAS (McNally et al. 2017), face similar challenges. Relatively small precipitation biases can accumulate within these models’ soil moisture reservoirs. IMERG Late exhibits a modest tendency to overestimate, except for the northwestern United States and portions of Iran and western Turkey. The GPI and PERSIANN-CCS overestimate rainfall in parts of the western United States and Iran. The PERSIANN-CCS and GPI bias patterns are quite similar.

Ratio of mean absolute error to observed mean results. Figure 5 shows maps displaying 2001–16 seasonal mean absolute errors (MAE) expressed as ratios of the average REGEN wet season precipitation. Pentadal results are listed in Table 2. MAE will be influenced by both systematic and random errors (Willmott et al. 1985). These results, therefore, reflect both the strength of seasonal correlations (Fig. 2) and systematic distribution biases (Figs. 3 and 4). The CHIME performance is consistently good across all four regions, with typical percent MAE values of less than 30%. The IMERG and CHIRP2 appear to be the next-best performers, though probably for different reasons, with the CHIRP2 benefitting from strong mean

![Image](image.jpg)
performance and moderate correlation performance, and the IMERG benefitting from excellent correlation performance and moderate mean performance. The GPI and PERSIANN-CCS's large mean biases translate into very large MAE values (Fig. 5). Capturing the mean precipitation structure is an important aspect of having overall low MAE values.

The pentadal MAE statistics (Table 2) indicate global percent errors ranging from 48% (CHIME) to 75% (PERSIANN-CCS, GPI) of the long-term mean, with IMERG late and CHIRP2 falling in the middle of this range (60% and 64%). Unbiasing reduces the overall IMERG errors by ~20%. Regionally, the IMERG late to CHIME adjustment seems most beneficial in mid-latitude regions. The U.S. and European validation sites saw the largest improvements, as MAE values were reduced by 20%–30%.

One striking feature of the CHIME is its robust performance across all the regions. All products tended to have the highest MAE values in the two western U.S. regions, southern Europe, and Iran. CHIME performance was substantially better in these regions, due to the bias-correction procedure. The CHIME percent MAE errors only exceeded 50% in one region (Iran). The other products exceeded this error level in 8–12 regions.

**Dry spell analysis.** While a detailed analysis is beyond the scope of this article, to briefly explore CHIME performance vis-à-vis agricultural applications, we evaluate hit ratios based on dry events (Table 3), with dry events defined as wet-season pentads receiving less than 10 mm of precipitation. All products performed quite well, with hit rates over 75%. Satellites' ability
to observe dry conditions from space is tremendously valuable in the context of agricultural drought monitoring. IMERGlate, however, clearly performed better than CHIRP2, and CHIME performed better than IMERG. CHIRP2’s variance underestimation leads to poor dry-spell detection, especially in Brazil and Mexico. CHIME performance, on the other hand, is strong in every region, reinforcing the consistently high levels of correlation, low bias, and variance fidelity shown in Figs. 2–5 and Table 2. This analysis underscores the strong performance of IMERGlate and CHIME across a very diverse collection of geographies and climate. Please also note that the IMERGlate/CHIME data being evaluated are completely independent of the station-based REGEN archive. So this high level of agreement indicates that both of these independent data sources are accurate, which leads to a high level of agreement.

### Evaluating the relative benefits of rain gauge enhancement

This section uses monthly in situ rain gauge observations and a take-one-away cross-validation procedure to assess the relative benefits of blending station data with CHIME. It is extremely important for readers to know that the density of precipitation gauge data varies dramatically from region to region. While space limitation precludes a detailed description here, resources

<table>
<thead>
<tr>
<th>Region</th>
<th>CHI</th>
<th>AGM</th>
<th>IME</th>
<th>PER</th>
<th>GPI</th>
<th>CHI</th>
<th>AGM</th>
<th>IME</th>
<th>PER</th>
<th>GPI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation</td>
<td>Ratio of std dev (satellite/obs)</td>
<td></td>
<td></td>
<td></td>
<td>CHI</td>
<td>AGM</td>
<td>IME</td>
<td>PER</td>
<td>GPI</td>
</tr>
<tr>
<td>SW U.S.</td>
<td>0.60</td>
<td>0.79</td>
<td>0.79</td>
<td>0.59</td>
<td>0.60</td>
<td>0.53</td>
<td>0.87</td>
<td>1.18</td>
<td>1.16</td>
<td>0.97</td>
</tr>
<tr>
<td>SE U.S.</td>
<td>0.65</td>
<td>0.80</td>
<td>0.80</td>
<td>0.61</td>
<td>0.65</td>
<td>0.60</td>
<td>0.86</td>
<td>1.24</td>
<td>0.88</td>
<td>0.83</td>
</tr>
<tr>
<td>NE U.S.</td>
<td>0.60</td>
<td>0.78</td>
<td>0.78</td>
<td>0.58</td>
<td>0.60</td>
<td>0.66</td>
<td>0.93</td>
<td>1.23</td>
<td>0.86</td>
<td>0.90</td>
</tr>
<tr>
<td>NW U.S.</td>
<td>0.60</td>
<td>0.72</td>
<td>0.72</td>
<td>0.53</td>
<td>0.60</td>
<td>0.63</td>
<td>0.93</td>
<td>1.13</td>
<td>1.04</td>
<td>1.06</td>
</tr>
<tr>
<td>Mexico</td>
<td>0.62</td>
<td>0.66</td>
<td>0.66</td>
<td>0.59</td>
<td>0.62</td>
<td>0.61</td>
<td>0.75</td>
<td>0.83</td>
<td>0.87</td>
<td>0.90</td>
</tr>
<tr>
<td>CenAm</td>
<td>0.45</td>
<td>0.57</td>
<td>0.57</td>
<td>0.29</td>
<td>0.45</td>
<td>0.68</td>
<td>0.78</td>
<td>0.88</td>
<td>1.13</td>
<td>0.73</td>
</tr>
<tr>
<td>Ecuador</td>
<td>0.41</td>
<td>0.62</td>
<td>0.62</td>
<td>0.41</td>
<td>0.41</td>
<td>0.79</td>
<td>0.91</td>
<td>0.99</td>
<td>0.91</td>
<td>0.90</td>
</tr>
<tr>
<td>Brazil</td>
<td>0.65</td>
<td>0.74</td>
<td>0.74</td>
<td>0.65</td>
<td>0.65</td>
<td>0.70</td>
<td>0.96</td>
<td>0.99</td>
<td>1.09</td>
<td>0.90</td>
</tr>
<tr>
<td>South EU</td>
<td>0.68</td>
<td>0.79</td>
<td>0.79</td>
<td>0.57</td>
<td>0.68</td>
<td>0.51</td>
<td>0.90</td>
<td>1.32</td>
<td>0.82</td>
<td>0.87</td>
</tr>
<tr>
<td>North EU</td>
<td>0.50</td>
<td>0.77</td>
<td>0.77</td>
<td>0.53</td>
<td>0.50</td>
<td>0.54</td>
<td>0.95</td>
<td>1.20</td>
<td>0.62</td>
<td>0.83</td>
</tr>
<tr>
<td>Iran</td>
<td>0.45</td>
<td>0.68</td>
<td>0.68</td>
<td>0.37</td>
<td>0.45</td>
<td>0.63</td>
<td>0.92</td>
<td>0.93</td>
<td>1.06</td>
<td>1.28</td>
</tr>
<tr>
<td>Aust</td>
<td>0.75</td>
<td>0.82</td>
<td>0.82</td>
<td>0.60</td>
<td>0.74</td>
<td>0.46</td>
<td>0.69</td>
<td>0.89</td>
<td>0.74</td>
<td>0.51</td>
</tr>
<tr>
<td>All</td>
<td>0.60</td>
<td>0.76</td>
<td>0.76</td>
<td>0.56</td>
<td>0.60</td>
<td>0.59</td>
<td>0.87</td>
<td>1.13</td>
<td>0.93</td>
<td>0.87</td>
</tr>
</tbody>
</table>

**Ratio of means (obs/satellite) Percent MAE (MAE/obs mean)**

<table>
<thead>
<tr>
<th>Region</th>
<th>CHI</th>
<th>AGM</th>
<th>IME</th>
<th>PER</th>
<th>GPI</th>
<th>CHI</th>
<th>AGM</th>
<th>IME</th>
<th>PER</th>
<th>GPI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ratio of means (obs/satellite)</td>
<td>Percent MAE (MAE/obs mean)</td>
<td></td>
<td></td>
<td></td>
<td>CHI</td>
<td>AGM</td>
<td>IME</td>
<td>PER</td>
<td>GPI</td>
</tr>
<tr>
<td>SW U.S.</td>
<td>0.97</td>
<td>0.98</td>
<td>1.34</td>
<td>1.42</td>
<td>1.29</td>
<td>0.72</td>
<td>0.50</td>
<td>0.68</td>
<td>0.94</td>
<td>0.98</td>
</tr>
<tr>
<td>SE U.S.</td>
<td>0.99</td>
<td>0.94</td>
<td>1.29</td>
<td>0.99</td>
<td>1.12</td>
<td>0.58</td>
<td>0.44</td>
<td>0.58</td>
<td>0.64</td>
<td>0.68</td>
</tr>
<tr>
<td>NE U.S.</td>
<td>1.02</td>
<td>0.96</td>
<td>1.21</td>
<td>0.87</td>
<td>1.01</td>
<td>0.62</td>
<td>0.46</td>
<td>0.58</td>
<td>0.65</td>
<td>0.67</td>
</tr>
<tr>
<td>NW U.S.</td>
<td>0.98</td>
<td>0.98</td>
<td>1.14</td>
<td>1.38</td>
<td>1.14</td>
<td>0.64</td>
<td>0.50</td>
<td>0.63</td>
<td>0.95</td>
<td>0.94</td>
</tr>
<tr>
<td>Mexico</td>
<td>0.99</td>
<td>0.93</td>
<td>1.03</td>
<td>1.05</td>
<td>1.17</td>
<td>0.52</td>
<td>0.46</td>
<td>0.52</td>
<td>0.60</td>
<td>0.72</td>
</tr>
<tr>
<td>CenAm</td>
<td>1.10</td>
<td>1.00</td>
<td>1.17</td>
<td>1.43</td>
<td>1.27</td>
<td>0.45</td>
<td>0.38</td>
<td>0.47</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>Ecuador</td>
<td>1.02</td>
<td>1.01</td>
<td>1.00</td>
<td>0.69</td>
<td>0.33</td>
<td>0.52</td>
<td>0.49</td>
<td>0.51</td>
<td>0.62</td>
<td>0.35</td>
</tr>
<tr>
<td>Brazil</td>
<td>1.04</td>
<td>0.99</td>
<td>0.99</td>
<td>1.14</td>
<td>0.99</td>
<td>0.52</td>
<td>0.46</td>
<td>0.49</td>
<td>0.61</td>
<td>0.55</td>
</tr>
<tr>
<td>South EU</td>
<td>0.98</td>
<td>0.94</td>
<td>1.30</td>
<td>1.09</td>
<td>1.02</td>
<td>0.82</td>
<td>0.54</td>
<td>0.78</td>
<td>0.89</td>
<td>0.76</td>
</tr>
<tr>
<td>North EU</td>
<td>1.03</td>
<td>0.98</td>
<td>1.16</td>
<td>0.68</td>
<td>0.80</td>
<td>0.66</td>
<td>0.45</td>
<td>0.59</td>
<td>0.64</td>
<td>0.79</td>
</tr>
<tr>
<td>Iran</td>
<td>1.33</td>
<td>1.04</td>
<td>0.93</td>
<td>1.45</td>
<td>1.64</td>
<td>1.00</td>
<td>0.65</td>
<td>0.73</td>
<td>1.14</td>
<td>1.33</td>
</tr>
<tr>
<td>Aust</td>
<td>0.86</td>
<td>0.84</td>
<td>1.05</td>
<td>0.91</td>
<td>0.76</td>
<td>0.70</td>
<td>0.49</td>
<td>0.58</td>
<td>0.79</td>
<td>0.51</td>
</tr>
<tr>
<td>All</td>
<td>1.00</td>
<td>0.96</td>
<td>1.18</td>
<td>1.09</td>
<td>1.06</td>
<td>0.64</td>
<td>0.48</td>
<td>0.60</td>
<td>0.75</td>
<td>0.75</td>
</tr>
</tbody>
</table>
at the CHC website support detailed visualization of time series\(^4\), national maps of stations locations,\(^5\) and global geotiffs\(^5\) that quantify station density.

For each year and month between 2001 and 2019, each monthly station observation is withheld, and at-station gauge/satellite adjustment factors are calculated at surrounding stations. In the previous section, the 2016 end point of the REGEN dataset limited us to 2016; here we can use a longer period because we use the CHC’s routinely updated station archive. A modified inverse distance weighting (IDW) interpolation is then used to estimate the weighted adjustment value at the station location [as in Eqs. (B12)–(B14) in appendix B].

This procedure allows us to evaluate the decreases in MAE and the increases in variance associated with the gauge-blending procedure. Some fraction of the total variability of the global station dataset will be explained by the original IMERG\(_{\text{final}}\) archive. Unbiasing with CHP\(_{\text{clim}}\) will then explain some additional variance. Finally, gauge-based station enhancement can explain another additional amount of variance. The remaining variance is the mean-squared error (MSE).

Figure 6 presents our cross-validation results in terms of variance explained for the entire domain (60°S–60°N, 180°–180°) and by continent.\(^6\) Three sources of information (variance explained) stack: IMERG\(_{\text{late}}\), the incremental improvement due to unbiasing, and the information added when stations are blended with CHIME. The remaining variance is the MSE. Please note that no attempt was made to spatially weight these results, so densely instrumented regions will tend to dominate the global values. Overall, IMERG\(_{\text{late}}\) performs well, explaining 50% of the station variance. IMERG\(_{\text{late}}\) performance in Europe, North Asia, and Australia, however, is lower than 27%.

Unbiasing in Africa, North America, Europe, and North Asia increases the variance explained, respectively, by 19%, 24%, 30%, and 33%—large improvements in performance. In Australia, South Asia, and South America, gains due to bias adjustment were minor or modest, 1%–7%. The original IMERG\(_{\text{late}}\), exhibited little bias in these areas, so unbiasing adds little benefit.

Station enhancement adds additional information. The largest benefits occurred in Australia, North America, and Europe, with increases of 47%, 20%, and 17% variance explained. South America and North Asia saw modest 13% enhancements. Africa and South Asia exhibited limited enhancements; the results in these regions underrepresent the potential merits of enhancement, because the interstation distances are too large to well support cross-validation-based analyses. Note that spatial gauge distributions will also influence these results. For example, gauges in Australia are much more frequent near the populated wetter coastal regions, and the cross-validation procedure

### Table 3. Dry pentad detection results based on comparison of pentadal satellite-based precipitation estimates with gridded REGEN rain gauge observations. Dry events were defined as pentads with less than 10 mm of precipitation. Validation seasons and regions are shown in Fig. 1d. Hit ratios are calculated as \(100 \times \frac{\#\text{hit}}{\#\text{hit} + \#\text{miss}}\). Guatemala and Ecuador were excluded because they very rarely exhibited REGEN pentads with less than 10 mm of precipitation.

<table>
<thead>
<tr>
<th>Region</th>
<th>CHI</th>
<th>AGM</th>
<th>IME</th>
<th>PER</th>
<th>GPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SW U.S.</td>
<td>78</td>
<td>88</td>
<td>80</td>
<td>72</td>
<td>67</td>
</tr>
<tr>
<td>SE U.S.</td>
<td>56</td>
<td>79</td>
<td>74</td>
<td>69</td>
<td>66</td>
</tr>
<tr>
<td>NE U.S.</td>
<td>57</td>
<td>83</td>
<td>74</td>
<td>77</td>
<td>72</td>
</tr>
<tr>
<td>NW U.S.</td>
<td>75</td>
<td>86</td>
<td>84</td>
<td>57</td>
<td>64</td>
</tr>
<tr>
<td>Mexico</td>
<td>38</td>
<td>63</td>
<td>57</td>
<td>63</td>
<td>62</td>
</tr>
<tr>
<td>Brazil</td>
<td>19</td>
<td>73</td>
<td>65</td>
<td>74</td>
<td>83</td>
</tr>
<tr>
<td>South EU</td>
<td>42</td>
<td>73</td>
<td>77</td>
<td>65</td>
<td>66</td>
</tr>
<tr>
<td>North EU</td>
<td>75</td>
<td>89</td>
<td>86</td>
<td>73</td>
<td>76</td>
</tr>
<tr>
<td>Iran</td>
<td>44</td>
<td>84</td>
<td>73</td>
<td>78</td>
<td>73</td>
</tr>
<tr>
<td>Australia</td>
<td>69</td>
<td>88</td>
<td>92</td>
<td>66</td>
<td>63</td>
</tr>
<tr>
<td>All</td>
<td>76</td>
<td>86</td>
<td>81</td>
<td>84</td>
<td>78</td>
</tr>
</tbody>
</table>
identifies large increases in variance, explained here (Fig. 6). The number of available monthly stations varies dramatically by country. In 2021, the United States had around 6,000 monthly observations, while Tanzania had 10. This heterogeneity underscores why the consistency of CHIME is important; our results suggest that CHIME is likely to perform reasonably well, even in areas without gauges. These results can also be expressed as changes in MAE (Table 4). Globally, the unbiasing and gauge blending reduces MAE values from 46 to 39 to 27 mm month$^{-1}$, implying that CHIMES's MAE
values are almost half of IMERGlate MAE values, with the caveat that our “global” results tend to favor data-rich regions like Europe, Australia, and North America. While the contribution from unbiasing and station enhancement varies from region to region, CHIMES variance explained is substantially better in every case.

Finally, it should be noted that this evaluation approach will underestimate the value of stations in data-sparse areas. The actual value of station observations is probably greatest where they are least dense, as in Africa, but because the actual interstation distances are so great, cross-validated assessments of these benefits are less conclusive.

Discussion and conclusions
There are two major findings in the results presented here. The first finding has been that the IMERGlate product does a very good job of representing pentadal and wet season precipitation variability, both in terms of correlation and variance. The second finding, which comes as no surprise, is that high-resolution climatologies and in situ gauge observations can be used to make substantial enhancements to the satellite-only IMERGlate product.

CHIMES effectively marries two approaches to precipitation estimation—a complex, physically based estimation process (Huffman et al. 2020) driven by GPM-CO, PMW, and GEO-IR observations (Hou et al. 2014; Skofronick-Jackson et al. 2017), as interpreted by GPROF (Kummerow et al. 2015), PERSIANN-CCS (Hong et al. 2004), and the IMERG algorithm (Huffman et al. 2020)—and the CHC’s geostatistical modeling approach (Funk et al. 2015a,b), which can be seen as extension of presatellite approaches to “smart interpolation” (Willmott and Matsuura 1995; Willmott and Robeson 1995) and current efforts led by the Global Precipitation Climatology Centre (GPCC) (Becker et al. 2013; Schneider et al. 2017) and the Climatic Research Unit (CRU) of East Anglia (Harris et al. 2020). The satellite-only CHIME combines the strong temporal fidelity of IMERGlate, which ultimately originates in the ability of space-borne microwave and radar sensors to actually sense hydrometeors, with the high-spatial accuracy of CHP clim and CHIRP2. While sparse in their coverage, microwave and radar observations can be used to tune thermal infrared (TIR) estimates. These estimates are the only source of information when the surface is icy or snowy. To IMERGlate’s strong temporal fidelity, CHIME can add very low systematic bias and consistently low MAEs (Fig. 5, Table 1).

One striking attribute of the version 6 IMERGlate algorithm is very consistent correlation performance across most of the REGEN validation regions (Fig. 2). The median pentadal correlation for the 12 study regions was 0.75 (Table 2), a value substantially higher than the 0.58 median correlation for the GEO-IR-based CHIRP2 and PERSIANN-CCS. While the difference between 0.58 and 0.75 may seem small, this can represent a substantial change in variance explained. Another positive aspect of the IMERGlate performance was quite consistent estimates of seasonal (Fig. 3) and pentadal precipitation standard deviations (Table 2). While IMERGlate overestimated the variance slightly, the estimated standard deviations were within ±25% of the observed REGEN standard deviations in all of the study regions. The PERSIANN-CCS performance was also quite good, with 10 out of 12 regions within ±25%. CHIRP2 performance, by this metric, was very poor. The CHC has tracked this issue back to the use of intercept

<table>
<thead>
<tr>
<th></th>
<th>Globe</th>
<th>North America</th>
<th>South America</th>
<th>Europe</th>
<th>Africa</th>
<th>North Asia</th>
<th>South Asia</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMERG</td>
<td>46</td>
<td>43</td>
<td>68</td>
<td>41</td>
<td>37</td>
<td>24</td>
<td>53</td>
<td>35</td>
</tr>
<tr>
<td>CHIME</td>
<td>39</td>
<td>35</td>
<td>64</td>
<td>34</td>
<td>29</td>
<td>17</td>
<td>46</td>
<td>29</td>
</tr>
<tr>
<td>CHIMES</td>
<td>27</td>
<td>20</td>
<td>49</td>
<td>25</td>
<td>27</td>
<td>13</td>
<td>40</td>
<td>16</td>
</tr>
</tbody>
</table>
values in the pentadal CHIRP2 cold cloud duration (CCD) estimates (CHIRP2 = $b_0 + b_1 \cdot \text{CCD}$). The constant $b_0$ terms incorrectly reduce the year-to-year variance.

Despite its systematic underrepresentation of precipitation variance, CHIRP2’s performance was fairly comparable to IMERG$_{\text{late}}$ (Fig. 5), largely due to CHIRP2’s strong compensating performance in terms of low bias (Fig. 4). In the United States, Europe, and Central America, IMERG$_{\text{late}}$ means were 14%–34% too large, while CHIRP2 means were within a range from −2% to 3% (Table 2). Recent evaluations of IMERG (Nguyen et al. 2020) indicate that this may relate to the overestimation of the wettest precipitation events.

Another rather striking result was the relatively strong performance of CHIRP2, relative to PERSIANN-CCS, in the United States, Iran, and Europe. Given that the products have similar inputs—GEO-IR data—it is interesting to note how CHIRP2’s more consistent seasonal correlations (Fig. 2) and lower bias (Fig. 4) translate into lower MAE values (Fig. 5). Given that the CHIRP2 algorithm is much simpler than the PERSIANN-CCS, which uses a neural net (Hong et al. 2004) and cloud classifications (Hong et al. 2004) to estimate precipitation intensities, these results provide support for spatially explicit algorithms, i.e., relatively simple algorithms which are tuned to specific locations and times.

While the advantages of reducing systematic errors in satellite-only QPE are obvious, an equally important aspect of bias reduction involves the spatial interpolation of intermittent station data. Station networks are not static, and gauge observations appear and disappear, especially in many data-sparse regions. This gauge intermittency predated the satellite era, and led to the widespread use of combinations of static climatologies and time-varying station anomalies, beginning in the late 1990s (Willmott and Matsuura 1995; Willmott and Robeson 1995). Consider, for example, interpolated precipitation estimates halfway between a wet mountain and a dry coastal region (Funk and Shukla 2020). If the mountain gauge stops reporting, one might anticipate a spurious “drought” emerging at the point between the mountain and the coast. Modern interpolation schemes, therefore, interpolate station anomalies, and then combine these anomalies with static climatologies to get rainfall in millimeters. Now, allow that static climatology to shift based on information from IMERG, and we have CHIME, a very good basis for “SMART” interpolation (Willmott and Matsuura 1995) of precipitation gauge observations, which results in CHIMES.

Station quality control is also critical. The CHC science team’s extensive experience with station data, and sophisticated data management and quality control procedures (appendix A), help ensure that gauge enhancement avoids most serious quality issues associated with poor-quality station observations. The CHC process and inputs are quite similar to the GPCC’s, but they are more timely, and comparisons indicate high levels of correlation between the GPCC and CHIRPS (Funk et al. 2015b). The IMERG$_{\text{final}}$ has a latency of about two months, the CHIMES$_{\text{prelim}}$ and CHIMES$_{\text{final}}$ will have latencies of 2 and 18 days.

The cross-validation results, presented in the “Evaluating the relative benefits of rain gauge enhancement” section and Fig. 6, highlight the value of station blending. Using station data, as opposed to gridded station data, as a basis for evaluation is demanding. So the fact that CHIMES can explain 75% of monthly at-station variance is very promising (Fig. 6). In the United States and Africa, IMERG$_{\text{late}}$ performance is fairly good, 38% and 48%, but CHIMES is substantially better (82% and 63%). Another aspect of gauge enhancement is the benefit that stations provide in data-sparse areas, especially when those regions are highly vulnerable. For example, the CHIRPS product only had four or five stations in northern Ethiopia in 2015, but those stations helped capture the worst drought in 50 years, and in conjunction with the Famine Early Warning System Land Data Assimilation System (FLDAS) (McNally et al. 2017), captured the extreme and persistent soil moisture deficits that contributed to creating extreme

---

8 Chapter 8 in the copyright free chapters at https://data.chc.ucsb.edu/people/chris/DroughtEarlyWarningBook/.
food insecurity conditions for more than 11 million people. Supported by, and supporting, the United Nations Framework for Climate Services, the CHC engages directly with many national agencies, leading to timely and very valuable data contributions (Table 1). Where station densities are low or declining, the CHIME results presented here suggest the satellite-only CHIME represents a substantial improvement over CHIRP2, PERSIANN-CCS, and GIF.

Thus, in summary, the very promising performance of CHIMES should be understood as arising from dozens of contributions: the GPM satellites, approximately homogenized using the GPM-CO via the numerous algorithms that have been painstakingly assembled by the GPM science team, the CHC efforts focused on building world-class high-resolution climatologies, station archives, and interpolation strategies, which, in turn, rely on the exceptional global, regional, and national data collection and coordinating systems for satellite and in situ observations.

Acknowledgments. Primary support for this work came from the National Aeronautics and Space Administration (NASA) GPM mission Grant 80NSSC19K0686 and the U.S. Agency for International Development (USAID) Cooperative Agreement 72DFFP19CA00001. Additional support for CF was provided by the U.S. Geological Survey Drivers of Drought and the Famine Early Warning Systems Network programs, U.S. Geological Survey (USGS) Cooperative Agreement G14AC00042. Support for GJH and CP-L was provided by NASA under ISFM WBSs 573945.04.80.01.01 and 573945.04.80.01.03.

Data availability statement. The CHI climatology and CHIMES data will be freely available at data.chc.ucsb.edu, and routinely updated on pentadal and monthly basis. The updated schedule will be identical to the CHIRPS processing stream. Once every pentad, with a delay of approximately two days, a “preliminary” product will be created using submonthly stations. Once a month, following the CHC curation process, a “final” monthly CHIMES field will be produced, and “final” pentadal CHIMES be produced. CHIMES will be licensed under a Creative Commons license.
Appendix A: The CHC station curation process

The CHC Station Climatology Database (CSCD) contains over 200,000 (not necessarily unique) stations with over a billion daily records from 65 different data sources. This appendix summarizes CSCD curation and quality-control procedures. In general, each data source has two or more linked relational databases: a station database containing station metadata, and daily and/or monthly tables containing time-varying observations. The metadata contains an important “use rank” identifier. Some stations are marked as low quality and are not used. Other stations, however, are marked as being of intermediate quality. These are typically from rapidly updated automatic reporting systems like the Global Telecommunication System (GTS) and the Global Summary of the Day archive (GSOD). Finally, stations from carefully curated sources, such as the Global Historical Climatology Network (GHCN) and national meteorology agencies, are given high-use rank values. Such observations are given priority, but augmented as necessary by intermediate-quality gauge totals.

GTS and GSOD screening procedures. While the GTS and GSOD archives are valuable sources of daily precipitation, careful analyses of these sources reveal the occurrence of “false zeros.” False zeros arise when missing data are coded as zero. False zeros can create very misleading and inconsistent results, such as large, erroneous rainfall deficits. To increase the reliability of the GTS and GSOD archives, a comparison between the daily station data and local daily CHIRP values is used to screen and exclude likely false-zero candidates. CHIRP is used so that the station archive can extend back to 1981. This screening dramatically improves performance in some data-sparse regions.

Initial station screening procedures. The following steps are used to identify and isolate questionable rain gauge stations.

1) Station locations are sanity checked to ensure they are in reasonable locations.
2) Stations are compared visually and statistically with neighbors. These evaluations are based on precipitation totals for the wettest three months. For each station, the following process is carried out.
   (i) Four high-quality neighboring stations are identified.
   (ii) A time series of distance weighted mean (DWM) of these neighbors is calculated, and the correlation (R) and differences between the individual station and DWM is calculated. The mean of the absolute differences is divided by the mean of the DWM time series to create a fraction of average error (FAE) value. The R and FAE are then combined into a single quality index:

   \[
   QI = \left(1.0 - \max\left(R, 0.0\right) + \min\left(FAE, 10.0\right)\right) \times 50.
   \]

3) The sum of the monthly QI values for the three wettest months is used to calculate a “total quality index” (TQI) for each station. Locations with TQI values above 100 are examined visually. Time series plots for each station and its neighbors are evaluated. Maps of elevation and climatological average precipitation provide context. CHC scientists assign each station a “high-,” “intermediate-,” or “low-” quality metric. High-quality stations are used to create the anchor station list (described below). Intermediate stations are considered valid “fill sources.” Low-quality stations are not used in any way.

Creating the anchor station list. The next step in the CHC processing creates a unique set of “anchor stations” for each month. All stations that pass the screening, as described above,
are used to define, for each month, a unique set of “anchor station” locations. These anchor stations are then used in two ways:

1) to build as complete-as-possible time series for each given location, which supports the identification of suspicious values, and
2) to support the precomputation of neighbors and neighbor distances, which supports rapid interpolation.

In the first step, all available data sources are used to produce a continuous series of values. A search radius of 0.05° is used to identify potential fill values. In the infilling procedure, higher-quality curated station sources (like the GHCN or national-level meteorological agency data) are preferred over lower-quality sources like the GTS or GSOD. In the second step, for every 0.05° land pixel, a set of nearest neighboring stations is identified, and the pixel-to-station distances are calculated.

**Automatic statistical quality control procedures.** Gamma-distribution-based standardized precipitation index (SPI) and absolute magnitude criteria are used to identify, and mark as problematic, suspicious observations. For monthly data, stations failing the following checks are flagged and excluded:

1) \( \text{CHIRP} > 75 \) AND \( [(\text{CHIRP} - \text{station})/\text{CHIRP}] > 1 \)
2) \( \text{CHIRP} > 75 \) AND climatology > CHIRP AND \( 4 \times \text{CHIRP} < \text{station} \)
3) \( (\text{CHIRP} - \text{station})/\text{CHIRP} < -1 \) AND CHIRP > 75
4) \( \text{[Station} > (5 \times \text{CHIRP}) \text{AND CHIRP} < 20] \text{OR station} > 2000 \)
5) \( \text{CHIRP} > 20 \) AND \( \text{station} > 0 \) AND source = (GTS OR GSOD)
6) The time series of station at anchor location has at least 100 observation and SPI > 4.

Pentad stations are excluded if any of the following criteria are met:

1) \( \text{[Station} > (5 \times \text{CHIRP}) \text{AND CHIRP} > 20] \text{OR station} > 300 \)
2) \( \text{CHIRP} > 7 \) AND \( \text{station} = 0 \) AND source = (GTS OR GSOD)
3) Time series of station at anchor location has at least 100 observation AND SPI > 4

Suspicious observations are not used in further calculations. These automatic procedures are applied every pentad and every month.

**Visual quality control via Reality Checks.** In the middle of every month, a set of monthly station totals from the previous month is rigorously evaluated by the “Reality Checks” team at the CHC. In addition to statistical quality parameters, special “Reality Check” images are generated and examined using an interactive web mapping tool, the Early Warning Explorer (EWX). In these monthly Reality Checks images, draft CHIRPS and CHIMES values are overlain with station values, with station values represented as boxes. Individual station values can be rapidly compared to neighboring station observations, as well as the satellite precipitation fields. The comparisons examine absolute values (monthly totals), relative differences (arithmetic anomalies), and statistical likelihood (SPI). The EWX also displays topography and a wide variety of remotely sensed datasets, which facilitates cross-checking extreme precipitation outcomes and identifying stations with false zeros. Analysts also take into account weather model output, placing potential extremes in climatic context. The team discusses suspicious station observations and removes the most likely problematic stations. Additional ancillary data sources, such as NOAA and WMO resources, news reports, and government
meteorological reports, are frequently used in the Reality Checks process. These discussions and outcomes are documented on a CHC wiki: https://wiki.chc.ucsb.edu/CHIRPS_Reality_Checks.

**Resources for viewing the locations and counts of CHC station archives.** The CHIMES product will build on the same CSCD stations as those used in the CHIRPS product. Readers interested in learning more about the spatial and temporal distributions can use the resources located at http://data.chc.ucsb.edu/products/CHIRPS-2.0/diagnostics/, as described here—https://chc.ucsb.edu/data/chirps/diagnostics.

**Global station density.** Every month, we provide GeoTiff files of the number of stations within each pixel for 0.05° resolution and 0.25° resolution. This product is useful for the identification of station-rich and station-poor regions of the globe. The GeoTiff files are available for download https://data.chc.ucsb.edu/products/CHIRPS-2.0/diagnostics/global_monthly_station_density/tifs/, where you can select the resolution of interest (0.05° or 0.25°).

In addition to the spatial representation of the global data, we also make available the locations of all the stations that make it into the CHIRPS final product. These files are available at https://data.chc.ucsb.edu/products/CHIRPS-2.0/diagnostics/list_of_stations_used/monthly/. In this directory you can find files named “global.stationsUsed.YYYY.MM.csv,” which cover the full time series of CHIRPS, and only stations which were available in the historical record at the time of the release of CHIRPS in 2015. Additional stations made available after the initial release are also available in a similar format, but named “extra.stationsUsed.YYYY.MM.csv.” For periods after 2015, you can combine the two files to find the complete list of station locations making their way into CHIRPS.

**Stations by country.** There are two diagnostic products to evaluate stations going into CHIRPS for an individual country. The first is a map of the stations’ locations falling within a country’s boundaries, as well as neighboring areas, for every month of the historical record. You can find a listing of the mapped countries at http://data.chc.ucsb.edu/products/CHIRPS-2.0/diagnostics/chirps-n-stations_byCountry/ and then click on the country of interest to find the map for each historical month. At the top of each graphic there is also a count of the number of stations within the country, and in the mapped area (including the country’s surrounding regions) for reference.

In addition to the spatial map, there is also a graphic showing the time series of station input to CHIRPS. This can show the trend in how many stations are reporting over time and reveal where station support may be increasing or decreasing. These plots for each country can be found at http://data.chc.ucsb.edu/products/CHIRPS-2.0/diagnostics/stations-perMonth-byCountry/pngs/. Additional versions of this same plot are made over a geographic region and can be found at https://data.chc.ucsb.edu/products/CHIRPS-2.0/diagnostics/stations-perMonth-byRegion. The process described here generates two sets of station values: preliminary and final. The preliminary station values are produced every pentad, with a latency of about two days. Automatic quality control procedures are applied to these values. The final station values are produced monthly, with a latency of two to three weeks. Automatic statistical analyses and visual curation results in reliable, timely, high-quality gauge archives.
Appendix B: The CHIMES process in equations

Creating the $CHP_{clim}$ climatology. Version 2 of Climate Hazards Precipitation Climatology ($CHP_{clim}$) is an update of version 1. Version 1 used moving window regression to translate mean satellite precipitation (the average of the TMPA and CMORPH datasets), elevation, and station normals into high-resolution (0.05°) global climatologies that perform well in data-sparse areas with complex terrain (Funk et al. 2015a). The $CHP_{clim}$ update incorporates IMERG$_{final}$ mean fields and a larger set of gauge undercatch-corrected in situ climate normals. In $CHP_{clim}$, climate normals were not undercatch corrected, leading to a modestly low bias in some mountainous regions (Beck et al. 2017). Monthly climate surfaces are derived as follows. In the first step, a local regression model is derived at each 0.05° pixel for each month:

$$p_{x,y} = b_{0,x,y} + b_{1,x,y} \text{IMERG}_{final} + b_{2,x,y} \text{Elevation.}$$  \hspace{1cm} (B1)

This regression equation is fit to a sequence of 9 × 9 moving windows, centered on every pixel from 60°S to 60°N. The predictand is the $CHP_{clim}$ climatology. The predictors are $\text{IMERG}_{final}$ and ETOPO30 elevation values. This process creates a continuous field similar to $CHP_{clim}$, but without the artifacts associated with the borders between individual $CHP_{clim}$ tiles.

For the next step, a large set of gauge-undercatch-corrected climate normals ($n$) is used to define a set of correction ratios: $r = (n + \varepsilon)/(p + \varepsilon)$; $r$, $n$, and $p$ are vectors of monthly ratios, in situ means, and regression estimates. The latter are derived as described above in Eq. (B1). The $\varepsilon$ term is a small constant, set to 7 mm. Standard inverse distance weighting interpolation is then used generate a grid of adjustment ratios ($a_{x,y}$):

$$a_{x,y} = \sum_{i=1}^{5} w_{i} r_{i},$$ \hspace{1cm} (B2)

where the $w_{i} r_{i}$ represent inverse distance weights, which sum to 1, and a set of five neighboring ratios ($r_{1...5}$).

While the interpolation process creates a spatially continuous set of adjustment ratios, these adjustment factors can extend far beyond the location of the station, overemphasizing the importance of a station in data-sparse regions. To address this problem, two grids of weights ($a_{x,y}$ and $\beta_{x,y}$) are created, based on the distance to the closest station ($d_{x,y}$) and a constant $d_{max}$, which is assumed to be 500 km:

If $d_{min} < d_{max}$, then $a_{x,y} = 1 - (d_{min} / d_{max})^{\alpha}$, \hspace{1cm} (B3)

and $\beta_{x,y} = 1 - a_{x,y}$. \hspace{1cm} (B4)

If $d_{min} > d_{max}$ then $a_{x,y}$ and $\beta_{x,y}$ are set to 0 and 1. A continuous grid of scalars ($s_{x,y}$) that relaxes to 1 as $d_{min}$ approaches $d_{max}$ is then produced:

$$s_{x,y} = a_{x,y} a_{x,y} + \beta_{x,y} \times 1.$$ \hspace{1cm} (B5)

At distances greater than $d_{max}$, $s_{x,y}$ will relax to 1. At small distances, $s_{x,y}$ will equal $a_{x,y}$. The $CHP_{clim}$ is then created as follows:

$$CHP_{clim,x,y} = s_{x,y} p_{x,y}.$$ \hspace{1cm} (B6)

Thus, $CHP_{clim}$ incorporates mean IMERG$_{final}$ and elevation fields [Eq. (B1)] and station normals [Eqs. (B2)--(B6)] to produce a continuous high-resolution climatology. These are created for each month, and then the 12 monthly values are disaggregated to daily and pentadal means.
using a nonparametric smoothing procedure. Visual quality control was also performed on all the monthly fields by the CHC team, in a manner similar to the Reality Checks process described above. Questionable stations were removed, and the gauge enhancement rerun. This process was repeated twice.

**Temporal versus spatial homogeneity.** It should be noted that there is a mismatch between the 1981 and 2010 time period used to produce the CSCD climate normals and the 2001–present time period of the IMERG\textsubscript{late}. This mismatch arises from the declining number of stations.\textsuperscript{b1} We selected a 1981–2010 baseline to capture the much greater station density over that time period. We selected this time period to maximize the number of stations used in our climatology. In general, precipitation means change much more rapidly from place to place than from decade to decade. CHIMES has been designed to work well in countries like Ethiopia, in which mean precipitation values can change by 200% or 300% over very short distances. See, for example, (Funk et al. 2015a) which describes our climatology-building process. The boreal spring season in eastern Ethiopia has been declining, but that decline is on the order of 15%–20%, this is also discussed in our CHIRPS paper (Funk et al. 2015b). As a general principle, spatial variability in mean precipitation rates is much stronger than decadal variability, so we have tried to be as contemporaneous as feasible while including a large number of stations. Another factor relates to the shrinking number of available station observations, which can be visualized for all stations (https://data.chc.ucsb.edu/products/CHIRPS-2.0/diagnostics/stations-perMonth-byRegion/pngs/all.station.count.CHIRPS-v2.0.png) and for stations in Africa (https://data.chc.ucsb.edu/products/CHIRPS-2.0/diagnostics/stations-perMonth-byRegion/pngs/Africa.station.count.CHIRPS-v2.0.png). Unfortunately, there are many fewer stations available over the 2001–present time period, and this is why we selected a 1981–2010 mean period. Note also that if there are station data that have shifted between 1981–2010 and 2001–20, then this information would show up in the blending procedure and be incorporated in the CHIMES dataset.

**Reducing bias in IMERG\textsubscript{late}.** For each month and each location, the ratio of \text{CHP}_{clim,x,y} and \text{IMERG}_{late,x,y} can be calculated as

\[
r_{x,y} = \frac{\text{CHP}_{clim,x,y} + \varepsilon}{\text{IMERG}_{late,x,y} + \varepsilon},
\]

where a value of 7 mm is used for \(\varepsilon\). Experimentation suggests that a value near 7 performed reasonably well in areas with low rainfall, while also not distorting \(r_{x,y}\) values too much. The \(r_{x,y}\) terms are produced for each month, and then the 12 monthly values are disaggregated to daily values using a nonparametric smoothing procedure. Daily CHIME values can then be produced via a simple scaling of the daily IMERG\textsubscript{late} fields:

\[
\text{CHIME}_{x,y,day} = r_{x,y,day} \times \text{IMERG}_{late,x,y,day}.
\]

**Calculating decorrelation slopes.** A unique aspect of the station-blending procedure used in CHIMES is that location and season-dependent precipitation decorrelation slopes are used to guide the station-blending process. In areas and seasons with more spatially coherent precipitation, the information associated with a given station observation is given a greater area of influence. The IMERG\textsubscript{final} provides this coherence information. This process begins by converting monthly IMERG\textsubscript{final} values into percent anomalies:
\[ PCT_{x,y} = \frac{\text{IMERG}_{\text{final},x,y} + \varepsilon}{\text{IMERG}_{\text{final},x,y} + \varepsilon}, \]  

where a value of 7 mm is used for \( \varepsilon \). Experimentation suggests that a value near 7 performed reasonably well in areas with low rainfall, while also not distorting PCT \( x,y \) values too much. This process removes the long-term mean information. Then for each location and month, a ring of 72 neighboring pixels 1.5° away is identified. For each month from 2001 to 2019, these 72 neighboring percentage values are extracted and stored in a 19 × 72 element vector \( \text{n} \). A similar 19 × 72 element vector is created \( \text{c} \), but using replicates of the central pixel, i.e., the center value is repeated 72 times in 2001, and then again in 2002, etc. The correlation \( \text{c} \) between \( \text{n} \) and \( \text{c} \) represents one measure of the rate at which information dissipates with increasing distance. Assuming that the “true” correlation is 1 at the center pixel, and \( d \) is the distance associated with 1.5°, then the decorrelation slope can be estimated as

\[ \tau_{x,y} = \frac{1 - C_{x,y}}{d}, \]  

where \( \tau_{x,y} \) has units of per kilometer \( (\text{km}^{-1}) \). This is equivalent to a linear covariogram model with a nugget of zero (Isaaks and Srivastava 1989).

**The station blending procedure.** At monthly and pentadal time scales, station observations and decorrelation slopes [Eq. (B10)] can be used to perturb/improve a precipitation grid \( p_{x,y} \). For each pixel, we locate the five nearest neighboring gauge observations \( g \). Note that different stations are likely be identified at different times in data-sparse areas, due to incomplete reporting. For each of these stations, we calculate inverse distance weights based on the associated distances:

\[ w_i = \begin{cases} 1 & \text{if } 1 - d_i \tau_{x,y} > 0, \\ \frac{1}{(1 - d_i \tau_{x,y})^2} & \text{else} \\ 0 & \end{cases} \]  

This formulation is different from naive inverse distance weighting schemes, because the influence of distance is predicated on the local IMERG decorrelation structure [Eq. (B10)]. A critical component of the gauge enhancement computation is the precalculation of the pixel-to-anchor neighbors and distances. This precludes a very intensive search/sort procedure for every pixel, for every interpolated field. Assuming that we have a set of five perturbation adjustments, \( a = (g + \varepsilon)/(p + \varepsilon) \), the weights defined in Eq. (B11) are then used to create a typical inverse distance weighted scalar anomaly:

\[ S_{x,y} = \frac{\sum_{i=1}^{5} w_i a_i}{\sum_{i=1}^{5} w_i}. \]  

The next step of the process takes into account the anticipated information decay associated with the remote location of the stations. The rigorous treatment of this problem, i.e., kriging (Journel and Huijbregts 1978), involves defining all the expected spatial covariances and inverting the associated covariance matrix. Inverting matrices for every 0.05° pixel is not computationally feasible. Kriging-like behavior, however, can be approximated by using two correlation values: the expected correlation of the nearest station, based on the decorrelation slope \( r_{\text{neighbor}} \), and an assumed correlation between station anomalies and satellite-estimated anomalies \( r_s \). The \( r_{\text{neighbor}} \) is calculated as \( 1 - d_{\text{min}} \tau_{x,y} \) and set to zero if negative.
This term is quite sophisticated in the sense that it incorporates a direct measure of spatial coherence based on IMERG observations, as described above. Assumptions for $r_S$ are based on diagnostic validation studies examining the correlation of gauge observations and the IMERG$_{lat}/$CHIME. While $r_S$ could vary by location and season, it is assumed to be constant in this version of CHIMES. These correlation assumptions are translated into weight fields that sum to one, and represent our relative confidence in the station-adjustment ratios:

$$\alpha_{x,y} = \frac{r_S^2}{r_S^2 + r_{\text{neighbor}}^2},$$

$$\beta_{x,y} = \frac{r_{\text{neighbor}}^2}{r_S^2 + r_{\text{neighbor}}^2}. \quad (B13)$$

As we move away from the closest station, $\alpha$ will go to zero. When we are collocated with a station, $\alpha_{x,y} = \frac{r_S^2}{r_S^2 + 1}$. A pixel containing a station will still be a combination of the CHIME estimate. Hence, in interpolation parlance, this scheme is not a “perfect” interpolator, because it will not perfectly reproduce the station values. Rather, it provides a reasonable blend of these two sources of information. The final CHIMES estimates are then obtained:

$$\text{CHIMES}_{x,y} = \alpha_{x,y} \text{CHIME}_{x,y} + \beta_{x,y} s_{x,y} \text{CHIME}_{x,y}. \quad (B14)$$
Aon-Benfield Reinsurance Company, 2021: Weather, climate and catastrophe in- 
https://www.aon.com/global-weather-catastrophe-natural-disasters-costs-
medium=clients&utm_campaign=natcat21.
rainfall and cold cloud over the Western Hemisphere during 1982–84. Mon. 
TRBLC>2.0.CO;2.
Arsenault, K. R., and Coauthors, 2020: The NASA hydrological forecast system for 
https://doi.org/10.1175/BAMS-D-18-0264.1.
datasets using gauge observations and hydrological modeling. Hydrocl. 
Becker, A., P. Finger, A. Meyer-Christoffer, B. Rudolf, K. Schamm, U. Schneider, and 
M. Ziese, 2013: A description of the global land-surface precipitation data 
products of the Global Precipitation Climatology Centre with sample applications 
including centennial (trend) analysis from 1901–present. Earth Syst. Sci. 
Contractor, S., and Coauthors, 2020: Rainfall Estimates on a Gridded Network 
(REGEN)—A global land-based gridded dataset of daily rainfall from 
Dinku, T., C. Funk, P. Peterson, R. Maidment, T. Tadesse, H. Gadain, and P. Ceccato, 
2018: Validation of the CHIRPS satellite rainfall estimates over eastern of Afri- 
Funk, C., 2021: Drought, Flood, Fire: How Climate Change Contributes to Recent 
——, and S. Shukla, 2020: Drought Early Warning and Forecasting: Theory and 
global satellite assisted precipitation climatology. Earth Syst. Sci. Data, 7, 
——, and Coauthors, 2015b: The Climate Hazards Infrared Precipitation with 
Stations—A new environmental record for monitoring extremes. Sci. Data, 2, 
150066, https://doi.org/10.1038/sdata.2015.66.
——, and Coauthors, 2019: Recognizing the Famine Early Warning Systems 
Network: Over 30 years of drought early warning science advances and 
partnerships promoting global food security. Bull. Amer. Meteor. Soc., 100, 
Harris, I., T. J. Osborn, P. Jones, and D. Lister, 2020: Version 4 of the CRU TS monthly 
doi.org/10.1038/s41597-020-0453-3.
Hong, Y., K.-L. Hsu, S. Sorooshian, and X. Gao, 2004: Precipitation estimation 
from remotely sensed imagery using an artificial neural network cloud classi-
M2173.1.
Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mis-
Hsu, K., X. Gao, S. Sorooshian, and H. V. Gupta, 1997: Precipitation estimation from 
>2.0.CO;2.
Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project 
(GPPC) combined precipitation dataset. Bull. Amer. Meteor. Soc., 78, 5–20, 
——, and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis 
(TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at 


