



Article

Evaluating the Performance of Satellite Derived Temperature and Precipitation Datasets in Ecuador

Rachel Huber Magoffin ¹, Riley C. Hales ¹ , Bolívar Erazo ^{2,3} , E. James Nelson ^{1,*}, Karina Larco ^{1,4,5} and Taylor James Miskin ¹

¹ Department of Civil and Construction Engineering, Brigham Young University, Provo, UT 84602, USA; rhuber6@byu.edu (R.H.M.); rchales@byu.edu (R.C.H.); karinaml@student.byu.edu (K.L.); tjm98@byu.edu (T.J.M.)

² Instituto Nacional de Meteorología e Hidrología (INAMHI), Quito 170517, Ecuador; bolivar.erazo@gmail.com

³ Departamento de Gestión de Recursos Hídricos, Empresa Pública Metropolitana de Agua Potable y Saneamiento de Quito, EPMAAPS Agua de Quito, Quito 170519, Ecuador

⁴ Fundación EcoCiencia, Quito 170517, Ecuador

⁵ SERVIR-Amazonia, Cali 76001, Colombia

* Correspondence: jimn@byu.edu

Abstract: Temperature and precipitation data are crucial for hydrology and meteorology. In 2014, Ecuador started an automatic gauge network which monitors these variables. The measurements are not publicly available. Global gridded datasets from numerical models and remote sensors were previously the only way to obtain measurements for temperature and precipitation. Now that in situ measurements are beginning to be available in significant quantities, we assessed the performance of IMERG, CHIRPS, GLDAS and ERA5 for both temperature and precipitation using the in situ data. We used the Pearson R correlation coefficient, ME (Mean Error), MAE (Mean Absolute Error), and RMSE (Root Mean Square Error). We found that global gridded data were more suited for determining averages over time rather than for giving exact values at specific times for in situ gauges. The Pearson R values increased for all datasets when we used monthly aggregations in place of daily aggregations, suggesting that the monthly values are more correlated than the daily. The Pearson R value for temperature increased from 0.158 to 0.719 for the ERA5 dataset. Additionally, we show the statistical values for each of the three regions in Ecuador. We found that the IMERG and CHIRPS datasets, which contain station data, performed significantly better for both RMSE and MAE. Both IMERG and CHIRPS have a RMSE value a little over 260, whereas ERA5 and GLDAS had values over 300. We discuss the short comings of these datasets as being related to their relatively coarse resolution, lack of in situ data in Ecuador to calibrate against, and the rapidly varying terrain of Ecuador. We recommend using higher temporal and spatial resolution datasets for immediate applications. We recommend repeating this analysis in the future when more automatic gauges and longer time periods are available to facilitate a more detailed analysis which is presently not possible.

Keywords: precipitation; temperature; hydrology; remote sensing; reanalysis datasets



Citation: Huber Magoffin, R.; Hales, R.C.; Erazo, B.; Nelson, E.J.; Larco, K.; Miskin, T.J. Evaluating the Performance of Satellite Derived Temperature and Precipitation Datasets in Ecuador. *Remote Sens.* **2023**, *15*, 5713. <https://doi.org/10.3390/rs15245713>

Academic Editor: Yuanjian Yang

Received: 9 November 2023

Revised: 5 December 2023

Accepted: 10 December 2023

Published: 13 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Accurate temperature and precipitation data play an important role in hydrology and meteorology. These important climate variables are used in numerical models, environmental status and outlook products, climate change measurements, and detection of hydrologic patterns over time [1–4]. In situ measurements, or gauge data, are the typical source for this information, but many countries lack local capacity and funding to build, maintain, and operate a national gauge network [5]. Even where funding is present, it is not practical to monitor in situ measurements throughout the whole country. Kidd et al. found that only about 1% of the Earth's surface is covered with rain gauges [6] proving that there is a substantial gap in the data provided by rain gauges alone. Gauge measurements often

have incomplete spatial and temporal coverage [7,8]. Rain gauges are particularly prone to errors in regions with changing topography, such as Ecuador, where precipitation has high spatiotemporal variability [9,10]. Gauge measurements in such places likely do not accurately represent the surrounding region.

Global satellite observations and modeled data are critical to fill gaps in poorly gauged regions [11]. The global temperature and precipitation data are typically in a gridded format and either collected at or resampled to consistent temporal intervals. The data are desirable for their uniform and consistent spatiotemporal coverage. Several such global gridded datasets exist, and this paper focuses on IMERG, CHIRPS, ERA5, and GLDAS. These datasets have been evaluated throughout the world and are well documented in published research. Ecuador began a national automatic gauge network for hydrometeorological variables in 2014. The network is relatively young but has a growing number of gauges and period of record for the maintained gauges. Until this gauge network was established, nationwide in situ measurements were not available for comprehensive evaluation of remotely sensed values or calibration of models covering Ecuador.

Ecuador poses unique challenges for gauge networks because of its complex topography which includes three main regions, namely: the Andes Mountains, the Pacific coastal region, and the Amazon rainforest. Condom et al. studied in situ data in the Andean region and found only 49 stations were reported from Ecuador to OSCAR (Observing Systems Capability Analysis and Review Tool) in the WMO (World Meteorological Organization). In the same study, researchers also identified that there are significant gaps in data availability in these regions [12]. Ecuador is not adequately represented in global studies due to the limited information provided and the neighboring countries which contain the Andes mountains not contributing large amounts of data to global analyses. Furthermore, Condom et al. found that most of the stations were concentrated in the mountains and on the coast with limited data availability in the Amazon basins. This shows that there is placement bias with the gauges in Ecuador and the country is not equally represented in the available in situ data. It has been shown that the spatial distribution of gauges can impact the ability to estimate water resources [13]. Therefore, it is essential to have data distributed throughout the country of Ecuador. Ecuador is in South America which is considered a data sparse area [14]. The performance of global gridded datasets varies depending on the area studied and the conditions of the study. This shows that it is necessary to evaluate these products on a case-by-case basis. The results from one area of the world cannot be assumed to apply to other areas even if similar topographical features exist.

One validation study of global gridded datasets in the Andes observed the lack of in situ data in that region. It showed that global climate models could be downscaled to meet the need for data in ungauged areas. One of the limitations observed for a dataset that incorporates stations, such as CHIRPS, is that it is based on in situ data. This data is extremely limited in the Andean region, which affects the quality of CHIRPS in this region [12]. In northeastern Brazil, CHIRPS performed better during the wet season, and in general it overestimated low rainfall values and underestimated high rainfall values [15]. CHIRPS was compared to in situ data in Colombia. In Colombia there is an uneven distribution of rain gauges due to the complex topography. It was found that CHIRPS can be useful in understanding the rainfall in Colombia [16].

IMERG has been tested as an input into a hydrologic model in the Amazon Basin of Peru and Ecuador. It was found that the data was useful in predicting the streamflow in this data sparse region [17]. In the Andean and Amazonian subregions, IMERG performs better with larger spatiotemporal aggregations [18]. TRMM, which is a satellite that contributes to IMERG, has been tested on the Pacific Slope and Coast of Ecuador showing low bias over lowlands but more limitations over the Andes [19]. IMERG has been evaluated in various regions of the world [20]. In Europe, it was shown that IMERG underestimated precipitation in the mountainous regions [21]. In Australia, IMERG outperformed many other satellite precipitation products [22]. In Turkey, the precipitation was underestimated on the coasts [11].

ERA5 has also been validated using in situ data. A study performed in China compared ERA5 as a modeled dataset to satellite-based datasets, including IMERG, across mainland China [23]. The results show similar distributions of the mean annual precipitation between the datasets. However, generally the satellite-based products outperformed the ERA5 data [23]. GLDAS performed poorer when compared to other precipitation products in the Biliu basin in China [24].

Several temperature analyses have also been performed using these datasets. A study in the Tibetan Plateau assessed how well air temperature reanalysis datasets represented the air temperature values. It found that GLDAS 2.1 and ERA5 performed the best [25]. Ji et al. studied how the air temperature from GLDAS compared to global weather station data. They found that the air temperature in South America and Africa tended to be less accurate than in other areas of the world, and that the temperature estimates are less accurate in high mountain areas [26]. ERA5 air temperature was evaluated in Turkey and it was found to have a correlation coefficient of 0.97 [27].

Some studies compared the performance of both precipitation and temperature datasets. A study performed in Africa used both temperature and precipitation datasets as the input into the model. Different combinations of datasets performed differently, but overall, the precipitation contributes to a lot more uncertainty than the temperature does. Temperature has much smaller spatial and temporal variability than precipitation [28].

The global gridded datasets evaluated in this paper can help meet the needs of Ecuador. We chose these datasets because they contain information about temperature and/or precipitation globally, with frequent time steps. They also have data that are current and available on GEE (Google Earth Engine). As global datasets have become more widely available, GEE has made it easier to access and use large quantities of geospatial data [29,30]. GEE has been used to work with precipitation and temperature datasets in many studies [31–33]. GEE was chosen for this study because of ongoing projects in Ecuador by teams also using GEE. This makes our findings more useful and relevant to Instituto Nacional de Meteorología e Hidrología de Ecuador (National Institute of Meteorology and Hydrology, INAMHI). In a survey of 300 journal papers written using GEE, only one of them involved the country of Ecuador [27]. It is possible that studies were performed using alternative methods to analyze global gridded data, but this statistic does suggest that Ecuador is underrepresented in studies comparing global gridded datasets to in situ data. Additionally, the gauge network of Ecuador is also relatively new, so until recently there was not enough data to perform this type of study. The automatic gauge network in Ecuador was started in 2014, and therefore is less than ten years old. There is a network of conventional gauges with data series from around 1960, but for this study, we will use the automatic network because it provides us with hourly data. The gauge network used in this study is not publicly available data. This case study could be performed because we had access to more data than has been historically available for Ecuador. While similar studies exist in other areas of the world, this paper assesses the results in Ecuador.

This paper is a case study that aims to assess the quality of global gridded datasets over Ecuador using the observations from the new automatic gauge network that was initiated in 2014. Understanding the quality of the global datasets across Ecuador is now possible and can enable using the datasets with greater confidence to fill the gaps in recent precipitation and temperature observations as well as further into the past. We compare precipitation data from IMERG, CHIRPS, ERA5, and GLDAS as well as air temperature data from GLDAS and ERA5. We compare values from each of these datasets with a matching gauge from the Ecuador automatic network. The goal is to understand if the global datasets provide adequate coverage and accuracy to be used to fill spatial and temporal gaps in the in situ data for Ecuador. In this study, we compare the datasets using performance metrics. We perform comparisons at hourly (where available), daily, and monthly temporal aggregations. By evaluating the performance of various datasets throughout Ecuador, we provide insights into understanding the potential of global gridded data to bridge gaps in in situ data, especially in poorly gauged contexts.

2. Materials and Methods

2.1. Study Overview

We collected and preprocessed in situ data from Ecuador. Section 2.2 introduces our study area. Section 2.3 explains the quality assurance preprocessing measures we applied to the observations. Section 2.4 presents the procedures for preparing the global datasets from GEE.

Next, we computed error and performance metrics to compare the gridded and in situ datasets. The data from GEE were gridded and the in situ data were point data. We approximated the gridded data as points. Then we calculated error metrics and the Pearson R correlation coefficients.

These processes are described in Figure 1.

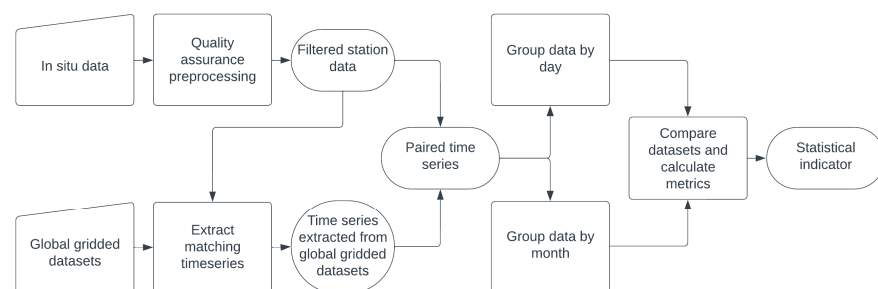


Figure 1. Flowchart showing the methods section.

2.2. Study Area

Our study area for this paper is the country of Ecuador, excluding the Galapagos islands since we had no access to data from the islands. Ecuador is located on the equator in the northwest of South America. The Andean Mountains run through the center of Ecuador from north to south. The Amazon is to the east and the Pacific coast is to the west. There is a large range of elevation from the coast to the top of the Andes mountains at elevations around 6300 m. The climatology is significantly influenced by the Andes. The topography creates substantial variability and serves as a climatic barrier between the west and east side of the country. Rainfall patterns are influenced by the movement of the intertropical convergence zone, changes in sea surface temperatures, transport of moisture from the Amazon, trade winds, and the impact of Hadley and Walker circulation cells. Precipitation along the coast has an annual average around 1600 mm while certain stations in the Andes record less than 900 mm. A total of 80% of rainfall occurs during the wet season between December and May [34].

2.3. Collecting and Preprocessing In Situ Data

We obtained hourly in situ precipitation and temperature data for our analysis from the INAMHI. The data were provided as part of a collaborative project and covers nine full years from January 2014 to December 2022.

We selected data from an initial database of 132 stations with temperature and precipitation data from INAMHI. Based on data quality, consistency in station coordinates, and data series records, we initially kept the data from 80 stations for precipitation and 83 for temperature, as shown in Figure 2. The gauges are not evenly distributed. Most gauges are located near Quito. There are only a few gauges in the Amazon Rainforest, to the east of the Andes. This is shown in Table 1.

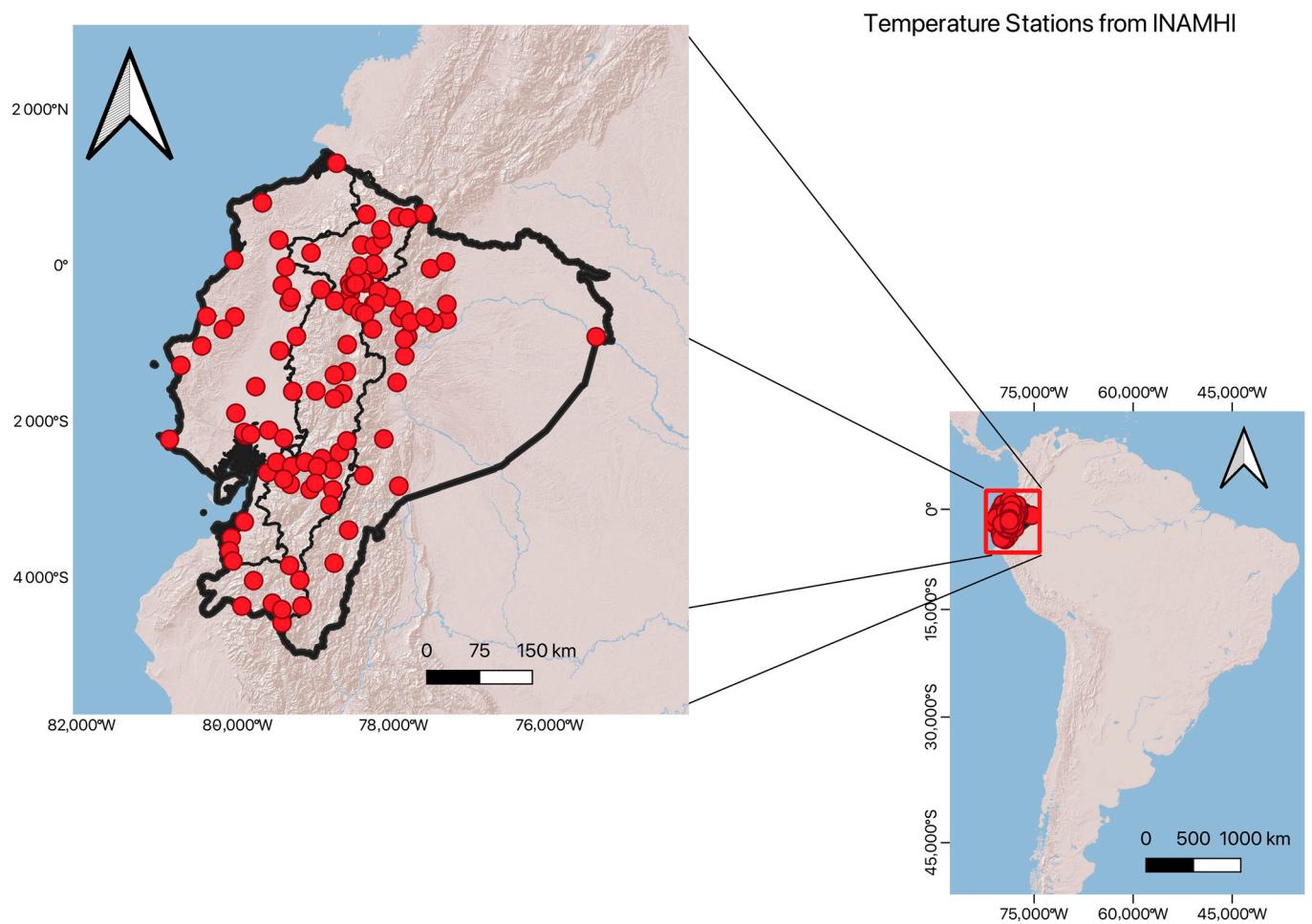


Figure 2. Locations of the stations provided by INAMHI.

Table 1. Precipitation stations with data in each of the subregions.

	Number of Precipitation Stations
Amazon	16
Coast	21
Mountains	38

The in situ devices recorded hourly measurements, but several had occasional gaps in coverage. These temporal gaps spanned anywhere from a couple of hours to a couple of months. Each station covered different date ranges, as shown in Figure 3. In this chart, each bar on the x-axis represents a station and the y-axis indicates the available data ranged from 2014 to 2022. The bar starts at the first recorded date of the station and extends to the last recorded date in the station. The graphic does not show the temporal gaps in the data. Each line in the figure represents a station and the length represents the duration of the station. Some stations had large amounts of available data whereas others only had data for short time periods. The start of INAMHI's automatic weather stations (AWS) network began in the year 2014, and this is the reason for the data starting then. Figure 3 also reveals that around 30 stations have data covering the entire period from 2014 to 2022. More than half of the in situ data covers shorter intervals. Several were discontinued near the beginning of 2018 or began later in the same year. These changes are due to shifting priorities and funding adjustments for AWS maintenance. The period with the most stations available was in 2016.

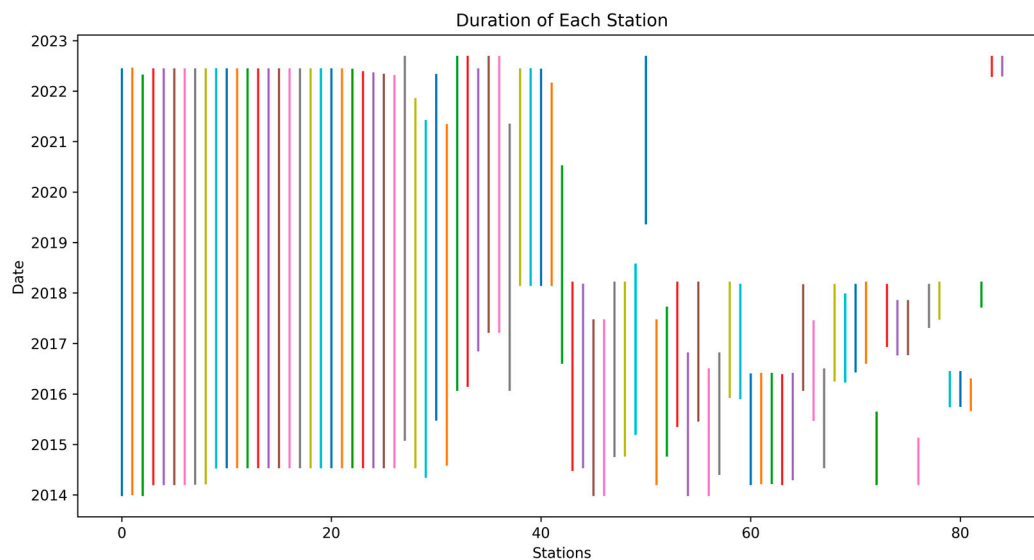


Figure 3. A bar graph where each bar represents a station, and the height represents the length of coverage from each station's first and last reported measurements.

We are aware that there are limitations to using a gauge network with fewer years of data. Due to this limitation, data were not grouped by annual or seasonal values because there are not enough available data to be able to properly compare at these scales. Furthermore, with the limited data, outliers have a chance of having a larger impact on the data. We are also aware that this could lead to limitations in the accuracy of our results.

Additionally, portions of the data were unusable and needed to be removed. The in situ data had inconsistencies, including values that were outside a possible range in Ecuador. Therefore, we applied techniques to filter the data. First, we removed all temperature values that were exactly equal to 0. Due to the abundance of 0 values in our dataset, we suspected that 0 denoted a missing value rather than an actual temperature reading. We assumed that instances of exactly 0 degrees Celsius were rare, as opposed to values such as 0.1 or -0.1 , and thus the loss of data due to this exclusion was minimal. We also removed any temperature values that were below -30 degrees Celsius or above 45 degrees Celsius. These values are outside a possible range of temperature values in the country of Ecuador.

We excluded negative values and those exceeding 320 mm/h for precipitation. We chose this value because it is slightly above the world record rainfall recorded according to the Army Corps of Engineers [35]. Some stations recorded values where hourly precipitation was over 1000 mm. We assumed that if Ecuador was recording values higher than the world record in their rain gauges it would have been documented as a record-breaking storm. In some of the stations, most values were above 320 mm or below zero. Therefore, we chose to remove them from the study because we could not verify that the other numbers in the stations were correct. Three stations were removed entirely from the precipitation data comparisons because of this filtering.

2.4. Collecting Gridded Data

IMERG, CHIRPS, GLDAS and ERA5 were used as the global gridded datasets to be examined in this study. Each uses slightly different methods to collect its values. IMERG (Integrated Multi-Satellite Retrievals for GPM), produced by NASA (National Aeronautics and Space Administration) and JAXA (Japan Aerospace Exploration Agency), uses information from GPM (Global Precipitation Measurement) to produce global precipitation data. IMERG aims to combine precipitation estimates from microwave sensors with infrared-based observations. These measurements are validated using ground-based observations [36]. CHIRPS (Climate Hazards Infrared Precipitation with Stations) dataset uses the Cold Cloud Duration (CCD) method to approximate rainfall. CHIRPS is a blend of

station data and satellite data. CHIRPS uses TMPA (TRMM Multi-Satellite Precipitation Analysis) to calibrate global CCD products and interpolated gauge products. It aims to fill the gaps in datasets by being low latency, having a high resolution, and a long period of record [37]. ERA5 (ECMWF Re-Analysis) is the successor of ECMWF's (European Centre for Medium-Range Weather Forecasts) ERA-Interim reanalysis dataset and contains data from 1950 to the present. ERA5 is built from data that is based on the IFS (Integrated Forecast System) model cycle Cy41r2. It also uses LDAS (land data assimilation) to analyze land surface variables [38]. GLDAS (Global Land Data Assimilation System) uses ground and space-based observation systems to constrain the data. The land surface model is forced with meteorological data that drives the land surface model [39]. GLDAS produces data for many different variables, but we use only temperature and precipitation. Both variables are forcing variables of the model and therefore are not an output from the model.

We used GEE to obtain and process the remote sensing data for this project. The information for the datasets we used is presented in Table 2. The specific GEE identifiers and band names we chose to use in this study are given and further details are provided by GEE at the GEE data catalog: <https://developers.google.com/earth-engine/datasets> (accessed on 15 June 2023).

Table 2. Summary of Datasets from GEE.

	GEE Identifier	Band Name	Spatial Resolution	Temporal Resolution	Citation
GLDAS	NASA/GLDAS/V021/NOAH/G025/T3H	Tair_f_inst, Rain_f_tavg	¼ degrees latitude	3 h	[39]
ERA5	ECMWF/ERA5_LAND/HOURLY	temperature_2m, total_precipitation_hourly	0.1 degrees latitude	1 h	[40]
IMERG	NASA/GPM_L3/IMERG_V06	HQprecipitation	0.1 degrees latitude	30 min	[41]
CHIRPS	UCSB-CHG/CHIRPS/DAILY	precipitation	0.05 degrees latitude	daily	[37]

2.5. Performance Metrics at Discrete Points

To compare the raster data to point data, we extracted values from the gridded data at the cells containing the in situ stations. If multiple stations were contained within the same cell, we compared each station separately to the values at that cell. Then, we filtered the data so that it matched the date range that was available for the in situ data. The result is a time series of values from each of the GEE datasets for the corresponding location and times of each of the gauges. This allowed us to calculate error and performance metrics based on the time series data. We calculated the sum for each time step used for precipitation and the averages for precipitation.

There were many gaps in the time steps of our data. When doing monthly comparisons, we removed all months where the in situ data did not have values from at least 15 days. This ensures that we are using months where the in situ values are representative of the whole month. This applies when we did monthly comparisons for all types of tests.

We chose to focus on mean error (ME), root mean square error (RMSE), mean absolute error (MAE), and Pearson correlation coefficient (R). We also calculated the *p*-value using the paired *t*-test for temperature. We used the SciPy package to calculate for Pearson R and paired *t*-test values [42]. We calculated each of these values using different temporal groupings. For temperature, we used hourly for ERA-5 and 3-hourly for GLDAS, daily averages, and monthly averages. For the precipitation data, we used daily and monthly sums.

The following equations were used:

$$R = \frac{\sum (x - m_x)(y - m_y)}{\sqrt{\sum (x - m_x)^2 \sum (y - m_y)^2}} \quad (1)$$

$$ME = \frac{1}{n} \sum_{i=0}^n x_{\text{global data}} - x_{\text{in situ}} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=0}^n |x_{\text{global data}} - x_{\text{in situ}}| \quad (3)$$

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=0}^n (x_{\text{global data}} - x_{\text{in situ}})^2} \quad (4)$$

$$p \text{ value} = 2 * \min(P(T \leq t), P(T \geq t)) \quad (5)$$

Using temperature data, we performed a paired *t*-test with the hypothesis of equal means to show whether the differences in the datasets were significant. We did not perform this test with the precipitation data because the large number of zeros in the precipitation data creates significant skew. The precipitation data are not normally distributed and are not easily transformed to fit a normal distribution. RMSE, MAE, and ME are measured in Celsius for temperature and millimeters for precipitation. They indicate the typical magnitude of error between each dataset and the in situ data. Values closer to zero indicate better performance. We chose ME to show the bias of the dataset. RMSE shows the overall spread of the errors with errors that are greater being weighted more. The Pearson R correlation coefficient shows the correlation of the data and whether the datasets are identifying the same trends. The Pearson R values could range from negative one to one and values closer to one indicate better performance. Values closer to one show that the data is highly positively correlated, suggesting that it is more likely that the two datasets are showing the same trends over time. We created maps of the gauge locations colored by their values in each statistic to see how the global gridded data performed in different areas of Ecuador.

We feel that the metrics we chose helped to portray the error, bias, and correlation between the datasets. These performance metrics were also chosen based on our available data. In this study, we compare monthly and daily aggregations for precipitation data instead of hourly values. This is due to the missing values in the in situ data and different temporal resolutions of the global gridded data. Therefore, it was not reasonable for us to perform a false alarm ratio or probability of detection, even though it was prevalent in other studies [23,24].

As we began to evaluate the in situ data, we noticed that a couple of additional stations needed to be removed from the evaluation process. We calculated the RMSE, and six stations had much larger RMSE than the rest of the stations for all four of the datasets. The values ranged between 1000 and greater than 10,000. The RMSE for these six stations were outliers for every dataset. We found that many of the values for those six stations were removed already due to the parameters explained. Many of the remaining values had repeated hourly values that were quite high. While these values were not removed by themselves because they were beneath our threshold, the total sum for daily precipitation was higher than is possible. These stations were removed from the precipitation in situ data because they appear to be inaccurate in representing true precipitation values based on their consistently high RMSE values when compared against the global gridded data.

Figure 4 shows a box and whisker plot of the RMSE values once the stations were removed. These stations were all used in the analysis. As we proceeded with the analysis, the values from the removed stations were not included in the results for precipitation.

Once these stations were removed, we proceeded with the analysis of comparing the datasets. Information about the error metrics and Pearson R values are presented. Finally, we present scatterplots where we graphed the in situ data versus the global datasets. This allowed us to visualize the spread of the data and to see how the data compared.

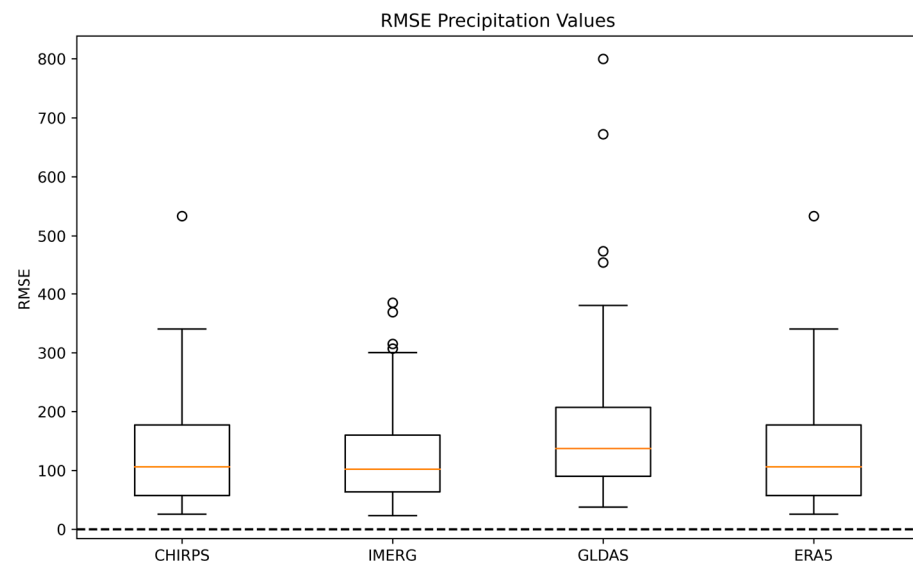


Figure 4. RMSE values in mm of precipitation for each of the four precipitation datasets after the six outliers were removed. The circles represent the remaining outliers in the dataset.

3. Results

3.1. Performance Metrics of Temperature

3.1.1. Paired *t*-Test *p*-Values

We performed several tests to compare values at discrete points between in situ data and the global gridded datasets as described in Section 2.4. The paired *t*-test returns a *p*-value for the temperature comparison for each station. Table 3 shows the average *p*-values returned from this test and the number of stations with a *p*-value > 0.05. Stations with *p* > 0.05 means those stations have no statistically significant difference at the 95% confidence level. The table shows that the *p*-value increased as the temporal aggregation of the data increased. The *p*-values are lower for daily average comparisons and larger for monthly average comparisons. This means the difference is less statistically significant at coarser time steps.

Table 3. Results from the paired *t*-test for temperature measurements.

	ERA 5		GLDAS	
	<i>p</i> -Value	Stations with <i>p</i> -Value > 0.05	<i>p</i> -Value	Stations with <i>p</i> -Value > 0.05
Hourly (ERA5) or 3-hourly (GLDAS)	<0.001	0	0.00186	0
Daily Averages	0.0007	0	0.0106	2
Monthly Averages	0.0253	6	0.0676	13

3.1.2. Pearson R Correlation Coefficient

We computed the Pearson R correlation coefficient between the in situ data and the gridded data at each available station. The Pearson R was calculated using all data pairs, daily average pairs, and monthly average pairs. Values closer to one show that the values between in situ data and the global data are highly positively correlated. We found the Pearson R at each station individually and then averaged the results (Table 4). The average correlation coefficient increased with greater temporal aggregations, like the results from the paired *t*-test. Figure 5 shows boxplots of the correlation coefficient values for each of the different stations. Each plot in the figure represents the Pearson R values at different temporal groupings. The top plot shows the results for the GLDAS data, and the bottom shows the results from ERA5. The boxplots show the average Pearson R value increasing as

the temporal groupings increased. The GLDAS plot has a larger spread between the values as the temporal aggregations increase. Both datasets show that there are some values with negative correlation coefficients.

Table 4. Pearson R values for temperature.

	ERA5	GLDAS
All Available Data	0.158	0.267
Daily Averages	0.593	0.490
Monthly Averages	0.719	0.556

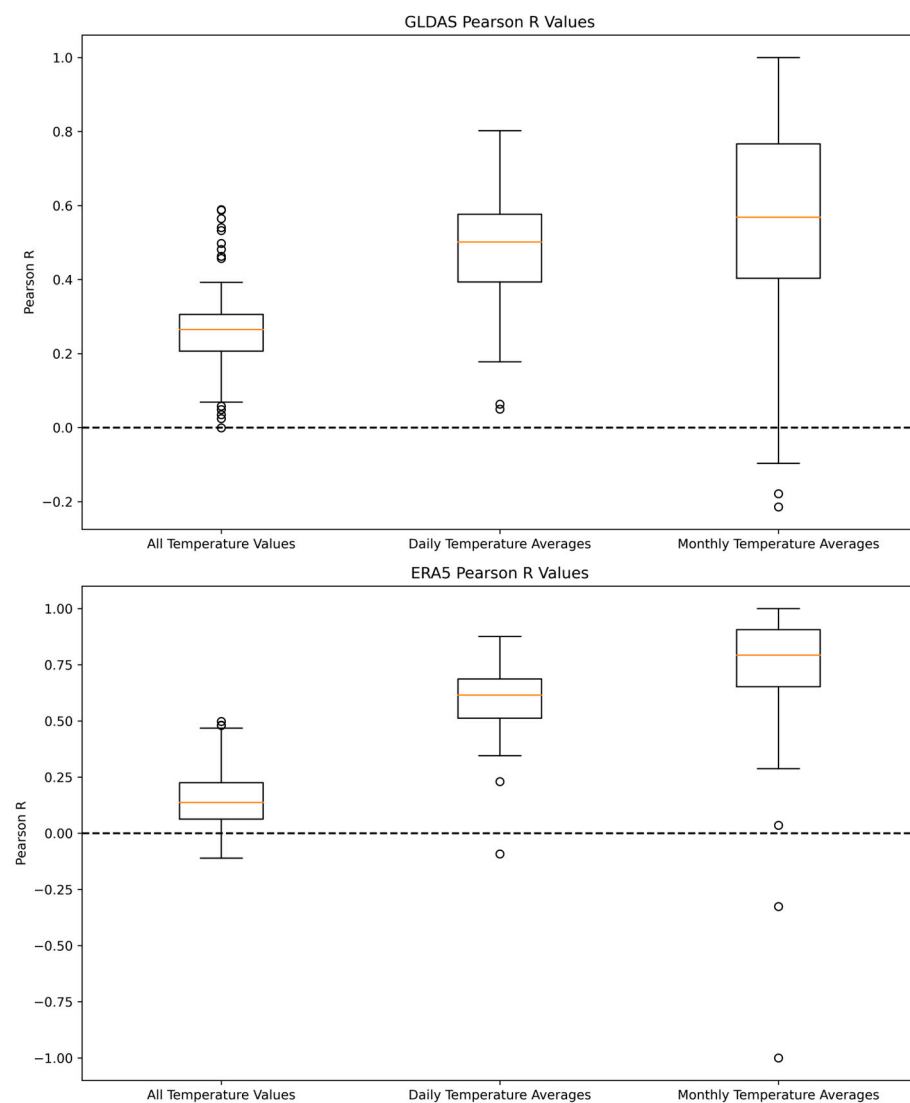


Figure 5. Boxplots of Pearson R values showing different temporal groupings for both ERA5 and GLDAS. The circles show the remaining outliers in the dataset.

3.1.3. Error Metrics

We calculated the ME, RMSE, and MAE error metrics between gridded data and in situ data for temperature measurements. Table 5 shows the results from the error metrics for both daily and monthly temporal aggregations for temperature data. These values are in degrees Celsius.

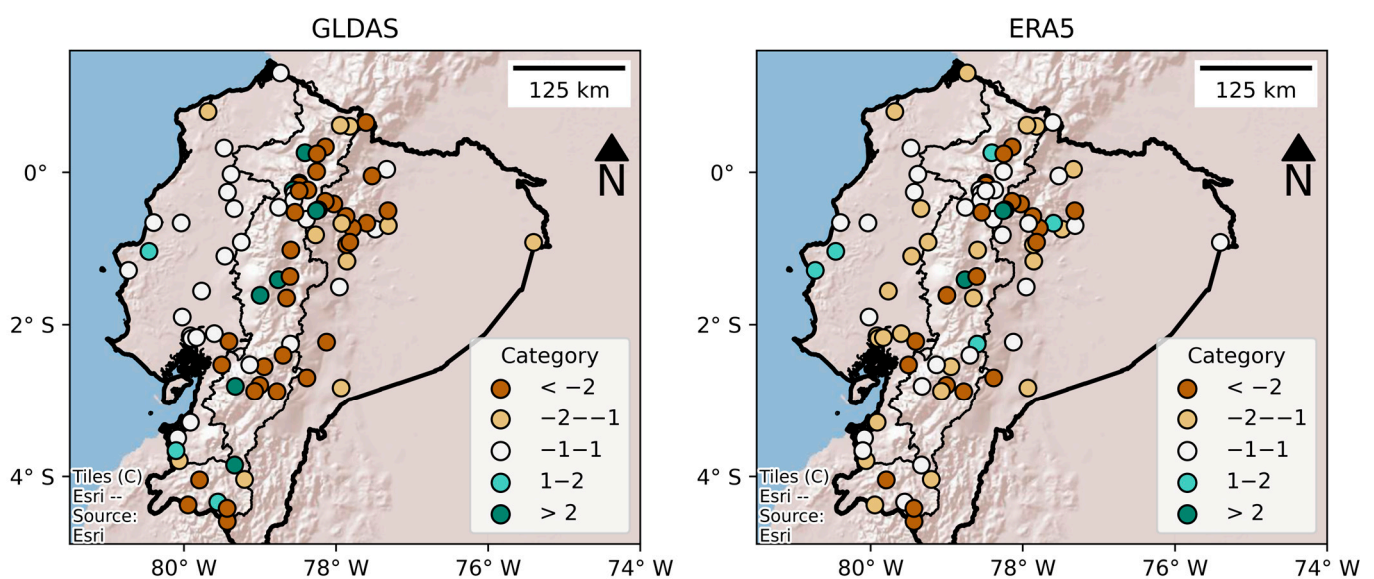
Table 5. Error metrics for temperature measurements in Celsius based on daily and monthly averages.

	Daily		Monthly	
	ME	RMSE	ME	RMSE
ERA5	−1.144	2.020	−1.147	1.797
GLDAS	−1.419	2.784	−1.423	2.567

Additionally, Table 6 was created to show the error metrics spread out through Ecuador. The metrics are shown based on the results from each region in Ecuador. To help visualize this data, a map was created to show the spread of the ME throughout Ecuador. Figure 6 presents a map depicting ME values for temperature at each station. Each dot represents a station, and they are colored to match the ME value at the corresponding station. In the map showing the GLDAS error, the stations by the coast all have a mean error closer to zero when compared with other regions.

Table 6. Performance metrics for temperature measurements in Celsius grouped by regions of Ecuador.

Dataset	Statistics	Amazon	Mountain	Coast	AVERAGE
GLDAS	MAE	2.80	2.96	1.24	2.44
	ME	−2.63	−1.39	−0.38	−1.42
	R	0.54	0.56	0.56	0.56
	RMSE	2.95	3.05	1.39	2.57
	<i>p</i> -values	0.05	0.03	0.12	0.06
ERA5	MAE	1.67	1.93	1.30	1.69
	ME	−1.26	−1.22	−0.92	−1.15
	R	0.77	0.70	0.71	0.72
	RMSE	1.78	2.01	1.45	1.80
	<i>p</i> -values	0.0003	0.02	0.05	0.03
AVERAGE	MAE	2.23	2.45	1.27	2.07
	ME	−1.95	−1.30	−0.65	−1.29
	R	0.65	0.63	0.64	0.64
	RMSE	2.37	2.53	1.42	2.18
	<i>p</i> -values	0.03	0.03	0.08	0.04

**Figure 6.** ME values compared to the in situ temperature data in Celsius.

3.1.4. Scatterplots

To visualize the difference between the different datasets, we created scatterplots of gridded temperature values and corresponding in situ values. We did this with all available data as well as with daily and monthly groupings. Figure 7 shows the scatterplots created for temperature. In situ measurements are given on the x-axis and the global dataset's value is on the y-axis. The red line is a least squares trend line, or line of best fit, for the data. The equation is shown in the upper left corner. The black line shows the values for the line $y = x$, which represents what the values would be for the data if the in situ data and global gridded data reported the same values. The graphs show the overall spread of the data. As the temporal grouping of the data increases, the overall spread of the data decreases.

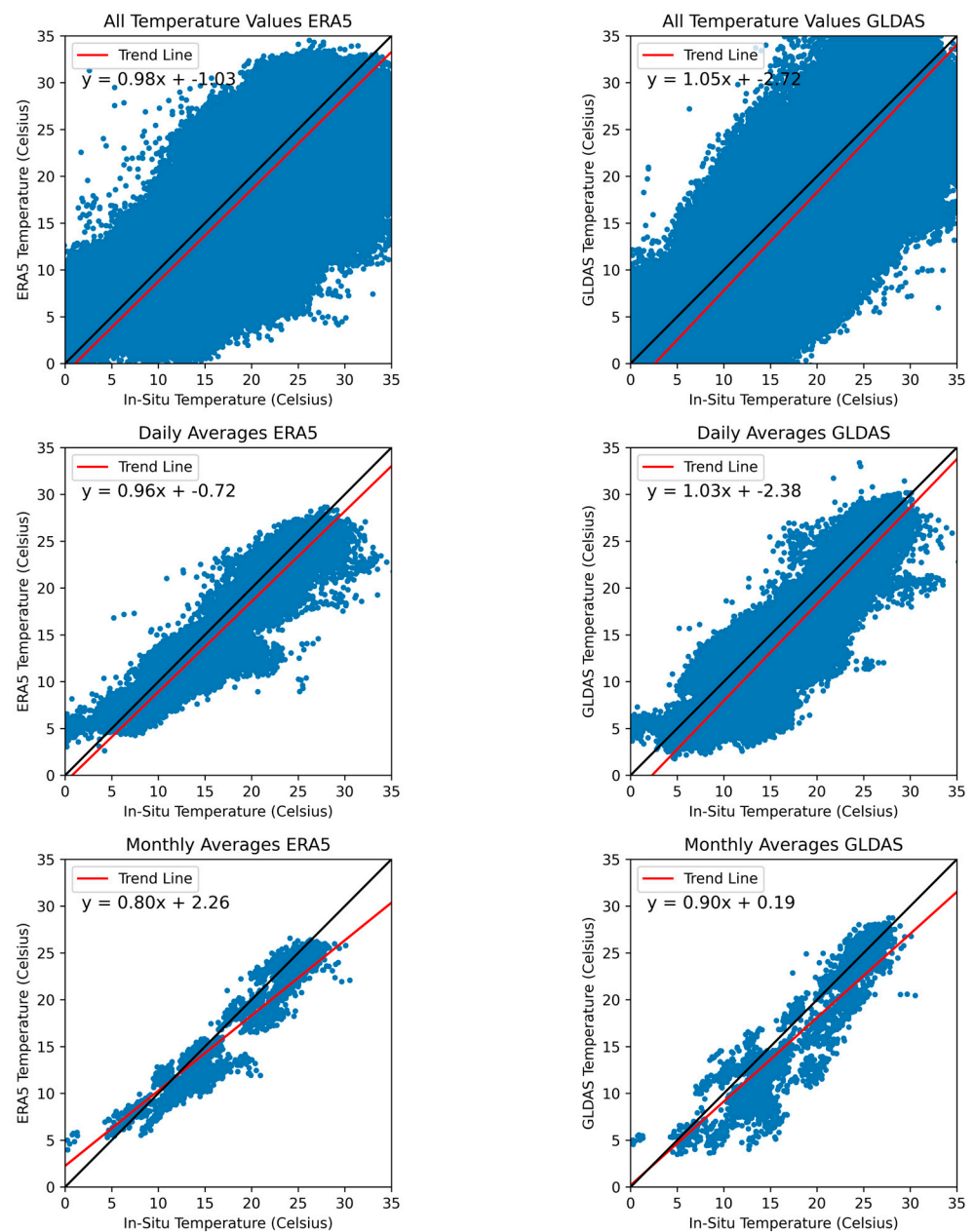


Figure 7. Scatterplots for temperature comparing the in situ values and their paired gridded dataset values using all the stations. The x-axis position indicates in situ data values, while the y-axis position indicates values from global data. The black line represents the line $y = x$.

3.2. Performance Metrics of Precipitation

3.2.1. Pearson R Correlation Coefficient

We calculated the Pearson R correlation coefficient values for precipitation. Table 7 displays Pearson R values for precipitation, calculated using daily and monthly precipitation sums. Figure 8 shows the boxplots with the values of Pearson R for each of the different datasets. There are four plots in the figure, each representing the results from a different precipitation dataset. The boxplots show that the range of values from the stations also increased as the temporal aggregation increased. We observed that the Pearson R values increased as the temporal aggregations increased from daily to monthly. Figure 9 shows a map of the Pearson R values throughout Ecuador. The values in the map are based on the monthly sums.

Table 7. Pearson R values for each precipitation dataset with different temporal groupings.

	ERA5	GLDAS	IMERG	CHIRPS
Daily Sums	0.313	0.269	0.386	0.239
Monthly Sums	0.50	0.422	0.524	0.473

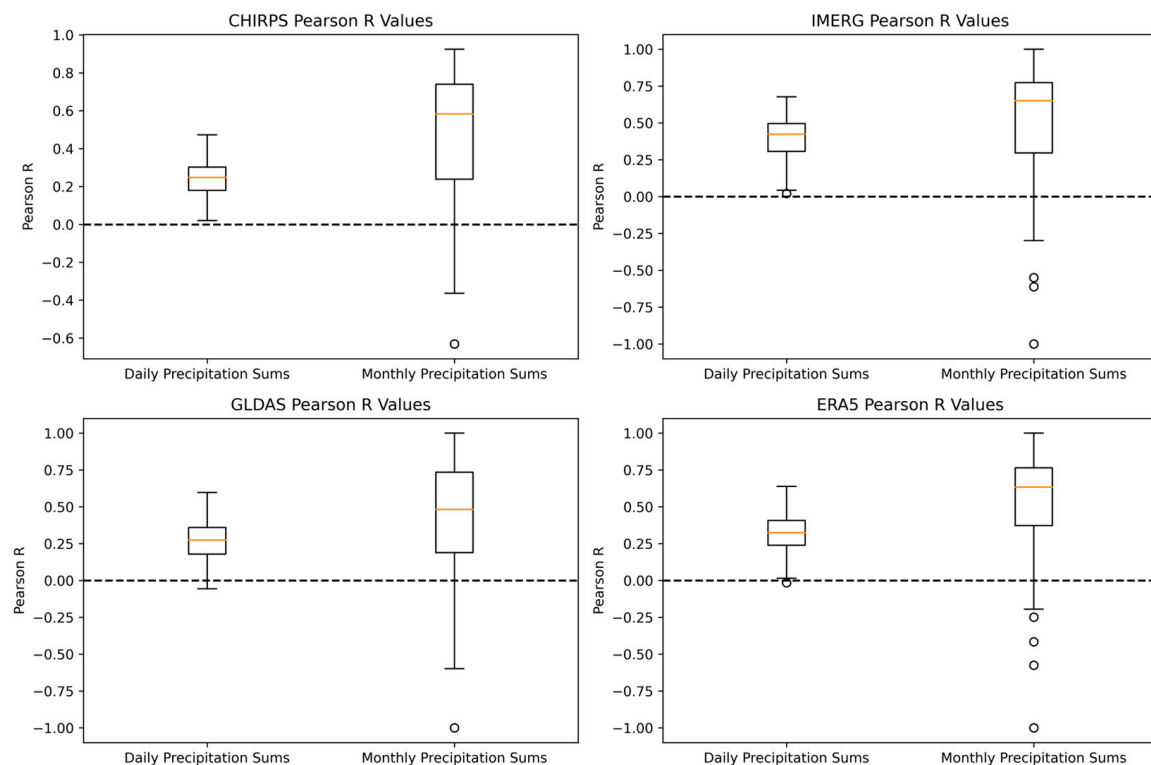


Figure 8. Boxplots of the Pearson R values based on daily and monthly sums of precipitation at each station for the four different datasets. The dots represent the remaining outliers in the dataset.

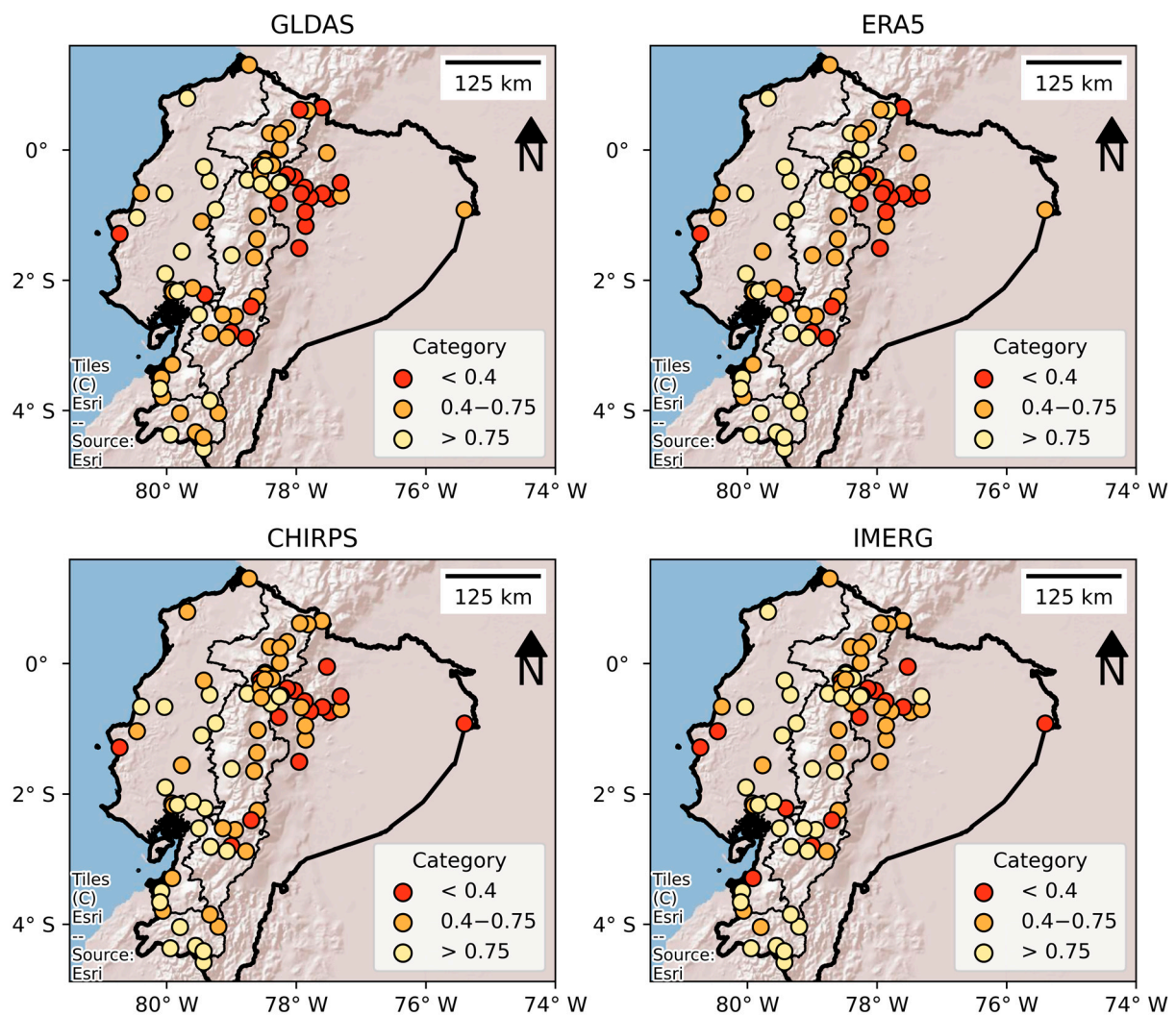


Figure 9. Pearson R values for monthly precipitation at each of the stations in Ecuador.

3.2.2. Error Metrics

We also calculated the error metrics based on the location of each station. This shows how the error metrics are spread throughout the country of Ecuador. Table 8 shows the error metrics in mm for precipitation. The values are based on monthly cumulative values.

Table 8. Error metrics for the four precipitation datasets in millimeters based on monthly sums grouped by region of Ecuador.

Dataset	Statistic	Amazon	Mountain	Coast	AVERAGE
CHIRPS	MAE	161.28	54.54	254.08	133.18
	ME	68.35	10.67	−122.07	−14.19
	R	0.10	0.54	0.61	0.47
	RMSE	201.10	76.68	653.58	264.76
IMERG	MAE	137.32	81.34	247.87	139.91
	ME	−79.94	35.35	−162.60	−44.67
	R	0.23	0.60	0.59	0.52
	RMSE	167.95	104.65	615.96	261.32
ERA5	MAE	176.16	167.45	344.85	218.98
	ME	71.65	157.59	−27.89	87.32
	R	0.14	0.60	0.61	0.50
	RMSE	206.57	189.69	716.05	340.67

Table 8. Cont.

Dataset	Statistic	Amazon	Mountain	Coast	AVERAGE
GLDAS	MAE	155.24	132.28	288.32	180.87
	ME	−29.98	102.25	−113.24	13.70
	R	0.10	0.45	0.60	0.42
	RMSE	189.32	174.77	660.47	313.87
AVERAGE	MAE	157.51	108.37	283.78	168.24
	ME	7.52	76.46	−106.45	10.54
	R	0.14	0.55	0.60	0.48
	RMSE	191.24	136.45	661.51	295.16

To visualize the results from the table, maps were created showing the values at each station spread throughout Ecuador. Figure 10 is a map of the RMSE values for each station throughout Ecuador. The values are shown in millimeters of precipitation. The light black line shows the three different regions in Ecuador. The furthest west in the coast, the center region is the mountain region, and the east region is the Amazon. We used the monthly sums to compare the maps. The map shows how the values of the error metrics are spread through Ecuador.

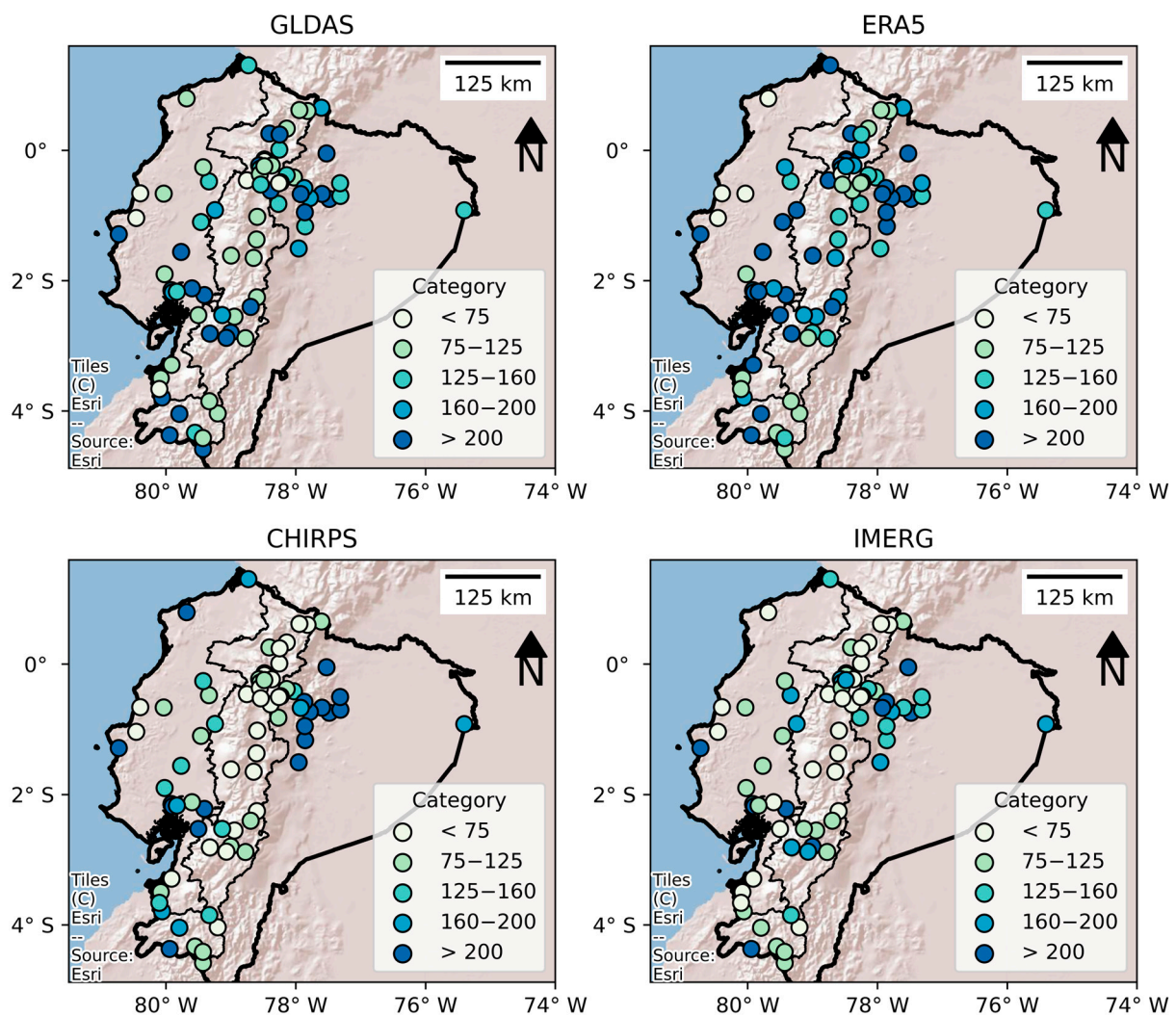


Figure 10. RMSE of monthly precipitation sums at each of the stations in Ecuador.

3.2.3. Scatterplots

Additionally, we created scatterplots showing the comparison of the in situ and global gridded data. Figure 11 shows the scatterplots produced for the daily sums in millimeters of precipitation. Figure 12 shows the scatterplots based on monthly sums of precipitation in millimeters.

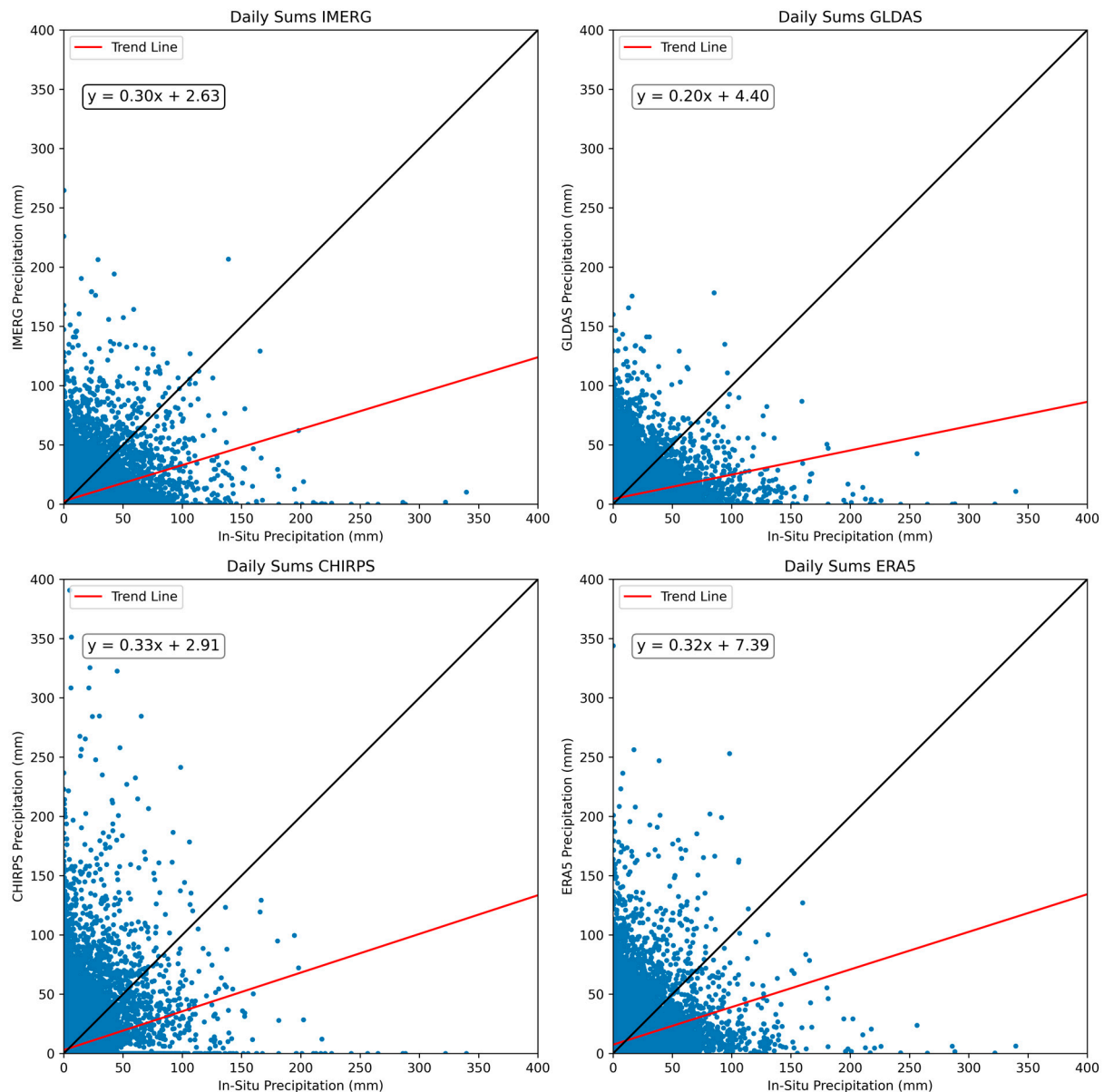


Figure 11. Precipitation scatter plots showing the in situ data compared global gridded data using daily sums of precipitation in mm. The x-axis position of each dot indicates in situ data values, while the y-axis position indicates values from global data. The black line represents the line $y = x$.

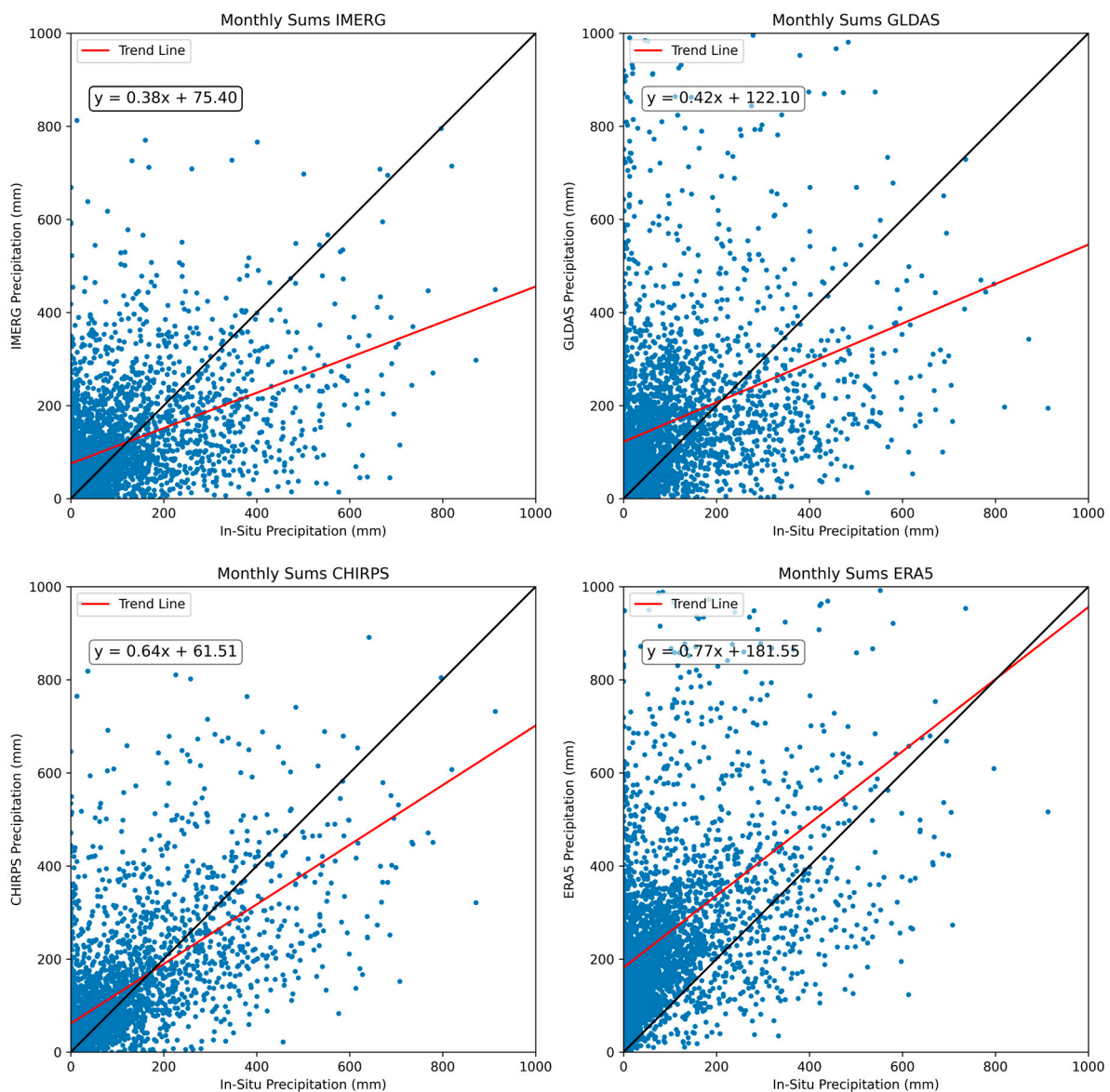


Figure 12. Scatter plots showing the in situ data compared to global gridded data using monthly sums in mm of precipitation. The x-axis position of each dot indicates in situ data values, while the y-axis position indicates values from global data. The black line represents the line $y = x$.

4. Discussion

4.1. Limitations

We identified gauges and individual measurements where the measurements were suspect. These errors may be due to errors in recording or measuring the data. We attempted to filter the data to remove all unrealistic values, including negative precipitation values and impossibly high temperatures. It is possible that other values were also incorrectly recorded, or the measurement devices were inaccurate. Furthermore, numerous stations exhibited temporal gaps and missing values at certain intervals. These gaps could impact the outcomes of daily or monthly aggregation analyses. We attempted to identify these stations and exclude values or stations to avoid unfair comparisons and inaccurate representations of the gauge or global data.

The challenges with in situ data is not a problem unique to this study [7,9]. Researchers have also highlighted that remote areas of the world and developing countries have sparse

rain gauges [14,26]. We found this was true in the Ecuadorian gauge network where there are significantly fewer gauges in the Amazon region which is less populated.

Furthermore, the gauges were sparsely located, with only about 30 of the gauges having data that expanded throughout the entire duration of the study (2014 to 2022). However, this sort of problem exists in many other countries and situations as well [6,16,17]. Areas of the world are not equally represented in the global gridded datasets [12]. This highlights the importance of more data collection in data sparse areas. This is similar to the conclusions made in a study performed in the tropical Andes which highlights the need for a good gauge network to evaluate satellite products [18].

The global gridded datasets used in this study also have limitations due to the size and changing terrain in Ecuador. Ecuador includes areas in the Amazon rainforest, the coast, and the Andes mountains. These different geographical features affect the temperature and precipitation of a given area. Many studies have noted how temperature and precipitation vary greatly with space and time [11,43]. The resolution of the global gridded data is coarse compared with the rapidly varying terrain in Ecuador. The gridded datasets likely had difficulty capturing the complexity of the terrain. Therefore, remote sensing-based and modeled data may perform better when showing a general overview of an area rather than values at specific points.

Additionally, each of these gauges has different spatial resolutions. For example, CHIRPS has a resolution of 0.05 degrees whereas GLDAS only has a resolution of 0.25 degrees. The spatial resolution of the datasets was not altered as part of this study. Therefore, the different spatial resolution affected the performance of the datasets. A dataset with a higher spatial resolution has an advantage when matching data at discrete points.

Our analysis proves the datasets perform well at coarser temporal aggregations and have higher correlations over time. This means that specific values at a single time step may be less accurate, but the overall averages from the global gridded products are suitable for monitoring as well as status and outlook product applications. We can use the global gridded datasets when in situ data is unavailable to show the trends of temperature and precipitation over time. This is important to understand as global gridded datasets are used in producing status and outlook products and this shows how they can be used in decision making situations.

It has been suggested in other studies that satellite data can help provide answers in ungauged areas, especially when there are limited gauges available. However, the satellite data has a large amount of variability [7]. We agree with these existing conclusions that satellite data can be useful under certain conditions and show that these same ideas extend to Ecuador. Because of the coarseness of the gridded data and relative sparsity of gauge data, our conclusions in this study should be limited. We acknowledge that the resolution and availability of data and the size of Ecuador forced us to compare data that may not fairly assess the skill of modeled or remotely sensed data in general. These conclusions apply specifically to Ecuador and are not necessarily reflective of the global performance of these datasets.

Our study also underscores the importance of increased spatial and temporal resolution of globally available gridded datasets for temperature and precipitation. We focused on datasets available on GEE because they were most likely to be used given the other current and future research projects. New datasets are being produced which may perform better than the datasets analyzed here. The higher resolution datasets likely have greater potential to fill in gaps and otherwise complement the in situ data because they do not have spatial or temporal gaps and more sufficiently measure variation across the unique terrain constraints in Ecuador.

4.2. Temperature

The temperature data was available on hourly time steps, and when compared against the 3-hourly GLDAS data and hourly ERA5 data, the p -value of the paired t -test was under 0.05 at every station. The difference between the in situ stations and the global gridded

datasets was considered statistically significant at every station. However, the p -value of the paired t -test increased along with the plausibility of no statistically significant difference when the data was instead compared using daily and monthly averages. The change is at least partially due to the decreased sample size of months to compare. Nevertheless, the result agrees with trends seen in error metrics which improved at coarser time steps, such as is shown in Figure 7. Those results corroborate our conclusions that the global datasets perform better at longer time steps. As we group the data by daily and monthly averages, the spread of our data greatly decreased. This shows that there are fewer outliers and that the global gridded datasets are estimating values that are closer to the observed values.

The error metrics from the temperature data show that on average the temperature was underestimated with both GLDAS and ERA5 due to the negative ME values. They had an ME of -1.419 and -1.144 , respectively, when comparing daily averages. The MAE value showed a lower accuracy than the ME values. This shows that some of the ME values are being cancelled out by having both positive and negative values. However, when looking at the MAE, it is still only 2.07 on average when comparing monthly values (see Table 6). We expect this to be an acceptable amount of error for various applications such as producing status and outlook summaries using these datasets. We are still able to understand the temperature trends in the country and tell if the temperature is warmer or colder than normal. This aligns with what has been found in other studies which show that GLDAS generally has accurate temperature values [26].

It was found that the error metrics are lower on the coast than in the Amazon or mountains as shown in Table 5. Land surface temperature has been found to vary as much as 10 degrees Celsius over just a few centimeters [43]. It would therefore make sense that the air temperature over a complex topography, such as the mountains, would make it more difficult to obtain accurate numbers.

The GLDAS sensor had a higher average p -value and a higher average correlation coefficient for temperature when every value was considered. However, temporal aggregations improved the ERA5 correlation coefficient much more than it improved the GLDAS correlation coefficient. It should be noted that GLDAS uses 3-hourly values, so this difference could be in part due to the lower temporal resolution of the GLDAS data when comparing all given values. ERA5 had a higher correlation coefficient when looking at daily and monthly averages. ERA5 also had a lower RMSE and ME. This suggests that on average, the ERA5 temperature data is more accurate for daily and monthly averages in Ecuador. This could in part be attributed to the fact that GLDAS has a much lower spatial resolution than ERA5.

4.3. Precipitation

When observing Table 8, we see that the ME values for precipitation are negative in the coastal region and positive in the mountainous areas. This means that in general, precipitation was being underestimated in the coastal areas and overestimated in the mountains. This is similar to what was found in Colombia when comparing daily IMERG data [44]. Table 8 also shows that the ME, MAE and RMSE are worse for the coastal regions than the other regions. This is similar to what was found in India, where it was found that biases were higher on the periphery of the country [31]. The absolute value of the ME is always lower than the RMSE and MAE. This shows that the errors are both positive and negative depending on the month, causing them to cancel each other out.

Some of the stations have Pearson R values close to one, while others have values that are negative. This could be caused in part by the missing time steps in the in situ data. Even though months with less than half the days were filtered out when comparing monthly data, months could still be missing many values causing the in situ data to be much lower than it would be otherwise. Overall, it appears that the global datasets had better error metrics in the mountainous areas, but a better Pearson R value in the coastal areas. The Pearson R value is significantly lower in the Amazon areas.

IMERG has the highest average correlation coefficient. IMERG has often been found to outperform other products, including in Australia [22], and China, where it outperformed ERA5 data specifically [23].

CHIRPS had the lowest MAE of all precipitation datasets in every region. CHIRPS has a higher spatial resolution which could in part account for this. In this study, however, CHIRPS performed significantly worse than it has in other studies of the Amazon where the MAE and RMSE were found to be much lower [45]. Several factors could account for this, including the way the data was aggregated, the missing in situ values and the fact that only 16 stations in the Amazon region were available for use in this study.

CHIRPS and IMERG were best at several metrics showing the added benefit of assimilating gauge data regardless of their difference in spatial and temporal resolutions. Both IMERG and CHIRPS use station data when providing precipitation values [36,37]. We are unaware of how many of the Ecuador stations were used in the IMERG and CHIRPS data production. Therefore, we recommend Ecuador continue pursuing data sharing practices that allow precipitation products to improve in Ecuador.

4.4. Comparison of Temperature and Precipitation Data

In this study, the temperature datasets performed better than the precipitation datasets. The average Pearson R value for monthly values is 0.48 for precipitation and 0.64 for temperature. This suggests the temperature data is more correlated than the precipitation. This same thing has been found in other studies as well [28]. This suggests that global gridded data is doing a better job at estimating the values for temperature than for precipitation.

5. Conclusions

In this paper, we aimed to compare global gridded datasets to in situ data in Ecuador. We present four major conclusions: (1) The limited networks and gaps in the data highlight the need for global gridded datasets, such as those being examined in this paper, to provide a viable alternative for decision makers as well as the need to increase the in situ coverage nationwide; (2) Monthly aggregations tend to better represent the results than daily aggregations; (3) Geographical features impact the performance of the global gridded datasets; (4) The global gridded datasets may be sufficient to fill gaps in the in situ data.

Our study was limited by the availability of data and the gaps present in the in situ data. We also had to consider the different temporal resolutions of the global gridded data. Therefore, we were unable to perform the false alarm ratio and probability of detection due to focusing more on monthly aggregations than on hourly timesteps. With more data available, assessing these things could help add to our understanding of how the global gridded datasets are performing. The gaps in our data also highlighted the need for continual improvement of global gridded data that can help fill these gaps in for decision makers.

We found the RMSE, ME, and Pearson R correlation coefficient at specific points and reported the metrics based on their geographic region within Ecuador. For temperature, the coastal area had the lowest error with a ME value of only -0.65 and an RMSE value of 1.42 . We found that geographical features do impact the performance of the global gridded datasets, especially the mountains, which impacts both temperature and precipitation. When comparing precipitation values, the correlation coefficient was highest on the coast, with a value of 0.6 . The Amazon had the lowest correlation coefficient of 0.14 . We show that global gridded datasets perform better when showing the trends of data at daily or monthly temporal aggregations. Overall, our results suggest that the global gridded datasets do a poor job approximating exact values at specific points but can be representative of monthly averages. Therefore, the global gridded datasets can begin to fill gaps in data for decision makers in this way.

We recommend that similar studies be performed in other data sparse regions so that global temperature and precipitation dataset producers can better understand the performance of their models in places they previously could not evaluate. We recommend

that similar evaluations be performed in Ecuador as the automatic gauge network matures to cover more areas and have longer periods of record. More in situ measurements will allow for analysis, which is currently not possible. We recommend these global datasets for applications in evaluating historical and real-time trends in precipitation and temperature to create status and outlook summaries. We showed that IMERG, CHIRPS, GLDAS and ERA5 can all be used to show how current conditions compare to past conditions using monthly averages since they had the highest performance measures at coarser time steps.

Author Contributions: Conceptualization, R.C.H.; methodology, R.C.H. and R.H.M.; software, R.C.H., R.H.M. and T.J.M.; validation, R.C.H. and R.H.M.; formal analysis, R.H.M.; investigation, R.C.H. and R.H.M.; resources, R.C.H. and E.J.N.; data curation, R.H.M.; writing—original draft preparation, R.H.M.; writing—review and editing, R.C.H., R.H.M., B.E., K.L. and T.J.M.; visualization, R.H.M.; supervision, R.C.H. and E.J.N.; project administration, R.C.H. and E.J.N.; funding acquisition, E.J.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by SERVIR-Amazonia Program AST-3, grant number 80NSSC20 K0157 and NASA Applied Sciences Water Resources, grant number 80NSSC22K0927.

Data Availability Statement: We used four datasets available on Google Earth Engine for our global gridded data as can be seen in Table 2. The in situ data used in this study was received as part of a project through SERVIR-Amazonia. It cannot be published or redistributed because it is the property of INAMHI. This paper was published in collaboration with INAMHI.

Acknowledgments: We acknowledge the support of Fundación EcoCiencia and INAMHI for providing the data necessary to perform this study. We also want to acknowledge the support of Brigham Young University Civil and Construction engineering department for providing laboratory space and computer resources for the research.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Larson, L.W.; Peck, E.L. Accuracy of Precipitation Measurements for Hydrologic Modeling. *Water Resour. Res.* **1974**, *10*, 857–863. [\[CrossRef\]](#)
2. Gosset, M.; Kunstmann, H.; Zougmore, F.; Cazenave, F.; Leijnse, H.; Uijlenhoet, R.; Chwala, C.; Keis, F.; Doumounia, A.; Boubacar, B.; et al. Improving Rainfall Measurement in Gauge Poor Regions Thanks to Mobile Telecommunication Networks. *Bull. Am. Meteorol. Soc.* **2016**, *97*, ES49–ES51. [\[CrossRef\]](#)
3. Siepielski, A.M.; Morrissey, M.B.; Buoro, M.; Carlson, S.M.; Caruso, C.M.; Clegg, S.M.; Coulson, T.; DiBattista, J.; Gotanda, K.M.; Francis, C.D.; et al. Precipitation Drives Global Variation in Natural Selection. *Science* **2017**, *355*, 959–962. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Price, K.; Purucker, S.T.; Kraemer, S.R.; Babendreier, J.E.; Knightes, C.D. Comparison of Radar and Gauge Precipitation Data in Watershed Models across Varying Spatial and Temporal Scales: RADAR AND GAUGE PRECIPITATION. *Hydrol. Process* **2014**, *28*, 3505–3520. [\[CrossRef\]](#)
5. Grimes, D.I.F.; Pardo-Igúzquiza, E.; Bonifacio, R. Optimal Areal Rainfall Estimation Using Raingauges and Satellite Data. *J. Hydrol.* **1999**, *222*, 93–108. [\[CrossRef\]](#)
6. Kidd, C.; Becker, A.; Huffman, G.J.; Muller, C.L.; Joe, P.; Skofronick-Jackson, G.; Kirschbaum, D.B. So, How Much of the Earth's Surface Is Covered by Rain Gauges? *Bull. Am. Meteorol. Soc.* **2017**, *98*, 69–78. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Sun, Q.; Miao, C.; Duan, Q.; Ashouri, H.; Sorooshian, S.; Hsu, K. A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Rev. Geophys.* **2018**, *56*, 79–107. [\[CrossRef\]](#)
8. Yoon, S.-S.; Phuong, A.T.; Bae, D.-H. Quantitative Comparison of the Spatial Distribution of Radar and Gauge Rainfall Data. *J. Hydrometeorol.* **2012**, *13*, 1939–1953. [\[CrossRef\]](#)
9. Devine, K.A.; Mekis, É. Field Accuracy of Canadian Rain Measurements. *Atmos.-Ocean* **2008**, *46*, 213–227. [\[CrossRef\]](#)
10. Sieck, L.C.; Burges, S.J.; Steiner, M. Challenges in Obtaining Reliable Measurements of Point Rainfall: RELIABLE MEASUREMENTS OF POINT RAINFALL. *Water Resour. Res.* **2007**, *43*. [\[CrossRef\]](#)
11. Derin, Y.; Yilmaz, K.K. Evaluation of Multiple Satellite-Based Precipitation Products over Complex Topography. *J. Hydrometeorol.* **2014**, *15*, 1498–1516. [\[CrossRef\]](#)
12. Condom, T.; Martínez, R.; Pabón, J.D.; Costa, F.; Pineda, L.; Nieto, J.J.; López, F.; Villacis, M. Climatological and Hydrological Observations for the South American Andes: In Situ Stations, Satellite, and Reanalysis Data Sets. *Front. Earth Sci.* **2020**, *8*, 92. [\[CrossRef\]](#)

13. Lee, J.; Kim, S.; Jun, H. A Study of the Influence of the Spatial Distribution of Rain Gauge Networks on Areal Average Rainfall Calculation. *Water* **2018**, *10*, 1635. [\[CrossRef\]](#)
14. Wilby, R.L. A Global Hydrology Research Agenda Fit for the 2030s. *Hydrol. Res.* **2019**, *50*, 1464–1480. [\[CrossRef\]](#)
15. Paredes-Trejo, F.J.; Barbosa, H.A.; Lakshmi Kumar, T.V. Validating CHIRPS-Based Satellite Precipitation Estimates in Northeast Brazil. *J. Arid Environ.* **2017**, *139*, 26–40. [\[CrossRef\]](#)
16. Ocampo-Marulanda, C.; Fernández-Álvarez, C.; Cerón, W.L.; Canchala, T.; Carvajal-Escobar, Y.; Alfonso-Morales, W. A Spatiotemporal Assessment of the High-Resolution CHIRPS Rainfall Dataset in Southwestern Colombia Using Combined Principal Component Analysis. *Ain Shams Eng. J.* **2022**, *13*, 101739. [\[CrossRef\]](#)
17. Zubietta, R.; Getirana, A.; Espinoza, J.C.; Lavado-Casimiro, W.; Aragon, L. Hydrological Modeling of the Peruvian–Ecuadorian Amazon Basin Using GPM-IMERG Satellite-Based Precipitation Dataset. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3543–3555. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Manz, B.; Páez-Bimos, S.; Horna, N.; Buytaert, W.; Ochoa-Tocachi, B.; Lavado-Casimiro, W.; Willems, B. Comparative Ground Validation of IMERG and TMPA at Variable Spatiotemporal Scales in the Tropical Andes. *J. Hydrometeorol.* **2017**, *18*, 2469–2489. [\[CrossRef\]](#)
19. Erazo, B.; Bourrel, L.; Frappart, F.; Chimborazo, O.; Labat, D.; Dominguez-Granda, L.; Matamoros, D.; Mejia, R. Validation of Satellite Estimates (Tropical Rainfall Measuring Mission, TRMM) for Rainfall Variability over the Pacific Slope and Coast of Ecuador. *Water* **2018**, *10*, 213. [\[CrossRef\]](#)
20. Delgado, D.; Sadaoui, M.; Ludwig, W.; Méndez, W. Spatio-Temporal Assessment of Rainfall Erosivity in Ecuador Based on RUSLE Using Satellite-Based High Frequency GPM-IMERG Precipitation Data. *Catena* **2022**, *219*, 106597. [\[CrossRef\]](#)
21. Navarro, A.; García-Ortega, E.; Merino, A.; Sánchez, J.; Kummerow, C.; Tapiador, F. Assessment of IMERG Precipitation Estimates over Europe. *Remote Sens.* **2019**, *11*, 2470. [\[CrossRef\]](#)
22. Islam, M.A.; Yu, B.; Cartwright, N. Assessment and Comparison of Five Satellite Precipitation Products in Australia. *J. Hydrol.* **2020**, *590*, 125474. [\[CrossRef\]](#)
23. Xu, J.; Ma, Z.; Yan, S.; Peng, J. Do ERA5 and ERA5-Land Precipitation Estimates Outperform Satellite-Based Precipitation Products? A Comprehensive Comparison between State-of-the-Art Model-Based and Satellite-Based Precipitation Products over Mainland China. *J. Hydrol.* **2022**, *605*, 127353. [\[CrossRef\]](#)
24. Qi, W.; Zhang, C.; Fu, G.; Sweetapple, C.; Zhou, H. Evaluation of Global Fine-Resolution Precipitation Products and Their Uncertainty Quantification in Ensemble Discharge Simulations. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 903–920. [\[CrossRef\]](#)
25. Liu, L.; Gu, H.; Xie, J.; Xu, Y. How Well Do the ERA-INTERIM, ERA -5, GLDAS -2.1 and NCEP-R2 Reanalysis Datasets Represent Daily Air Temperature over the Tibetan Plateau? *Int. J. Climatol.* **2021**, *41*, 1484–1505. [\[CrossRef\]](#)
26. Ji, L.; Senay, G.B.; Verdin, J.P. Evaluation of the Global Land Data Assimilation System (GLDAS) Air Temperature Data Products. *J. Hydrometeorol.* **2015**, *16*, 2463–2480. [\[CrossRef\]](#)
27. Hasan Karaman, Ç.; Akyürek, Z. Evaluation of Near-Surface Air Temperature Reanalysis Datasets and Downscaling with Machine Learning Based Random Forest Method for Complex Terrain of Turkey. *Adv. Space Res.* **2023**, *71*, 5256–5281. [\[CrossRef\]](#)
28. Tarek, M.; Brissette, F.; Arsenault, R. Uncertainty of Gridded Precipitation and Temperature Reference Datasets in Climate Change Impact Studies. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 3331–3350. [\[CrossRef\]](#)
29. Kumar, L.; Mutanga, O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. *Remote Sens.* **2018**, *10*, 1509. [\[CrossRef\]](#)
30. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [\[CrossRef\]](#)
31. Dubey, S.; Gupta, H.; Goyal, M.K.; Joshi, N. Evaluation of Precipitation Datasets Available on Google Earth Engine over India. *Int. J. Clim.* **2021**, *41*, 4844–4863. [\[CrossRef\]](#)
32. Elnashar, A.; Zeng, H.; Wu, B.; Zhang, N.; Tian, F.; Zhang, M.; Zhu, W.; Yan, N.; Chen, Z.; Sun, Z.; et al. Downscaling TRMM Monthly Precipitation Using Google Earth Engine and Google Cloud Computing. *Remote Sens.* **2020**, *12*, 3860. [\[CrossRef\]](#)
33. Sazib, N.; Bolten, J.; Mladenova, I. Exploring Spatiotemporal Relations between Soil Moisture, Precipitation, and Streamflow for a Large Set of Watersheds Using Google Earth Engine. *Water* **2020**, *12*, 1371. [\[CrossRef\]](#)
34. Morán-Tejeda, E.; Bazo, J.; López-Moreno, J.I.; Aguilar, E.; Azorín-Molina, C.; Sanchez-Lorenzo, A.; Martínez, R.; Nieto, J.J.; Mejía, R.; Martín-Hernández, N.; et al. Climate Trends and Variability in Ecuador (1966–2011): CLIMATE TRENDS AND VARIABILITY IN ECUADOR. *Int. J. Climatol.* **2016**, *36*, 3839–3855. [\[CrossRef\]](#)
35. Krause, P. Kathleen Flood Weather and Climate Extremes. In *Weather and Climate Extremes*; US Army Corps of Engineers Topographic Engineering Center: Alexandria, VA, USA, 1997; p. 89.
36. Hou, A.Y.; Kakar, R.K.; Neeck, S.; Azarbarzin, A.A.; Kummerow, C.D.; Kojima, M.; Oki, R.; Nakamura, K.; Iguchi, T. The Global Precipitation Measurement Mission. *Bull. Am. Meteor. Soc.* **2014**, *95*, 701–722. [\[CrossRef\]](#)
37. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Husak, G.; Rowland, J.; Harrison, L.; Hoell, A.; et al. The Climate Hazards Infrared Precipitation with Stations—A New Environmental Record for Monitoring Extremes. *Sci. Data* **2015**, *2*, 150066. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 Global Reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [\[CrossRef\]](#)

39. Rodell, M.; Houser, P.R.; Jambor, U.; Gottschalck, J.; Mitchell, K.; Meng, C.-J.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M.; et al. The Global Land Data Assimilation System. *Bull. Am. Meteor. Soc.* **2004**, *85*, 381–394. [CrossRef]
40. Copernicus Climate Change Service ERA5-Land Monthly Averaged Data from 2001 to Present 2019. Available online: https://developers.google.com/earth-engine/datasets/catalog/ECMWF_ERA5_LAND_HOURLY#description (accessed on 15 June 2023).
41. Precipitation Processing System (PPS) At NASA GSFC GPM IMERG Final Precipitation L3 Half Hourly 0.1 Degree $\times$ 0.1 Degree V06 2019. Available online: https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGHH_06/summary (accessed on 15 June 2023).
42. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef]
43. Prata, A.J.; Caselles, V.; Coll, C.; Sobrino, J.A.; Otlé, C. Thermal Remote Sensing of Land Surface Temperature from Satellites: Current Status and Future Prospects. *Remote Sens. Rev.* **1995**, *12*, 175–224. [CrossRef]
44. Palomino-Ángel, S.; Anaya-Acevedo, J.A.; Botero, B.A. Evaluation of 3B42V7 and IMERG Daily-Precipitation Products for a Very High-Precipitation Region in Northwestern South America. *Atmos. Res.* **2019**, *217*, 37–48. [CrossRef]
45. de Moraes Cordeiro, A.L.; Blanco, C.J.C. Assessment of Satellite Products for Filling Rainfall Data Gaps in the Amazon Region. *Nat. Resour. Model.* **2021**, *34*, e12298. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.