remote sensing

MDPI

Article

The Accuracy of Precipitation Forecasts at Timescales of 1–15 Days in the Volta River Basin

Mekonnen Gebremichael [1,*], Haowen Yue [1] and Vahid Nourani [2,3]

[1] Department of Civil and Environmental Engineering, University of California, Los Angeles, CA 90095-1593, USA; yuehaowen@g.ucla.edu
[2] Center of Excellence in Hydroinformatics, Faculty of Civil Engineering, University of Tabriz, 29 Bahman Ave, Tabriz 51666, Iran; nourani@tabrizu.ac.ir
[3] Faculty of Civil and Environmental Engineering, Near East University, Near East Boulevard, via Mersin 10, Nicosia 99138, Turkey
* Correspondence: mekonnen@seas.ucla.edu

Abstract: Medium-range (1–15 day) precipitation forecasts are increasingly available from global weather models. This study presents evaluation of the Global Forecast System (GFS) for the Volta river basin in West Africa. The evaluation was performed using two satellite-gauge merged products: NASA's Integrated Multi-satellitE Retrievals (IMERG) "Final Run" satellite-gauge merged rainfall observations, and the University of California Santa Barbara's Climate Hazard's group Infrared Precipitation with Stations (CHIRPS). The performance of GFS depends on the climate zone, with underestimation bias in the dry Sahel climate, overestimation bias in the wet Guinea Coastal climate, and relatively no bias in the moderately wet Savannah climate. Averaging rainfall over the watershed of the Akosombo dam (i.e., averaging across all three climate zones), the GFS forecast indicates low skill (Kling-Gupta Efficiency KGE = 0.42 to 0.48) for the daily, 1-day, lead GFS forecast, which deteriorates further as the lead time increases. A sharp decrease in KGE occurred between 6 to 10 days. Aggregating the forecasts over long timescales improves the accuracy of the GFS forecasts. On a 15-day accumulation timescale, GFS shows higher skills (KGE = 0.74 to 0.88).

Keywords: weather forecasts; medium-range precipitation forecasts; Volta; global forecast system

Citation: Gebremichael, M.; Yue, H.; Nourani, V. The Accuracy of Precipitation Forecasts at Timescales of 1–15 Days in the Volta River Basin. *Remote Sens.* **2022**, *14*, 937. https://doi.org/10.3390/rs14040937

Academic Editor: Christopher Kidd

Received: 13 January 2022
Accepted: 11 February 2022
Published: 15 February 2022

1. Introduction

Medium-range precipitation forecasts are needed for a number of applications, such as early warning systems for floods and droughts, reservoir management and operations, and decisions in agriculture [1–6]. Medium-range precipitation forecasts are becoming increasingly available from global weather forecasting models, such as the Global Forecast System (GFS) [7], the NCEP climate forecast system (NSF CFS) [8], the European Centre for Medium-Range Weather Forecasts (ECMWF) [9], and the Global Spectral Model (GSM) [10]. However, assessments of the accuracy of these forecasts is relatively rare, particularly in Africa, where ground validation sites, as well as access to data from such sites, are limited.

In the Waikato River basin of New Zealand, GFS was able to identify synoptic-scale weather systems up to +7.5 days' lead time, although the position and exact timing of these systems can vary between successive model runs, especially at long lead times [11]. In India, the GFS forecast showed some skills at 1-day and 2-day lead times, but lower skills from 3 days and over [12]. In China, the U.S., and Australia, the GFS overestimated light rain and underestimated moderate and heavy rain, which was attributed to errors arising from not accounting for the aerosol–cloud interactions in the GFS model [13]. In a GFS forecast sensitivity experiment, it was found that the impact of the inclusion of field data on the root-mean-squared-error (RMSE) of the forecasts was large at short lead times (about 1 day), whereas the impact of the type of Kalman Filter data assimilation method considered lasted throughout the 7-day period [14]. Lien et al. [15] compared the global statistical properties of

GFS forecasts and Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) [16] observations, and reported that the GFS model had positive bias for precipitation amount compared to TMPA observations, and that the GFS forecasts had large random errors at higher resolutions, especially for convective precipitation. Based on a numerical experiment, it was reported that, while the overall precipitation forecast uncertainty increased gradually with lead time, the corresponding weather-related power outage prediction uncertainty in the northeast United States exhibited a lower dependence on lead times up to 3 days and a stepwise increase in 4- and 5-day lead times [17].

The evaluation of satellite-gauged precipitation forecasts in West Africa shows that these products may be used, with caveats, for evaluating weather forecasts. Dezfuli et al. [18] evaluated the performance of NASA's Integrated Multi-satellitE Retrievals (IMERG) "Final Run" (IMERG Final) (version 4) [19,20] in comparison with two high-resolution, experimental rain gauge station data provided by the Trans-African Hydro-Meteorological Observatory (TAHMO) [21] and reported the capability of IMERG Final to accurately represent the diurnal cycle of rainfall. Using the same dataset, Dezfuli et al. [22] showed that IMERG Final is able to capture the propagation of large Mesoscale Convective Systems (MCSs), a significant advantage over its predecessor's (TMPA) 3-h temporal resolution, which misses the time evolution of most of these systems. Satge et al. [23] evaluated the accuracy of a number of gridded precipitation datasets over West Africa through comparison against rain gauge station data, and reported that the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) and TMPA (the predecessor to IMERG) provided reliable estimates at both daily and monthly timescales, while the remaining satellite products considered (CMORPH, PERSIANN, GSMaP, ARC, and TAMSAT) and all the atmospheric reanalysis products considered (MERRA and JRA) were deemed unreliable. Furthermore, they found out that satellite products that incorporated rain gauge information outperformed satellite-only products. Maranan et al. [24] compared IMERG Final products against experimental rain gauge station data in the moist forest region of Ghana, West Africa, and showed that IMERG Final estimates are able to capture monthly rainfall with a correlation coefficient close to unity.

The primary objective of this study was to evaluate the accuracy of medium-range precipitation forecasts derived from the Global Forecast System (GFS) over the Volta River basin, focusing on its major reservoir dam (Akosombo Dam) through comparison against IMERG Final. The basin encapsulates the three climatic regimes of West Africa, namely, the Sahel, Savannah, and Guinea Coast regimes. We chose the GFS model due to its relatively high spatial ($0.25^\circ \times 0.25^\circ$) and temporal resolution (3 h to 6 h), as well as its free-of-charge data availability to users. The main questions addressed in this study are as follows: What is the accuracy of the GFS forecast over the watershed of the Akosombo Dam? How does the accuracy of the forecast depend on lead time and temporal aggregation scale?

In the event of GFS forecasts containing large errors, post-processing techniques (e.g., calibration and bias-correction (e.g., Stellingwerf et al. [25]) using near-real-time satellite products could be considered to reduce the errors in the forecasts. Although satellite precipitation products that incorporate rain gauge information (usually referred to as "research-version" products), such as IMERG Final and CHIRPS, may seem appropriate for validation, they are not suitable for the dynamic calibration of GFS forecasts due to the long data latency period before they are produced and made available to users [18]. Near-real-time satellite precipitation products, such as, IMERG Early, are more suitable for calibrating GFS forecasts given their short latency period (<4 h) [19]. However, such near-real-time products may have larger errors compared to research-version products as the former do not incorporate rain gauge information for bias correction [26]. For the near-real-time products to be useful for the calibration of GFS forecasts, their accuracy must be higher than the accuracy of GFS forecasts. Therefore, our secondary objective was to compare the performance of GFS (validated against IMERG Final) to the performance of IMERG Early (validated against IMERG Final). Finally, we acknowledge that the reference dataset used in our evaluation (i.e., IMERG Final) have their own estimation errors. Another research-

quality satellite rainfall product, CHIRPS, is also available for validation. Therefore, we conducted an additional assessment evaluating GFS using CHIRPS as reference.

2. Materials and Methods

2.1. Study Region

The Volta River basin (Figure 1) in West Africa covers an area of 409,000 km^2 and is shared by six riparian countries, of which Ghana and Burkina Faso make up the largest areas, of 42% and 43%, respectively; the remaining 15% is distributed between Benin, Côte d'Ivoire, Mali, and Togo [27,28]. The physical relief of the basin is generally low with a predominantly undulating topography, except for the Akwapim-Togo mountain ranges located in the eastern part. The climate of the basin is strongly influenced by the movement of the Inter-Tropical Convergence Zone (ITCZ), which produces low rainfall in the northern part of the basin and heavy rainfall in the southern part of the basin [29].

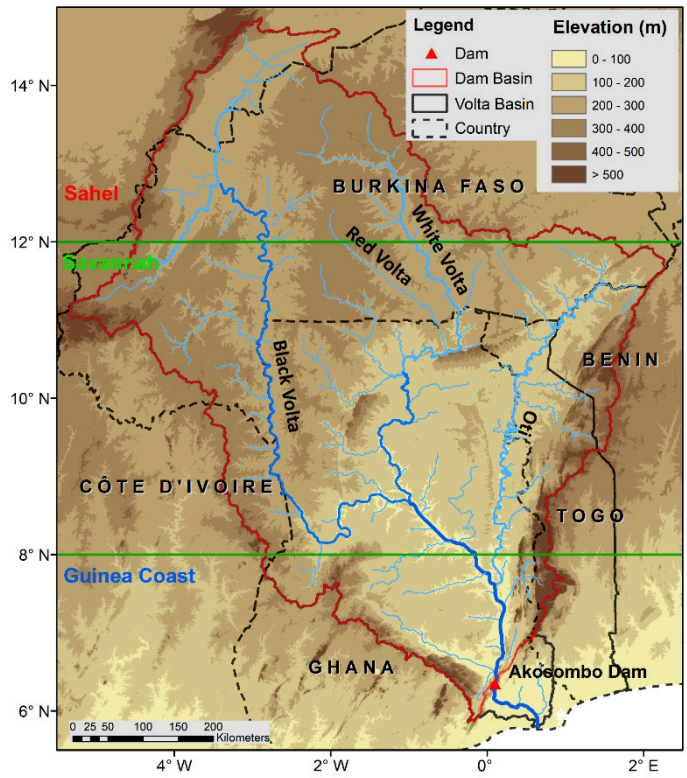

Figure 1. Digital elevation model for the Volta River Basin, with the location of the Akosombo dam.

The Volta River system consists of the Black Volta in the west, the White and Red Volta (each flowing from Burkina Faso), and the Oti River in the east (draining the highlands of Togo and Benin), all of which join in Ghana to form the lower Volta before flowing into the Gulf of Guinea. The Akosombo dam on the Lower Volta formed the Lake Volta, one of the largest man-made reservoirs in the world, covering 8500 km^2, or 4% of Ghana's land area [30], with a gross storage capacity of 148 Bm^3 and an active capacity of 78 Bm^3. Akosombo Dam, with a drainage area of 407,093 km^2, is a hydropower dam that generates 1012 MW, of which around 20% is supplied to the national grid and the remaining 80% is supplied to the aluminum industry.

2.2. Datasets

The Global Forecast System (GFS) is a global numerical weather prediction system run by the U.S. National Weather Service (NWS). The GFS forecasts at a resolution of 0.25° by 0.25° are obtained from National Center for Atmospheric Research (NCAR) Research Data Archive (RDA) GFS Historical Archive [7]. The GFS is run four times a day at UTC 00,

UTC 06, UTC 12, and UTC 18 h. One of the GFS model output variables is accumulated precipitation, where the precipitation forecasts are accumulations starting from the model run time. We obtained the 1-day lead daily rainfall forecast by subtracting the 24 h rainfall accumulation forecast from the 48 h rainfall accumulation forecast. Similarly, in order to obtain the 5-day lead daily rainfall forecast, we subtracted the 120 h rainfall accumulation forecast from the 144 h rainfall forecast. We only considered the model runs at UTC 00 h. We used the version-15 GFS forecasts, which have been available since 12 June 2009, following a major upgrade of the GFS model. Information on the GFS model's structure and inputs is available from various sources [31–39].

IMERG Final rainfall products were used in this study as references to evaluate the performance of GFS precipitation forecasts. IMERG Final is a satellite-gauge product, available at a resolution of 30 min and 0.10°, and with a data latency of 3.5 months. We also explored the accuracy of IMERG Early, which is a satellite-only product (i.e., no rain gauges are not included), available in near-real time at a resolution of 30 min and 0.10°. Comparison of the performance of IMERG Early with the performance of GFS indicated to what extent the IMERG Early products could be used for calibration of GFS forecasts. The latest version (V6B) of IMERG Final and IMERG Early datasets were used in this study. Details on the IMERG algorithm are available from Huffman et al. [19,20]. CHIRPS is another satellite-gauge product, available at a spatial resolution of 0.05° and a temporal resolution of 1 day, with a data latency period of about 3 weeks. Details on the CHIRPS algorithm are available from Funk et al. [40]. Agreement between the reference (IMERG Final) and CHIRPS indicated that the IMERG Final estimates were robust.

2.3. Evaluation Methodology

IMERG Final satellite rainfall products were used in this study as references to evaluate the performance of GFS precipitation forecasts. The comparison period was 15 June 2019 to 15 June 2020, to match the period for which version-15 of the GFS model forecasts is available. The spatial resolutions of the forecast and satellite products are different: 0.25° (GFS), 0.10° (IMERG Final and IMERG Early), and 0.05° (CHIRPS). The temporal resolutions of the satellite products are: 30 min (IMERG Final and IMERG Early) and daily (CHIRPS). Our comparison was mostly based on watershed-averaged values, in which case we averaged all the datasets to the sub-basin spatial scale. In cases where we compared the spatial patterns of rainfall, we resampled both IMERG products and CHIRPS to 0.25° using the bilinear interpolation technique to match the spatial resolution of the GFS.

For evaluation metrics, we used the modified Kling–Gupta Efficiency (*KGE*) method [41,42] and its components: bias ratio (*BR*), correlation (*R*), and variability ratio (γ). *KGE* measures the goodness-of-fit between estimates of precipitation forecasts and reference observations as:

$$KGE = 1 - \sqrt{(R-1)^2 + (BR-1)^2 + (\gamma-1)^2}, \quad (1)$$

$$BR = \frac{\mu_f}{\mu_o}, \quad (2)$$

$$\gamma = \frac{CV_f}{CV_o} \quad (3)$$

where R is the linear correlation coefficient between forecasted and observed precipitation, BR is the bias ratio, γ is the variability ratio, μ is the mean precipitation, CV is the coefficient of variation, and the indices f and o represent forecasted and observed precipitation values, respectively. *KGE* values range from $-\infty$ to 1, with values closer to 1 indicating better model performance. Towner et al. [43] suggested the following classifications: "Good" (KGE $\geq$ 0.75), "Intermediate" (0.75 $\geq$ KGE $\geq$ 0.5), "Poor" (0.5 $\geq$ KGE > 0), and "Very poor" (KGE $\leq$ 0). The *BR* values greater than 1 indicate a positive bias, whereby forecasts overestimate precipitation relative to the observed data, while values less than 1 represent an underestimation. The γ values greater than 1 indicate that the variability in the forecast time series is higher than that observed, and values less than 1 show the opposite effect.

The *R* measures the strength and direction of the linear relationship between the forecast and observed values, and to what extent the temporal dynamics of the observed rainfall are captured in the forecasts. Correlation values of 0.6 or more are considered to be skillful. In addition, the root mean-square-error normalized by reference precipitation mean (NRMSE) was used.

3. Results and Discussion

3.1. Annual Spatial Variability and Seasonal Characteristics

Figure 2 presents the spatial map of annual (15 June 2019 to 15 June 2020) rainfall derived from the GFS and three satellite precipitation products. According to the reference rainfall product (i.e., IMERG Final), the spatial rainfall pattern is characterized by a strong north–south gradient: 815 mm yr^{-1} in the Sahel to 1190 mm yr^{-1} in the Savannah to 1300 mm yr^{-1} in the Guinea Coast zone. The 1-day lead GFS forecast captures the overall spatial pattern, but with large underestimation in the dry Sahel (656 mm yr^{-1} on average) and large overestimation in the Guinea Coast (1300 mm yr^{-1} on average), while producing a relatively accurate result in the Savannah (1130 mm yr^{-1} on average).

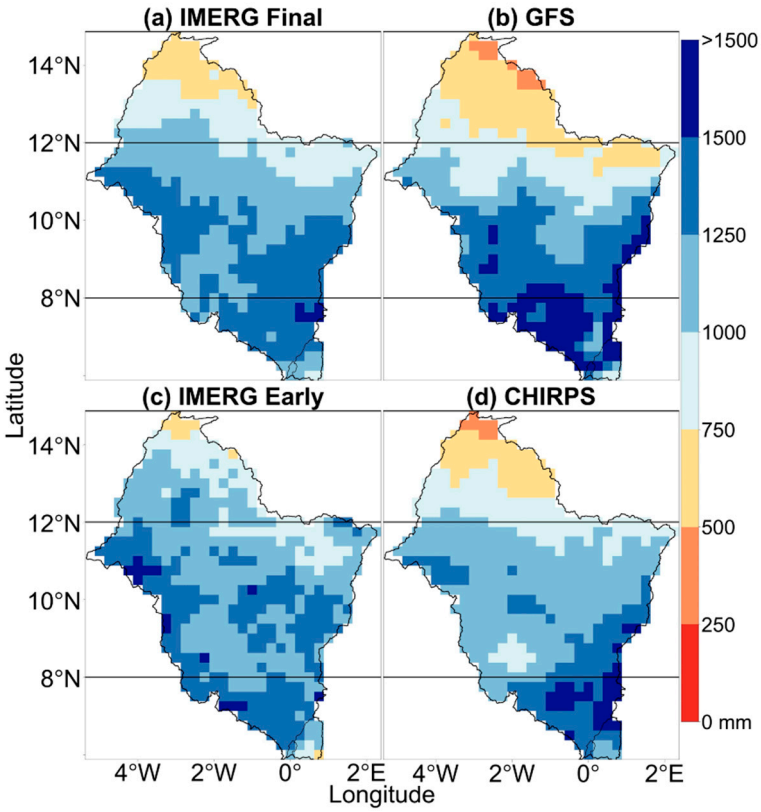

Figure 2. Spatial map of annual rainfall (in mm), for the period 15 June 2019 to 15 June 2020, derived from (**a**) IMERG Final, (**b**) GFS (1-day lead time), (**c**) IMERG Early, and (**d**) CHIRPS.

The uncalibrated IMERG Early estimates capture the overall spatial pattern, but result in large overestimation in the Sahel (1000 mm yr^{-1} on average), while providing more accurate results in both the Savannah and Guinea Coast areas. A comparison of the performance of GFS and IMERG Early, when both are validated against IMERG Final, shows that GFS outperforms IMERG Early in the Guinea Coast area, but that both provide highly biased estimates in the dry Sahel region with contradicting bias characteristics (i.e., large negative bias by GFS, but large positive bias by IMERG Early). The CHIRPS estimates accurately capture not only the overall spatial pattern, but also the actual magnitudes.

Figure 3 presents the monthly time series of climatological-zone-averaged rainfall derived from the 1-day lead GFS and three satellite precipitation products. Let us first

look at results in the Sahel region. According to the reference rainfall (i.e., IMERG Final), the rainy season is six months long, from May until November, and the rainfall peak is about 220 mm month^{-1} during July and August. The 1-day lead GFS forecast shows a similar seasonal pattern but tends to underestimate in all months, except in August. By contrast, IMERG Early is close to the reference product in almost all months but has large overestimation bias in two months (August and September). The CHIRPS estimates accurately capture the seasonal pattern as well as the actual magnitude of rainfall.

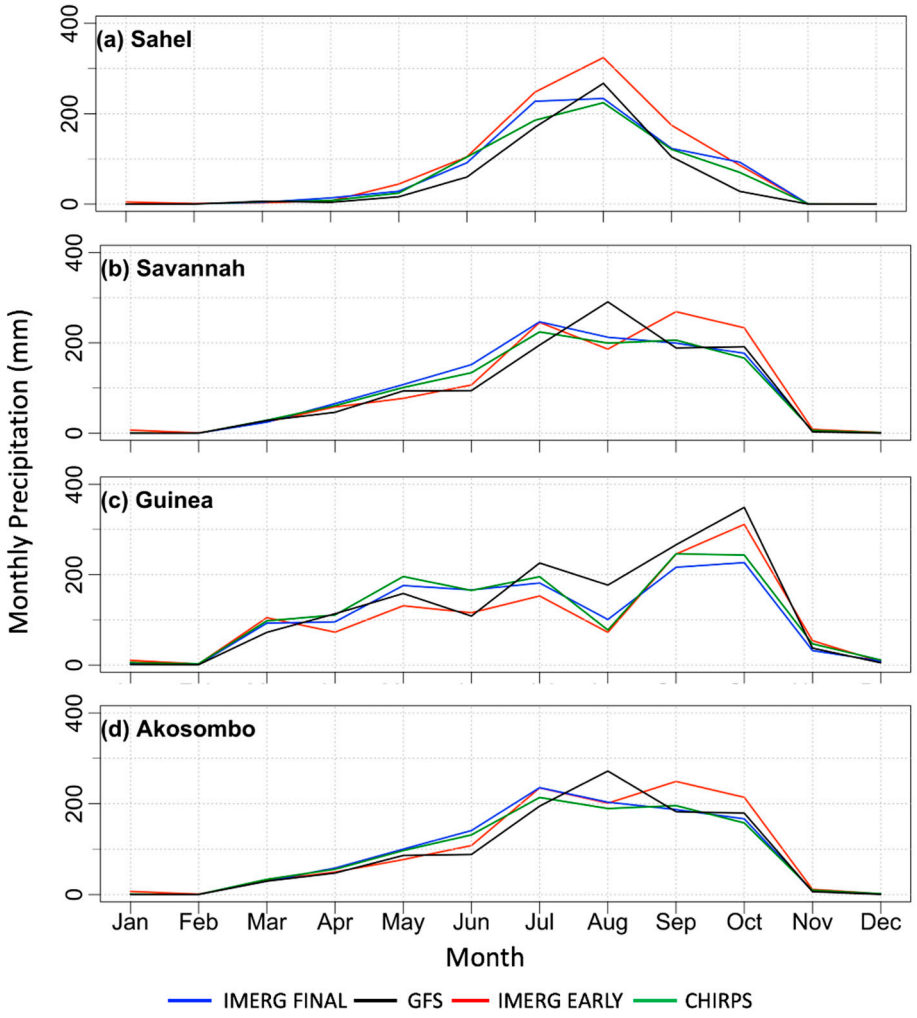

Figure 3. Monthly precipitation regime for the following regions of the Volta River Basin: (**a**) Sahel climate, (**b**) Savannah climate, (**c**) Guinea Coastal climate, and (**d**) watershed of the Akosombo dam. Analyses are based on rainfall fields derived from IMERG Final, 1-day lead GFS, IMERG Early, and CHIRPS. The time period covers from 15 June 2019 to 15 June 2020.

In the Savannah region, the rainfall starts early in March, and large rainfall accumulations occur over four months (July to October). The GFS underestimates rainfall in the first half of the rainy season (March through July), overestimates it in August, but estimates it well in the second half (September through November). By contrast, IMERG Early provides more accurate results in the first half of the rainy season, including August, but overestimates rainfall in September and October. Therefore, in the first half of the rainy season, including August, IMERG Early outperforms the GFS, but in the second half, GFS outperforms IMERG Early. The CHIRPS estimates are almost identical to the IMERG Final estimates.

Guinea Coastal zone is characterized by a similar rainfall season to Savannah but with large rainfall accumulations starting from as early as March. The GFS performs well in the first half of the rainy season but tends to overestimate in the second half. By contrast,

IMERG Early performs well in almost all months, except for October, when it overestimates. The CHIRPS monthly estimates are almost identical to IMERG Final.

Averaging the rainfall over the watershed of the Akosombo dam (i.e., averaging across all three climate zones) shows that the GFS forecasts accurately capture the seasonal rainfall pattern, but with some underestimation in the first few months of the rainy season, some overestimation in August, and almost perfect agreement from September until November. The GFS' overestimation bias in August was the result of overestimation bias in all climatological zones. The GFS' underestimation bias in the first few months of the rainy season was the result of similar underestimation bias in all the climatic zones. However, the GFS' almost perfect agreement with IMERG Final in the second half of the rainy season (September to November) arises as a result of the cancellation of the negative bias in the dry Sahel zone by the positive bias in the wet Guinea Coastal zone. In contrast to the GFS, IMERG Early provides accurate results in the first half of the rainy season, including August, but overestimates in the second half. Therefore, while IMERG Early outperforms GFS in the first half of the rainy season, GFS outperforms IMERG Early in the second half.

3.2. Dependence of Forecast Performance on Precipitation Rate

Figure 4a,b present the evaluation metrics of 1-day lead GFS using IMERG Final as the reference for the first half (March through July) and the second half (August through October) of the rainy season, using basin-averaged values for the Akosombo watershed. The GFS shows different bias characteristics during the first half and the second half of the rainy season.

During March through July, GFS tends to underestimate by 21% overall. Moreover, the bias depends on the rain rate: the GFS is almost unbiased and close to the reference at low rain rates (<5 mm/day), but has large underestimation bias for moderate and high rain rates. The overall RMSE of GFS is 113% of the mean rain rainfall rate. The linear correlation coefficient between GFS and IMERG Final is 0.59. The KGE of GFS is 0.48, which indicates almost "intermediate" skills for GFS in terms of its ability to forecast the basin-averaged daily rainfall values at a lead time of 1 day.

During August through October, GFS tends to overestimate by 14% overall; moreover, the bias varies depending on the rain rate: the GFS has large overestimation bias at low rain rates (<5 mm/day) and large underestimation bias at high rain rates (>15 mm/day), but is unbiased at moderate rain rates (5–15 mm/day). The overall RMSE is 89% of the mean rainfall rate, the correlation between GFS and IMERG Final is 0.44, and the KGE of GFS is 0.44.

3.3. What Is the Effect of Lead Time on the GFS Forecast Performance?

Figure 4 presents the evaluation metrics of the daily rainfall forecasts for 1-day, 5-day, 10-day, and 15-day lead times, for the first half and second half of the rainy season. During March through July, KGE decreases from 0.48 for 1-day lead to 0.25 for 15-day lead; R decreases from 0.59 (1-day lead) to 0.26 (15-day lead); and NRMSE deteriorates from 113% (1-day lead) to 159% (15-day lead). Similarly, during August through October, KGE decreases from 0.42 (1-day lead) to 0.18 (15-day lead); R decreases from 0.44 (1-day lead) to 0.21 (15-day lead); and NRMSE deteriorates from 87% (1-day lead) to 93% (15-day lead). Therefore, for a given timescale (daily rainfall), as the lead time increases, the forecast performance (as measured by KGE, R, and NRMSE) decreases. Other studies have also reported similar results for different regions of the world [11,12,17]. The difference in performance between short- and long-lead forecasts may be partly attributed to the data assimilation used in the GFS (which improves the initial conditions for short-lead forecasts) and its lack of coupling with an ocean model as GFS uses only prescribed sea surface temperature (this would affect the performance of long-lead forecasts).

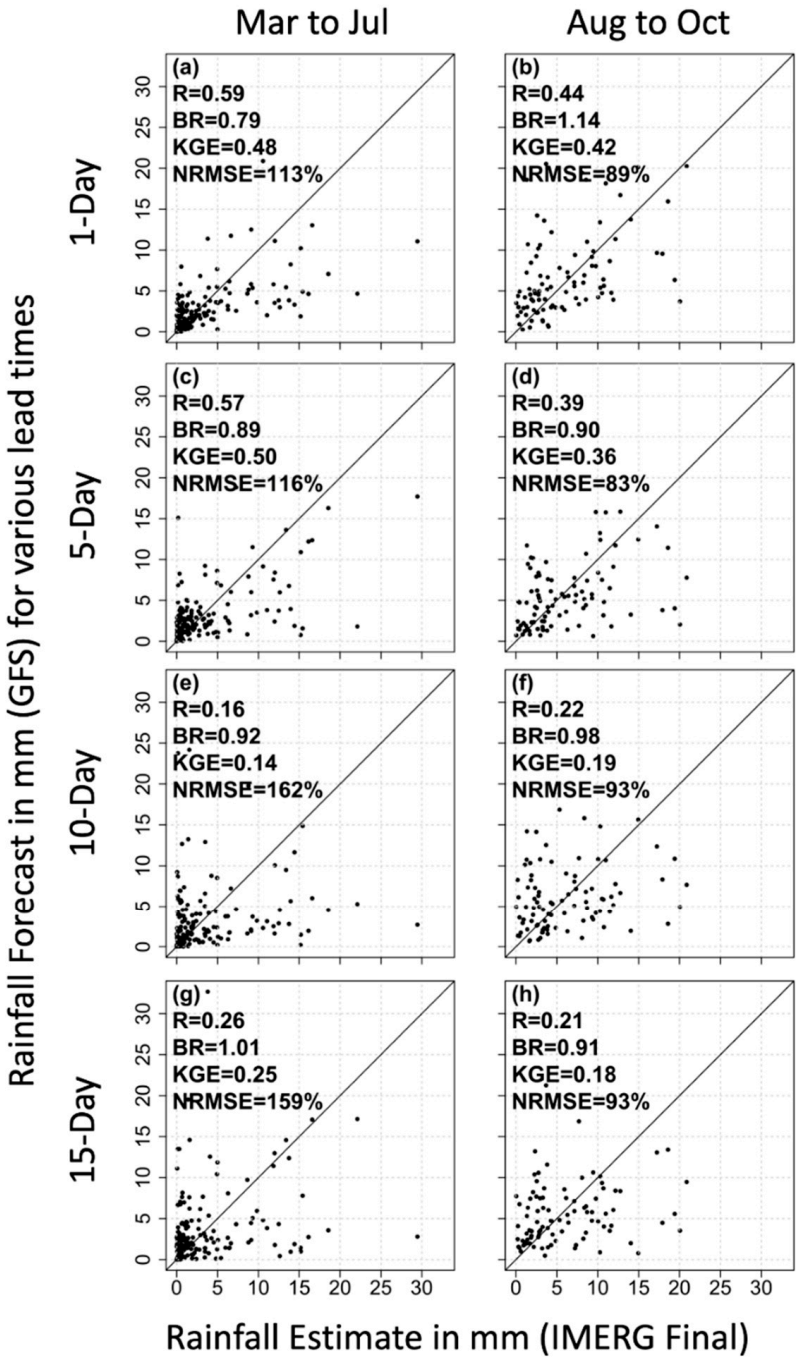

Figure 4. Scatterplot of sub-basin averaged accumulated precipitation forecast obtained from GFS forecast against corresponding values from IMERG Final, for various forecast lead times and periods of the rainy season. The lead times are: 1-day (**a**,**b**), 5-day (**c**,**d**), 10-day (**e**,**f**), and 15-day (**g**,**h**). The periods of the rainy season are: March through July (**a**,**c**,**e**,**g**), and August through October (**b**,**d**,**f**,**h**).

The performance measures between the 1 day and 5-day lead times are close to each other, and so are the forecasts at 10-day and 15-day leads. However, there is a sharp decrease in performance somewhere between 6 and 10 days. The reduction in performance from the 1-day lead time to the 5-day lead time is very small, but is substantial at 10-day lead time forecasts and beyond. Compared to the 1-day lead forecast, the performance measures of the 15-day lead show the following results for the first period of rainy season (March to July): NRMSE increases by 46%, R decreases by 0.33, and KGE decreases by 0.23. However, the bias statistics tend to improve with increasing lead time. As the lead time increases form 1-day lead to 15-day lead, the bias ratio improves from 0.79 to 1.01 (March to July),

and from 1.14 to 0.91 (August to October). Dravitzki and Mcgregor [11] reported little variation of bias with lead time.

3.4. What Is the Effect of Accumulation Timescale on Forecast Performance?

Figure 5 presents the performance statistics of the GFS rainfall forecast for 1-day, 5-day, 10-day, and 15-day rainfall accumulation timescales. For the March through July period, KGE increases from 0.48 for daily rainfall to 0.74 for 15-day accumulated rainfall; R increases from 0.59 (daily) to 0.89 (15-day accumulation); and NRMSE improves from 113% (daily) to 32 (15-day accumulation). Similarly, for the August through October period, KGE increases from 0.42 (daily), to 0.88 (15-day accumulation); R increases from 0.44 (daily) to 0.89 (15-day accumulation); and NRMSE improves from 87 (daily) to 15 (15-day accumulation). Therefore, the GFS forecast performance (as measured by KGE, R, and NRMSE) substantially increases with the temporal aggregation scale. We point out that aggregation over longer temporal scales involves the use of data from both short-lead and long-lead forecasts, a situation that leads to two contradictory error natures. On the one hand, introducing data from long-lead times brings larger random errors but, on the other hand, the temporal aggregation tends to average out the random errors. This study shows that the effect of averaging out the errors outweighs the additional error introduced by the use of long-lead forecast.

Let us compare the effects of increasing lead time (for the same timescale) with the effects of increasing the accumulation timescale. For daily rainfall, the ratio between the NRMSE for the 15-day lead and the 1-day lead is 1.407 (see Figure 4). For the accumulated precipitation, the ratio between the NRMSE for 1 day and the 15-day accumulated forecast is 3.53 (see Figure 5). This indicates that the reduction in the total error obtained by averaging the forecasts over a 15-day period was much higher than the error introduced by long-lead forecasts used in the averaging process. Visual inspection also indicates that the 15-day accumulated forecast has much better agreement with the reference (correlation of 0.89) compared to the 1-day forecast, indicating that the daily forecasts are dominated by large random errors, which average out during multi-day timescales.

The rate of improvement in forecast accuracy with increasing lead time is greater at short timescales than at long timescales. As the aggregation timescale increases from 1 day to 5 days, the NRMSE drops from 113% to 50%, and KGE jumps from 0.48 to 0.69. On the other hand, as the aggregation timescale increases from 10 days to 15 days, the NRMSE changes only slightly, from 36% to 32%, and KGE drops from 0.75 to 0.74.

3.5. Comparison of the Performance of IMERG Early and GFS

The GFS performance statistics indicate the need for the further improvement of GFS forecasts, particularly at short timescales. One way to achieve this could be through the use of post-processing techniques that involve rainfall estimates that have better accuracy and are available in near-real time. IMERG Early is a near-real time product, but is it more accurate than GFS? Figure 6 provides the scatterplot of IMERG Early against IMERG Final for 1-day, 5-day, 10-day, and 15-day aggregation temporal scales. As expected, due to the similarity of both algorithms, the correlation between the two estimates is very high, in excess of 0.90. However, IMERG Early tends to underestimate by about 20% in the first half of the rainy season, and to overestimate by about 20% in the second half of the rainy season. The underestimation by IMERG Early in the first half of the rainy season is primarily due to the underestimation of the medium rainfall rate. On the other hand, the overestimation of IMERG Early in the second half of the rainy season is due to its overestimation of the high rain rate. A comparison of the performance statistics of GFS and IMERG Early indicates that the KGE performance statistics are higher than those of GFS, which is attributable to the high correlation (and variability ratio) of IMERG Early and IMERG Final due to the similarity of both algorithms. In terms of the bias statistics, IMERG Early outperforms GFS during the first half of the rainy season, but GFS outperforms IMERG Early during the second half.

IMERG Early therefore has a better ability to reproduce the temporal dynamics of rainfall; however, in terms of bias, there is no clear winner between GFS and IMERG Early.

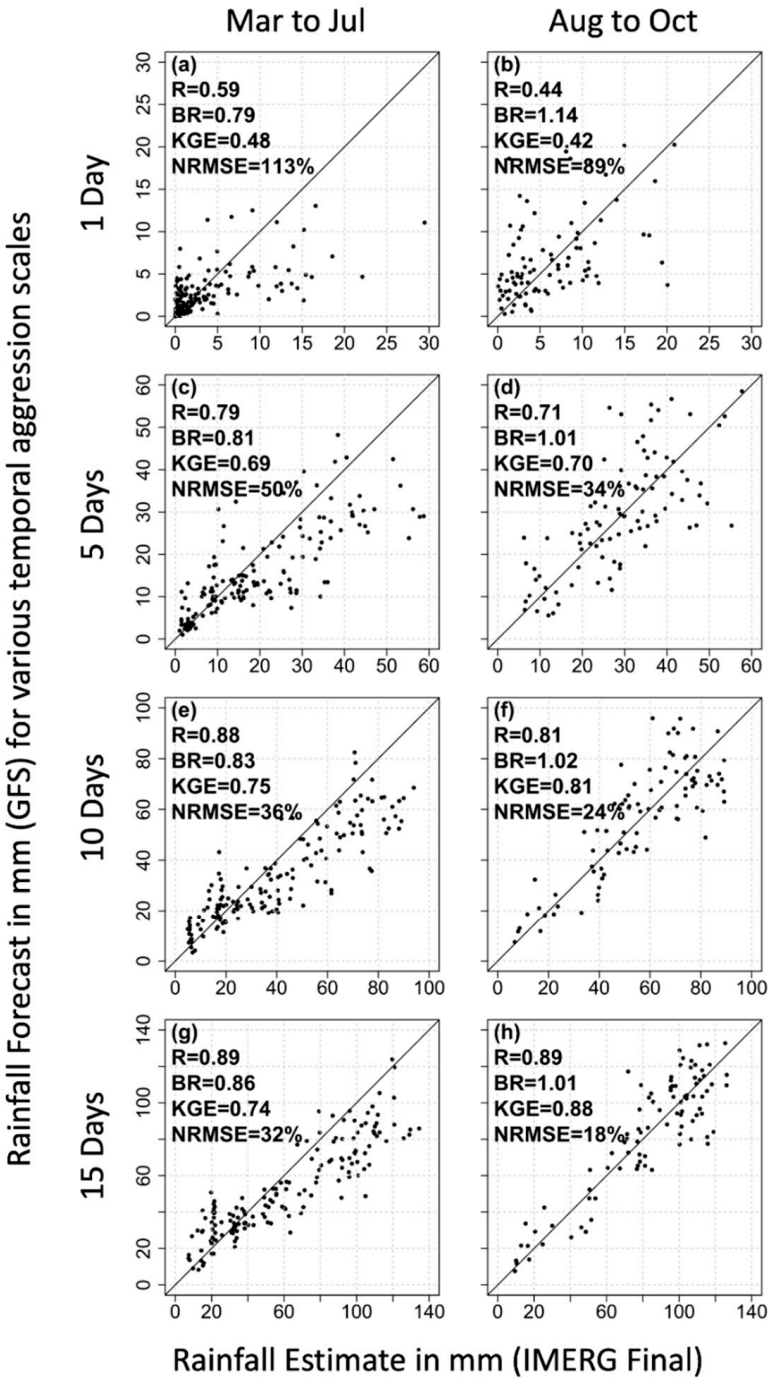

Figure 5. Scatterplot of sub-basin averaged accumulated precipitation forecast obtained from GFS forecast against corresponding values from IMERG Final, for various aggregation temporal scales and periods of the rainy season. The aggregation temporal scales are: 1-day (**a**,**b**), 5-day (**c**,**d**), 10-day (**e**,**f**), and 15-day (**g**,**h**). The periods of the rainy season are: March through July (**a**,**c**,**e**,**g**), and August through October (**b**,**d**,**f**,**h**).

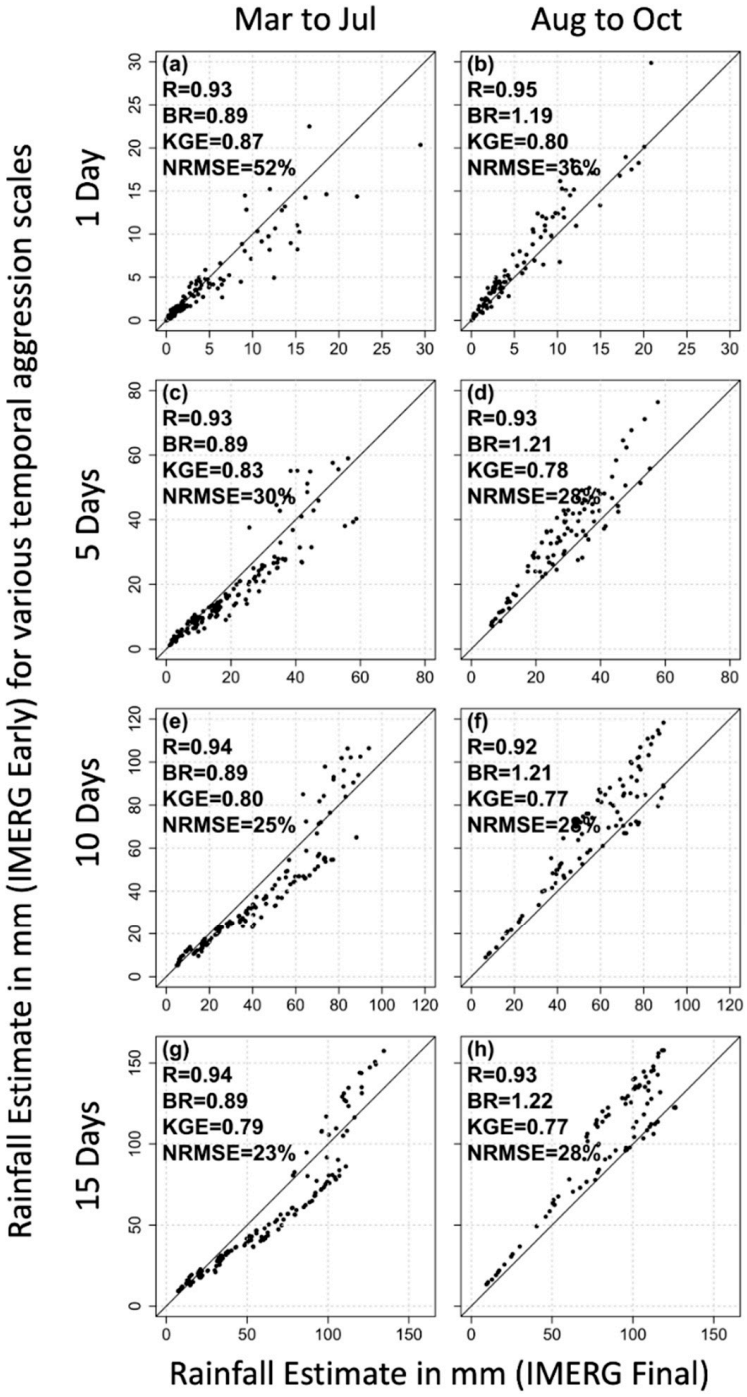

Figure 6. Scatterplot of sub-basin averaged accumulated precipitation forecast obtained from GFS forecast against corresponding values from IMERG Early evaluation, for various aggregation temporal scales and periods of the rainy season. The aggregation temporal scales are: 1-day (**a**,**b**), 5-day (**c**,**d**), 10-day (**e**,**f**), and 15-day (**g**,**h**). The periods of the rainy season are: March through July (**a**,**c**,**e**,**g**), and August through October (**b**,**d**,**f**,**h**).

3.6. How Is the Performance of GFS Affected If the Reference Product Is Changed from IMERG Final to CHIRPS?

We acknowledge that the reference dataset used in our evaluation (i.e., IMERG Final) has its own estimation errors. We conducted an additional assessment to evaluate the performance of GFS using CHIRPS rainfall products as references. Table 1 shows the performance statistics of GFS for different lead times, using IMERG FINAL and CHIRPS, separately, as reference. The pattern of GFS performance with respect to lead time is similar

when either rainfall product is used as a reference. The correlation and KGE statistics of GFS obtained using IMERG Final as reference are very close to the GFS obtained using CHIRPS as the reference. The NRMSE statistics computed using CHIRPS as the reference are lower than those obtained using IMERG Final as reference. The bias statistics computed using CHIRPS as the reference could be better or worse than those obtained using IMERG Final as the reference, depending on the period of the rainy season. Therefore, our results show that the overall performance of GFS remains the same if the reference product were to be changed from IMERG Final to CHIRPS, indicating the robustness of IMERG Final as the reference product.

Table 1. Performance statistics of GFS forecast for various lead times (1-day, 5-day, 10-day, and 15-day) using IMERG Final (CHIRPS) rainfall products as references, in terms of correlation, bias ratio, and NRMSE.

Lead Time of GFS Forecast	Correlation	Bias Ratio	KGE	NRMSE (%)
Time Period: March through July				
1 day	0.59 (0.57)	0.79 (0.65)	0.48 (0.44)	113 (87)
5 day	0.57 (0.50)	0.89 (0.73)	0.50 (0.42)	116 (91)
10 day	0.16 (0.17)	0.92 (0.76)	0.14 (0.11)	162 (121)
15 day	0.26 (0.30)	1.01 (0.83)	0.25 (0.22)	159 (118)
Time Period: August through October				
1 day	0.44 (0.50)	1.14 (1.00)	0.42 (0.49)	89 (71)
5 day	0.39 (0.39)	0.90 (0.79)	0.36 (0.36)	83 (72)
10 day	0.22 (0.18)	0.98 (0.86)	0.19 (0.17)	93 (81)
15 day	0.21 (0.21)	0.91 (0.80)	0.18 (0.19)	93 (80)

4. Summary and Conclusions

This study evaluated the accuracy of medium-range (1-day to 15-day lead time) forecasts available from the Global Forecast System (GFS), for the watershed of the Akosombo dam in the Volta basin, West Africa, using satellite-gauge merged IMERG Final and CHIRPS rainfall products as references. Additionally, the performance of GFS was compared to the performance of the satellite-only IMERG Early product to determine whether the latter can potentially be used in the post-processing of GFS.

The performance statistics of the 1-day lead watershed-averaged GFS forecast are: KGE = 0.42 to 0.49, correlation = 0.44 to 0.59, NRMSE = 71% to 114%. The performance of the GFS depends on the climate zone, month, rain intensity, lead time, and temporal aggregation scale. The GFS tends to underestimate (by 20%) in the dry Sahel region and overestimate (by 16%) in the Guinea Coastal region, while producing a relatively accurate result in the Savannah region. Averaging rainfall over the watershed of the Akosombo dam (i.e., averaging across all three climate zones) shows that the GFS forecasts accurately capture the seasonal rainfall pattern, but tend to underestimate rainfall (by 21%) in the first half of the rainy season (March through July) and overestimate rainfall (by 14%) in the second half of the rainy season (August through October). In the first half the rainy season, the GFS is unbiased at low rain rates, but has large underestimation bias for moderate and high rain rates. In the second half of the rainy season, the GFS has overestimation bias at low rain rates and large underestimation bias at high rain rates, but is unbiased at moderate rain rates.

For a given timescale, the GFS forecast performance decreases with increasing lead time. The reduction in forecast performance from the 1-day lead to the 5-day lead is very small, but it is substantial at the 10-day lead time forecast and beyond. The GFS forecast performance substantially increases with the temporal aggregation scale. The rate of improvement in forecast accuracy with increasing lead time is greater at short timescales than at long timescales. The 15-day total GFS forecast has the following performance statistics: KGE = 0.74 to 0.88, correlation = 0.89, and bias ratio = 0.86 and 1.01.

We also evaluated the performance of uncalibrated, near-real-time-available, IMERG Early rainfall products compared with IMERG Final in order to explore the potential possibility of utilizing IMERGE Early forecasts (because they are available in near-real time) for improving GFS forecasts through a post-processing technique. IMERG Early has worse bias than GFS during August to October; therefore, the inclusion of IMERG Early in GFS would not improve the bias of GFS forecasts. However, by contrast, IMERG Early products have similar ability to IMERG Final at capturing the spatial and temporal dynamics of rainfall. The performance statistics (correlation, KGE, and normalized RMSE) show that IMERG Early has better performance than GFS. Therefore, the inclusion of IMERG Early in GFS in post-processing could lead to the improvement of GFS forecast in its ability to predict the spatio-temporal dynamics of rainfall. However, further study is recommended to confirm this.

Our main conclusions are as follows: GFS has "poor" skills for predicting daily precipitation for the Akosombo dam watershed. Efforts toward improving the accuracy of the GFS forecasts should focus on reducing the large random error. Averaging GFS forecasts over longer accumulation timescales improves the accuracy of GFS forecasts. At the 15-day accumulation timescale, GFS shows "good" prediction skills, but has underestimation bias at moderate and high rain rates during March through July.

Author Contributions: M.G.: manuscript writing and conceptual design; H.Y.: data processing and analysis; V.N.: method design, contribution to manuscript text. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NASA Precipitation Measurement Mission, grant number 80NSSC19K0688.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: We acknowledge the National Center for Atmospheric Research (NCAR) for providing public access to the GFS rainfall forecast data products (https://rda.ucar.edu/datasets/ds084.1/) (accessed on 11 February 2022), NASA for providing public access to IMERG Final and IMERG Early rainfall data products (https://disc.gsfc.nasa.gov) (accessed on 11 February 2022), and the University of California Santa Barbara's (UCSB) Climate Hazard's group for providing public access to CHIRPS rainfall data (https://www.chc.ucsb.edu/data) (accessed on 11 February 2022).

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. Patt, A.G.; Ogallo, L.; Hellmuth, M. Learning from 10 Years of Climate Outlook Forums in Africa. *Science* **2007**, *318*, 49–50. [CrossRef]
2. Brueuer, N.E.; Fraisse, C.W.; Cabrera, V.E. The cooperative extension service as a boundary organization for diffusion of climate forecasts: A 5-year study. *J. Ext.* **2010**, *48*, 1–5.
3. Mase, A.S.; Prokopy, L.S. Unrealized Potential: A Review of Perceptions and Use of Weather and Climate Information in Agricultural Decision Making. *Weather Clim. Soc.* **2014**, *6*, 47–61. [CrossRef]
4. Pandya, R.; Hodgson, A.; Hayden, M.H.; Akweongo, P.; Hopson, T.; Forgor, A.A.; Yoksas, T.; Dalaba, M.A.; Dukic, V.; Mera, R.; et al. Using Weather Forecasts to Help Manage Meningitis in the West African Sahel. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 103–115. [CrossRef]
5. Koppa, A.; Gebremichael, M.; Zambon, R.C.; Yeh, W.W.-G.; Hopson, T.M. Seasonal Hydropower Planning for Data-Scarce Regions Using Multimodel Ensemble Forecasts, Remote Sensing Data, and Stochastic Programming. *Water Resour. Res.* **2019**, *55*, 8583–8607. [CrossRef]
6. Alexander, S.; Yang, G.; Addisu, G.; Block, P. Forecast-Informed Reservoir Operations to Guide Hydropower and Agriculture Allocations in the Blue Nile Basin, Ethiopia. *Int. J. Water Resour. Dev.* **2021**, *37*, 208–233. [CrossRef]
7. National Centers for Environmental Prediction; National Weather Service; NOAA; U.S. Department of Commerce. *NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive*; 109.074 TB; Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory: Boulder, CO, USA, 2015. [CrossRef]
8. Saha, S.; Moorthi, S.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Behringer, D.; Hou, Y.-T.; Chuang, H.; Iredell, M.; et al. The NCEP Climate Forecast System Version 2. *J. Clim.* **2014**, *27*, 2185–2208. [CrossRef]

9. Palmer, T.N.; Buizza, R.; Leutbecher, M.; Hagedorn, R.; Jung, T.; Rodwell, M.; Vitart, F.; Berner, J.; Hagel, E.; Lawrence, A.R.; et al. *The Ensemble Prediction System—Recent and Ongoing Developments*; ECMWF Technical Memoranda: Reading, UK, 2007. [CrossRef]
10. JMA. *Outline of the Operational Numerical Weather Prediction at the Japan Meteorological Agency (Appendix to WMO Numerical Weather Prediction Progress Report)*; Japan Meteorological Agency: Tokyo, Japan, 2019; Available online: https://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2019-nwp/index.htm (accessed on 1 July 2021).
11. Dravitzki, S.; McGregor, J. Predictability of Heavy Precipitation in the Waikato River Basin of New Zealand. *Mon. Weather Rev.* **2011**, *139*, 2184–2197. [CrossRef]
12. Sridevi, C.; Kumar, A.; Singh, K.K.; Durai, V.R.; Suneetha, P. Rainfall Forecast Skill of Global Forecasting System (GFS) Model over India during Summer Monsoon 2015. *Geofizika* **2018**, *35*, 39–52. [CrossRef]
13. Jiang, M.; Feng, J.; Li, Z.; Sun, R.; Hou, Y.-T.; Zhu, Y.; Wan, B.; Guo, J.; Cribb, M. Potential Influences of Neglecting Aerosol Effects on the NCEP GFS Precipitation Forecast. *Atmos. Chem. Phys.* **2017**, *17*, 13967–13982. [CrossRef]
14. Wang, J.-W.A.; Sardeshmukh, P.D.; Compo, G.P.; Whitaker, J.S.; Slivinski, L.C.; McColl, C.M.; Pegion, P.J. Sensitivities of the NCEP Global Forecast System. *Mon. Weather Rev.* **2019**, *147*, 1237–1256. [CrossRef]
15. Lien, G.-Y.; Kalnay, E.; Miyoshi, T.; Huffman, G.J. Statistical Properties of Global Precipitation in the NCEP GFS Model and TMPA Observations for Data Assimilation. *Mon. Weather Rev.* **2016**, *144*, 663–679. [CrossRef] [PubMed]
16. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Wolff, D.B.; Adler, R.F.; Gu, G.; Hong, Y.; Bowman, K.P.; Stocker, E.F. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [CrossRef]
17. Yang, F.; Cerrai, D.; Anagnostou, E. The Effect of Lead-Time Weather Forecast Uncertainty on Outage Prediction Modeling. *Forecasting* **2021**, *3*, 501–516. [CrossRef]
18. Dezfuli, A.K.; Ichoku, C.M.; Huffman, G.J.; Mohr, K.I.; Selker, J.S.; van de Giesen, N.; Hochreutener, R.; Annor, F.O. Validation of IMERG Precipitation in Africa. *J. Hydrometeorol.* **2017**, *18*, 2817–2825. [CrossRef]
19. Huffman, G.J.; Stocker, E.F.; Bolvin, D.T.; Nelkin, E.J.; Jackson, J. *GPM IMERG Final Precipitation L3 1 Day 0.1 Degree × 0.1 Degree V06*; Savtchenko, A., Ed.; Goddard Earth Sciences Data and Information Services Center (GES DISC): Greenbelt, MD, USA, 2019. [CrossRef]
20. Huffman, G.J.; Stocker, E.F.; Bolvin, D.T.; Nelkin, E.J.; Jackson, J. *GPM IMERG Early Precipitation L3 1 Day 0.1 Degree × 0.1 Degree V06*; Savtchenko, A., Ed.; Goddard Earth Sciences Data and Information Services Center (GES DISC): Greenbelt, MD, USA, 2019. [CrossRef]
21. van de Giesen, N.; Hut, R.; Selker, J. The Trans-African Hydro-Meteorological Observatory (TAHMO). *WIREs Water* **2014**, *1*, 341–348. [CrossRef]
22. Dezfuli, A.K.; Ichoku, C.M.; Mohr, K.I.; Huffman, G.J. Precipitation Characteristics in West and East Africa from Satellite and in Situ Observations. *J. Hydrometeorol.* **2017**, *18*, 1799–1805. [CrossRef]
23. Satgé, F.; Defrance, D.; Sultan, B.; Bonnet, M.-P.; Seyler, F.; Rouché, N.; Pierron, F.; Paturel, J.-E. Evaluation of 23 Gridded Precipitation Datasets across West Africa. *J. Hydrol.* **2020**, *581*, 124412. [CrossRef]
24. Maranan, M.; Fink, A.H.; Knippertz, P.; Amekudzi, L.K.; Atiah, W.A.; Stengel, M. A Process-Based Validation of GPM IMERG and Its Sources Using a Mesoscale Rain Gauge Network in the West African Forest Zone. *J. Hydrometeorol.* **2020**, *21*, 729–749. [CrossRef]
25. Stellingwerf, S.; Riddle, E.; Hopson, T.M.; Knievel, J.C.; Brown, B.; Gebremichael, M. Optimizing Precipitation Forecasts for Hydrological Catchments in Ethiopia Using Statistical Bias Correction and Multi-Modeling. *Earth Space Sci.* **2021**, *8*, e2019EA000933. [CrossRef]
26. Huffman, G.J.; Bolvin, D.T.; Braithwaite, D.; Hsu, K.; Joyce, R.; Xie, P.; Yoo, S.H. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm Theor. Basis Doc. (ATBD) Vers.* **2015**, *4*, 26.
27. Mul, M.; Obuobie, E.; Appoh, R.; Kankam-Yeboah, K.; Bekoe-Obeng, E.; Amisigo, B.; Logah, F.Y.; Ghansah, B.; McCartney, M. *Water Resources Assessment of the Volta River Basin*; International Water Management Institute (IWMI): Colombo, Sri Lanka, 2015; Volume 166, pp. 1–3.
28. Williams, T.O.; Marloes, M.; Biney, C.A.; Smakhtin, V. (Eds.) *The Volta River Basin: Water for Food, Economic Growth and Environment*; Routledge: London, UK, 2016. [CrossRef]
29. Baah-Kumi, B.; Ward, F.A. Poverty Mitigation through Optimized Water Development and Use: Insights from the Volta Basin. *J. Hydrol.* **2020**, *582*, 124548. [CrossRef]
30. Barry, B.; Obuobie, E.; Andreini, M.; Andah, W.; Pluquet, M. The Volta River Basin. Available online: http://www.iwmi.cgiar.org/assessment/files_new/research_projects/river_basin_development_and_management/VoltaRiverBasin_Boubacar.pdf (accessed on 7 November 2021).
31. National Centers for Environmental Prediction. National Centers for Environmental Prediction: The Global Forecast System (GFS) -Global Spectral Model (GSM). 2021. Available online: https://www.emc.ncep.noaa.gov/emc/pages/numerical_forecast_systems/gfs/documentation.php (accessed on 1 July 2021).
32. National Centers for Environmental Prediction. National Centers for Environmental Prediction: FV3: The GFDL Finite-Volume Cubed-Sphere Dynamical Core. 2021. Available online: https://www.gfdl.noaa.gov/fv3/ (accessed on 1 July 2021).
33. Putman, W.M.; Lin, S.-J. Finite-Volume Transport on Various Cubed-Sphere Grids. *J. Comput. Phys.* **2007**, *227*, 55–78. [CrossRef]

34. Mlawer, E.J.; Taubman, S.J.; Brown, P.D.; Iacono, M.J.; Clough, S.A. Radiative Transfer for Inhomogeneous Atmospheres: RRTM, a Validated Correlated-k Model for the Longwave. *J. Geophys. Res.* **1997**, *102*, 16663–16682. [CrossRef]
35. Iacono, M.J.; Mlawer, E.J.; Clough, S.A.; Morcrette, J.-J. Impact of an Improved Longwave Radiation Model, RRTM, on the Energy Budget and Thermodynamic Properties of the NCAR Community Climate Model, CCM3. *J. Geophys. Res.* **2000**, *105*, 14873–14890. [CrossRef]
36. Clough, S.A.; Shephard, M.W.; Mlawer, E.J.; Delamere, J.S.; Iacono, M.J.; Cady-Pereira, K.; Boukabara, S.; Brown, P.D. Atmospheric Radiative Transfer Modeling: A Summary of the AER Codes. *J. Quant. Spectrosc. Radiat. Transf.* **2005**, *91*, 233–244. [CrossRef]
37. Chen, F.; Janjić, Z.; Mitchell, K. Impact of Atmospheric Surface-Layer Parameterizations in the New Land-Surface Scheme of the NCEP Mesoscale Eta Model. *Bound.-Layer Meteorol.* **1997**, *85*, 391–421. [CrossRef]
38. Arakawa, A.; Schubert, W.H. Interaction of a Cumulus Cloud Ensemble with the Large-Scale Environment, Part I. *J. Atmos. Sci.* **1974**, *31*, 674–701. [CrossRef]
39. Grell, G.A. Prognostic Evaluation of Assumptions Used by Cumulus Parameterizations. *Mon. Weather Rev.* **1993**, *121*, 764–787. [CrossRef]
40. 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]
41. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling. *J. Hydrol.* **2009**, *377*, 80–91. [CrossRef]
42. Kling, H.; Fuchs, M.; Paulin, M. Runoff Conditions in the Upper Danube Basin under an Ensemble of Climate Change Scenarios. *J. Hydrol.* **2012**, *424*, 264–277. [CrossRef]
43. Towner, J.; Cloke, H.L.; Zsoter, E.; Flamig, Z.; Hoch, J.M.; Bazo, J.; Coughlan de Perez, E.; Stephens, E.M. Assessing the Performance of Global Hydrological Models for Capturing Peak River Flows in the Amazon Basin. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 3057–3080. [CrossRef]