Lawrence Berkeley National Laboratory
LBL Publications

Title
Performance Evaluation of Weather@home2 Simulations over West African Region

Permalink
https://escholarship.org/uc/item/7p393439

Journal
Atmosphere, 16(4)

ISSN
0004-6973

Authors
Lawal, Kamoru Abiodun
Akintomide, Oluwatosin Motunrayo
Olaniyan, Eniola
et al.

Publication Date
2025-03-28

DOI
10.3390/atmos16040392

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

atmosphere

MDPI

Article

Performance Evaluation of Weather@home2 Simulations over West African Region

Kamoru Abiodun Lawal [1,2,3,*], Oluwatosin Motunrayo Akintomide [3,4], Eniola Olaniyan [3,4], Andrew Bowery [5], Sarah N. Sparrow [5], Michael F. Wehner [6] and Dáithí A. Stone [7]

1 African Centre of Meteorological Applications for Development (ACMAD), Niamey 13184, Niger
2 African Climate and Development Initiative (ACDI), Department of Environmental and Geographical Science, University of Cape Town, Cape Town 7700, South Africa
3 Department of Meteorology, African Aviation and Aerospace University, Abuja 900102, Nigeria; m.akintomide@nimetagency.org.ng (O.M.A.); eniola.ajani@aaau.edu.ng (E.O.)
4 Numerical Weather and Climate Prediction Unit, Nigerian Meteorological Agency, Abuja 900102, Nigeria
5 Oxford e-Research Centre, Department of Engineering Science, University of Oxford, Oxford OX1 2JD, UK; andy.bowery@oerc.ox.ac.uk (A.B.); sarah.sparrow@oerc.ox.ac.uk (S.N.S.)
6 Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; mfwehner@lbl.gov
7 National Institute of Water & Atmospheric Research Ltd. (NIWA), Evans Bay Parade, Hataitai, Wellington 6021, New Zealand; daithi.stone@niwa.co.nz
* Correspondence: lawal@csag.uct.ac.za

Abstract: Weather and climate forecasting, using climate models, have become essential tools and life-savers in the West African region; in spite of the fact that climate models do not fully comply with attributes of forecast qualities—RASAP: reliability, association, skill, accuracy, and precision. The objective of this paper is to quantitatively evaluate, in comparison to CRU and ERA5 datasets, the RASAP compliance-level of the weather@home2 modeling system (w@h2). Findings from some statistical evaluations show that, to a moderately significant extent, w@h2 model provides useful information during the monsoon seasons; skills to capture the Little Dry Season over the Guinea zone; predictive skills for the onset season; ability to reproduce all the annual characteristics of the surface maximum air temperature over the region; as well as skill to detect heat waves that usually ravage West Africa during the boreal spring. The model displays traces of attributes that are needed for seasonal climate predictions and applications. Deficiencies in the quantitative reproducibility point to the facts that the model does provide a reliability akin to that of regional climate models. This paper further furnishes a prospective user with information on whether the model might be "useful or not" for a particular application.

Keywords: West Africa; forecast quality; weather@home2; seasonal climate; evaluation

check for updates

Academic Editor: Eugene Rozanov

Received: 20 February 2025
Revised: 23 March 2025
Accepted: 26 March 2025
Published: 28 March 2025

Citation: Lawal, K.A.; Akintomide, O.M.; Olaniyan, E.; Bowery, A.; Sparrow, S.N.; Wehner, M.F.; Stone, D.A. Performance Evaluation of Weather@home2 Simulations over West African Region. *Atmosphere* **2025**, *16*, 392. https://doi.org/10.3390/atmos16040392

Copyright: © 2025 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

Over the years, especially in the West African region, researchers' and stakeholders' confidence in the use of climate models is increasing. This is due to improvements in nearly all aspects of climate models' fidelity and skill, as well as more detailed understanding of the degree of fidelity and skill [1–5]. Consequently, information from climate models is extensively being used by the region's policy makers and various socio-economic sectors (e.g., water resources management, agriculture, engineering, environmental management, health, insurance, researchers, etc.) either for risk management or for day-to-day, season-to-season, or long-term strategies and planning [6–9]. Hence, there is a proliferation of climate models, which calls for caution among researchers and stakeholders. To calm the fears and

concerns of prospective users, thorough performance evaluations of climate models have to be carried out before their eventual utilization.

The performance of a climate model can be evaluated according to these five attributes of forecast qualities, hereafter known as RASAP: R (*reliability*), A (*association*), S (*skill*), A (*accuracy*), and P (*precision*) [10–13]. In short terms, *reliability* can be referred to as the ability of a forecast to provide an unbiased estimate. According to [12,14], it is a key quality of a probabilistic long-range forecast. Refs. [15,16] described association as a measure of linear relationship between forecast and observation. Skill is a comparative quantity that shows if a set of forecasts is better than a reference set, e.g., climatology, persistence, etc. It is a measure of the relative ability of a set of forecasts with respect to some set of standard reference forecasts [14,17–19]. Accuracy can be referred to as the overall correspondence or level of agreement between model and observation. According to [13] it summarizes the overall quality of a forecast, while precision, a measure of uncertainty, is simply the absence of random error, i.e., a measure of statistical variance of an estimation that is independent of a true value [20]. Precision is described as the spread of the data whenever sampling is involved [21].

Climate models do not fully pass thresholds for these measures over many regions of the world, including the West African region, and hence they are not fully RASAP compliant. Assessing their degree of RASAP compliance therefore provides a quantitative evaluation of their ability to represent regional climate. Performances of several climate models have been evaluated over the West African region. While some of these evaluations have been motivated by the importance of the West African monsoon and its circulation features, others have been interested in mechanisms and processes responsible for rainfall regimes [2,3,22,23]. There have also been some evaluations to improve the understanding of the nature of the interactions across the different dynamical systems within the West African monsoon [1,4,5,24,25].

A major challenge in evaluating RASAP performance is that many of the measures require large initial-condition ensembles of simulations, which can be computationally prohibitive. In this paper, we focus on evaluating the RASAP performance of a modeling system that has produced an exceptionally large number of simulations, thus providing material for robust tests against the RASAP measures—the weather@home2 modeling system (hereafter w@h2). w@h2 is a successor to the well-known weather@home modeling system (hereafter w@h1) [26,27]. Generally, the w@h2 modeling system can generate very large ensembles of simulations that allow denser sampling of the climate distributions. This is corroborated by [28]; they are of the opinion that a single-model initial-condition large ensemble (SMILE), if less than 30 members, may analytically underestimate precipitation variability, vis-à-vis distribution.

The w@h2 modeling system has been designed for the investigation of the behavior of extreme climate under anthropogenic climate change. This means that measures of the performance of the model in terms of climate variability are more relevant than measures of the mean climatology [29–31]. If the w@h2 modeling system is to be used to understand changes in extreme climate over West Africa, then it is pertinent to evaluate the performance of the w@h2 simulations over the region. Therefore, this paper aims to utilize a series of statistical metrics to calculate some selected attributes of forecast qualities, i.e., RASAP, to provide insights on the nature of the w@h2 simulations. Specifically, we are asking the following questions—1. Are the w@h2 simulations, over West Africa, reliable? 2. Does any linear association exist between the simulations and observations/reanalysis over West Africa? 3. Do these simulations have skill over West Africa? 4. Are the simulations accurate, as well as precise over this region? These questions are asked with a view to understanding

whether the w@h2 simulations may be useful for extreme event attribution analysis over West Africa.

While this section introduces the motivations and concept of this study, including the description of the study domain and its complexities, Section 2 will discuss the datasets analyzed in this paper, and the adopted analysis procedures. Section 3 will describe the results, while Section 4 will provide the summary and conclusions.

2. Materials and Methods

2.1. Description of Study Area

This study focuses on West Africa, a unique region of atmospheric complexities. It is a tropical land mass located roughly within longitudes 20° W to 20° E, and latitudes 0° to about 25° N of the African continent (Figure 1). The region comprises three climatic zones, namely: Guinea—a tropical rain forest along the Atlantic coast; Savannah—a transition zone of short trees and grasses; and the Sahel—an Arid desert in the northern inlands [32–34].

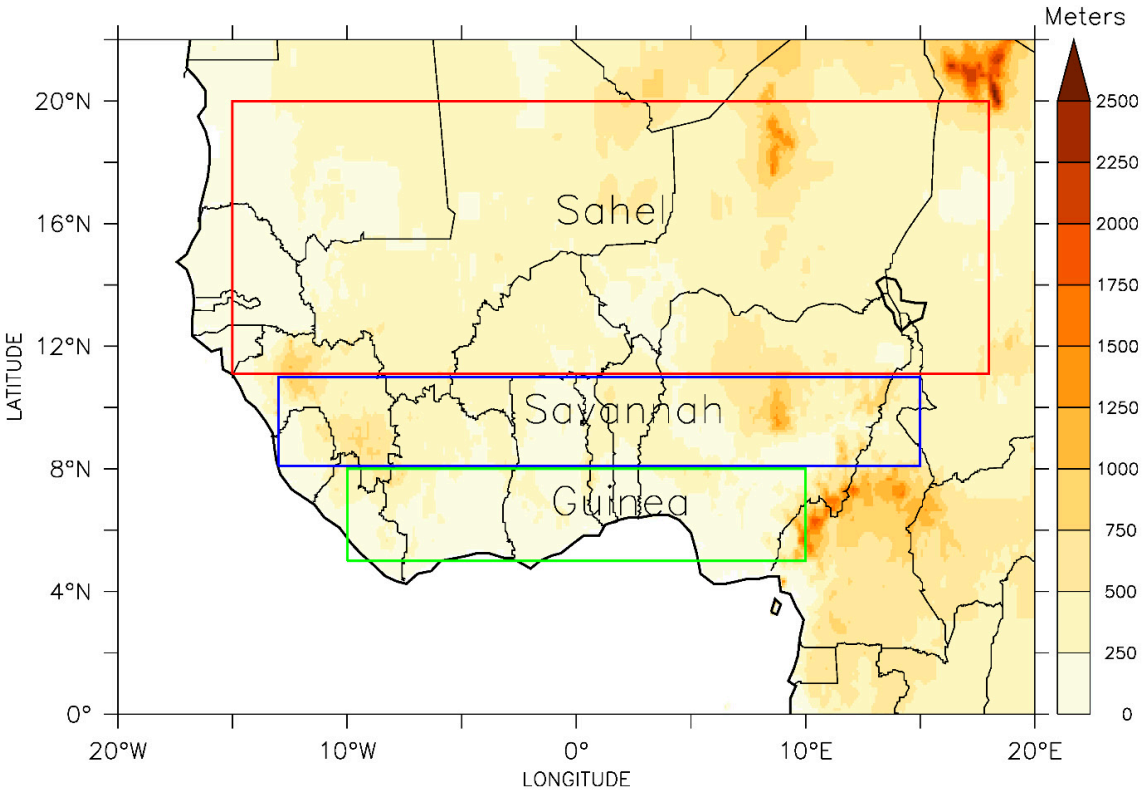

Figure 1. The domain of West Africa showing the topography of the surface (shaded; meters) and highlights of climatological zones—Guinea (green box), Savannah (blue box), and Sahel (red box).

West African climates result from the interactions of two migrating air masses: tropical maritime and tropical continental air masses. At the surface, these two air masses meet at a belt of variable width and stability called the Inter-Tropical Discontinuity (ITD) [35] or the Inter-Tropical Convergence Zone (ITCZ) if at the upper level. The north and south migration of ITD, which follows the annual cycle, influences the climate of the region [35,36]. Besides ITD, there are other key climate modification mechanisms over West Africa. Most relevant are the El Niño Southern Oscillation (ENSO) [37–39], the sea surface temperature (SST) anomalies over the Gulf of Guinea (GOG) [34,40], the African Easterly Jet (AEJ) [41–43], and the thermal lows [44–47]. The region's climate is classified into two seasons driven by the position of the ITD—the dry season and the rainy season. The period of the dry season runs approximately from November to March/April. It is a time of hot and dry

tropical continental air mass driven by the ridges from the northern hemispheric mid-latitude high-pressure system. During these periods, the prevailing northeasterly winds, north of the ITD, bring dry and dusty conditions across the region with the southernmost extension of this air mass occurring in January between latitudes 5° and 7°N. A tropical maritime southwesterly air mass is found at the southern ends of the ITD. The moist air mass dominates during the periods of the rainy seasons. The region's rainy seasons run from April/May to October depending on the climatic zone of interest (Figure 1) [34,48,49]. The northernmost penetration of the wet air mass is in August, usually between latitudes 19° and 22°N. With all these atmospheric complexities, the use of dynamical climate models for the forecasting of weather and projection of climate are indispensable over the region. Therefore, performance evaluation of meteorological forecasts and/or simulations is crucial for understanding the errors embedded within the modeling systems, and also to monitor the accuracy attained and the progress made in the climate modeling systems [12].

2.2. Datasets—Observation, Reanalysis, and Simulation Datasets

This study used monthly precipitation and near-surface (2 m) maximum air temperature from three categories of datasets—gridded observational, reanalysis, and w@h2 simulation. The observation datasets are from the University of East Anglia Climate Research Unit (CRU TS version 4.03 (CRU-TS4) [50–52]). This is based on the analysis of records of observations from over 4000 weather stations. The reanalysis datasets are from the European Centre for Medium-Range Weather Forecasts (ECMWF—ERA version 5 (ERA5) [53–55]). Simulated datasets are from the w@h2 modeling system run by the *climateprediction.net* project (CPDN) [56] at the University of Oxford. The w@h2 modeling system uses a one-way nesting of atmospheric models with the global HadAM3P-N96 and regional HadRM3P models [57]. HadAM3P-N96 is run at ~150 km resolution globally and drives HadRM3P at ~25 km resolution over a pan-African domain, that encompasses West Africa. These hydrostatic models are both run with 19 vertical levels. As detailed in [27], land surface processes are represented by the Met. Office Surface Exchange Scheme 2 (MOSES2) land surface scheme [58], while the greenhouse gas concentrations and aerosol burdens are as prescribed in [26]. Daily sea surface temperatures are imposed from the Operational Sea Surface Temperature and Ice Analysis (OSTIA) [59]. Simulations are made possible by the enlistment of thousands of volunteers around the world who, on their personal computers, run simulations starting from different initial conditions. The results are then uploaded onto the CPDN [56] server facility hosted by the University of Oxford [60]. Here, a prospective user of the datasets needs to register by creating an account and log-in on the CPDN website (i.e., server) in order to gain controlled access to the datasets. This distributed computing capacity is made possible by the Berkeley Open Infrastructure for Network Computing (BOINC) open-source infrastructure [61]. Details of the improvements made in w@h2 in comparison to w@h1 are discussed in [27].

These datasets have different spatial resolutions. The observed variables (CRU) are on a horizontal grid resolution of 0.5° × 0.5° longitude–latitude, while the reanalysis (ERA5) datasets have a horizontal resolution of 30 km grid. The horizontal resolution of w@h2 simulations is about 0.22° (25 km), compared to about 0.44° (50 km) in w@h1. For uniformity, the horizontal resolutions of all the simulated (w@h2) and the reanalysis (ERA5) datasets were re-gridded to match that of the observation (CRU) dataset before they were analyzed. All monthly simulated variables from w@h2 used in this study are from a sub-set of 71 ensemble members per year. Each ensemble member differs only slightly in their initial conditions and we focus on the 31-year period from January 1987 to December 2017.

2.3. Methodology and Analysis Procedures

Recall that this paper aims to evaluate the performances of the w@h2 simulations over West Africa in line with the qualities of selected forecast attributes—RASAP. In comparison to CRU and ERA5 datasets, w@h2 simulations are therefore subjected to a series of quantitative statistical metrics to calculate RASAP measures. As depicted in Table 1, temporal and spatial analyses of these statistical metrics are carried out and then presented in various graphical formats for interpretation. We also place some figures in the supplementary domain of this paper for clarity of purpose.

Results and analyses from this study will be presented on the basis of calendar months, in reflection of their common usage in climate services throughout the region and of the typical monthly duration of noteworthy extreme events in the region (e.g., [62]).

Table 1. List of descriptive statistical metrics used to calculate the attributes of forecast qualities in this study. ** A measure of statistical significance, such as *p*-value [63], is also assessed for the correlations that were evaluated in this study. Statistical significance was estimated using a two-tailed experiment at the $p = 0.1$ level, assuming uncorrelated Gaussian noise.

Attributes	Descriptive Statistics	Inference
Reliability	Climatology	To determine the monthly, seasonal, or annual cycle of a variable [64].
	Bias (B)	A measure of over- (positive bias) or under-estimations (negative bias) of variables. Generally, bias gives marginal distributions of variables [11].
	Mean bias error (MBE)	A measure to estimate the average bias in the model. It is the average forecast or simulation error representing the systematic error of a model to under- or over-forecast [11].
	Scatter diagrams	Provides information on bias, outliers, error magnitude, linear association, peculiar behaviors in extremes, misses, and false alarms. Perfect simulation points in comparison to observation should be on the 45° diagonal line [17,65].
Association	** Correlation coefficient (r)	A statistical measure of the strength of a linear relationship between the paired variables, i.e., simulations and observation/reanalysis datasets. By design, it is constrained as $-1 \leq r \leq 1$. Positive values denote direct linear association; negative values denote inverse linear association; a value of 0 denotes no linear association; while the closer the value is to 1 or -1, the stronger the linear association. Perfect relationship is denoted by 1. It is not sensitive to the bias but sensitive to outliers that may be present in the simulations [10,15,16].
	Coefficient of determination (CoD)	CoD is a measure of potential skill, i.e., the level of skill attainable when the biases are eliminated. It is also a measure of the fit of regression between forecast and observation. It is a non-negative parameter with a maximum value of 1. For a perfect regression, CoD = 1. CoD tends zero for a non-useful forecast [10,16].
Skill	Ranked probability skill score (RPSS)	Measures the forecast accuracy with respect to a reference forecast (e.g., observed climatology). Positive values (maximum of 1) have skill while negative values (up to negative infinity) have no skill [10,14,17–19].

Table 1. *Cont.*

Attributes	Descriptive Statistics	Inference
Accuracy	Mean absolute error (MAE)	A measure of how big of an error we can expect from the forecast on average, without considering their directions. MAE measures the accuracy of a continuous variable. Though, just like the root mean square error (RMSE), it also measures the average magnitude of the errors in a set of forecasts; however, while RMSE utilizes a quadratic scoring rule, MAE is a linear score—which means that all the individual differences are weighted equally in the average. MAE ranges from zero to infinity. Lower values are better [10,66,67].
	Root mean square error (RMSE)	It measures the magnitudes of the error, weighted on the squares of the errors. Though, it does not indicate the direction of the error; however, it is good in penalizing large errors. It is sensitive to large values (e.g., in precipitation) and outliers. This is very useful when large errors are undesirable. Ranges from zero to infinity. Lower values are better [10,68,69].
	Synchronization (Syn)	Synchronization focuses on the predictive capabilities of a model. It shows how much a simulated value agrees with an observed value in the signs of their anomalies without taking magnitudes into consideration. Therefore, the evaluated synchronization, in a probabilistic sense, is similar to accuracy. The best synchronization is 100% [10,13,70,71].
Precision	Standard deviation (Std)	Std helps to determine the spread of simulations and/or observations from their respective means, i.e., how far from the mean a group of numbers is. It has the same unit as the mean [10,21,72,73].
	Coefficient of variation (CoV)	It is used for comparing the degree of variation from one data series to another (in this case between forecast or simulation and observation where the means are significantly different from one another). A lower CoV implies a low degree of variation while a higher CoV implies a higher variation. Therefore, the higher the CoV, the greater the level of spreading around the mean [10,21].
	Normalized standard deviation (NSD)	This makes it possible to access the statistics of different fields (observations and simulations) on the same scale. Here, Taylor diagrams are used to depict the normalized standard deviation in line with correlation coefficients. The diagrams are able to measure how well observations and simulations match each other in terms of 1. similarity as measured by correlation coefficients, and 2. deviation factors as measured by normalized standard deviations. Taylor diagrams are able to provide a summarizing evaluation of model performance in simulating atmospheric parameters [10,21].

3. Results

3.1. Seasonality (and Reliability)

Here, the ability of the w@h2 model to replicate seasonality and its deviations from it are investigated, bearing in mind that the reliability of a probabilistic forecast is the

statistical consistency between each class of forecasts and the corresponding distribution of observations that follows such forecasts [12]. Statistical metrics used to support the evaluation of reliability, in this paper, are climatology, mean bias, and the use of scatter diagrams. More details are depicted in Table 1.

3.1.1. Precipitation

The w@h2 model is able to capture the monthly mean distributions of rainfall spatially and temporally (Figures 2 and S1). As the rain band transverses hundreds of kilometers from south inland to north during the first half of the calendar year, w@h2 is able to capture the maximum rainfall along the coastal Guinea areas as well as the tropical aridity climates over the Sahel (Figures 2 and S2a–c). The spatial correlations (r) between w@h2 simulations and CRU/ERA5 observations range from 0.68 to 0.85 (Figure S1a). While the model is able to reliably simulate the characteristic of rainfall in August over both Savannah and Sahel, when these zones normally experience their rainfall peaks, it is also able to capture the pause in rainfall intensities along the coastal Guinea areas in August—the little dry season (LDS: Figures 2 and S1a–c). However, the LDS as simulated by w@h2 extends from Sierra Leone to southern Cameroon (Figure S1a), contrary to the Cote d'Ivoire to southeastern Nigeria extent as observed by CRU and ERA5 (Figure S1b,c).

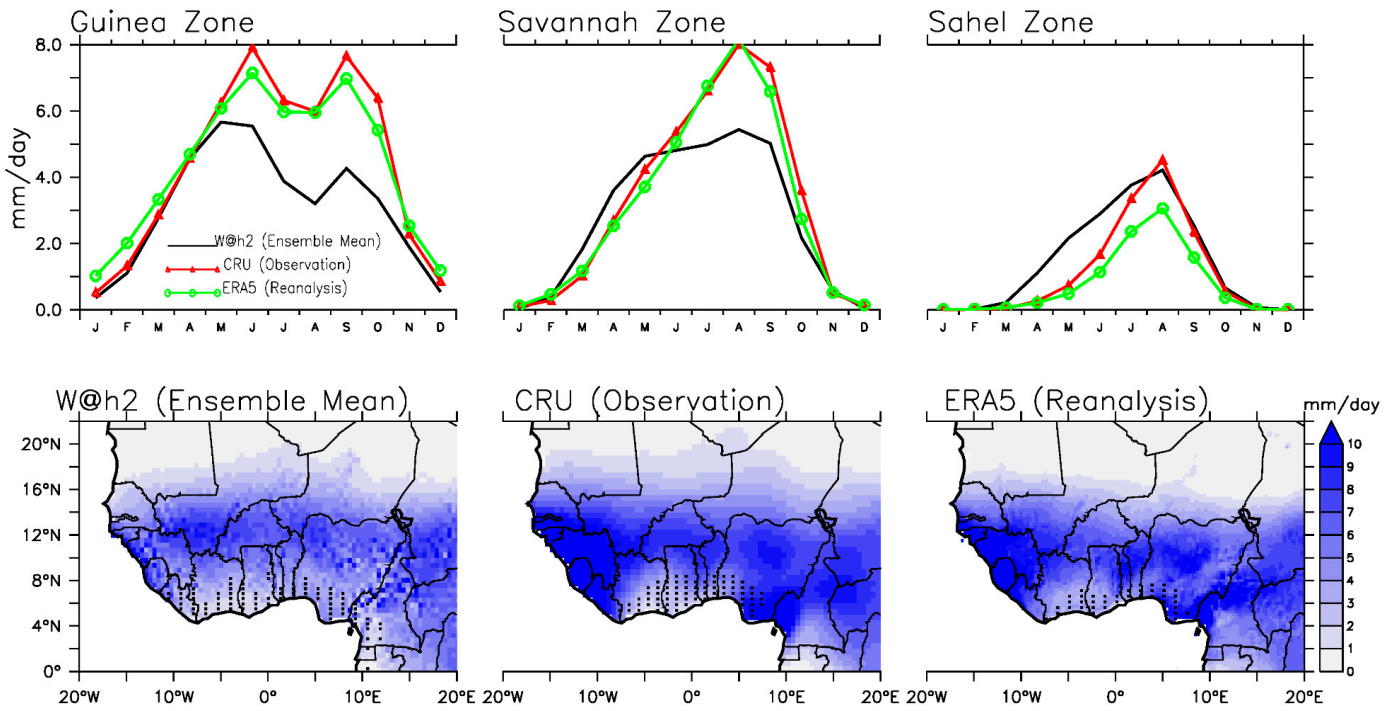

Figure 2. (**Top panels**) Areal averages of monthly mean distributions of rainfall (mm day^{-1}) for each climatological zone: (**left**) Guinea, (**middle**) Savannah, and (**right**) Sahel. (**Bottom panels**) Mean spatial distributions of rainfall (shaded; mm day^{-1}) over West Africa for the month of August: (**left**) w@h2 ensemble mean simulation, (**middle**) CRU-observation, (**right**) ERA5-reanalysis. Stippling on the bottom panels indicate areas, over West Africa, that usually experience the little dry season (LDS) in August.

Figure S2a shows that w@h2 rainfall over Guinea is consistently too low from June to October. Savannah rainfall is too high during March–May and too low during June–October (Figure S2b), while Sahel rainfall is too high from April to September (Figure S2c). The bias ranges of ± 5 mm day^{-1} (Figure S2d,e) are small in comparison to rainfall totals over most of the region.

3.1.2. Temperature

The spatial correlations (r) of the monthly temperature climatology between w@h2 simulations and CRU/ERA5 observations are generally greater than 0.9 (Figure S3a–c). w@h2 under-estimates the temperature in all the climatic zones by 0.5–2.0 °C (Figure S4a–e), though with patches of inconsistent over-estimations over the Sahel.

In addition, the w@h2 model captures the four main characteristics of the seasonal cycle of near-surface maximum temperature over West Africa. First, the model captures the two peaks of maximum air temperature exhibited annually in all the climatic zones, with the primary peak being from February to May, and the secondary peak being in October–November (Figures 3, S3a–c and S4a–c). Second, the model also agrees with observations that the Sahel region is always warmer than both the Savannah and coastal Guinea regions, except during the boreal winters. Third, the model agrees that there is a dip in the annual maximum temperatures over all the climatic zones during the peak of the rainy season (i.e., in August: Figures 3 and S3a–c). Lastly, the annual north–south oscillation of the thermal depression is also captured by the w@h2 model (Figure S3a–c), this being a large expanse of areas where the lowest atmospheric pressure coincides with surface temperature maximum (Figure 3) [44,46,47].

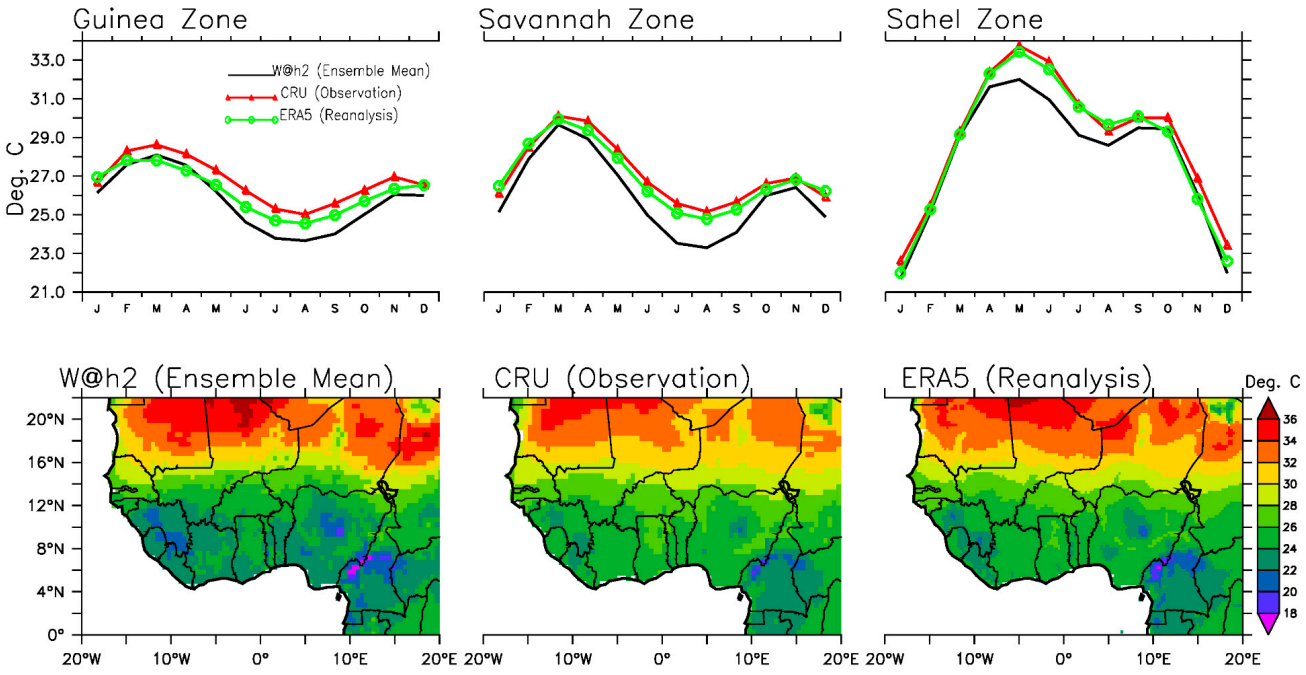

Figure 3. (**Top panels**) Areal averages of monthly mean distributions of near-surface maximum temperature (°C) for each climatological zone: (**left**) Guinea, (**middle**) Savannah, and (**right**) Sahel. (**Bottom panels**) Mean spatial distributions of near-surface maximum temperature (shaded; °C) over West Africa for the month of August: (**left**) w@h2 ensemble mean simulation, (**middle**) CRU-observation, and (**right**) ERA5-reanalysis.

3.2. Association

Association, a statistical measure of the strength of a linear relationship between a paired simulation and observation/reanalysis datasets, is evaluated here by the use of spatio-temporal Pearson's Product-Moment Correlation Coefficient (r) (Table 1). To a low extent, we also utilize the coefficient of determination (CoD) which is simply the square of r. CoD measures the level of skill attainable when the biases are eliminated.

The inter-annual variability of Savannah rainfall for August and Sahel near-surface maximum temperature for May are shown in Figure 4 (see Figure S5 for other months and

zones). The observed (CRU/ERA5) values generally fall within the spread, notably during the unusually wet August 1999 over the Savannah and hot May 1998, 2010, and 2016 over the Sahel. There are some cases though when observed values are outside the spread of the ensemble members, such as the cool May 1991 over the Sahel.

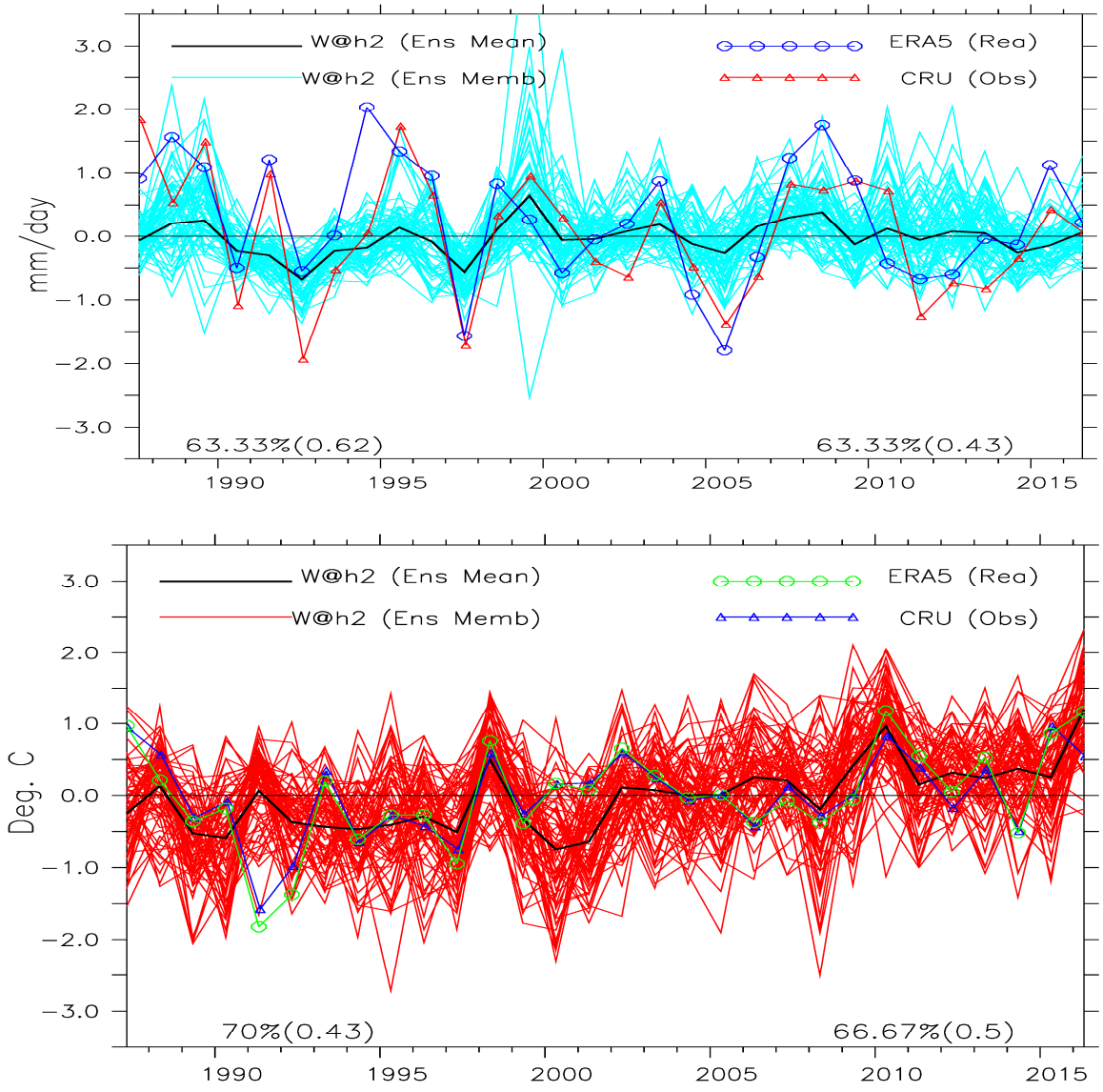

Figure 4. Areal averages of inter-annual variations of **top panel**—precipitation anomalies (mm day^{-1}) over Savannah in August and **bottom panel**—near-surface maximum air temperature anomalies (°C) Sahel in May. Values of synchronization (%) and the temporal correlation, r (in brackets), between the w@h2 ensemble mean precipitation and temperature and CRU (left) and ERA5 (right) are written at the bottom of each panel.

The linear relationship between the w@h2 model's temperature simulations and observations are strongly direct, while it is less strong for precipitation simulations. For correlations, *r*, values as large as 0.78 and 0.89 were evaluated for precipitation and temperature, respectively (Figure 5); however, cases of weak relationships with *r* as low as ≈ -0.4 are also present for individual simulations (Figures 5 and S5–S7). Cases of weak relationships are more noticeable in the inter-annual variabilities of monthly precipitation simulations than in temperature (Figures S5–S7). For both precipitation and temperature simulations, the strength of linear associations diminishes as we move to the drier north towards the Sahel (Figures 5 and S5–S8).

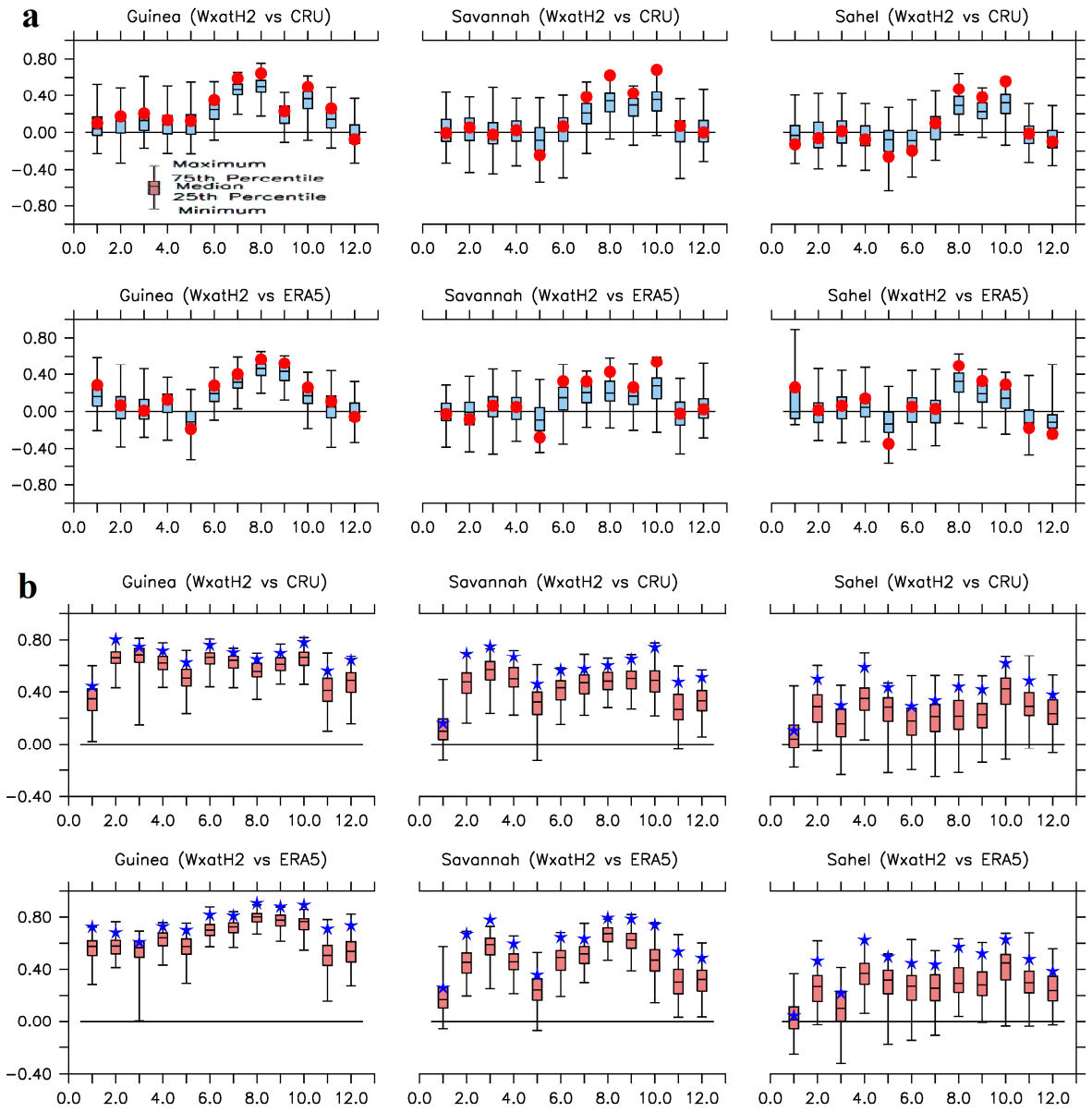

Figure 5. Monthly spreads of the correlation coefficient, r, between the w@h2 simulations and the observed and reanalyzed (**a**) area-averaged precipitation [ensemble mean (red circles) and ensemble members (box and whisker plots)], and (**b**) areal-averaged near-surface maximum air temperature [ensemble mean (blue stars) and ensemble members (box and whisker plots)] over the climatological zones. Months are on the horizontal axes, e.g., 2.0 represents February.

Irrespective of the magnitudes, the ability of the ensemble means of the w@h2 model to capture the anomaly sign of the observed precipitation and temperature is generally greater than 40%, and at most 90% (Figure S5a–f). In other words, the model's ensemble mean will adequately predict the sign of 2 out of 5 observations correctly, and will, at most, simulate about 9 out of 10 observations correctly (synchronization $\approx$ 90%).

The normalized standard deviations (NSD) of the majority of the ensemble members are greater than those of the ensemble means (Figures S6 and S7). This is because of the averaging that filters out the simulated variabilities of the ensemble means [60,69]. These imply that the discrepancies between the ensemble means and observations, CRU/ERA5, are smaller than the discrepancies between individual ensemble members and observations.

Furthermore, there are noticeable differences and similarities in the way w@h2 model's precipitation and near-surface maximum temperature simulations associate with obser-

vations (CRU/ERA5). Figure S8a–d shows that r between precipitation simulations and observations contain both direct and weak linear relationships, while cases of strong direct linear relationships dominate the r between the temperature simulations and observations. For instance, the correlations exhibited by the precipitation ensemble means are $-0.4 < r < 0.78$ while those of temperature are $0 < r < 0.8$ (Figure 5a,b). Some of the precipitation ensemble means and members exhibited weak linear relationship with observations on monthly basis, except in July and August for CRU, and July, August, and September for ERA5 over coastal Guinea (Figure 5a). This is, however, different for temperature simulations where all the ensemble means exhibited direct linear relationship, of various strength, with observations on monthly basis (Figure 5b). The best performance here is over Guinea where none of the temperature ensemble members had a negative linear relationship with observations, i.e., $0 < r < 1$.

Four similarities are typical to the associations of the w@h2 model's precipitation and temperature simulations with observations. Firstly, the CoD for both precipitation and temperature simulations are generally less than 0.5. Higher values, $0.5 < CoD < 0.8$, are recorded during the peaks of the monsoon seasons. This corroborates the values of r, and implies that the w@h2 model may also be skillful when biases are absent. Secondly, the spatio-temporal linear associations seem to strengthen with observations as rainfall seasons set in and stabilize. These are very obvious during the months of July, August, and September (Figure S8a–d). Thirdly, the strength of linear associations, for both precipitation and temperature simulations, diminishes as we move north towards the Sahel. And lastly, all values of the associations of the ensemble means are enveloped by the spreads of the ensemble members' associations (Figure 5). However, while the values of associations are generally greater than the 75th percentiles of the spreads in temperature simulations, they do not have any agreeably defined positions in precipitation simulations. The implications here are that the w@h2 model exhibits more significant associations during the peak of the West African monsoon seasons than the rest of the year. However, cautions are encouraged in terms of significant associations when applying the simulations over the Sahel.

Summarily for temperature, the ensemble mean always has a stronger correlation with observations than do most of the simulations; for precipitation, the rule seems to hold but maintain the sign of the correlation, i.e., a stronger anti-correlation when most simulations have negative r. The temperature's positive correlation may be attributed to the strong warming trend over the experimental period [74], while the weak correlations for precipitation may primarily reflect the inter-annual variability [75,76].

3.3. Skill

The ranked probability skill score (RPSS) is used here to evaluate the ability of the w@h2 model to reproduce the observed monthly inter-annual variations in precipitations and near-surface maximum temperature over West Africa (Table 1). RPSS measures the forecast accuracy with respect to a reference observation (e.g., observed climatology) as the scores reflect discrimination, reliability, and resolution.

Positive skills, $0 < RPSS < 1$, dominate Guinea and Savannah zones in all the months for both precipitation and temperature. However, the reverse is the case over Sahel in precipitation simulations (Figures 6 and S9). Nevertheless, all the values of RPSS from the ensemble means are within the spreads of the ensemble members' RPSS; though, the spreads are of diverse thickness, the broadest being exhibited over Guinea zone. The ensemble means of the w@h2 model, with reference to the two observations (CRU/ERA5), returned positive values of RPSS for precipitation over Guinea throughout the year and positive values of RPSS for temperature over all the climatological zones, also throughout the year (except in January with reference to ERA5 over Sahel: Figure 6). Generally, while

the skills of the w@h2 model with respect to precipitation simulations, over Sahel, may not be significantly impressive, the model may however have skills to detect heat waves that usually ravage West Africa during the boreal springs as well as skills to capture the LDS over Guinea zone.

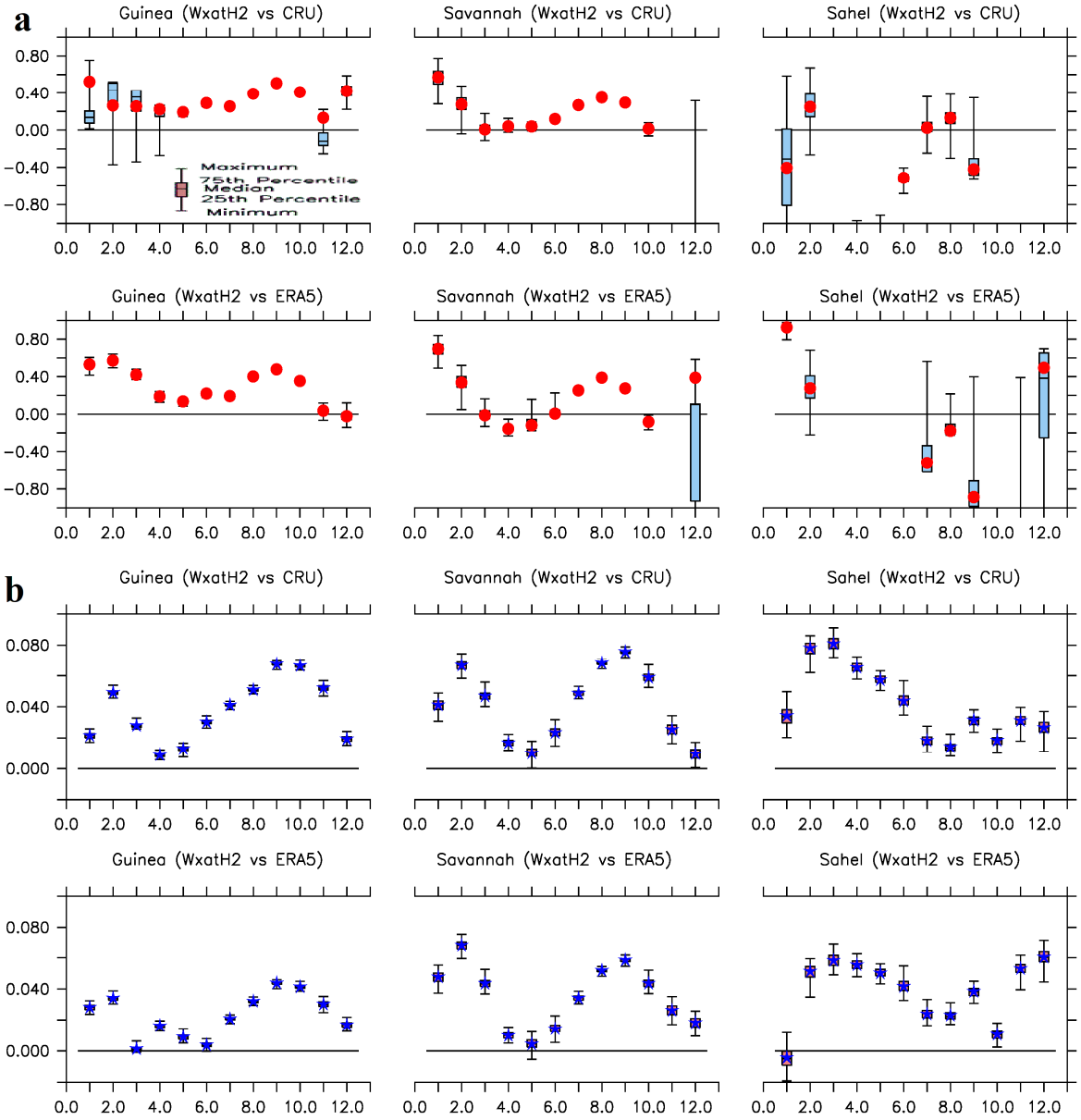

Figure 6. Monthly spreads of the ranked probability skill score (RPSS) for w@h2 simulations with respect to observed and reanalyzed (**a**) precipitation [ensemble mean (red circles) and ensemble members (box and whisker plots)], and (**b**) near-surface maximum air temperature [ensemble mean (blue stars) and ensemble members (box and whisker plots)] over the climatological zones. Months are on the horizontal axes, e.g., 2.0 represents February. Missing red circles, blue stars, and/or error bars (either in parts or wholly) indicate that RPSS < −1.0 for the month. Note the different vertical scales for the precipitation and temperature panels.

3.4. Accuracy

As suggested by [10], we utilized mean absolute error (MAE), root mean square error (RMSE), and synchronization as measures to estimate accuracy in this paper. As tabulated in Table 1, MAE is a measure of the average magnitude of largest error that can be expected from a forecast without considering their directions. It is a linear score, meaning that all

the individual differences are weighted equally in the average. Similar to MAE, RMSE also does not indicate the direction of the error, but it penalizes large errors. In contrast, synchronization shows how much a simulated value agrees with an observed value in the signs of their anomalies, without taking magnitudes into consideration.

The maximum average difference, as depicted by MAE, between the w@h2 model simulations and the observed (CRU/ERA5) precipitation over West Africa is about 5 mm day^{-1} (Figure S10a,b). The average differences grow in values as rainfall season is setting in. High values of MAE, like 3 to 5 mm day^{-1}, are more vivid between the months of March to October and are more present in the southern coast of Guinea. In line with annual characteristics of rainfall, these high values migrate northward in a rainfall-like pattern and annual oscillation. Interestingly, they start to retreat southward in August/September. The relatively low values of MAE from November to February do not imply higher accuracy in rainfall estimation by the w@h2 model than the other months (Figure S10a,b); these are months of relatively very low precipitation (Figures 2 and S1). The error magnitudes in precipitation do not represent up to 50% of over- or under-estimations in most parts of the sub-region. Therefore, the w@h2 model cannot be labeled as a biased estimator of rainfall. Nevertheless, as recommended by [77], we may need to apply caution when utilizing ±30% of rainfall estimations as an indicator of biasness and accuracy.

The maximum average difference between the w@h2 model simulations and the observed (CRU/ERA5) near-surface maximum temperature over West Africa is about 4 °C (Figure S10c,d). Generally, these differences are less than 2.8 °C. The higher values of MAE tend to occur in the early monsoon months of May, June, July, and August, predominantly, over northern Savannah and southern parts of Sahel. The lower differences, MAE, that dominate the larger spatial expanse of West Africa, in all months, presumably make the w@h2 model an accurate estimator of near-surface maximum temperature.

There are sharp differences in RMSE and MAE produced from the precipitation simulations of the w@h2 model. For example, the maximum RMSE in precipitation simulation is about 10 mm day^{-1} (Figure S11a,b), while the maximum MAE is about 5 mm day^{-1} (Figure S10a,b). This difference is, though, not large enough to indicate the presence of very large errors in the simulation; it, nevertheless, signals that the precipitation simulations have large variance in the individual errors of its samples due to the existence of extreme precipitation values (the outliers). It is the outliers that introduces the random errors, i.e., the variability and/or noise in the internal system of the precipitation simulations.

All errors are possibly of the same magnitudes in the near-surface maximum temperature simulated by the w@h2 model. This is because RMSE and MAE are almost equal in magnitudes (panels c and d of Figures S10 and S11). In similarity to MAE, RMSE is generally <2.8 °C. This implies that it is the bias errors that are predominant here, meaning that the deviations in the temperature simulations, from observations, are not due to chance alone. They are, rather, systemic in nature. The exception here is in October–December over Sahel (Figure S11d), where there are sharp disagreements between CRU and ERA5 datasets (RMSE > 4 °C). Investigating the causes of these disagreements is outside the scope of this work.

While corroborating the depictions on Figure S5a–f, Figures 7 and S12 show that the abilities of the w@h2 model to simulate the actual anomaly signs of the observed precipitation and near-surface maximum temperature correctly is generally between 20 and 80% for precipitation ensemble members, and between 25 and 95% for temperature ensemble members. These imply that any ensemble member of the w@h2 model, picked at random, will at worst/best simulate 1 out of 5 (synchronization ≈ 20%)/4 out of 5 (synchronization ≈ 80%) actual signs of the anomalies correctly for precipitation, while, also at random, they will simulate at least 1 out of 4 (synchronization ≈ 25%) actual signs

of the anomalies and at most more than 9 out of 10 (synchronization ≈ 95%) actual signs of the anomalies correctly for temperature. When combined, the model's ensemble means of precipitation and temperature synchronize between 40% and 90% (Figures 7 and S12). This shows that at worst they (i.e., the ensemble means) will simulate 2 out of 5 actual anomalies correctly, and will, at best, simulate 9 out of 10 actual signs of the anomalies correctly. Conclusively, while the w@h2 model may not be accurate in terms of getting the magnitudes of climate parameters right due to inherent presence of biases of different types, it may however be accurate in simulating the actual anomaly signs of the climate parameters rightly, to some significant extent. This is good because the ability to reliably/accurately simulate the actual anomaly signs of observed climate parameters is one of the special attributes of a model that is needed for seasonal climate predictions and applications.

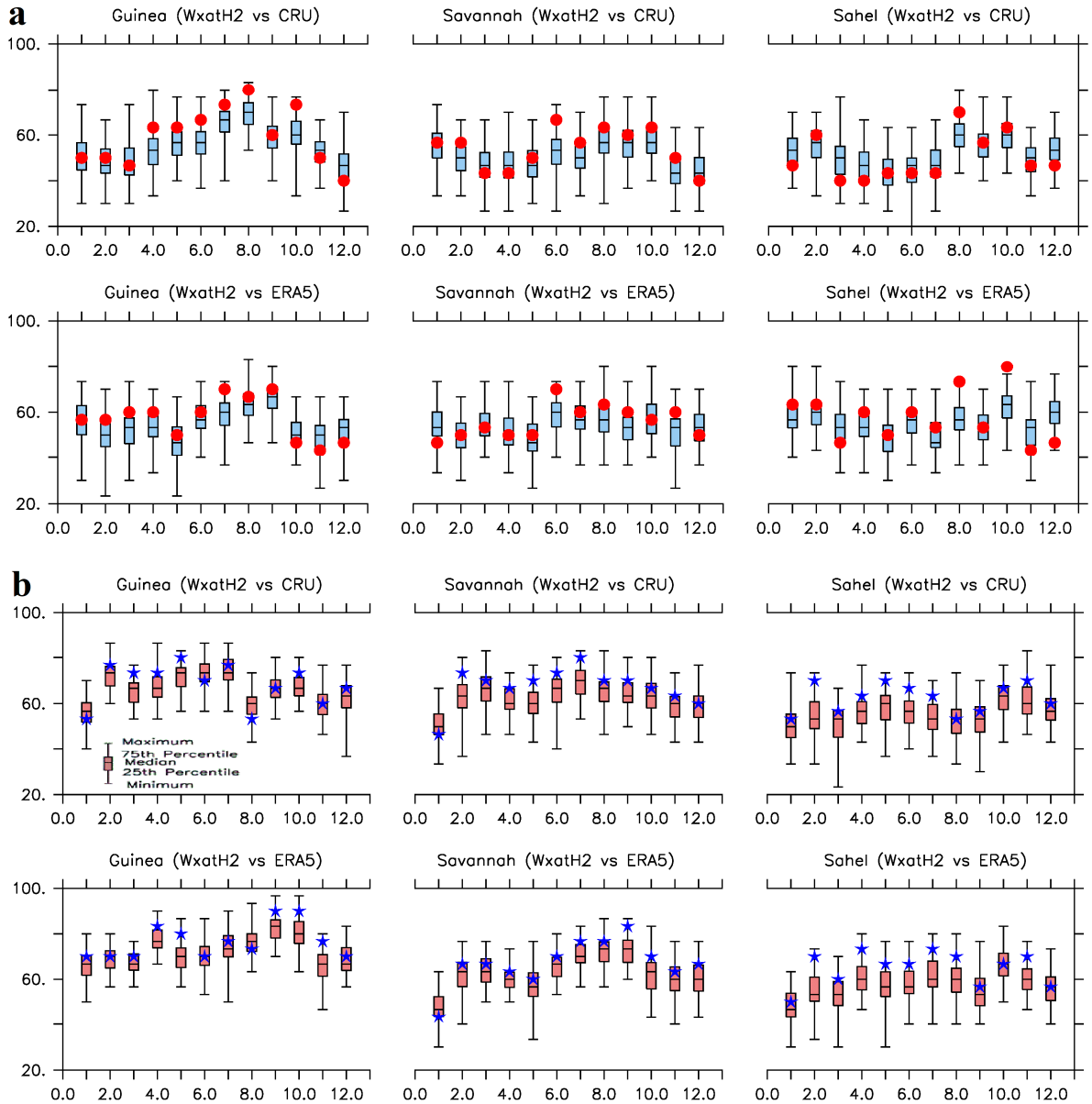

Figure 7. Monthly spreads of synchronization (%) of w@h2 simulations with the observed and reanalyzed (**a**) precipitation [ensemble mean (red circles) and ensemble members (box and whisker plots)], and (**b**) near-surface maximum air temperature [ensemble mean (blue stars) and ensemble members (box and whisker plots)] over the climatological zones. Months are on the horizontal axes, e.g., 2.0 represents February.

3.5. Precision

As shown in Table 1, precision is here evaluated with the use of the coefficient of variation (CoV) and the normalized standard deviations (NSD). CoV, the ratio of standard deviations as measures of spreads to the mean of the sample populations, is used to determine the degrees of variability within the simulated and the observed climate parameters. Specifically, we employed the bias produced by CoV (i.e., CoV_{model} minus $CoV_{observation}$) to really know which of the two (simulations or observations) produces more spatio-temporal variabilities. NSD is used to measure the deviation factors between the simulations and the observations.

The degrees of spatio-temporal variability within the simulated and the observed climate parameters (precipitation and near-surface maximum temperature) are depicted in Figure S13. The w@h2 model produces largely lesser spatio-temporal variabilities in precipitation in comparison to observations and almost normal deviations in spatio-temporal variabilities for temperature simulations. The largely sub-zero ($\approx -50\%$) spatio-temporal variabilities are clearly evident in precipitation simulations as depicted by the biases of CoV (Figure S13a,b) over Savannah and Sahel zones, except during the months of July-August (peaks of raining season) when the CoV biases do not significantly deviate from zero ($\pm 10\%$), meaning that, during the monsoon seasons, the degrees of spatio-temporal variabilities around the mean is almost the same for precipitation simulations and observations.

Biases of CoV in temperature simulations range between $\pm 1\%$ (Figure S13c,d). Exceptions here are during the dry months of December–March, when largely negative CoV biases are visible over Savannah and Sahel zones. The implications of CoV biases, here, are that the degrees of spatio-temporal variabilities around the mean are almost the same for both temperature simulations and their observations during the wet seasons. Therefore, generally speaking, the w@h2 model simulations perform precisely well during monsoon seasons in terms of simulating precipitation and near-surface maximum temperature. Summarily, on the average, the inter-annual variabilities of simulated w@h2 precipitation and near-surface maximum temperature do not significantly exceed those of observations.

The discrepancies, as depicted by NSD, are majorly less than a factor of 1.0. As depicted in Figure 8 and Figures S6 and S7, the discrepancies between the ensemble means and observations (i.e., CRU/ERA5) are smaller than the discrepancies between individual ensemble members and observations. For precipitation simulations, NSD is growing larger (>1) as we move northwards towards the Sahel zone (Figure 8a). Meanwhile, NSD tends to a factor of 1.0 as the monsoon seasons are approached for temperature simulations (Figure 8b). This shows that there are little or negligible deviations between simulated and observed temperatures during the monsoon seasons. Generally, the majority of the ensemble means' NSD values are outside the spreads of the ensemble members' NSD; specifically, below the first percentiles (the minimum on the error bars). These behaviors on the path of the ensemble means confirm the "precise nature" of the ensemble means over the members as already documented, e.g., [78–83].

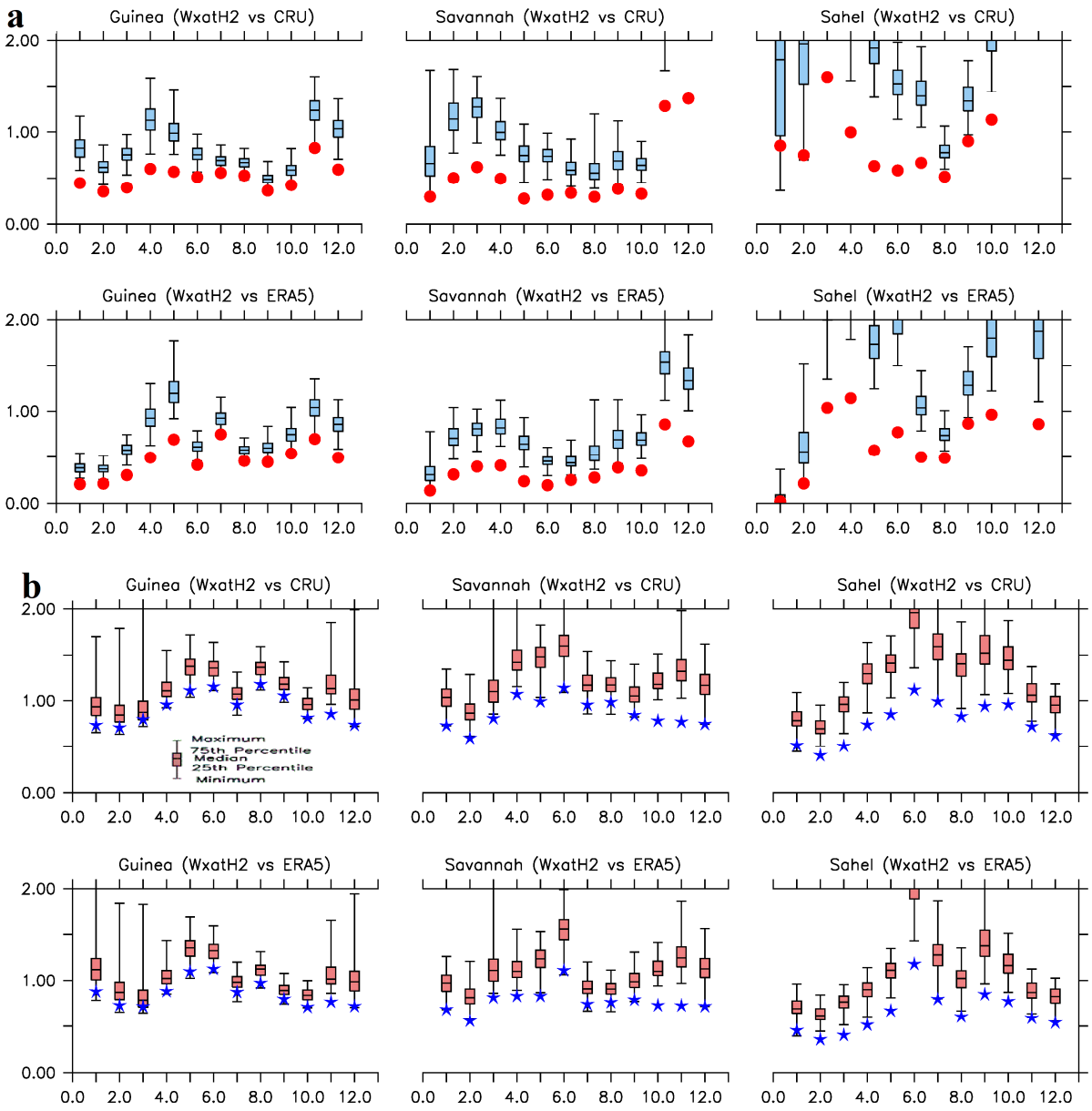

Figure 8. Monthly spreads of normalized standard deviations (NSD) of w@h2 simulations with respect to observed and reanalyzed (**a**) precipitation [ensemble mean (red circles) and ensemble members (box and whisker plots)], and (**b**) near-surface maximum air temperature [ensemble mean (blue stars) and ensemble members (box and whisker plots)] over the climatological zones. Months are on the horizontal axes, e.g., 2.0 represents February. Missing red circles, blue stars, and/or error bars (either in parts or wholly) indicate that NSD > 2.0 for the month.

4. Discussion

This study is motivated by the generation of a remarkably huge ensemble of simulations, from which a sub-set of 71 ensemble members per year were utilized. The achievement allows denser sampling of the climate distributions, allowing more precise calculation of climate model properties [84]. We seek to provide a performance evaluation of the w@h2 simulations over West Africa using the framework of RASAP (*reliability*, *association*, *skill*, *accuracy*, and *precision*) measures.

The results show that, to a moderately significant extent, w@h2 model provides little, if any, predictive information for precipitation during the dry season, but may provide useful information during the monsoon seasons. This means that a prospective user gets to

decide whether it is "useful" for his/her particular application. This evaluation provides a prospective user with information that he/she can use to decide whether the model might be "useful or not". For instance, a prospective user may ignore rainfall simulations during the dry seasons (when there are no activities to predict), but consider it for the wet seasons.

Contrary to the results for precipitation, the w@h2 model provides sufficient predictive information for the near-surface maximum temperature over West Africa throughout the year. For example, the model is able to reproduce all the annual characteristics of maximum air temperature, such as (1) the two peaks of maximum air temperature over all the climatic zones; (2) the Sahel being the warmest of all the zones, except during the boreal winters; (3) the dip in the annual maximum temperatures over all climatic zones during the peak of the rainy season; and (4) the annual north–south oscillation of the thermal depression.

The analyses carried out in this paper have provided some statistical insights to the nature of the w@h2 simulations over West African region. The w@h2 modeling system was designed for the investigation of the behavior of extreme climate under anthropogenic climate change, i.e., event attribution. For event attribution, as earlier stated, measures of the performance of a model in terms of climate variability may be more relevant than measures of the mean climatology [23,29–31]. In addition, ref. [28] opines that a single-model initial-condition large ensemble (SMILE), if less than 30 members, may analytically underestimate precipitation variability, vis-à-vis distribution. Therefore, the w@h2 model is unique in being able to produce large sample sizes that are able to show that sampling quality of the tails of the distribution is no longer the primary constraint/source of uncertainty.

Refs. [29,31] point out that if the unforced variability of a model, in comparison to observation, is too small/large then the model will be too keen/not keen enough to attribute an event's occurrence to emissions. Here, the seemingly substantial bias and low variability in its precipitation and temperature simulations present the model as too keen to attribute an event to emissions. In addition, high skills, especially during the monsoon seasons, probably mean that there is a lot of predictability in the system. Therefore, for an SST-forced system like w@h2, this means that event attribution conclusions are conditional on the occurrence of the observed SST state [85].

Furthermore, the lack of obvious quality of the w@h2 model in terms of rainfall simulations during the dry seasons may not mean that it has a bias for event attribution analysis. It may only mean that there is no evidence that strongly supports the notion that w@h2 is accurately simulating the appropriate processes for extremes. But, on the contrary, predictive skills for the onset season suggest that w@h2 may be getting processes right.

Overall, the performance achieved in terms of RASAP suggests that the w@h2 model could be suitable for forecasting in various socio-economic sectors, e.g., agriculture, water resource management, health, etc. In addition, the model's ability to robustly simulate surface maximum temperatures is particularly important for adaptive ecosystem management and the protection of species sensitive to changing climate. With this moderate performance, w@h2 thus joins the growing population of models that may be utilized to support science and decision-making [9]. However, the model's deficiencies (in terms of the underestimation of rainfall in the Guinea region and the overestimation in the Sahel) are noted. These deficiencies, in terms of biases, in the quantitative reproducibility of the temperature and precipitation point to the fact that the w@h2 model does provide a reliability akin to that of the regional climate models [86]. Nevertheless, this calls for further calibration and refinement of the model to better represent extreme climate conditions; but, investigating the reasons for the model's deficiencies is beyond the scope of this work. The investigation shall be attended to in the second part of this work, where we intend to consider the model's reproducibility of atmospheric dynamics that influence and modulate

West African weather and climate. Then, we will fully be able to say if the model is doing a reasonable job of capturing processes over West Africa.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/atmos16040392/s1: **Figure S1.** Monthly mean spatial distributions of rainfall (shaded; mm day^{-1}) over West Africa for (a) w@h2 ensemble mean simulation, (b) CRU-observation, and (c) ERA5-reanalysis. The values of the spatial correlation, r, between rainfall distribution of the ensemble mean of w@h2 simulation and CRU (black texts) and ERA5 (red texts) are written at the bottom of each sub-panel in panel **a**. Stippling in August of all panels indicate areas, over West Africa, that usually experience the little dry season (LDS) in August; **Figure S2**. Scatter plots of the monthly cycles of rainfall distributions over the climatological zones of West Africa: (a) Guinea, (b) Savannah, and (c) Sahel, where observations (CRU/ERA5) are expected to align on the 45° diagonal lines in the case of a perfect simulation, and the monthly mean spatial distributions of rainfall bias (shaded; mm day^{-1}) for (d) w@h2 minus CRU observation and e) w@h2 minus ERA5 reanalysis over West Africa; **Figure S3.** Monthly mean spatial distributions of near-surface maximum temperature (shaded; °C) over West Africa for (a) w@h2 ensemble mean simulation, (b) CRU-observation, and (c) ERA5-reanalysis. The values of the spatial correlation, r, between the temperature distributions of the w@h2 ensemble mean simulation and CRU (black texts) and ERA5 (red texts) are written at the bottom of each sub-panel in panel a. **Figure S4.** Scatter plots of the monthly cycles of near-surface maximum temperature distributions over the climatological zones of West Africa: (a) Guinea, (b) Savannah, and (c) Sahel, where observations (CRU/ERA5) are expected to align on the 45° diagonal lines in the case of a perfect simulation, and the monthly mean distributions of temperature bias (shaded; °C) for (d) w@h2 minus CRU observation and (e) w@h2 minus ERA5 reanalysis over West Africa. **Figure S5.** Areal averages of inter-annual variations of monthly: top row—precipitation anomalies (mm day^{-1}) over (a) Guinea, (b) Savannah, and (c) Sahel; and bottom row—near-surface maximum air temperature anomalies (°C) over (d) Guinea, (e) Savannah, and (f) Sahel. Values of synchronization (%) and the temporal correlation, r (in brackets), between the w@h2 ensemble mean precipitation and temperature and CRU (left) and ERA5 (right) are written at the bottom of each sub-panel. **Figure S6.** Taylor diagrams, of Figure S5a–c, showing the normalized standard deviations (NSD: horizontal axes) and the correlation coefficients, r (curved axes), between individual precipitation simulations (w@h2 ensemble mean: red circles, and ensemble members: blue circles) and observations—top row: CRU (black semi-circle) over (a) Guinea, (b) Savannah, and (c) Sahel; and bottom row: ERA5 (black semi-circle) over d) Guinea, (e) Savannah, and (f) Sahel. Triangles represent correlations that are less than zero. NSD of November and December over Sahel are to be multiplied by factors of 3 and 2 for CRU and ERA5, respectively, to get their true values. **Figure S7.** Taylor diagrams, of Figure S5d–e, showing the normalized standard deviations (NSD: horizontal axes) and the correlation coefficients, r (curved axes), between individual near-surface maximum temperature simulations (w@h2 ensemble mean: red circles, and ensemble members: blue circles) and observations—top row: CRU (black semi-circle) over (a) Guinea, (b) Savannah, and (c) Sahel; and bottom row: ERA5 (black semi-circle) over (d) Guinea, (e) Savannah, and (f) Sahel. Triangles represent correlations that are less than zero. **Figure S8.** Monthly spatial distributions of correlation coefficients, r, between w@h2 ensemble means of (top panels) precipitation and (bottom panels) near-surface maximum temperature and observations—(a,c) CRU, and b, d) ERA5 over West Africa. Stippling in all the panels indicate areas with significant correlations at the 2-sided 10% significance level. **Figure S9.** Monthly spatial distributions of ranked probability skill score (RPSS) for w@h2 ensemble means of (top panels) precipitation and (bottom panels) near-surface maximum temperature and observations—(a,c) CRU, and (b,d) ERA5 over West Africa. Note the different scale for temperature and precipitation panels. **Figure S10.** Monthly spatial distributions of mean absolute errors (MAE) for w@h2 ensemble means of (top row) rainfall (mm day^{-1}) and (bottom row) near-surface maximum temperature (°C) simulations with respect to (a,c) CRU (w@h2 minus CRU observations) and (b,d) ERA5 (w@h2 minus ERA5 reanalysis) over West Africa. **Figure S11.** Monthly distributions of root mean square errors (RMSE) for w@h2 ensemble means of (top row) rainfall (mm day^{-1}) and (bottom row) near-surface maximum temperature (°C) simulations with respect

to (a,c) CRU observations and (b,d) ERA5 reanalysis over West Africa. **Figure S12.** Monthly spatial distributions of synchronization (%) for w@h2 ensemble means of (top row) rainfall and (bottom row) near-surface maximum temperature simulations with respect to (a,c) CRU-observations and (b,d) ERA5-reanalysis over West Africa. **Figure S13.** Monthly spatial distributions of the bias of coefficient of variation (CoV) ($CoV_{w@h2}$ minus $CoV_{CRU/ERA5}$) for (top row) rainfall and (bottom row) near-surface maximum temperature simulations with respect to (a,c) CRU-observations and (b,d) ERA5-reanalysis over West Africa.

Author Contributions: All the authors listed have approved this work for publication, having made a substantial direct and intellectual contribution to this work in the order of activities listed against their names, as follows: K.A.L. conceptualization, investigation, methodology, validation, resources, writing (original draft, review, and editing), visualization, software, formal analysis, data processing, and supervision. O.M.A. investigation, writing (editing), visualization, and data processing. E.O. methodology, writing (review), software, formal analysis, visualization, and data processing. A.B. and S.N.S. resources, validation, writing (review and editing), and project administration. M.F.W. and D.A.S. conceptualization, investigation, methodology, validation, resources, writing (review and editing), supervision, and project administration. All authors have read and agreed to the published version of the manuscript.

Funding: K.A.L. was supported by the Intra-ACP Climate Services and Related Applications (ClimSA: https://www.climsa.org/) Program in Africa, an initiative funded by the 11th European Union's Development Fund (Contribution Agreement 2019/410-300) and implemented by the African Centre of Meteorological Applications for Development through a grant with the African Union Commission as the contracting authority (Contract ACP/FED/038-833), and the BNP Attribution Project of the African Climate and Development Initiative (ACDI: www.acdi.uct.ac.za) of the University of Cape Town, South Africa. M.F.W. was supported by the Director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy as part of the Regional and Global Model Analysis program under Contract No. DE340AC02-05CH11231. D.A.S. was supported by the Whakahura project, funded through the Endeavour Program of the Ministry of Business, Innovation, and Employment of Aoteaora, New Zealand.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request, without undue reservation.

Acknowledgments: The authors gratefully acknowledge the computational, technical, and infrastructural supports provided by the Climate Sciences Analysis Group (CSAG: www.csag.uct.ac.za), University of Cape Town, South Africa. We also thank all the reviewers whose comments helped to improve the quality of this manuscript.

Conflicts of Interest: Author Dáithí A. Stone is employed by the National Institute of Water & Atmospheric Research Ltd. (NIWA). The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Mariotti, L.; Coppola, E.; Sylla, M.B.; Giorgi, F.; Piani, C. Regional climate model simulation of projected 21st century climate change over an all-Africa domain: Comparison analysis of nested and driving model results. *J. Geophys. Res.* **2011**, *116*, D15111. [CrossRef]
2. Nikulin, G.; Jones, C.; Giorgi, F.; Asrar, G.; Büchner, M.; Cerezo-Mota, R.; Christensen, O.B.; Déqué, M.; Fernandez, J.; Hänsler, A.; et al. Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations. *J. Clim.* **2012**, *25*, 6057–6078. [CrossRef]
3. Diallo, I.; Sylla, M.B.; Camara, M.; Gaye, A.T. Inter-annual variability of rainfall over the Sahel based on multiple regional climate models simulations. *Theor. Appl. Climatol.* **2013**, *113*, 351–362. [CrossRef]

4. Klein, C.; Heinzeller, D.; Bliefernicht, J.; Kunstman, H. Variability of West African monsoon patterns generated by a WRF Multiphysics ensemble. *Clim. Dyn.* **2015**, *45*, 2733–2755. [CrossRef]
5. Sylla, M.B.; Giorgi, F.; Pal, J.S.; Gibba, P.; Kebe, I.; Nikiema, M. Projected changes in the annual cycle of high intensity precipitation events over West Africa for the late 21st century. *J. Clim.* **2015**, *28*, 6475–6488. [CrossRef]
6. Tall, A.; Mason, S.J.; Van Aalst, M.; Suarez, P.; Ait-Chellouche, Y.; Diallo, A.A.; Braman, L. Using seasonal climate forecasts to guide disaster management: The Red Cross experience during the 2008 West Africa floods. *Int. J. Geophys.* **2012**, *2012*, 986016. [CrossRef]
7. Niang, I.; Ruppel, O.; Abdrabo, M.; Essel, A.; Lennard, C.; Padgham, J.; Urquhart, P. Africa. In *Climate Change 2014: Impacts, Adaptation and Vulnerability—Contributions of the Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2014; pp. 1199–1265.
8. Nkiaka, E.; Taylor, A.; Dougill, A.; Antwi-Agyei, P.; Fournier, N.; Bosire, E.; Konte, O.; Lawal, K.A.; Mutai, B.; Mwangi, E.; et al. Identifying user needs for weather and climate services to enhance resilience to climate shocks in sub-Saharan Africa. *Environ. Res. Lett.* **2019**, *14*, 123003. [CrossRef]
9. Lundstad, E.; Brugnara, Y.; Pappert, D.; Kopp, J.; Samakinwa, E.; Hürzeler, A.; Andersson, A.; Chimani, B.; Cornes, R.; Demarée, G.; et al. The global historical climate database HCLIM. *Sci. Data* **2023**, *10*, 44. [CrossRef]
10. Storch, H.V.; Zwiers, F.W. *Statistical Analysis in Climate Research*; Cambridge University Press: London, UK, 2003.
11. Walther, B.A.; Moore, J.L. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography* **2005**, *28*, 815–829. [CrossRef]
12. Ebert, E.; Wilson, L.; Weigel, A.; Mittermaier, M.; Nurmi, P.; Gill, P.; Gober, M.; Joslyn, S.; Brown, B.; Fowlerh, T.; et al. Progress and challenges in forecast verification. *Meteorol. Appl.* **2013**, *20*, 130–139. [CrossRef]
13. Wilson, L.J.; Giles, A. A new index for the verification of accuracy and timeliness of weather warnings. *Meteorol. Appl.* **2013**, *20*, 206–216. [CrossRef]
14. Mason, S.J. On using "climatology" as a reference strategy in the brier and ranked probability skill scores. *Mon. Weather Rev.* **2004**, *132*, 1891–1895.
15. Murphy, A.H. Skill score on the mean square error and their relationship to the correlation coefficient. *Mon. Weather Rev.* **1988**, *116*, 2417–2424.
16. Murphy, A.H. The coefficients of correlation and determination as measures of performance in forecast verification. *Weather Forecast.* **1995**, *10*, 681–688.
17. Wilks, D.S. *Statistical Methods in Atmospheric Sciences: An Introduction*, 2nd ed.; Academic Press: San Diego, CA, USA, 1995.
18. Weigel, A.P.; Liniger, M.A.; Appenzeller, C. The discrete brier and ranked probability skill scores. *Mon. Weather Rev.* **2006**, *135*, 118–124.
19. Kim, G.; Ahn, J.-B.; Kryjov, V.N.; Sohn, S.-J.; Yun, W.-T.; Graham, R.; Kolli, R.K.; Kumar, A.; Ceron, J.-P. Global and regional skill of the seasonal predictions by WMO Lead Centre for Long-Range Forecast Multi-Model Ensemble. *Int. J. Climatol.* **2016**, *36*, 1657–1675. [CrossRef]
20. Debanne, S.M. The planning of clinical studies: Bias and precision. *Gastrointest. Endosc.* **2000**, *52*, 821–822.
21. West, M.J. Stereological methods for estimating the total number of neurons and synapses: Issues of precision and bias. *Trends Neurosci.* **1999**, *22*, 51–61. [PubMed]
22. Xue, Y.; De Sales, F.; Lau, K.M.W.; Bonne, A.; Feng, J.; Dirmeyer, P.; Guo, Z.; Kim, K.M.; Kitoh, A.; Kumar, V.; et al. Intercomparison of West African Monsoon and its variability in the West African Monsoon Modelling Evaluation Project (WAMME) first model inter-comparison experiment. *Clim. Dyn.* **2010**, *35*, 3–27. [CrossRef]
23. Verdin, A.; Funk, C.; Peterson, P.; Landsfeld, M.; Tuholske, C.; Grace, K. Development and validation of the CHIRTS-daily quasi-global high-resolution daily temperature data set. *Sci. Data* **2020**, *7*, 303. [CrossRef]
24. Zaroug, M.A.H.; Sylla, M.B.; Giorgi, F.; Eltahir, E.A.B.; Aggarwal, P.K. A sensitivity study on the role of the Swamps of Southern Sudan in the summer climate of North Africa using a regional climate model. *Theor. Appl. Climatol.* **2013**, *113*, 63–81. [CrossRef]
25. Diallo, I.; Bain, C.L.; Gaye, A.T.; Moufouma-Okia, W.; Niang, C.; Dieng, M.D.B.; Graham, R. Simulation of the West African monsoon onset using the HadGEM3-RA regional climate model. *Clim. Dyn.* **2014**, *45*, 575–594. [CrossRef]
26. Massey, N.; Jones, R.; Otto, F.E.L.; Aina, T.; Wilson, S.; Murphy, J.M.; Hassell, D.; Yamazaki, Y.H.; Allen, M.R. weather@home—Development and validation of a very large ensemble modelling system for probabilistic event attribution. *Q. J. Roy. Meteor. Soc.* **2015**, *141*, 1528–1545. [CrossRef]
27. Guillod, B.P.; Jones, R.G.; Bowery, A.; Haustein, K.; Massey, N.R.; Mitchell, D.M.; Otto, F.E.L.; Sparrow, S.N.; Uhe, P.; Wallom, D.C.H.; et al. weather@home 2: Validation of an improved global–regional climate modelling system. *Geosci. Model Dev.* **2017**, *10*, 1849–1872. [CrossRef]
28. Wood, R.R.; Lehner, F.; Pendergrass, A.G.; Schlunegger, S. Changes in precipitation variability across time scales in multiple global climate model large ensembles. *Environ. Res. Lett.* **2021**, *16*, 084022. [CrossRef]

29. Bellprat, O.; Doblas-Reyes, F. Attribution of extreme weather and climate events overestimated by unreliable climate simulations. *Geophys. Res. Lett.* **2016**, *43*, 2158–2164. [CrossRef]
30. Lott, F.C.; Stott, P.A. Evaluating Simulated Fraction of Attributable Risk Using Climate Observations. *J. Clim.* **2016**, *29*, 4565–4575. [CrossRef]
31. Bellprat, O.; Guemas, V.; Doblas-Reyes, F.; Donat, M.G. Towards reliable extreme weather and climate event attribution. *Nat. Commun.* **2019**, *10*, 1732. [CrossRef]
32. Nicholson, S.E.; Palao, I.M. A Re-evaluation of rainfall variability in the Sahel. Part I. Characteristics of rainfall fluctuations. *Int. J. Climatol.* **1993**, *13*, 371–389.
33. Nicholson, S.E. Sahel, West Africa. *Encycl. Environ. Biol.* **1995**, *3*, 261–275.
34. Omotosho, J.B.; Abiodun, B.J. A numerical study of moisture build-up and rainfall over West Africa. *Meteorol. Appl.* **2007**, *14*, 209–225. [CrossRef]
35. Omotosho, J.B. Pre-rainy season moisture build-up and storm precipitation delivery in the West Africa Sahel. *Int. J. Climatol.* **2007**, *28*, 937–946. [CrossRef]
36. Nicholson, S.E. An overview of African rainfall fluctuations of the last decade. *J. Clim.* **1993**, *6*, 1463–1466.
37. Latif, M.; Grotzner, A. On the equatorial Atlantic oscillation and its response to ENSO. *Clim. Dyn.* **2000**, *16*, 213–218.
38. Camberlin, P.; Janicot, S.; Poccard, I. Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature: Atlantic vs. ENSO. *Int. J. Climatol.* **2001**, *21*, 973–1005. [CrossRef]
39. Newman, M.; Sardeshmukh, P.D.; Winkler, C.R.; Whitaker, J.S. A study of sub-seasonal predictability. *Mon. Weather Rev.* **2003**, *131*, 1715–1732. [CrossRef]
40. Odekunle, T.O.; Eludoyin, A.O. Sea surface temperature patterns in the Gulf of Guinea: Their implications for the spatio-temporal variability of precipitation in West Africa. *Int. J. Climatol.* **2008**, *28*, 1507–1517. [CrossRef]
41. Diedhiou, A.; Janicot, S.; Viltard, A.; de Felice, P. Evidence of two regimes of easterly waves over West Africa and the tropical Atlantic. *Geophys. Res. Lett.* **1998**, *25*, 2805–2808.
42. Grist, J.P.; Nicholson, S.E. Easterly waves over Africa. Part II: Observed and modeled contrasts between wet and dry years. *Mon. Weather Rev.* **2001**, *130*, 212–225. [CrossRef]
43. Afiesimama, E.A. Annual cycle of the mid-tropospheric easterly jet over West Africa. *Theor. Appl. Climatol.* **2007**, *90*, 103–111. [CrossRef]
44. Parker, D.J.; Thorncroft, C.D.; Burton, R.; Diongue-Niang, A. Analysis of the African easterly jet, using aircraft observations from the JET 2000 experiment. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 1461–1482. [CrossRef]
45. Lavaysse, C.; Diedhiou, A.; Laurent, H.; Lebel, T. African easterly waves and convective activity in wet and dry sequences of the West African monsoon. *Clim. Dyn.* **2006**, *27*, 319–332. [CrossRef]
46. Lavaysse, C.; Flamant, C.; Janicot, S.; Parker, D.J.; Lafore, J.-P.; Sultan, B. Seasonal evolution of the West African heat low: A climatological perspective. *Clim. Dyn.* **2009**, *33*, 313–330. [CrossRef]
47. Lavaysse, C.; Flamant, C.; Janicot, S.; Knippertz, P. Links between African easterly waves, midlatitude circulation and intraseasonal pulsations of the West African heat low. *Q. J. R. Meteorol. Soc.* **2010**, *136*, 141–158. [CrossRef]
48. Nicholson, S.E.; Grist, J.P. The seasonal evolution of the atmospheric circulation over West Africa and equatorial Africa. *J. Clim.* **2003**, *16*, 1013–1030.
49. Redelsperger, J.-L.; Thorncroft, C.D.; Diedhiou, A.; Lebel, T.; Paker, D.J.; Polcher, J. African Monsoon Multidisciplinary Analysis: An international research project and field Campaign. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 1739–1746.
50. CRU TS Version 4.03. Available online: https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.03/ (accessed on 31 January 2025).
51. New, M.; Hulme, M.; Jones, P.D. Representing twentieth century space-time climate variability. Part 2: Development of 1901-96 monthly grids of terrestrial surface climate. *J. Clim.* **2000**, *13*, 2217–2238. [CrossRef]
52. Harris, I.; Osborn, T.J.; Jones, P.; Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* **2020**, *7*, 109. [CrossRef]
53. European Centre for Medium-Range Weather Forecasts (ECMWF—ERA Version 5 (ERA5). Available online: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5 (accessed on 31 January 2025).
54. Hersbach, H.; de Rosnay, P.; Bell, B.; Schepers, D.; Simmons, A.; Soci, C.; Abdalla, S.; Balmaseda, M.A.; Balsamo, G.; Bechtold, P.; et al. *Operational Global Reanalysis: Progress, Future Directions and Synergies with NWP*; ERA Report Series, No. 27; ECMWF: London, UK, 2018.
55. Steinkopf, J.; Engelbrecht, F. Verification of ERA5 and ERA-Interim precipitation over Africa at intra-annual and interannual timescales. *Atmos. Res.* **2022**, *280*, 106427. [CrossRef]
56. CPDN. Available online: https://www.climateprediction.net (accessed on 31 January 2025).

57. Jones, R.G.; Noguer, M.; Hassell, D.C.; Hudson, D.; Willson, S.S.; Genkins, G.J.; Mitchell, J.F.B. *Generating High Resolution Climate Change Scenarios Using PRECIS*; Technical report; Met Office Hadley Centre: Exeter, UK, 2004; 40p.
58. Essery, R.; Clark, D.B. Developments in the MOSES 2 land surface model for PILPS 2e. *Glob. Planet. Change* **2003**, *38*, 161–164. [CrossRef]
59. Donlon, C.J.; Martin, M.; Stark, J.; Roberts-Jones, J.; Fiedler, E.; Wimmer, W. The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.* **2012**, *116*, 140–158. [CrossRef]
60. Anderson, D.P. Boinc: A system for public-resource computing and storage. In Proceedings of the Fifth IEEE/ACM International Workshop on Grid Computing, Pittsburgh, PA, USA, 8 November 2004; IEEE: Piscataway, NJ, USA, 2004; pp. 4–10.
61. Black, M.T.; Karoly, D.J.; Rosier, S.M.; Dean, S.M.; King, A.D.; Massey, N.R.; Sparrow, S.N.; Bowery, A.; Wallom, D.; Jones, R.G.; et al. The weather@home regional climate modelling project for Australia and New Zealand. *Geosci. Model Dev.* **2016**, *9*, 3161–3176. [CrossRef]
62. Lawal, K.A.; Abiodun, B.J.; Stone, D.A.; Olaniyan, E.; Wehner, M.F. Capability of CAM5.1 in simulating maximum air temperature anomaly patterns over West Africa during boreal spring. *Model. Earth Syst. Environ.* **2019**, *5*, 1815–1838. [CrossRef]
63. Mason, S.J. Understanding forecast verification statistics. *Meteorol. Appl.* **2008**, *15*, 31–40. [CrossRef]
64. Hidore, J.J.; Oliver, J.E.; Snow, M.; Snow, R. *Climatology: An Atmospheric Science*, 3rd ed.; Pearson: San Francisco, CA, USA, 2009; ISBN-10: 0321602056/ISBN-13: 978-0321602053.
65. Jolliffe, I.T.; Stephenson, D.B. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 1st ed.; John Wiley & Sons, Ltd.: Berlin/Heidelberg, Germany, 2012; Print ISBN 9780470660713/Online ISBN 9781119960003. [CrossRef]
66. Pledger, S. Unified maximum likelihood estimates for cloud capture-recapture models using mixtures. *Biometrics* **2000**, *56*, 434–442. [CrossRef]
67. Pledger, S.; Schwarz, C.J. Modelling heterogeneity of survival in band-recovery data using mixtures. *J. Appl. Stat.* **2002**, *29*, 315–327. [CrossRef]
68. Rosenberg, D.K.; Everton, W.S.; Anthony, R.G. Estimation of animal abundance when capture probabilities are low and heterogenous. *J. Wildl. Manag.* **1995**, *59*, 252–261. [CrossRef]
69. Zelmer, D.A.; Esch, G.W. Robust estimation of parasite component community richness. *J. Parasitol.* **1999**, *85*, 592–594.
70. Misra, J. Phase synchronization. *Inf. Process. Lett.* **1991**, *38*, 101–105. [CrossRef]
71. Lawal, K.A. Understanding the Variability and Predictability of Seasonal Climates over West and Southern Africa Using Climate Models. Ph.D. Thesis, Faculty of Sciences, University of Cape Town, Cape Town, South Africa, 2015. Available online: https://open.uct.ac.za/handle/11427/16556?show=full (accessed on 31 January 2025).
72. Brose, U.; Martinez, N.D.; Williams, R.J. Estimating species richness: Sensitivity to sample coverage and insensitivity to spatial patterns. *Ecology* **2003**, *84*, 2364–2377. [CrossRef]
73. Melo, A.S.; Pereira, R.A.S.; Santos, A.J.; Shepherd, G.J.; Machado, G.; Medeiros, H.F.; Sawaya, R.J. Comparing species richness among assemblages using sample units: Why not use extrapolation methods to standardize different sample sizes? *Oikos* **2003**, *101*, 398–410. [CrossRef]
74. Cook, K.H.; Vizy, E.K. Detection and Analysis of an Amplified Warming of the Sahara Desert. *J. Clim.* **2015**, *28*, 6560–6580. [CrossRef]
75. Nicholson, S.E. Climatic and environmental change in Africa during the last two centuries. *Clim. Res.* **2001**, *17*, 123–144. [CrossRef]
76. Nicholson, S.E. On the factors modulating the intensity of the tropical rain belt over West Africa. *Int. J. Climatol.* **2009**, *29*, 673–689. [CrossRef]
77. Olaniyan, E.; Adefisan, A.E.; Oni, F.; Afiesimama, E.; Balogun, A.; Lawal, K.A. Evaluation of the ECMWF Sub-seasonal to Seasonal Precipitation Forecasts During the Peak of West Africa Monsoon in Nigeria. *Front. Environ. Sci.* **2017**, *6*, 1–15. [CrossRef]
78. Ehrendorfer, M. Predicting the uncertainty of numerical weather forecasts: A review. *Meteor Z* **1997**, *6*, 147–183.
79. Hamill, T.M.; Colucci, S.J. Verification of Eta–RSM short-range ensemble forecasts. *Mon. Weather Rev.* **1997**, *125*, 1312–1327.
80. Palmer, T.N. Predicting uncertainty in forecasts of weather and climate. *Rep. Prog. Phys.* **2000**, *63*, 71–116.
81. Stensrud, D.J.; Bao, J.-W.; Warner, T.T. Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Weather Rev.* **2000**, *128*, 2077–2107.
82. Stensrud, D.J.; Yussouf, N. Short-range ensemble predictions of 2-m temperature and dew-point temperature over New England. *Mon. Weather Rev.* **2003**, *131*, 2510–2524. [CrossRef]
83. Jankov, I.; Gallus, W.A., Jr.; Segal, M.; Shaw, B.; Koch, S.E. The impact of different WRF model physical parameterizations and their interactions on warm season MCS rainfall. *Weather Forecast.* **2005**, *20*, 1048–1060. [CrossRef]
84. Deser, C.; Lehner, F.; Rodgers, K.B.; Ault, T.; Delworth, T.L.; DiNezio, P.N.; Fiore, A.; Frankignoul, C.; Fyfe, J.C.; Horton, D.E.; et al. Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Clim. Change* **2020**, *10*, 277–286. [CrossRef]

85. Risser, M.D.; Stone, D.A.; Paciorek, C.J.; Wehner, M.F.; Angélil, O. Quantifying the effect of interannual ocean variability on the attribution of extreme climate events to human influence. *Clim. Dyn.* **2017**, *49*, 3051–3073. [CrossRef]
86. Tanimoune, L.I.; Smiatek, G.; Kunstmann, H.; Abiodun, B.J. Simulation of temperature extremes over West Africa with MPAS. *J. Geophys. Res. Atmos.* **2023**, *128*, e2023JD039055. [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.