Journal of Atmospheric Science Research
https://journals.bilpubgroup.com/index.php/jasr

RESEARCH ARTICLE

Assessing Subseasonal Forecasts of Dry Spells and Heatwaves at the Regional Scale in Brazil

Christopher Cunningham [1*], ***Nicholas P. Klingaman*** [2], ***Liana O. Anderson*** [1], ***Adriana Cuartas*** [1], ***Foster Brown*** [3,4], ***Paulo Henrique Valadares*** [3], ***Ianca Ribeiro*** [3], ***Luciana Londe*** [1]

[1] *Brazilian Center for Monitoring and Early Warning of Natural Disasters (Cemaden), São Jose dos Campos 12247-016, Brazil*

[2] *National Centre for Atmospheric Science and Department of Meteorology, University of Reading, Reading RG6 7BE, UK*

[3] *Post-Graduation Department, Acre's Federal University, Rio Branco 69915-900, Brazil*

[4] *Woodwell Climate Research Center, Falmouth, MA 02541, USA*

ABSTRACT

This study evaluates the performance of subseasonal forecasts for dry spells and heatwaves at a regional scale in Brazil. The forecasts' verification was designed to provide end-users with relevant information about the forecasts' quality. The U.K. Met Office model was assessed using a significant sample of weekly forecasts: 552 for dry spells and 240 for heatwaves. The analysis reveals that the overall performance of the forecasts is low, with a chance of detecting an event close to 0.2, indicating that only one out of five observed dry spells is accurately predicted on average. The application of quantile mapping corrections demonstrates improvements in predicting shorter dry spells (up to 5 days) and longer lead times, although the timing of these forecasts often remains inaccurate, leading to increased false alarms. A significant improvement in the forecast quality occurs when categorization by duration is disregarded. The detection chances increase to 0.5−0.7 for dry spells and 0.5 for heatwaves. The Brier Score indicates that the probabilistic forecasts issued by the model are equivalent or less skilful than climatological probabilities. Overall, the findings underscore the challenges in forecasting dry spells and heatwaves in Brazil and highlight the need for ongoing improvements in forecasting methodologies to

*CORRESPONDING AUTHOR:
Christopher Cunningham, Brazilian Center for Monitoring and Early Warning of Natural Disasters (Cemaden), São José dos Campos 12247-016, Brazil; Email: christopher.castro@cemaden.gov.br

ARTICLE INFO
Received: 31 July 2024 | Revised: 20 September 2024 | Accepted: 10 October 2024 | Published Online: 15 October 2024
DOI: https://doi.org/10.30564/jasr.v7i4.6973

CITATION
Cunningham, C., Klingaman, N.P., Anderson, L.O., et al., 2024. Assessing Subseasonal Forecasts of Dry Spells and Heatwaves at the Regional Scale in Brazil. Journal of Atmospheric Science Research. 7(4): 23–39. DOI: https://doi.org/10.30564/jasr.v7i4.6973

enhance their reliability and utility for regional decision-making. This research contributes to understanding subseasonal climate forecasting and its implications for managing climate-related risks in Brazil.

Keywords: Dry Spells; Heatwaves; Subseasonal Forecast; Early Warning System; Disaster Risk Reduction

1. Introduction

Earth's climate is warming fast, and the worldwide consequences of 0.99°C [0.84−1.10] of global warming since the industrial era are evident [1, 2]. Over South America, the climate projections under optimistic and pessimistic scenarios indicate a precipitation reduction in tropical areas (e.g., Chou et al. [3] and Marengo et al. [4]). Consistently, the warming in the region has also been depicted by the models, reaching up to 9°C in the most pessimistic scenario (RCP8.5) [5]. Apparently, what was projected to happen in the future is occurring in the present. The year 2023 was a worldwide record-breaking event in several aspects [6]: (1) 2023 is the warmest calendar year in global temperature records going back to 1850. (2) 2023 had a global average temperature of 14.98 °C, 0.17 °C higher than the previous highest annual value in 2016. (3) Each month from June to December 2023 was warmer than the corresponding month in any previous year. (4) In September 2023, the temperature deviation above the 1991–2020 average was larger than in any month in any year.

Long-term climate projections are undeniably vital to assessing potential scenarios and supporting long-term policies for mitigating climate emergencies. However, there is also an urgent need to forecast accurately in timescales that permit coping with the present-time adversities due to the warming climate. In this regard, a critical steering project was initiated in 2013: the Subseasonal-to-Seasonal prediction project (S2S) [7–11]. The subseasonal time scale of prediction, typically defined between 2 weeks to 2 months, is very important for several socio-economic sectors. One potential benefit of S2S forecasts is improving society's preparedness for extreme weather events leading to disruptive societal conditions, aka socio-environmental disasters [10, 12].

The International Strategy for Disaster Reduction (ISDR) in 2000 granted visibility to the concept of Early Warning Systems (EWS). During this decade, there were two International Conferences on Early Warning in Bonn, Germany (2003 and 2006), the first fostering the integration of early warning into public policies and the second emphasizing the implementation of the Hyogo Framework. After the 2004 tsunami in the Indian Ocean, many efforts have been made to develop and improve capacities for people-centred EWS. Current guidelines highlight strengthening capacities at the institutional foundations and local communities' levels. Warning systems for hazards used to be assumed to be top-down: to supply technology, data, and messages and then connect to the people affected as the "last mile" of the warning system. Yet lessons from past decades, alongside recent work, demonstrated that bringing in affected people last creates problems. Instead, warning systems need to be inclusive from the beginning.

Most scientific literature in atmospheric sciences on dry spells or heatwaves focuses on diagnostics, with few studies addressing forecast verification. For example, a search in American Meteorological Society journals from 1980 to 2020 using "dry spell" as a title keyword yielded 14 results, with only one related to forecasting [13]. The gap is less pronounced for heatwaves. A search in Royal Meteorological Society journals from the same period using "heatwave" returned 58 results, but only four focused on prediction or projection [14–17].

Krishnamurti et al. [13] were among the first to explore the idea that it would be possible to forecast at the subseasonal scale (1 month in advance). Their work was conducted as a case study of a dry spell over India during June 1979. They used a General Circulation Model (GCM), for which they eliminated the higher frequencies of variability in the initial state. The results were subjectively evaluated by comparing some forecasted atmospheric fields with the observed state. Vitart [14] also showed that some potentially useful skills might exist in a month's time scale. He used a GCM version from the European Center for Medium-Range Forecasts (ECMWF) to show that the model successfully predicted a chance of heatwave higher than in climatology more than ten days in advance. He studied the heatwave case that impacted Europe in August 2003. The evaluation was performed objectively using ROC and reliability diagrams. H. J. Lee, W.

S. Lee and Yoo[16] evaluated five GCMs from the Interactive Grand Global Ensemble (TIGGE) of the Observing system Research and Predictability Experiment (THORPEX). They used objective scores, such as the Critical Success Index, which was adopted in the present work as well. Their results showed that heatwaves are presumably predictable with 5 to 9 days lead times. However, the assessment was done only for four South Korean heatwave cases during July and August 2013.

In an original study, Lowe et al.[18] combined the S2S forecasts and a mortality model due to heatwaves to assess the potential of S2S forecasts in aiding preparedness and response in the health sector in Europe. They used the ECMWF subseasonal forecast system as input to a mortality model. Their work was based on a case study of the heatwave that impacted Europe during August 2003. Their results indicated that for short lead times (up to 4 days), the combined forecast skill (S2S forecasts + mortality model), given by the Area Under the ROC curve, is better than climatology.

Studies of heatwaves in Brazil are still exploratory, aiming to provide climatological details[19–22] but lack forecasting aspects and society-integration proposals. This study examines subseasonal forecasts of dry spells and heatwaves in two critical regions of Brazil, given that both events are expected to become more frequent and intense[23]. Those critical regions are characterized by decreased rainfall[24] or increased temperatures, and stakeholders could benefit from subseasonal forecasts for strategic planning.

The objective of the present study is to assess the quality of dry spell and heatwave forecasts at the subseasonal timescale, regarding the end-user perspective at a regional scale. It is essential to highlight that information for stakeholders regarding evaluating forecasts of dry spells and heat waves is scarce in Brazil. Dry spells were assessed over the Três Marias watershed in Southeastern Brazil and the heatwaves over Rio Branco's municipality in Southwestern Amazon. Supposing that an S2S forecast is part of a system to promote better decisions in Brazil's strategic socio-economic sectors, the relevant inquiry would be: considering a particular model, what is the usefulness of the raw and corrected forecasts predicting heatwaves and dry spells?

The paper is organized into five sections. In the second section, the datasets used are described, and the methods are explained in detail. Section 3 presents the results and discusses them. Section 4 presents the concluding remarks.

2. Materials and Methods

2.1. General Methods

The assessment approach was designed to provide relevant information about the forecasts' quality to users who might have to take critical decisions based on the forecasts' performance. For instance, we avoided evaluations over large areas (e.g., a whole hemisphere or an entire continent). While they are informative of a model's overall skill, therefore motivating from the modelling science point of view, they do not provide specific information about the forecasts' quality in the region where the decision process should be taken. In addition, the experiments' architecture was thought to mimic an operational condition on a hypothetical early warning system. Further, we used actual observed data instead of model-generated analysis. It is recognized that verification procedures that use analysis reduce the utility of the results for the users outside the modelling community and lead to an overestimation of the forecasts' quality[25].

Dry spell or heatwave predictions must be verified differently from precipitation or temperature extremes. First, you must select a threshold for defining a dry or hot day. In addition, we must consider that either dry spells or heatwaves are defined as uninterrupted sequences of days obeying such criterion. In other words, dry spells or heatwaves are also represented in the time dimension: the events' duration. Different durations imply different chances for occurrence and different impacts. For instance, Q. Zhao et al.[26] have found evidence in Brazil that heatwaves characterized by high temperatures and long durations were associated with a higher risk of hospitalization than lower and shorter heatwaves.

It is well known that GCM variables, such as precipitation or temperature, are not free of systematic errors[27]. Hence, postprocessing is a necessary step for any practical application. The present study adopted quantile mapping as the error correction technique. Quantile mapping is a statistical transformation method that finds a function to map the modelled variable such that its new distribution equals the distribution of the observed variable[28]. This method has been widely used to correct errors from GCM outputs in applications for hydrological prediction[29, 30] or meteorological studies[31]. One of the advantages of this method is that it can correct errors in the extremes of the distribution, which is a proper characteristic given that the present work deals with dry and hot days, both found in the tails of the precipitation and temperature distributions. As with any other method, it also presents disadvantages. T. Zhao et al.[32] have demonstrated that while quantile mapping is highly effective in correcting errors in model variables, it cannot ensure reliability in forecast ensemble spread or guarantee coherence, i.e., that forecasts are at least as skilful as climatology.

We used the R statistical computation software and the package "qmap" for such correction. Based on observational data, the mapping function was derived from the Empirical Cumulative Distribution Function (ECDF). When verifying the quantile mapping method results, we adopted a cross-validation approach to ensure the independence of the training and verification of data samples. Each week being evaluated is left out for verification, and the remaining weeks in the forecast time series compose the training dataset. The training dataset is then used to calculate the quantile-quantile relation.

Remembering that there is no universal approach to forecast assessment is valuable. The procedure adopted and the scores selected must match the study's specific objectives[33, 34]. The central perspective of the present assessment aims to answer the question: Is a dry spell/ heatwave predicted for the next week (or the second, third, or fourth week)? From this perspective, it is possible to verify the forecast as a non-probabilistic dichotomous forecast: "Yes, an event will happen" or "No, the event will not happen".

The model evaluated is the Met Office global seasonal prediction system (GloSea5)[35] in the scientific configuration Global Coupled 2.0[36]. In a recent study, Klingaman et al.[37] found that the Met Office model predicted well weekly mean rainfall anomalies across South America, validated against the same CHIRPS dataset used here. Of the four models examined in that study, the Met Office model performed second best, particularly in southeast Brazil. Such a result motivates examining its performance at shorter temporal and smaller spatial scales. From a practical standpoint, the Met Office forecast model output is readily available to decision-making agencies in Brazil, which allows the co-construction of a product design which can be operationalized in the future.

The available set of reforecasts covers 23 years (1993–2015), with seven (7) ensemble members. The model is integrated four times a month, on dates 1st, 9th, 17th and 25th. The total model's forecast horizon (45 days) was divided into 7-day weeks. The first week spans from forecast days 2 to day 8 (week 1), the second from days 9 to day 15 (week 2), the third from days 16 to 22 (week 3) and the last week from days 23 to 29 (week 4). Henceforward, those weeks may also be referred to as target-weeks.

The event durations considered for verification, either for dry spells or heatwaves, are 3 to 7 days. Given the present methodology's architecture, the verification experiments do not consider when, within the week, the dry spell occurred. For instance, for a forecast verification of a 3-day dry spell during the first week, a hit is considered if the model predicted the dry spell occurring at the beginning of the week, but the dry spell happened at the end.

The general procedure for verification of the forecasts is a multi-step process intended to mimic an operational condition when you have real-time predictions for the upcoming weeks (1 to 4). The scores calculated are then straightforwardly interpreted by the decision-makers. The process involves several key steps, including separating the model's initialization dates, selecting forecast dates within the assessed target-week range, adding observations to the dataset, and categorizing forecasts for each member within the target-week. The ensemble forecast is determined based on the majority prediction of the members, and this information is compared to the observations (**Table 1**). The outcome of this process is the creation of arrays representing the forecasts and observations, which are then used to build a contingency table for further analysis.

Several scores can be used when assessing dichotomous forecasts. We adopted the performance diagram approach[33]. With this diagram, it is possible to exploit the geometric relationship between four measures of dichotomous forecast performance: the probability of detection (POD), false alarm ratio or its opposite, the success ratio (SR), BIAS and critical success index (CSI, also known as the Threat Score). The BIAS is not the error in the forecasted values of temperature or precipitation. Instead, it is the frequency BIAS, which measures the ratio of the frequency of forecast events to the frequency of observed events. POD, SR, BIAS and CSI approach unity for good forecasts, such that a perfect

Table 1. Example of the categorization in dry days (1) or non-dry days (0) for a random week 1.

ICnDate	FctDate	FctM1	FctM2	FctM3	FctM4	FctM5	FctM6	FctM7	EnsFct	Obs
3/17/1993	3/18/1993	0	0	1	0	0	1	0	0	0
3/17/1993	3/19/1993	1	1	1	1	1	1	1	1	0
3/17/1993	3/20/1993	1	1	1	1	1	1	1	1	1
3/17/1993	3/21/1993	0	0	1	1	1	0	0	0	0
3/17/1993	3/22/1993	1	1	1	1	1	1	1	1	1
3/17/1993	3/23/1993	1	1	1	1	1	1	1	1	1
3/17/1993	3/24/1993	1	1	1	1	1	1	1	1	0

Note: The rightmost column indicates the observed (CHIRPS) categories, and the immediate column to the left is the EnsFct, calculated as the mode of the members.

forecast lies in the upper right of the diagram. Conversely, poor-performance forecasts tend to cluster in the lower left corner of the diagram.

Given a contingency table as below (**Table 2**), the definition for the Probability of Detection (POD), success ratio (SR), BIAS and critical success index (CSI) is:

Table 2. The joint distribution, i.e., the four combinations of forecasts (yes or no) and observations (yes or no).

	Observed Yes	Observed No
Forecast yes	Hits	False alarms
Forecast no	Misses	Correct negatives

$$\text{POD} = \frac{hits}{hits + misses} \tag{1}$$

$$\text{SR} = \frac{hits}{hits + false\ alarms} = 1 - \text{FAR} \tag{2}$$

where: $$\text{FAR} = \frac{false\ alarms}{hits + false\ alarms} \tag{3}$$

$$\text{BIAS} = \frac{hits + false\ alarms}{hits + misses} \tag{4}$$

$$\text{CSI} = \frac{hits}{hits + misses + false\ alarms} \tag{5}$$

When verifying the quality of dry spell predictions as non-probabilistic dichotomous forecasts, we assessed two experiment configurations that intend to suggest possible ways of presenting the dry spell forecasts for a decision-maker. The first experiment (Trial 1) is the most rigorous. It assesses the model's ability (as represented by the Ensemble Mode) to predict a dry spell of an exact given duration in the target-week. Hence, scores resulting from this trial are categorized according to the duration (3, 4, 5, 6 or 7 days), which results in five realizations of the verification procedure (see above) for the raw forecasts, plus five for the corrected forecasts.

The second experiment (Trial 2) makes the verification much more flexible since it considers hits if the EnsFct predicts an event (dry spell or heatwave) of any duration (3 to 7 days) and an event of any duration is observed. Hence, this trial assesses the forecast of an event's existence, disregarding the categorization by duration. Since all five durations are pooled in a single category, Trial 2 has only two realizations: raw and corrected forecasts.

Despite the apparent excessive flexibilization, this is a plausible forecast product from the end user's perspective. Representatives of the Center for Monitoring and Early Warning of Natural Disasters (Cemaden) regularly meet with the Ministry of Energy representatives. The goal is to support critical decisions on operating the generation and distribution of electricity in Brazil. Cemaden is a Brazilian federal institution that supports disaster risk reduction through operational activities. One learning from such interaction is that end users are frequently only interested in knowing if some extreme weather period is expected in the near future (for instance, a dry spell or a heatwave), disregarding their exact beginning or duration. Therefore, because of: (1) the scarcity of heatwaves and dry spells forecasts' verification for Brazil; and (2) considering the needed improvement in scientific and technical aspects of S2S predictions, the information generated is helpful and of interest.

Since the main advantage of the ensemble forecasts is to provide probabilistic estimates of the future, we have also calculated the Brier Score (BS). The Brier score is essentially the mean squared error of the probability forecasts, considering that the observation is one (1) if the event occurs and zero (0) if it does not. The formulation is as follows.

$$\text{BS} = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2 \tag{6}$$

in the present study, k varies from 1 to 552 (weeks) for the dry spells and 1 to 240 for heatwaves. Y represents the probability forecast, a fraction ranging from 0 to 1. The forecast probability was calculated by verifying for each member (1−7) if an event of a given duration is predicted during the forecasted week. The probability is the number of members forecasting an event divided by the total number of ensemble members.

It is common to present the BS as a skill score (BSS), which expresses the percentage improvement over a reference forecast. This work calculated the BS and BSS using the R-package "verification" from the National Center for Atmospheric Research (NCAR), Research Applications Laboratory. The BSS is expressed as follows.

$$\text{BSS} = \frac{BS - BS_{ref}}{0 - BS_{ref}} = 1 - \frac{BS}{BS_{ref}} \tag{7}$$

The BS_{ref} is a sample climatology calculated from the observations (see Sections 2.2 and 2.3 to learn the observation datasets). The following two sessions describe specific methodological aspects, particularly for dry spells or heatwaves.

2.2. Methods Concerning Dry Spell Analysis

The first critical area is the upper São Francisco River, the Três Marias basin, in Southeastern Brazil. This basin is crucial for hydropower generation, irrigation, water supply, and navigation[38]. The Três Marias reservoir regulates river flow and supports downstream Sobradinho and Itaparica reservoirs, which are essential for hydropower and water supply in the Brazilian Semiarid Region[39, 40]. Since 2014, the Southeast region has faced drought conditions[41]. The Três Marias basin has been significantly impacted, with mean discharge in summer 2014 at 357.0 $m^3 \cdot s^{-1}$ and in 2015 at 378.0 $m^3 \cdot s^{-1}$, considerably lower than the historical average of 988.0 $m^3 \cdot s^{-1}$ (1941–2015). By late November 2014, the reservoir level was approximately 2.6%. The National Centre for Early Warning of Natural Disasters (Cemaden, Brazil) conducts weekly strategic meetings with energy sector stakeholders to provide forecasts for priority hydropower basins. Operational decisions on sub-seasonal timescales are critical for water resources managers. However, existing S2S and seasonal predictions are seldom integrated into decision-making processes[42], highlighting the necessity to investigate and evaluate S2S forecasts and develop methodologies for their effective utilization by decision-makers and end-users.

The dataset used to validate the dry spells forecast was produced by the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS)[43]. The CHIRPS dataset has been previously validated for many regions in Brazil[44–47], including the Amazon[48] and at the national scale[49], showing robust association with observed weather station data.

The horizontal resolution of the UKMO model is 0.83 degrees longitude by 0.55 degrees latitude, which is approximately 90 km in longitude and 60 km in latitude. The original horizontal resolution of CHIRPS is 0.05 degrees (5 km) in both latitude and longitude. Both datasets were interpolated to a 1.5-degree grid by area-weighted bilinear interpolation. This grid matches the grid on which model output can be obtained from the S2S prediction project database, which means that the methods applied in this study could be easily extended to validate other forecast models.

The rainfall dataset was geographically and seasonally stratified. First, the CHIRPS daily rainfall was area-averaged over the seven grid points encompassing the basin (**Figure 1**). The delimited region presents a monsoonal climate regime[50, 51], with a well-defined dry season from April to September. As such, sequences of days without rain are expected during those months. For strategic sectors like food security or energy management, dry spells that matter are those that potentially may cause the most impact: those during the rainy season. Regarding the UKMO's retrospective forecast initializations, the dates span from 1993-Jan-01 to 2016-Dec-25, giving 552 forecasted weeks for any week considered (1 to 4).

The daily rainfall threshold differentiating a dry day from a non-dry day is 1 mm[52]. Such threshold corresponds to the 36th percentile of the CHIRPS time series, considering only the rainy period from October to March (ONDJFM). Consistently, the sample of assessed forecasts includes only initialization dates during ONDJFM.

Figure 2 shows the effect of the correction by the quantile mapping procedure. The quantile-quantile plot compares all the forecasted week 3 and the corresponding observed precipitation. The features are very similar for week 1, week 2 and week 4. The larger frame exhibits the whole range for precipitation variability. The smaller frame focuses on the lower values of precipitation, which are important for defining a dry day. The UKMO model presents a misrepresentation of days with no rainfall, as the vertical straight line at the beginning of **Figure 2a** demonstrates. It is well known that observed precipitation presents a significant fraction of days with zero precipitation. The UKMO misrepresented this fact, generating several days with very low rainfall but different from zero (**Figure 2a**). In the CHIRPS time series, the transition from zero rainfall to rainfall occurs approximately at 0.31st quantile. In the UKMO model, this same quantile corresponds to approximately 1.7 mm·day^{-1}. This discrepancy signifies fewer dry days in the model than in CHIRPS. The quantile mapping can correct this limitation (**Figure 2b**).

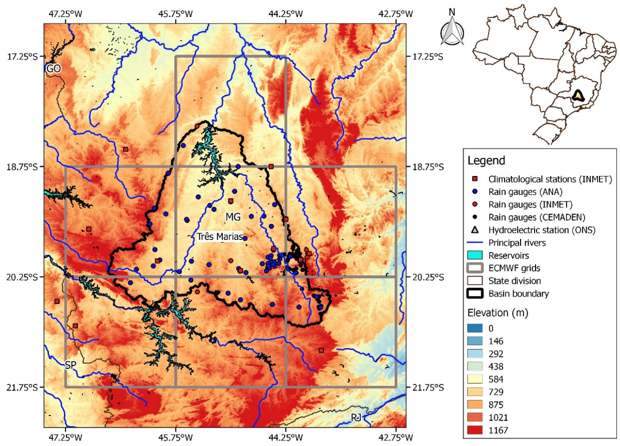

Figure 1. Três Marias watershed location and delimitation.

Note: The grey squares indicate the grid points used for averaging the CHIRPS daily precipitation. The dots indicate rainfall gauges that are used for operational purposes in Cemaden. They were not used in the present work.

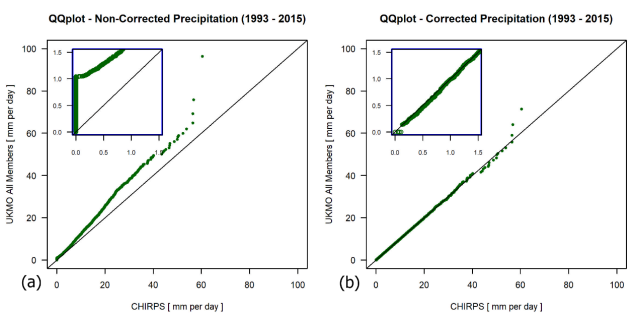

Figure 2. Quantile-Quantile plots comparing the UKMO model and the CHIRPS precipitation for week 3. (**a**) Raw (Non-Corrected) forecasts; (**b**) Corrected forecasts.

2.3. Methods Concerning Heatwave Analysis

The second study site is the Rio Branco municipality of the southwestern Brazilian Amazon (**Figure 3**). From 1973 to 2013, the Brazilian Amazon exhibited an upward trend in minimum, maximum, and average annual temperatures of respectively 0.038, 0.045 and 0.036 °C per year [53]. In the southwestern region, a temperature increase of approximately 1.72(±0.15) °C has been recorded since 1979, alongside a 20% reduction in rainfall during the driest months [54]. A notable warming trend has been documented both annually and seasonally, with both minimum and maximum temperatures rising more significantly during the dry season, thereby elevating the risk of temperature extremes adversely impacting population health and ecosystem stability in the Amazon [55–57].

We used the daily maximum temperature registered at the surface station in the Rio Branco municipality of Acre, southwestern Amazon, to validate the heatwave forecasts. This station is owned and maintained by the National Institute of Meteorology (INMET, in Portuguese), the Brazilian meteorological institution representing the World Meteorological Organization (WMO). The time series extends from 1979 to 2019, but there are some long periods of missing data at the beginning. The intersection of both datasets, the 23 years (1993–2016) of the UKMO model re-forecast and the non-missing years of the Rio Branco station, resulted in a period for verification from 1997 to 2016. The next step was to stratify the datasets, isolating the months when maximum temperatures are climatologically higher during the year: August, September and October (ASO). The outcome dataset contains 240 weeks to be verified.

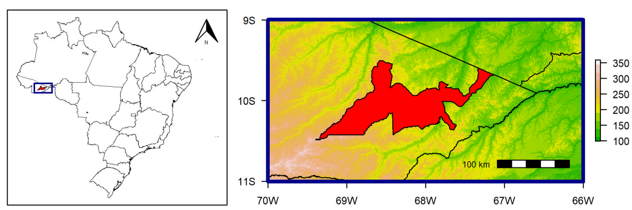

Figure 3. Rio Branco's location and boundaries (red polygon) in the Acre State, Southwestern Amazon.

Note: The left panel indicates the location. The right panel shows the municipality's boundaries and the elevation (meters) in the region (shades).

For heatwaves, a hot day is defined as exceeding 33.8 °C, representing the 80th percentile of the annual dataset from 1997 to 2016. The simplified wet bulb globe temperature (WBGT*) [58] indicates the effects of elevated temperatures on human health, as elevated air temperatures and humidity notably increase human stress. The application of the WBGT formula yields a temperature of 31.4 °C. Historically, significant heat-related mortality has occurred below this WBGT threshold. The European heatwave of 2003 caused around 45,000 deaths, while the Russian heatwave of 2010 resulted in approximately 54,000 fatalities, with WGBT values exceeding 25 °C in affected regions [58].

The ensemble forecast is represented by the ensemble mode precisely in the same way defined for the dry spells' analysis (Section 0). As was done for dry spells, two classes of experiments were performed. Trial 1 assesses the quality of the EnsFct to predict different categories of heatwaves, defined by their duration (3, 4, 5, 6 or 7 days). The second experiment (Trial 2) was designed to assess a heatwave's existence in the target-week without discrimination by duration.

The quantile mapping procedure effectively adjusts the UKMO's percentiles' deviation for weeks 2, 3, and 4 (**Figure 4**). After mapping the quantiles, the larger residual error occurs at the lower temperature's extremity, which is irrelevant to the present work. The correction in the higher temperatures could adjust the percentiles almost perfectly (**Figure 4**). However, applying the same mapping procedure to week 1 gives significantly poorer corrections (**Figure 5b**). Plotting the time evolution for all initialization dates clarifies what is happening (**Figure 5c**). The plot presents all forecast days, from 1 to 15 days lead-times. Each forecasted day is given as a boxplot. The variance represented by the boxplot concerns the differences among the initialization times. Viewing the forecast data from this perspective, it is clear that the UKMO's model presents a systematic error. Over the Rio Branco municipality region, the model exhibits 5 to 10 days of steady warming during the first lead times. As we shall see, such an error is likely related to the poor performance registered during week 1.

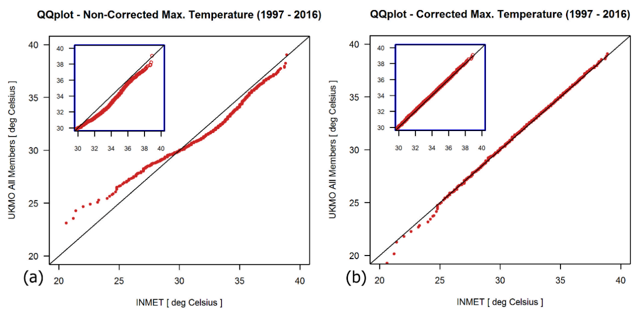

Figure 4. Quantile-Quantile plots comparing the INMET-Rio Branco maximum temperature and the UKMO forecast for week 3. (**a**) Raw (Non-Corrected) forecasts; (**b**) Corrected forecasts.

We repeated the boxplot procedure for a different gridpoint away from the southwestern Amazon to assess if this initial warming is specific to the Rio Branco region. The selected gridpoint is at the centre of the Três Marias watershed (45W, 19.5S). The analysis demonstrated that the systematic error pointed out in Rio Branco is not ubiquitous. Hence, there are reasons not to discredit the model or the study.

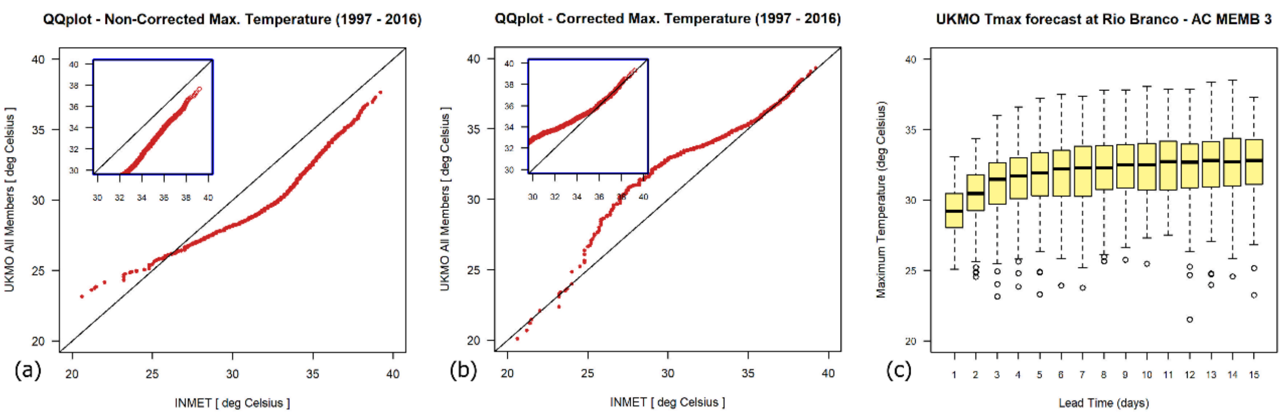

Figure 5. Drawback associated with the correction for week 1. Quantile-Quantile plots comparing the Inmet-Rio Branco maximum temperature and the Ukmo forecasts for (**a**) the Raw forecasts; and (**b**) the Corrected forecasts. (**c**) Box plots time evolution for all initialization dates during the first 15 days of forecasts.

Note: The **Figure 5c** plot corresponds to member 3 of the model, but the error is also found in other members.

3. Results and Discussion

3.1. Dry Spells

Cunningham [51] studied the climatological characteristics of dry spells over Southeastern Brazil, including the area of the Três Marias basin. The characteristic dry spells duration during the rainy season (October to April) is up to 9 days (80%). Dry spells ten days long or even longer are relatively uncommon. The dry spells' life cycle during the rainy season exhibits a markedly subseasonal variability. From October onwards, there is a progressive tendency to a minimum in occurrences and duration, culminating in December, the wettest month. After this point in the cycle, the tendency is to increase frequency and duration, peaking in February. During March, the rainy season recovers to some extent and demises in April. This characteristic behaviour determines that January and February are predisposed to dry spells, much more than December. For instance, the chance of a dry spell ten days long or longer in December is 9/100, but 18/100 in January and 26/100 in February.

Figure 6 shows the four scores presently considered, plotted as the performance diagram [33] for the first and second forecast weeks. Scores for weeks 3 and 4 are presented in **Table 3** and not as performance diagrams because they tend to cluster in the diagram's origin. The positioning of the points concerning Trial 1 and week 1 reflects low score values. The better performance occurs when the model forecasts 3-day dry spells (**Figure 6a**, cyan points). An equivalent performance is found for a raw forecast of dry spells 6-day long (yellow open square). Nevertheless, the overall performance, considering all duration categories, is low. The POD is roughly lower than 0.2, meaning that for every five observed dry spells, only one is correctly predicted on average. The SR presents values of the same order (0.2). Then, on average, one is a hit for a group of five forecasts predicting a dry spell.

The gold points in **Figure 6** represent the scores resulting from Trial 2. They show an increase in performance when the forecast system is verified without restricting duration categories. For week 1, POD and SR are around 60% (0.60 and 0.63, respectively) for the raw forecasts and 0.7 and 0.59 for the corrected forecasts. There is a drop in performance in week 2. The POD and SR are respectively 0.34 and 0.49 for the uncorrected forecasts and 0.40 and 0.41 for the corrected forecasts. Nevertheless, it is much better than the scores for Trial 1, all clustered with POD and SR below 20% (0.20). For week 3, the scores drop even further: POD and SR are respectively 0.12 and 0.30 for the raw forecasts and 0.30 and 0.40 for the corrected forecasts (not shown). At the longest lead time (week 4), performance is slightly increased compared to week 3. POD and SR are respectively 0.18 and 0.54 for the uncorrected forecasts and 0.28 and 0.48 for the corrected forecasts.

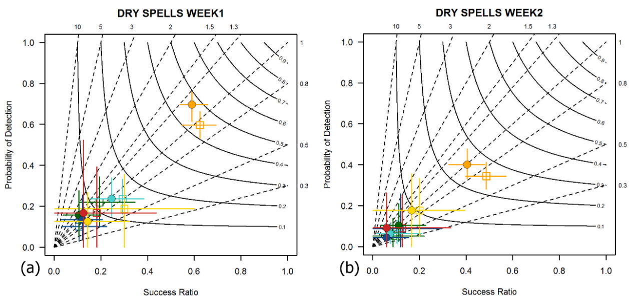

Figure 6. Performance diagram for the (**a**) first and (**b**) second weeks predicted in trial 1 and trial 2.

Note: Open squares denote scores for the raw forecasts, and filled circles denote scores after quantile mapping correction. For Trial 1, 3-day dry spells are cyan, 4-day are green, 5-day are blue, 6-day are yellow, and 7-day are red. The two gold points represent the results of Trial 2. Dashed lines represent BIAS scores with labels on the line's outward extension; while labelled solid contours are CSI. The crosshairs give sampling uncertainty.

In general, applying quantile mapping increases the BIAS (Equation (4)). This feature can be noticed in the diagrams (**Figure 6**), observing that, in general, for the same dry spell duration (the same colour), the corrected point (filled circle) is displaced leftward and upward relative to the non-corrected point (empty square). The BIAS is mostly below one for non-corrected (raw) forecasts (**Table 3**). Section 2.2. shows that the raw UKMO forecasts predict fewer dry days than observed. The BIAS scores (**Table 3**) show that this misrepresentation leads to fewer dry spells, which is a logical fact since dry spells are a chain of dry days. This underrepresentation happens virtually every week and in every duration category except for dry spells 4-day long forecasted during week 1. The quantile mapping causes an increase in forecasts' counts (hits + false alarms; not shown). The increased frequency can be appraised in Table 3 , where it can be seen that BIAS is increased for every week and duration category, and in many instances, the model is forecasting more dry spells than occurred.

After correction, the POD improves for shorter dry spells (up to 5 days) and longer lead times (weeks 3 and 4). The best improvement is 10% (from 0 to 0.10), predicting a dry spell 5-day long during week 3. The SR improvements tend to favour longer lead times. The best improvement is again for a dry spell 5-day long during week 3 (from 0 to 0.12, after correction). In a broader picture, there are no improvements, or there are worsenings in the scores. Since

Table 3. Dry spells' scores that compose the performance diagram for trial 1.

	POD *uncorrected*				SR *uncorrected*				BIAS *uncorrected*			
Weeks	1	2	3	4	1	2	3	4	1	2	3	4
Heatwaves duration												
3-days	0.24	0.06	0.03	0.04	0.29	0.11	0.08	0.17	0.81	0.59	0.38	0.27
4-days	0.22	0.05	0.03	0.05	0.19	0.09	0.05	0.13	1.13	0.59	0.50	0.41
5-days	0.10	0.09	0.00	0.03	0.11	0.12	0.00	0.09	0.93	0.74	0.33	0.34
6-days	0.19	0.18	0.06	0.10	0.30	0.20	0.25	0.25	0.63	0.88	0.22	0.40
7-days	0.17	0.09	0.00	0.00	0.18	0.13	0.00	0.00	0.92	0.73	0.64	0.71
	POD *difference after correction*				SR *difference after correction*				BIAS *difference after correction*			
Heatwaves duration												
3-days	0.00	0.00	0.02	0.06	−0.05	−0.03	−0.01	0.05	0.15	0.23	0.27	0.21
4-days	−0.06	0.05	0.08	0.00	−0.09	0.02	0.09	−0.05	0.34	0.33	0.26	0.24
5-days	0.03	−0.04	0.10	0.03	0.01	−0.06	0.12	0.03	0.17	0.00	0.48	0.19
6-days	−0.06	0.00	−0.06	0.00	−0.16	−0.03	−0.25	−0.08	0.25	0.18	0.50	0.20
7-days	0.00	0.00	0.00	0.00	−0.06	−0.07	0.00	0.00	0.42	0.82	0.09	0.14

Note: The scores presented are the Probability of Detection (POD), the Success Ratio (SR) and the frequency bias (BIAS). The top numbers indicate different target-weeks (1 to 4). The results are presented as function of the event duration.

the BIAS is generally increased, the results indicate that even though the model forecasts more dry spells, the timing (simultaneous occurrence in the week) is incorrect.

The BS and BSS are displayed in **Figure 7**. We compare the performance among weeks by comparing different colours, the performance between raw and corrected forecasts by comparing bars and lines for the same colour, and performance as a function of the dry spell's duration by comparing the categories (3 to 7 days). The most noticeable feature is a diminishing error as the duration becomes more prolonged, suggesting better skill in predicting longer dry spells. However, this might be a misleading result. A background characteristic of the BS is that it is easier to achieve a low BS for less frequent events without real skill [59]. We have not computed the number of forecasted events during the Brier score calculation process. However, we do know that for observed events, it is true that the longer they are, the less frequent they are. For instance, the experiment resulted in 72, 32, 30, 16 and 12 observed dry spells during week 1, 3, 4, 5, 6 and 7 days, respectively. The diminishing frequency of events as they become longer is maintained for the other weeks (not shown).

The skill scores (BSS) indicate that the UKMO's forecasts are less skilful than a climatological forecast since they are all negative (**Figure 7b**). Comparing bars (raw forecasts) and lines (corrected forecasts) for the same colour (dry spell duration), we see that the quantile mapping correction increases the forecasts' probability error. This increase is more evident in the BSS, where the lines are systematically and notoriously at lower values than the bars (raw forecasts). The only exception is when the forecast system predicts 3-day dry spells during week 3. Also, week 3 presented the best results: a small increase in error as the event duration increases and an overall error or the order of 10% or less.

A critical characteristic of the results is that there is no significant increase in probability errors with increasing lead time (**Figure 7a**). The BS indicates that the probability error predicting week 4 (red) is slightly lower than predicting week 1 (green). Usually, studies verifying the quality of forecast predicting, for instance, precipitation anomalies for weeks 1 to 4, report a sharp decrease in skill as the lead time increases. In general, it is well known for any variable that the skill of a forecast 20 days ahead is very low [27].

3.2. Heatwaves

Pallotta [60] has analyzed the meteorological agents associated with temperature extremes in Rio Branco. Most registered cases of extreme heat tend to occur in Rio Branco between the end of winter and spring (August to October). This time of the year corresponds to the demise of the dry season in the region, with a large availability of sensible heat in the layers close to the surface. Associated with this excess of heat, the author also found: (1) a large-scale subsidence pattern over the Acre state at 500 hPa level; and (2) an anomalous northwesterly wind pattern that combines warm advection from the hotter equatorial regions with katabatic winds descending the Andes over Acre.

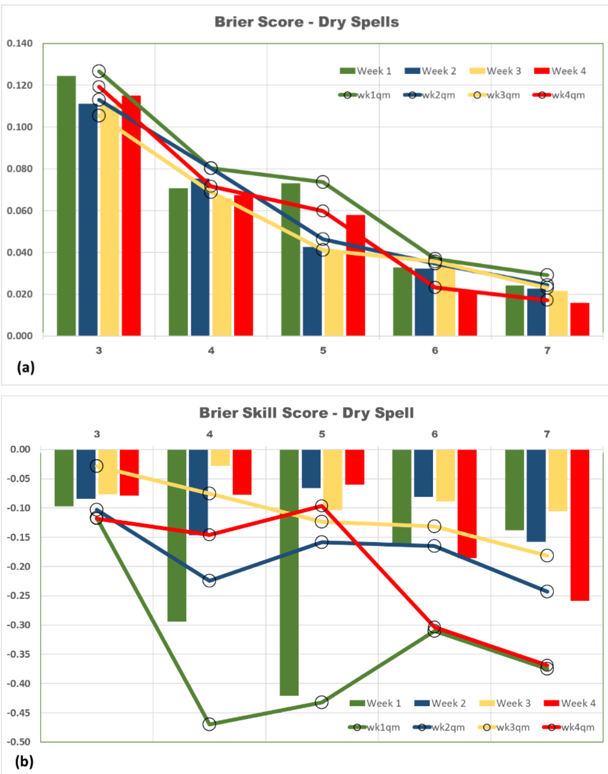

Figure 7. (**a**) B.S. and (**b**) BSS for dry spells' forecasts.

Note: Bars (lines) show the scores for the raw (corrected) forecasts. The scores for week 1 are represented as green, week 2 as blue, week 3 as yellow, and week 4 as red. The scores are categorized according to the dry spell duration.

The scores concerning Trial 1 are presented in **Table 4**. The SR scores show that the best raw model's performance is associated with longer heatwaves and lead times. The model shows a success of 4 in 10 predicting 3-day heatwaves during week 4. During week 2 the model hits 1 in 4 yes forecasts, considering heatwaves as long as 6 and 7 days. From the observed heatwaves' perspective, the raw forecasts best predict 5-day heatwaves during week 1 and week 4, though with chances not greater than 1 in 10. Also, during week 4, the model demonstrated a chance of 8 in 10 to hit a 3-day heatwave.

The correction has an improving effect over the POD and SR, different to what was seen in the analysis for dry spells (see Discussion in Section 3.1). The applied correction mostly increased the frequency of predicted heatwaves, as can be deduced by the majority of positive values in the BIAS difference. This improvement reflects better in the success rate, which is consistent since this score represents the fraction of the forecast "yes" events that were correctly observed.

Results from Trial 2 are displayed in **Figure 8**. As in the dry spells' assessment, the performance is substantially increased when the event being predicted is not discriminated by duration. The effect of the error minimization employing the quantile mapping is visible. All corrected weeks (filled circles) are displaced to positions near the BIAS diagonal equal to 1. Before quantile mapping, the BIAS values indicate a sub-representation of the heatwaves (values lower than 1), i.e., on average, the forecast system anticipates fewer heatwaves than observed.

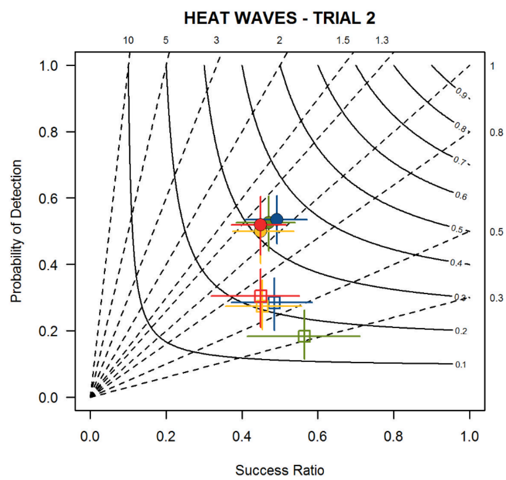

Figure 8. Performance diagram showing the scores resulting from trial 2.

Note: Open squares denote scores for the raw forecasts, and filled circles denote scores after quantile mapping correction. The scores for week 1 are represented as green, week 2 as blue, week 3 as yellow, and week 4 as red. Dashed lines represent BIAS scores with labels on the line's outward extension, while labelled solid contours are CSI. The crosshairs give sampling uncertainty.

The corrected forecasts exhibited a sharp increase in the POD, i.e., displacements along the Y-axis. Before the correction, the POD tends to cluster around 0.2 to 0.3. After the quantile mapping, POD scores increase to values over 0.5, which means that the forecast system predicts half of the observed heatwaves well. It is important to notice the discrepancy between week 1 (open green square) and the other lead times (**Figure 8**). This feature is likely due to the systematic warming error identified and associated with the first days of forecasts (see **Figure 5**).

An important feature of the present results is the small

Table 4. Heatwaves' scores that compose the performance diagram for trial 1.

	POD *uncorrected*				SR *uncorrected*				BIAS *uncorrected*			
Weeks	1	2	3	4	1	2	3	4	1	2	3	4
Heatwaves duration												
3-days	0.02	0.00	0.05	0.08	0.11	0.00	0.20	0.40	0.23	0.20	0.26	0.21
4-days	0.00	0.03	0.03	0.03	0.00	0.08	0.10	0.07	0.38	0.42	0.34	0.45
5-days	0.10	0.05	0.00	0.06	0.20	0.07	0.00	0.20	0.50	0.79	0.21	0.31
6-days	0.00	0.06	0.00	0.00	0.01	0.25	0.00	0.00	0.07	0.25	0.67	1.17
7-days	0.00	0.06	0.00	0.00	0.01	0.25	0.00	0.00	0.06	0.25	1.50	0.83
	POD *difference after correction*				SR *difference after correction*				BIAS *difference after correction*			
Heatwaves duration												
3-days	−0.02	0.12	−0.03	0.04	−0.11	0.54	−0.11	−0.08	−0.05	0.02	0.03	−0.19
4-days	0.00	0.00	0.03	0.06	0.00	0.01	0.15	0.16	0.03	−0.06	−0.07	−0.06
5-days	0.05	0.00	0.04	0.06	0.03	0.04	0.08	−0.02	0.15	−0.32	0.29	0.38
6-days	0.07	−0.06	0.17	0.00	0.06	−0.25	0.11	0.00	0.87	0.31	0.83	−0.17
7-days	0.06	0.00	0.00	0.00	0.01	−0.15	0.00	0.00	2.75	0.38	−0.25	1.00

Note: The scores presented are the Probability of Detection (POD), the Success Ratio (SR) and the frequency bias (BIAS). The top numbers indicate different target-weeks (1 to 4). The results are presented as function of the event duration.

variability in the scores as the lead time increases. All the lead times are clustered in a region with POD and SR of 50%. The BIAS score is less than 1.3, and the CSI score is between 0.3 and 0.4, respectively. This relative invariance was not present in Trial 2 concerning the prediction of dry spells. We saw a significant drop in performance from week 1 to week 2 (**Figure 6**) and subsequent weeks.

The probability error associated with the heatwaves' forecast showed results comparable to the dry spells assessment. **Figure 9** shows the B.S. and BSS for each week, discriminated by the heatwave duration. For some duration categories, the error decreases as the lead time becomes longer, like for heatwaves lasting 3, 4 and 5 days. For all duration categories, the error is lower in week 4.

Like in the dry spells analysis, the skill scores (BSS) are all negative, indicating that the forecasts are less skillful than the BS_{ref} (climatology). The BSS values tend to present small variability among the different categories of durations. Except for week 1, the error increment over climatology (BSS) is 20% or less. Week 3 and week 4 presented the best results. The average score over all the categories for raw or corrected forecasts is greater than −0.2.

Predictions for week 1 present a significant degree of degradation (loss of skill), both as a function of the predicted heatwave duration and when the quantile mapping is applied. This degradation is not so severe for the other predicted weeks. We attribute this feature to the systematic error identified during the first lead times (**Figure 5**).

Figure 9. (**a**) B.S. and (**b**) BSS for heatwaves' predictions.

Note: Bars (lines) show the scores for the raw (corrected) forecasts. The scores for week 1 are represented as green, for week 2 as blue, for week 3 as yellow and for week 4 as red. The scores are categorized according to the dry spell duration.

4. Conclusions

With the global population's exposure to heat waves and dry spells escalating due to climate change, the threat to societal stability and public health is becoming more severe. In this context, the urgent need to expand Early Warning Systems for heat-related and drought-related health risks across Brazil cannot be overstated. While these systems encompass more than just predicting the occurrence of hazards such as heatwaves, the importance of having an accurate forecasting system is paramount. Yet, both in general and specifically in Brazil, there is a noticeable lack of studies on the effectiveness of cutting-edge models in predicting these events.

The original contribution of this research paper lies in its focused assessment of subseasonal forecasts for dry spells and heatwaves in the Brazilian context, which unfolds into several aspects. First, the study addresses a significant gap in the literature by emphasizing forecast verification for dry spells and heat waves, which has been largely overlooked in atmospheric sciences. Most existing studies focus on diagnostics rather than on the accuracy of forecasts, making this work a valuable addition to the field. Second, the assessment architecture was designed after transferring knowledge acquired from regular interactions with water management and energy production stakeholders. The objective was to provide relevant information about the forecasts' quality for those making critical decisions. Third, by concentrating on essential regions of Brazil, the study contributes region-specific insights that can inform local decision-making and policy development related to climate impacts, particularly in sectors like agriculture, water management and disaster risk reduction. This regional focus is essential for tailoring early warning systems and improving community resilience to climate variability.

Concerning the most significant results, they indicate that the prevailing performance of the raw forecasts is low, particularly in predicting dry spells. Only one out of five observed events are correctly predicted on average. Applying quantile mapping corrections improves the chance for detection of dry spells and heatwaves. However, while the model forecasts more events after correction, the timing of these forecasts is often incorrect, leading to an increase in false alarms. The Brier Score provides insights into the mean squared error of the probability forecasts, indicating that while the ensemble forecasts offer probabilistic estimates, the overall accuracy remains a concern.

The present work suggests that while the models designed for predictions at the subseasonal timescale represent a crucial tool, there is still a need for better forecasts that can accurately predict dry spells and heatwaves, especially in the context of regional decision-making. There is a myriad of techniques that can be used in an attempt to improve the forecast quality. We highlight two that look promising and intend to explore them as the next step in developing the present research. The first is the association between atmospheric circulation patterns and extreme weather conditions. Several studies have associated periods of extreme temperature with unusual and persistent patterns of atmospheric circulation [61, 62]. More specifically, dry spells in southeastern Brazil have also been associated with these semi-permanent atmospheric circulation patterns. An anomalous westward displacement of the South Atlantic subtropical anticyclone is attributed as the cause of the summer monsoon dry spell in South America [51, 63]. Occasionally, this climatological dry spell becomes more intense and prolonged, causing severe droughts [64, 65].

Artificial Intelligence (AI) and data science technologies have been rising as techniques in the atmospheric and hydrological modelling itself or as supporting techniques [66–68]. Recently, their importance has been leveraged by some of the leading groups in atmospheric science. For instance, for Phase II of the S2S project, the experimentation of AI and machine learning techniques for model postprocessing and downscaling is explicitly encouraged. In 2021, a prize challenge was launched to improve sub-seasonal to seasonal prediction of 2-metre temperature and precipitation 3 to 6 weeks in advance using Artificial Intelligence and Machine Learning methods [69]. The European Centre for Medium-Range Forecast also pays attention to machine learning. In April 2021, the ECMWF launched the MAELSTROM project (MAchinE Learning for Scalable Meteorology and Climate). The three-year project aims to help prepare the weather and climate community for large-scale machine-learning applications (ECMWF, 2020).

Overall, this work fills a critical gap in the existing literature and provides practical tools and insights for improving subseasonal forecasting in Brazil. Doing so contributes to better preparedness and response strategies for extreme weather-related events, demonstrating the tangible impact of the research.

Author Contributions

Conceptualization and design, data collection, data analysis, writing and drafting, review and editing, visualization, literature review, supervision and coordination, C.C.; data collection, review and editing, N.P.K.; writing and drafting; review and editing; literature review, L.O.A.; writing and drafting, review and editing, literature review, A.C.; writing and drafting, F.B.; writing and drafting; literature review, P.H.V.; writing and drafting, literature review, I.R.; writing and drafting, literature review, L.L.

Funding

This work was partially supported by the Climate Science for Services Partnership for Brazil, part of the Newton Fund. Dr. Nicholas Klingaman was funded by an Independent Research Fellowship from the Natural Environment Research Council (NE/L010976/1). Dr. Liana Anderson was funded by the project: MAP-FIRE (IAI-SGP-HW016) National Council for Scientific and Technological Development–CNPq Productivity fellowship (314473/2020-3).

Institutional Review Board Statement

Not applicable

Informed Consent Statement

Not applicable

Data Availability Statement

The data used in this study are publicly available. The authors will be happy to share the data used in the manuscript upon request.

Acknowledgments

Dr. Christopher Cunningham acknowledges the Real-Time Pilot Project under the umbrella of the Subseasonal to Seasonal Prediction Project, for the opportunities provided. Dr. Christopher Cunningham acknowledges Dr. Mariana Pallotta for discussions and enlightening during the earlier stages of the study.

Conflict of Interest

The authors declare no sources of conflict of interest.

References

[1] Intergovernmental Panel on Climate Change (IPCC), 2022. 'Summary for Policymakers', in Global Warming of 1.5°C: IPCC Special Report on Impacts of Global Warming of 1.5°C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty. Cambridge: Cambridge University Press. pp. 1–24. DOI: https://doi.org/10.1017/9781009157940.001

[2] Intergovernmental Panel on Climate Change (IPCC), 2023. 'Summary for Policymakers', in Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press. pp. 3–32. DOI: https://doi.org/10.1017/9781009157896.001

[3] Chou, S.C., Marengo, J.A., Lyra, A.A., et al., 2012. Downscaling of South America present climate driven by 4-member HadCM3 runs. Climate Dynamics. 38, 635−653.

[4] Marengo, J.A., Jones, R., Alves, L.M., et al., 2012. Development of regional future climate change scenarios in South America using the Eta CPTEC/HadCM3 climate change projections: Climatology and regional analyses for the Amazon, São Francisco, and the Paraná River Basins. Climatic Dynamics. 38(9–10), 1829–1848. DOI: https://doi.org/10.1007/s00382-011-1155-5

[5] Chou, S.C., Lyra, A., Mourão, C., et al., 2014. Assessment of Climate change over South America under RCP 4.5 and 8.5 downscaling scenarios. American Journal of Climate Change. 3(5), 512–527. DOI: https://doi.org/10.4236/ajcc.2014.35043

[6] Copernicus Services, 2023. Global Climate Highlights 2023. Available from: https://climate.copernicus.eu/global-climate-highlights-2023 (cited 12 September 2024).

[7] Brunet, G., Shapiro, M., Hoskins, B., et al., 2010. Collaboration of the weather and climate communities to advance subseasonal-to-seasonal prediction. Bulletin of the American Meteorological Society. 91(10), 1397–1406. DOI: https://doi.org/10.1175/2010BAMS3013.1

[8] Robertson, A.W., Kumar, A., Peña, M., et al.,

2015. Improving and promoting subseasonal to seasonal prediction. Bulletin of the American Meteorological Society. 96(3), ES49–ES53. DOI: https://doi.org/10.1175/BAMS-D-14-00139.1

[9] Vitart, F., Ardilouze, C., Bonet, A., et al., 2017. The subseasonal to seasonal (S2S) prediction project database. Bulletin of the American Meteorological Society. 98(1), 163–173. DOI: https://doi.org/10.1175/BAMS-D-16-0017.1

[10] White, C.J., Carlsen, H., Robertson, A.W., et al., 2017. Potential applications of subseasonal-to-seasonal (S2S) predictions. John Wiley and Sons Ltd.: Hoboken, NJ, USA. pp. 315–325. DOI: https://doi.org/10.1002/met.1654

[11] Vitart, F., Robertson, A.W., 2018. The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. NPJ climate and atmospheric science. 1(1), pp. 1–7. DOI: https://doi.org/10.1038/s41612-018-0013-0

[12] Garcia-Herrera, R., Díaz, J., Trigo, R.M., et al., 2010. A review of the european summer heat wave of 2003. Critical Reviews in Environmental Science and Technology. 40(4), 267–306. DOI: https://doi.org/10.1080/10643380802238137

[13] Krishnamurti, T.N., Subramaniam, M., Daughenbaugh, G., et al., 1992. One-month forecasts of wet and dry spells of the monsoon. Monthly Weather Review. 120, 1191–1223. DOI: https://doi.org/10.1175/1520-0493(1992)120%3C1191:OMFOWA%3E2.0.CO;2

[14] Vitart, F., 2005. Monthly forecast and the summer 2003 heat wave over Europe: A case study. Atmospheric Science Letters. 6(2), 112–117. DOI: https://doi.org/10.1002/asl.99

[15] Kyselý, J., 2010. Recent severe heat waves in central Europe: How to view them in a long-term prospect? International Journal of Climatology. 30(1), 89–109. DOI: https://doi.org/10.1002/joc.1874

[16] Lee, H.J., Lee, W.S., Yoo, J.H., 2016. Assessment of medium-range ensemble forecasts of heat waves. Atmospheric Science Letters. 17(1), 19–25. DOI: https://doi.org/10.1002/asl.593

[17] Kim, S., Sinclair, V.A., Räisänena, J., et al., 2018. Heat waves in Finland: Present and projected summertime extreme temperatures and their associated circulation patterns. International Journal of Climatology. 38(3), 1393–1408. DOI: https://doi.org/10.1002/joc.5253

[18] Lowe, R., García-Díez, M., Ballester, J., et al., 2016. Evaluation of an early-warning system for heat wave-related mortality in Europe: Implications for sub-seasonal to seasonal forecasting and climate services. International Journal of Environmental Research and Public Health. 13(2), 206. DOI: https://doi.org/10.3390/ijerph13020206

[19] Jacondino, W.D., Nascimento, A.L.D.S., Nunes, A.B., et al., 2019. SYNOTIC ANALYSIS OF APRIL 2018 IN THE SOUTH REGION OF BRAZIL: AN EXTREME HEAT EPISODE. Revista Brasileira de Climatologia. 25, 182–203. (in Portuguese).

[20] dos Reis, N.C.S., Boiaski, N.T., Ferraz, S.E.T., 2019. Characterization and spatial coverage of heatwaves in subtropical Brazil. Atmosphere. 10(5), 284. DOI: https://doi.org/10.3390/atmos10050284

[21] Passos Alves, B.A., Cruz, S., de Cássia B.K., et al., 2016. Characterization of Strong Heat Wave 2014 in Santa Catarina. Ciência e Natura. 38 (1), 309–325. (in Portuguese). DOI: https://doi.org/10.5902/2179-460X15017

[22] Selmo, L., Franke, A.E., Passos, M., et al., 2016. Análise das Temperaturas Máximas do Ar em Florianópolis/SC em Abril de 2016: Onda de Calor? Analysis of Maximum Air Temperatures in Florianópolis/SC in April 2016: Heat Wave? Proceedings of Brazilian Conference of Disaster Risk Reduction; 12–15 October 2016; Curitiba, Paraná, Brazil. pp. 48–56.

[23] Hoegh-Guldberg, O., Jacob, D., Taylor, M., et al., 2019. The human imperative of stabilizing global climate change at 1.5°C. Science. 365(6459), eaaw6974. DOI: https://doi.org/10.1126/science.aaw6974

[24] Rao, V.B., Franchito, S.H., Santo, C.M.E., et al., 2016. An update on the rainfall characteristics of Brazil: Seasonal variations and trends in 1979–2011. International Journal of Climatology. 36 (1), 291–302. DOI: https://doi.org/10.1002/joc.4345

[25] Coelho, C.A.S., Brown, B., Wilson, L., et al., 2019. Forecast verification for S2S timescales. In: Robertson, A.W., Vitart, F. (eds). Sub-Seasonal to Seasonal Prediction: Filling the Gap Between Weather and Climate. Elsevier: London, UK. p. 344.

[26] Zhao, Q.,Li, S., Coelho, M.S.Z.S., et al., 2019. The association between heatwaves and risk of hospitalization in Brazil: A nationwide time series study between 2000 and 2015. PLoS Medicine. 16 (2), e1002753. DOI: https://doi.org/10.1371/journal.pmed.1002753

[27] Toth, Z., Buizza, R., 2019. Weather forecasting: What sets the forecast skill horizon?. In: Robertson, A.W., Vitart, F. (eds). Subseasonal to Seasonal Prediction: Filling the Gap between Weather and Climate. Elsevier: London, UK. p. 35.

[28] Gudmundsson, L., Bremnes, J.B., Haugen, J.E., et al., 2012. Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations—a comparison of methods. Hydrology and Earth System Sciences. 16(9), 3383–3390. DOI: https://doi.org/10.5194/hess-16-3383-2012

[29] Wood, A.W., Maurer, E.P., Kumar, A., et al., 2002. Long-range experimental hydrologic forecasting for the eastern United States. Journal of Geophysical Research: Atmospheres. 107(20), ACL 6-1–ACL 6-15. DOI: https://doi.org/10.1029/2001JD000659

[30] Lucatero, D., Madsen, H., Refsgaard, J.C., et al., 2018.

On the skill of raw and post-processed ensemble seasonal meteorological forecasts in Denmark. Hydrology and Earth System Sciences. 22(12), 6591–6609. DOI: https://doi.org/10.5194/hess-22-6591-2018

[31] Hirata, F.E., Grimm, A.M., 2018. Extended-range prediction of South Atlantic convergence zone rainfall with calibrated CFSv2 reforecast. Climate Dynamics. 50(9–10), 3699–3710. DOI: https://doi.org/10.1007/s00382-017-3836-1

[32] Zhao, T., Bennett, J.C., Wang, Q.J., et al., 2017. How suitable is quantile mapping for postprocessing GCM precipitation forecasts? Journal of Climate. 30(9), 3185–3196. DOI: https://doi.org/10.1175/JCLI-D-16-0652.1

[33] Roebber, P.J., 2009. Visualizing multiple measures of forecast quality. Weather and Forecasting. 24(2), 601–608. DOI: https://doi.org/10.1175/2008WAF2222159.1

[34] WMO, 2008. Recommendations for the Verification and Intercomparison of QPFs and PQPFs from Operational NWP Models. WMO/TD - No 1485.

[35] Maclachlan, C., Arribas, A., Peterson, K.A., et al., 2015. Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. Quarterly Journal of the Royal Meteorological Society. 141(689), 1072–1084. DOI: https://doi.org/10.1002/qj.2396

[36] Williams, K.D., Harris, C.M., Bodas-Salcedo, A., et al., 2015. The Met Office Global Coupled model 2.0 (GC2) configuration. Geoscientific Model Development. 88, 1509–1524. DOI: https://doi.org/10.5194/gmd-88-1509-2015

[37] Klingaman, N.P., Young, M., Chevuturi, A., et al., 2021. Subseasonal prediction performance for austral summer South American rainfall. Weather and Forecasting. 36(1), 147–169. DOI: https://doi.org/10.1175/WAF-D-19-0203.1

[38] Maneta, M.P., Torres, M., Wallender, W.W., et al., 2009. Water demand and flows in the São Francisco River Basin (Brazil) with increased irrigation. Agricultural Water Management. 96, 1191−1200.

[39] Stolf, R., De S Piedade, S.M., Da Silva, J.R., et al., 2012. Water transfer from São Francisco river to semiarid northeast of Brazil: Technical data, environmental impacts, survey of opinion about the amount to be transferred. Engenharia Agricola. 32, 998−1010.

[40] Cappio, D.L.F., 2008. São Francisco river transposition project. Estudos Avançados. 22, 191−194.

[41] Cunha, A.P.M.A., Zeri, M., Leal, K.D., et al., 2019. Extreme drought events over Brazil from 2011 to 2019. Atmosphere. 10(11), 642. DOI: https://doi.org/10.3390/atmos10110642

[42] Baker, S.A., Wood, A.W., Rajagopalan, B., 2019. Developing subseasonal to seasonal climate forecast products for hydrology and water management. Journal of the American Water Resources Association. 55(4), 1024–1037. DOI: https://doi.org/10.1111/1752-1688.12746

[43] Funk, C., Peterson, P., Landsfeld, M., et al., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific Data. 2, 150066. DOI: https://doi.org/10.1038/sdata.2015.66

[44] Nogueira, S.M.C., Moreira, M.A., Volpato, M.M.L., 2018. Evaluating precipitation estimates from Eta, TRMM and CHIRPS data in the south-southeast region of Minas Gerais state-Brazil. Remote Sens (Basel). 10(2), 313. DOI: https://doi.org/10.3390/rs10020313

[45] Dalagnol, R., Gramcianinov, C.B., Crespo, N.M., et al., 2022. Extreme rainfall and its impacts in the Brazilian Minas Gerais state in January 2020: Can we blame climate change?. Climate Resilience and Sustainability. 1(1), e15. DOI: https://doi.org/10.1002/cli2.15

[46] Paredes-Trejo, F.J., Barbosa, H.A., Lakshmi Kumar, T.V., 2017. Validating CHIRPS-based satellite precipitation estimates in Northeast Brazil. Journal of Arid Environments. 139, 26–40. DOI: https://doi.org/10.1016/j.jaridenv.2016.12.009

[47] de Oliveira-Júnior, J.F., da Silva Junior, C.A., Teodoro, P.E., et al., 2021. Confronting CHIRPS dataset and in situ stations in the detection of wet and drought conditions in the Brazilian Midwest. International Journal of Climatology. 41(9), 4478–4493. DOI: https://doi.org/10.1002/joc.7080

[48] Anderson, L.O., Neto, G.R., Cunha, A.P., et al., 2018. Vulnerability of Amazonian forests to repeated droughts. Philosophical Transactions of the Royal Society B: Biological Sciences. 373(1760), 20170411. DOI: https://doi.org/10.1098/rstb.2017.0411

[49] Costa, J., Pereira, G., Siqueira, M.E., et al., 2021. VALIDATION OF RECIPITATION DATA ESTIMATED TO BRAZIL. Revista Brasileira De Climatologia. 24. (in Portuguese). DOI: https://doi.org/10.5380/abclima.v24i0.60237

[50] Marengo, J.A., Liebmann, B., Grimm, A.M., et al., 2012. Recent developments on the South American monsoon system. International Journal of Climatology. 32(1), 1–21. DOI: https://doi.org/10.1002/joc.2254

[51] Cunningham, C., 2020. Characterization of dry spells in Southeastern Brazil during the monsoon season. International Journal of Climatology. 40(10), 4609–4621. DOI: https://doi.org/10.1002/joc.6478

[52] Nastos, P.T., Zerefos, C.S., 2009. Spatial and temporal variability of consecutive dry and wet days in Greece. Atmospheric Research. 94(4), 616–628. DOI: https://doi.org/10.1016/j.atmosres.2009.03.009

[53] Almeida, C.T., Oliveira-Júnior, J.F., Delgado, R.C., et al., 2017. Spatiotemporal rainfall and temperature trends throughout the Brazilian Legal Amazon, 1973–2013. International Journal of Climatology. 37(4), 2013–2026. DOI:

https://doi.org/10.1002/joc.4831

[54] Gatti, L.V., Basso, L.S., Miller, J.B., et al., 2021. Amazonia as a carbon source linked to deforestation and climate change. Nature. 595(7867), 388–393. DOI: https://doi.org/10.1038/s41586-021-03629-6

[55] Da Silva, P.E., Santos e Silva, C.M., Spyrides, M.H.C., et al., 2019. Precipitation and air temperature extremes in the Amazon and northeast Brazil. International Journal of Climatology. 39(2), 579–595. DOI: https://doi.org/10.1002/joc.5829

[56] Charette, M., Berrang-Ford, L., Coomes, O., et al., 2020. Dengue incidence and sociodemographic conditions in Pucallpa, Peruvian Amazon: What role for modification of the dengue-temperature relationship? American Journal of Tropical Medicine and Hygiene. 102(1), 180–190. DOI: https://doi.org/10.4269/ajtmh.19-0033

[57] Jiménez-Muñoz, J.C., Sobrino, J.A., Mattar, C., et al., 2013. Spatial and temporal patterns of the recent warming of the Amazon forest. Journal of Geophysical Research Atmospheres. 118(11), 5204–5215. DOI: https://doi.org/10.1002/jgrd.50456

[58] Li, D., Yuan, J., Kopp, R.E., 2020. Escalating global exposure to compound heat-humidity extremes with warming. Environmental Research Letters. 15(6), 064003. DOI: https://doi.org/10.1088/1748-9326/ab7d04

[59] JWGFVR, 2015. WWRP/WCRP/WGEN Joint Working Group on Forecast Verification Research Web Page. Available from: https://www.cawcr.gov.au/projects/verification/#Methods_for_probabilistic_forecasts (cited 15 March 2021).

[60] Pallotta, M., 2018. Extremos de temperatura no Centro-Sul do Brasil: Climatologia, padrões sinóticos e impactos ao conforto térmico. TEMPERATURE EXTREMES IN SOUTH CENTRAL BRAZIL: CLIMATOLOGY, SYNOPTIC PATTERNS AND IMPACTS TO THERMAL COMFORT [PhD Thesis]. São José dos Campos, SP: INPE. 226p. (in Portuguese).

[61] Della-Marta, P.M., Haylock, M.R., Luterbacher, J., et al., 2007. Doubled length of western European summer heat waves since 1880. Journal of Geophysical Research Atmospheres. 112 (D15). DOI: https://doi.org/10.1029/2007JD008510

[62] Xie, J., Yu, J., Chen, H., et al., 2020. Sources of subseasonal prediction skill for heatwaves over the Yangtze River basin revealed from three S2S models. Advances in Atmospheric Sciences. 37(12), 1435–1450. DOI: https://doi.org/10.1007/s00376-020-0144-1

[63] Kelly, P., Mapes, B., 2016. February drying in Southeastern Brazil and the Australian Monsoon: Global mechanism for a regional rainfall feature. Journal of Climate. 29(20), 7529–7546. DOI: https://doi.org/10.1175/JCLI-D-15-0838.1

[64] Marengo, J.A., Nobre, C.A., Seluchi, M.E., et al., 2015. Drought and water crisis during 2014-2015 in São Paulo State. Revista USP. 106, 31−44.

[65] Nobre, C.A., Marengo, J.A., Seluchi, M.E., et al., 2016. Some characteristics and impacts of the drought and water crisis in southeastern Brazil during 2014 and 2015. Journal of Water Resource and Protection. 8(2), 252–262. DOI: https://doi.org/10.4236/jwarp.2016.82022

[66] Diamantopoulou, M.J., Georgiou, P.E., Papamichail, D.M., 2006. Daily reservoir inflow forecasting using time delay artificial neural network models. Proceedings of the 2006 IASME/WSEAS International Conference on Energy & Environmental Systems; 8–10 May 2006; Chalkida, Greece. pp. 1–6.

[67] Belayneh, A., Adamowski, J., 2013. Drought forecasting using new machine learning methods. Journal of Water and Land Development. 18(9), 3–12. DOI: https://doi.org/10.2478/jwld-2013-0001

[68] Hwang, J., Orenstein, P., Cohen, J., et al., 2019. Improving subseasonal forecasting in the western U.S. with machine learning. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; July 2019; pp. 2325–2335. DOI: https://doi.org/10.1145/3292500.3330674

[69] WMO/WWRP, WCRP, S2S, 2021. Challenge to improve Sub-seasonal to Seasonal Predictions using Artificial Intelligence. Available from: https://s2s-ai-challenge.github.io (cited 23 September 2021).