https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

A Novel Framework for Analyzing Rainy Season Dynamics in semi-arid environments: A case study in the Peruvian Rio Santa Basin

Lorenz Hänchen[1], Emily Potter[2], Cornelia Klein[3], Pierluigi Calanca[4], Fabien Maussion[5], Wolfgang Gurgiser[6], and Georg Wohlfahrt[1]

[1]Universität Innsbruck, Institut für Ökologie, Innsbruck, Austria
[2]University of Sheffield, Department of Geography, Sheffield, UK
[3]UK Centre for Ecology and Hydrology, Wallingford, UK
[4]Agroscope Reckenholz, Zurich, Switzerland
[5]University of Bristol, School of Geographical Sciences, UK
[6]Universität Innsbruck, Institut für Atmosphären- und Kryosphärenwissenschaften, Innsbruck, Austria

Correspondence: Lorenz Hänchen (lorenz.haenchen@uibk.ac.at)

Abstract. In semi-arid regions, the timing and duration of the rainy season determine plant water availability, which directly impacts food security. Rainy season metrics, which aim to define and, in some cases, predict the onset and end of rains can support agricultural planning, such as scheduling planting dates and managing water resources. However, these metrics based on precipitation time series do not always accurately reflect plant water availability, and the variety of available metrics can complicate the selection of the most suitable one. This study demonstrates that rainy season metrics are more useful for agricultural purposes when their parameters are calibrated using local vegetation data. Furthermore, a metric's ability to capture observed vegetation variability can indicate its applicability over larger spatial or temporal scales. We test this hypothesis in the semi-arid Rio Santa basin in the Peruvian Andes by evaluating seven common rainy season metrics, both calibrated and uncalibrated, against land surface phenology data obtained from 18 years of satellite-derived Normalised Difference Vegetation Index (NDVI) data. Additionally, we introduce a new bucket-type metric that incorporates a simplified water balance, considering both accumulation and storage. To test the robustness of the metrics under future climate scenarios, we examine the sensitivity of these metrics to variations in rainfall intensity and frequency using statistically downscaled CMIP5 rainfall data for historical (1981–2018) and future (2019–2100) periods under RCP 4.5 and 8.5 scenarios. Our results show that calibrating metrics using vegetation data improves their consistency in capturing the start and end dates of the rainy season. The newly introduced bucket metric outperforms the other metrics in both accuracy and robustness. However, some established metrics exhibit sensitivities that raise concerns about their applicability under potential shifts in rainfall patterns due to climate change. Overall, CMIP5 projections reveal no consistent trends in rainy season onset and only a slight delay in rainy season end, with inter-annual variability and ensemble spread being the dominant factors. Our findings highlight the importance of calibrating metrics and stress-testing them across various climate conditions to ensure their agricultural relevance. The framework introduced here can easily be adapted for application in other semi-arid regions.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

1 Introduction

In semi-arid regions, people's livelihoods are closely linked to seasonal water availability, relying strongly on the timing of the rainy season (Warner et al., 2012). Forecasting the local to regional onset and end of the rainy season is a crucial requirement in agriculture, tourism, water management and hydro-electricity generation while changes to the timing of the rainy season are frequently used as a measure of climate change (e.g. Zampieri et al., 2023). Previously, a variety of approaches for numerically determining the onset and end of rainy seasons from precipitation time series have been used in regions with distinct seasonalities of rainfall, hereafter named metrics (e.g. Bombardi et al., 2019b; Fitzpatrick et al., 2015; Sedlmeier et al., 2023). Broadly, these metrics consist of threshold-based approaches (Sedlmeier et al., 2023) which must be configured for each region, or time series inflection point approaches (hereafter objective metrics) which, in theory, are applicable to any region with a distinct hygric seasonality (Liebmann et al., 2007; Liebmann and Marengo, 2001). The latter have been previously used to create a global dataset of rainy season dynamics (Bombardi et al., 2019a). Furthermore, specialized methods have been designed for regions with bimodal rainy seasons (e.g. Dunning et al., 2016) or for regions with high spatiotemporal variability of rainfall by employing data manipulation approaches such as Principal Components Analysis (Camberlin and Diop, 2003), Standard Precipitation Index (Silva et al., 2008), two-phase linear regression (Cook and Buckley, 2009) or a flexible definition of the hydrological year to account for spatial and interannual variability in certain regions (Ferijal et al., 2022; Seregina et al., 2018, to name a few).

The resulting onsets and ends of rainy seasons can vary considerably depending on whether the methods were tailored to specific rain-gauge data, crop requirements or larger-scale characterization of temporal monsoon developments (Fitzpatrick et al., 2015; Sedlmeier et al., 2023). Often, authors emphasize the importance of determining rainy season characteristics for either agricultural planning, monitoring of ecosystems, assessments of temporal water availability in the light of a changing climate or water management topics in general. However, to the best of our knowledge, strategies to validate rainy season metric outputs based on independent data are currently lacking. In this context, the usability of these rainy season metrics from the perspective of the needs of water users and managers remains unknown. Furthermore, other aspects such as the importance of legacy effects beyond one hydrological year or the sensitivity of rainy season metrics to the alteration of the hydrological cycle have so far not been assessed. Regarding such independent data to assess the validity of rainy season metrics in semi-arid regions, spectral vegetation indices (proxies for land surface greenness) from satellite data are a promising candidate for calibration due to their availability in high spatio-temporal resolution. Additionally, precipitation measurements are subject to uncertainties (e.g. Kidd et al., 2017; Pollock et al., 2018). This emphasizes the need for integration of a reliable validation strategy as otherwise timings derived from rainy season metrics might be inappropriate from a practitioner's perspective or deduced climatic changes might be interpreted in a misleading way. We argue that incorporating an independent metric validation scheme based on vegetation development provides three crucial advantages: Firstly, validated and calibrated rainy season metrics align directly with local vegetation responses to changes in water availability. Secondly, time series of precipitation or other water-related variables can be tested regarding their quality. Lastly, previously published metrics, often designed for

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

specific data and locations, can be assessed for their applicability in different regions.

In complex regions like the Rio Santa basin in the Tropical Peruvian Andes, where climate-related challenges are anticipated to affect both society and ecosystems such adaptability is key for addressing these challenges. Precipitation in the Rio Santa basin is projected to increase in the future (Potter et al., 2023), but the potential changes to the timings of the rainy season remain uncertain and have not yet been assessed. In this context, identifying robust methods to address this uncertainty is of high importance. Numerous studies on past rainy season dynamics in the broader region reveal high inter-annual variability in onset, with generally non-significant or weak longer-term trends (Garcia et al., 2007; Giráldez et al., 2020; Gurgiser et al., 2016; Sedlmeier et al., 2023). Across those studies, the end of the rainy season is notably less variable, while showing no significant changes historically. For the Rio Santa basin specifically, Hänchen et al. (2022) note a delayed end of the growing season between 2000 and 2020, indicating increased water availability difficult to detect from both satellite or gauge rainfall data due to the small rainfall totals during the early dry season. Additionally, the regional hydroclimate experiences a complex interaction with El Niño Southern Oscillation (ENSO), where the overall amount of rainy season precipitation, in most, but not all years, increases (decreases) with La Niña (El Niño) (e.g. Maussion et al., 2015; Vuille et al., 2008). At the same time, there are indications that El Niño conditions might cause seasonal rainfalls to start earlier, thus increasing overall plant water availability even though peak season rainfalls are reduced (Hänchen et al., 2022).

To test our rainy season metric calibration strategy, we focus on the Rio Santa basin, building upon previous studies in the region: We make use of the convection-permitting, bias-corrected Weather Research and Forecasting (WRF) precipitation data and statistically-downscaled CMIP5 precipitation (Potter et al., 2023) to calculate rainy season characteristics for the Rio Santa basin and two other precipitation datasets for comparison. To calibrate and validate the outputs of the metrics, we utilize Land Surface Phenology (LSP) data for the period 2000–2018, derived from the temporal development of the remotely sensed Normalized Difference Vegetation Index (NDVI) provided by Hänchen et al. (2022). Their research successfully established NDVI as a proxy for water availability in the Rio Santa basin. Furthermore, other studies have demonstrated the applicability of NDVI in understanding precipitation variability in the central Peruvian Andes (Quiroz et al., 2011; Yarleque et al., 2016). LSP was assessed by applying a relative threshold to the Rio Santa basin average seasonal cycle of vegetation greenness (c.f. Caparros-Santiago et al., 2021) to obtain the start and end of the growing season based on which we subsequently calibrate all rainy season metrics.

The principal objective of this study is to showcase a novel framework for characterizing the rainy season, emphasizing the importance of both using a calibration strategy for inferred rainy season onsets and ends. In addition, we also test the sensitivity of rainy season metrics to plausible changes in rainfall intensity and frequency, as may occur as a consequence of global warming. This framework is designed to help us better comprehend variations in water availability within semi-arid regions, with a broader applicability beyond the Rio Santa basin. Regarding the Rio Santa basin, we aim to provide insights into past and future changes. We achieve this by:

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

1. Establishing an approach to derive reliable rainy season metric outputs from several existing methodologies from precipitation time series by calibrating them using LSP data.

2. Introducing a novel methodology to the community to derive rainy season indicators, where we simulate water availability in a simplified fashion using only precipitation time series as input and a number of calibrated constants.

3. Testing the response and sensitivity of each metric to physically plausible changes in the rainy season.

4. Analyzing changes of the temporal evolution of the rainy season in the Rio Santa Basin. By making use of the aforementioned calibrated metrics, we explore past (since 1981) and future (until 2100) changes of the onset and end of the rainy season based on CMIP5 models for the region.

2 Material and methods

2.1 Study area

The Upper Rio Santa basin (also: Callejón de Huaylas), situated in the central Tropical Andes of Peru (an overview is presented in Figure 1), encompasses a complex hydroclimate system governed by the topography, the numerous abundance of glaciers on the Cordillera Blanca (eastern slopes of the valley), the temporal evolution of the South American monsoon (Espinoza et al., 2020; Garreaud, 2009; Klein et al., 2023a) and its interaction with ENSO (e.g. Hänchen et al., 2022; Maussion et al., 2015). The dynamics of the rainy season are crucial for regional water resources and agriculture, as it provides water for irrigation, energy production, and maintenance of ecosystems (e.g. Dextre et al., 2022; Drenkhan et al., 2022). There has been much attention on the past, present and future alteration of water availability in response to changes in glacial melt (e.g. Bury et al., 2010; Drenkhan et al., 2015; Fyffe et al., 2021). Small-scale farmers however often have no or limited access to glacier-fed river runoff and perceive increasing challenges related to precipitation seasonality (Gurgiser et al., 2016) and/or water quality (Rangecroft et al., 2023). Recently, more efforts to understand and monitor several aspects of precipitation changes in the Rio Santa basin have been undertaken, but it remains challenging to derive successful mitigation strategies (Hänchen et al., 2022; Klein et al., 2023b; Mateo et al., 2022; Potter et al., 2023).

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

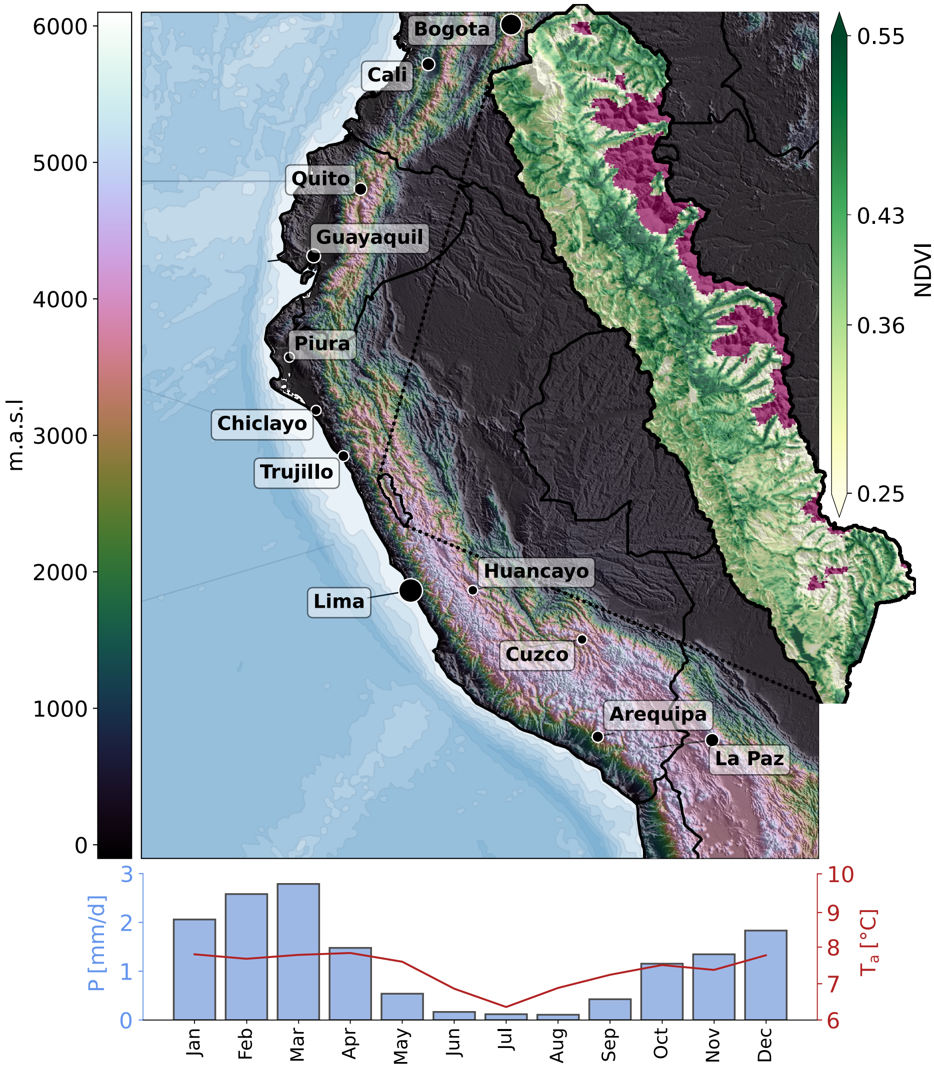

Figure 1. Overview of the study area. The large map shows the topography of the Andes (based on SRTM data, USGS EROS Archive, 2021), administrative borders and larger towns in the greater region. The enlarged area of the Rio Santa basin shows the long-term (2000-2018) average NDVI of each pixel; no-data areas, mostly referring to land-covers such as Ice or Bare rocks are shown in magenta color. Additionally, the Climograph at the bottom shows the seasonality of precipitation and temperature derived from spatially averaged WRF data for 1981-2018 (Potter et al., 2023) for the Rio Santa basin.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

2.2 Data

As a target dataset for calibration, we use Land Surface Phenology (LSP) data by Hänchen et al. (2022) from 2000 to 2018 derived from MOD13Q1 (Didan, 2015a) and MYD13Q1 (Didan, 2015b) NDVI time series in 250 m spatial resolution (c.f. Fig. 1). The data were i) filtered on quality assurance criteria, ii) gap-filled and smoothed using a Gaussian process regression algorithm (Belda et al., 2020), iii) masked based on unimodal seasonal vegetation development and land-cover data. The start (hereafter SOS_{NDVI}) and end (hereafter EOS_{NDVI}) of the growing season were derived as the day where the processed NDVI data reaches 30% of its seasonal amplitude (Hänchen et al., 2022). Here, we spatially averaged the data to the extent of the Rio Santa basin.

Additionally, we employ different precipitation data: Central to our analysis are WRF bias-corrected regional climate model data, published by Potter et al. (2023), providing consistent precipitation data at 4 km grid spacing from 1981 to 2018. We further use their corresponding statistically-downscaled CMIP5 projections for a high (RCP 8.5) and a stabilization (RCP 4.5) global emissions scenario from 2019 to 2100 to assess future changes in the rainy season in the Rio Santa basin and to have a large data basis for metric sensitivity analysis. In addition, we use the gridded Climate Hazards InfraRed Precipitation with Station data (CHIRPS, Funk et al., 2015), which are provided in $0.05° \times 0.05°$ spatial and daily temporal resolution between 1981 and 2018. The gridded data are compared to an average of three local weather stations operated by the National Meteorological and Hydrological Service of Peru (SENAHMI), which were sufficiently gap-free, all located along the valley floor of the Rio Santa basin (Hunziker et al., 2017). Both gridded precipitation datasets were restricted to the geographical coverage of the available NDVI pixels as seen in Fig. 1 within the Rio Santa basin to acknowledge that high precipitation sums in the elevated Cordillera Blanca regions (e.g., glacierized or bare ground land-cover) do not align with vegetation responses and then spatially averaged (i.e. no spatial dimension). We excluded leap year days (29^{th} Feb) and performed the analysis based on a hydrological year definition suitable for the Rio Santa basin starting from the 1^{st} Sep and ending on the 31^{st} of Aug of the subsequent year.

As vegetation responses to rainfall are not necessarily immediate, the lag between the NDVI and the precipitation data must be accounted for to allow to use the LSP data as targets. We therefore determined this lag between the spatial averages of smoothed NDVI and a 12-week rolling average for each of the three precipitation time series by utilizing a cross-correlation function to identify the lag with best alignment by the index (days) of the highest Pearson correlation coefficient. Finally, we subtracted the determined lags for each precipitation dataset from the LSP data before further analysis (c.f. Fig. A1).

2.3 Rainy season metric calculation

Here, we apply the same threshold-based metrics which Sedlmeier et al. (2023) compiled for the Southern Peruvian Andes, hereafter called Gurgiser (Gurgiser et al., 2016), Climandes (Sedlmeier et al., 2023), Garcia (Garcia et al., 2007), FP (Frere and Popov, 1986) and JD (Jolliffe and Sarria-Dodd, 1994) and tune them specifically for the Rio Santa basin and each precipitation dataset. The rationale of each rainy season metric can be found in Table 1. The first four metrics (Gurgiser, Climandes, Garcia

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

and JD) to derive the rainy season onset (hereafter RSO) all use some combination of four conditions: 1) The day of the onset has to have precipitation above a threshold value; 2) The total precipitation in a defined period after the onset must exceed a certain sum; 3) There must be a minimum number of wet days in a defined period after the onset; 4) There must be no continuous periods of dry days over a certain length within a defined period after the onset. Gurgiser uses conditions 1, 2 and 3, Climandes conditions 1, 2 and 4, Garcia conditions 2 and 4, and JD conditions 3, 2 and 4. FP use a different approach by dividing a 30-day period into terciles, where each tercile must exceed a certain total precipitation, similar to condition 2 of the other metrics. For calibration, our implementation of FP involves examining the first, second, and third tercile, while adjusting the length of the period and total precipitation thresholds.

While the FP and JD metrics only apply to the onset of the rainy season, the three remaining threshold-based metrics use a common technique to determine the end of the rainy season (hereafter RSE): 1) defining a precipitation threshold for the potential day of the rainy season end and 2) defining a threshold for the precipitation sum over a number of subsequent days. The Garcia metric omits the first criterion. For comparison, we also tested two other, non-threshold-based metrics: The widely established metric by Liebmann and Marengo (2001), hereafter named LM, which accumulates seasonal rainfall against the average of the hydrological year. Then, the days of the minimum and maximum are considered the onset and end of the rainy season. The method by Cook and Buckley (2009), hereafter CB, employs a change-point detection method, fitting a two-phase linear regression iteratively over i) the first 250 and ii) the last 200 days of the hydrological year independently. By minimizing the sum of squares of residuals, the best fit for the regressions is found and the changepoints determine the onset and end of the rainy season. We implemented this approach using the python package pwlf (Jekel and Venter, 2019). Approaches considering data other than rainfall time series, combining threshold-based approaches with fuzzy-logic (Laux et al., 2008) or Pentad-based approaches (e.g. Giráldez et al., 2020; Marengo et al., 2001) are beyond the scope of this study and thus not included.

Using the NDVI-derived SOS_{NDVI} and EOS_{NDVI} as targets, we calibrated each of the threshold-based metrics and the our novel metric (defined below) by changing their parameters (c.f. Table 1) for each of the three precipitation time series through the application of a Differential Evolution optimization algorithm (Storn and Price, 1997), with parameters constrained to physically plausible boundaries. Due to the limited number of growing seasons available (i.e., 18), splitting the data into calibration and validation periods would not have allowed to obtain robust correlation and was therefore omitted. To allow for the robust and efficient processing of a large number of timeseries regarding the threshold-based metrics, we generally start the iterative search for the RSE starting from the previously derived RSO. Additionally, when a RSE is found within the 90 subsequent days following the RSO, the iterative search is continued to account for erroneous detection of dry spells in the early rainy season.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

Name (Reference)	RSO (original)	RSO (adapted)	RSE (original)	RSE (adapted)
Gurgiser (Gurgiser et al., 2016)	$P_d > 0\,mm$ $\sum P_{d:d+7} > 10\,mm$ $N[P_{d:d+30} > 0\,mm] \geq 11$	$P_d > \alpha_p^1$ $\sum P_{d:d+\alpha_d^3} > \alpha_p^2$ $N[P_{d:d+\alpha_d^5} > \alpha_p^6] \geq \alpha_d^4$	$P_d = 0\,mm$ $\sum P_{d:d+46} < 10\,mm$	$P_d \leq \alpha_p^1$ $\sum P_{d:d+\alpha_d^3} < \alpha_p^2$
Climandes (Sedlmeier et al., 2023)	$P_d > 1\,mm$ $\sum P_{d:d+5} \geq 8\,mm$ $N_c[P_{d:d+30} < 0.1\,mm] \leq 6$	$P_d > \alpha_p^1$ $\sum P_{d:d+\alpha_d^3} \geq \alpha_p^2$ $N_c[P_{d:d+\alpha_d^5} < \alpha_p^6] \leq \alpha_d^4$	$P_d \leq 1\,mm$ $\sum P_{d:d+30} < 16\,mm$	$P_d \leq \alpha_p^1$ $\sum P_{d:d+\alpha_d^3} < \alpha_p^2$
Garcia (Garcia et al., 2007)	$\sum P_{d:d+3} > 20\,mm$ $N_c[P_{d:d+30} < 0.1\,mm] \leq 9$	$\sum P_{d:d+\alpha_d^2} > \alpha_p^1$ $N_c[P_{d:d+\alpha_d^4} < \alpha_p^5] \leq \alpha_d^3$	$P_{d:d+20} \leq 2^*\,mm$	$P_{d:d+\alpha_d^2} \leq \alpha_{pr}^1$
FP (Frere and Popov, 1986)	$\sum P_{d:d+10} \geq 25\,mm$ $\sum P_{d+10:d+20} \geq 20\,mm$ $\sum P_{d+20:d+30} \geq 20\,mm$	$\sum P_{d:d+\frac{1}{3}\alpha_d^2} \geq \alpha_p^1$ $\sum P_{d+\frac{1}{3}\alpha_d^2:d+\frac{2}{3}\alpha_d^2} \geq \alpha_p^3$ $\sum P_{d+\frac{2}{3}\alpha_d^2:d+\alpha_d^2} \geq \alpha_p^4$	Onset only	Onset only
JD (Jolliffe and Sarria-Dodd (1994), originally based on Stern et al. (1981))	$N[P_{d:d+5} > 0.1\,mm] \geq 3$ $\sum P_{d:d+5} \geq 25\,mm$ $N_c[P_{d:d+30} < 0.1\,mm] \leq 7$	$N[P_{d:d+\alpha_d^2} > \alpha_p^1] \geq \alpha_d^3$ $\sum Pr_{d:d+\alpha_d^2} > \alpha_p^4$ $N_c[P_{d:d+\alpha_d^6} < \alpha_p^1] \leq \alpha_d^5$	Onset only	Onset only
LM (Liebmann and Marengo, 2001)	d such that $A_d = \min A$ where $A = \sum_{n=1}^{d} P_n - P$		d such that $A_d = \max A$ where $A = \sum_{n=1}^{d} P_n - P$	
CB (Cook and Buckley, 2009)	Linear regression fitting and change point detection to determine when daily precipitation changes from decreasing each day to increasing each day (onset) and from increasing to decreasing (end).			

Table 1. Rainy season metric rationales where d is the day of the year marking the onset or end of the rainy season, P_d is the precipitation on day d (in mm), and for example, $\sum P_{d:d+6}$ is the sum of precipitation on each day from the onset to 6 days after the onset. $N[P_{d:d+30} > 0mm]$ is the number of days with precipitation over 0 mm in a 30 day period. Some metrics use N_c instead of N, which represents continuous dry days, e.g. $N_c[P_{d:d+30} < 0.1\,mm] < 7$ is the condition that no dry spells of more than 7 days occur in the 30 days after the onset. The parameters α^1 to α^n are the tuneable parameters of each metric, with α_p denoting a precipitation threshold in mm and α_d denoting an integer number of days. A is the cumulative sum of anomalous precipitation from day 1 to d and P is the annual average daily precipitation. * For $\text{RSE}_{\text{Garcia}}$, which is published as $P_{d:d+20} = 0\,mm$ we used a value of 2 mm instead as 20 consecutive days with zero precipitation are not present in any of the datasets.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

2.4 A new rainy season metric: The "bucket" metric

Finally, we introduce a novel approach, which attempts to simulate a simplified water balance by consecutive balancing of daily input through rainfall and output through constant evapotranspiration:

$$BWC(t) = \begin{cases} BWC(t-1) + \frac{BD}{\rho} \cdot (P(t) - ET), & \text{if } BWC_{\text{mn}} \leq BWC(t) \leq BWC_{\text{mx}} \\ BWC_{\text{mn}}, & \text{if } BWC(t) < BWC_{\text{mn}} \\ BWC_{\text{mx}}, & \text{if } BWC(t) > BWC_{\text{mx}} \end{cases} \quad (1)$$

where BWC [m^3/m^3] represents the Bucket water content at time t, BD is the bucket depth [m] and ρ is the water density, here constant as 1000 kg/m^3], P [mm/day] is the precipitation input and ET [mm/day] is the daily output.
Note that ET [mm/day] is inspired by evapotranspiration but does not represent the actual physical process as the simplistic design of the metric considers it to be constant over the whole hydrological year, partly integrates other hydrological components such as runoff and thus within- and between seasonal variation is not directly accounted for. The metric starts at day d = 0 at an initial BWC (BWC_{ini}). The model is constrained as no further evaporation occurs as soon as a minimum value (BWC_{mn}) is reached. Similarly, a maximum value (BWC_{mx}) is defined where no more water is accumulated – any surplus conceptually runs off or drains from the bucket. The parameters BWC_{in}, BWC_{mx}, BWC_{mn}, t_{RSO}, t_{RSE}, BD and ET need to be tuned and do not change over time. For each season, rainy season onset and end are then determined based on two previously calibrated BWC thresholds denoted as t_{RSO}, t_{RSE} in Figs. 2 & 3. In Fig. A2, an example of a full BWC and precipitation time series, along with the derived RSOs and RSEs, is shown.
In contrast to the metrics previously introduced which calculate each season independently, the bucket metric is able to calculate over the complete multi-year time series, allowing to incorporate legacy information about water availability prior to the rainy season of interest. As for the other approaches, we optimized all parameters according to the corresponding input data. While this approach is inspired by existing simple hydrologic bucket models and thus by actual hydrological processes, our aim is not to accurately represent these, but rather to account for parameters altering plant available water in a simplified fashion.

3 Results & Discussion

3.1 Evaluation of Rainy Season Metrics

We first compare the skill of all considered metrics in predicting the RSO close to respective reference SOS_{NDVI} across all years and datasets. Figure 2a-e illustrates that all calibrated threshold-based metrics consistently predict the lag-corrected SOS_{NDVI} across the three precipitation datasets, demonstrating a strong correspondence and outperforming the initial (i.e. uncalibrated) setup of the metric, which for all threshold-based metrics except JD is lacking correlation and showcasing higher RMSE values. The bucket metric stands out by exhibiting low errors (average RMSE = 8.7 days) and a robust correlation (r^2 = 0.79, on average) across all three input datasets (c.f. Fig. 2, Fig. 4). This is likely related to the fact that the bucket metric was

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

designed to directly determine the lag between rainfall inputs and vegetation responses while the other metrics make use of the cross-correlation maximization. Hence, the bucket metric does account for legacy information between seasons. Although we did not directly observe a deterioration of correlation when removing the legacy information from the metric before parameter optimization, the resulting BWC timeseries was highly unrealistic and unsuitable for transferability (not shown). LM and CB on the other hand demonstrate a relatively low agreement with SOS_{NDVI} with LM showing weak correlations which for CB are missing (c.f. Fig. 2g and h).

Regarding the RSEs, the three threshold-based metrics (Fig. 3a-c) demonstrate a relatively low RMSE (ranging between 8.8 and 14.4) albeit lacking correlation (maximum $r^2 = 0.25$), whereas the bucket metric (Fig. 3d) shows an even lower RMSE (5.6 on average) and a weak correlation ($r^2 = 0.4$ on average), likely related to the bucket metric incorporating non-plant-available water simulated as bucket overflow (see Section 2.4). Gurgiser and Climandes share the same criteria for RSE and thus the resulting calibration is identical (c.f. Table 1 & Fig. 3). The LM and CB metrics show an overall low skill in predicting the lag-corrected EOS_{NDVI} with high errors and LM showing a weak correlation for 2 of 3 datasets (c.f. Fig. 3f). The overall discrepancy across the metrics between skill on predicting SOS_{NDVI} and EOS_{NDVI} (see Fig. 4) may be linked to EOS_{NDVI} displaying low variability (standard deviation, σ = 6.63 days), unlike the higher variability of SOS_{NDVI} (σ = 17.61 days). Additionally, the coupling between precipitation and water availability tends to be more prominent at the onset of the rainy season due to depleted hydrological system storages, resulting in reduced predictive power of rainfall for vegetation development as rainfalls recede.

While each metric shows reasonably high skill for all three precipitation datasets after calibration, the substantial differences in the resulting optimization parameters (c.f. Tables in Fig. 2 a-f & 3 a-d) underscore the necessity of tuning and testing rainy season metrics according to local climatic conditions, specific datasets and target applications. Given proper tuning, the results are comparable even though the metrics follow different logic and use a different number of parameters. Interestingly, among the threshold-based metrics, those with more parameters do not necessarily perform better in terms of error and correlation. For example, RSO_{FP} and RSE_{Garcia}, which have the fewest parameters (four and two, respectively), still show a consistent performance. A systematic test of the relevance of individual parameters is beyond our scope here, especially given the high performance of the bucket metric which is our primary focus here. Finally, our generally skilful results after calibration also illustrate that existing concerns (e.g. MacLeod, 2018) regarding the sensitivity of thresholds-based metrics to rainfall dataset bias and resolution appear to be surmountable if independent reference data are taken into account, rendering these metrics more flexible than is currently appreciated.

Although other authors asserted a strong agreement between the LM method and local threshold-based approaches (e.g. Dunning et al., 2016), our results emphasize that agreement in metric outputs from the same time series alone does not necessarily guarantee the representativeness towards plant growth or suitability for practitioners of any kind. While we acknowledge the effectiveness of the LM method in larger scale climatological rainfall analysis, our analysis shows that it a) exhibits less

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

correspondence with vegetation development than calibrated methods (Figs. 2–4) and b) tends to produce delayed onsets in the specific setting of the Rio Santa basin during extended dry spells following early-season rains (not shown). Similarly, the two-phase regression method (CB) tends to compute late onsets in cases of prolonged dry spells and/or when the development of the rainy season follows a non-linear trajectory (i.e. rainfall increase from the onset towards the peak of the rainy season), making it unsuitable for accurate onset and end determination in the Rio Santa basin in many seasons. Furthermore, the objectivity of this method is limited as the rainy season needs to be split into two sub-seasons, potentially affecting resulting values. Here, we followed the same approach as the original authors, using the first 250 and the last 200 days of each season to determine the dates. Similarly, the metric demonstrates sensitivity to the definition of the hydrological year. For instance, shifting the start of the hydrological year back by two months significantly enhances the correspondence between RSO_{CB} and SOS_{NDVI}, while concurrently diminishing it between RSE_{CB} and EOS_{NDVI} (see Fig. A3 for an example). Given the high variability of the rainy season onset in the tropical Andes, coupled with the aforementioned sensitivity to the climatological year definition, we believe it is advisable to employ a flexible hydrological year approach (c.f. Ferijal et al., 2022) when exploring this method.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

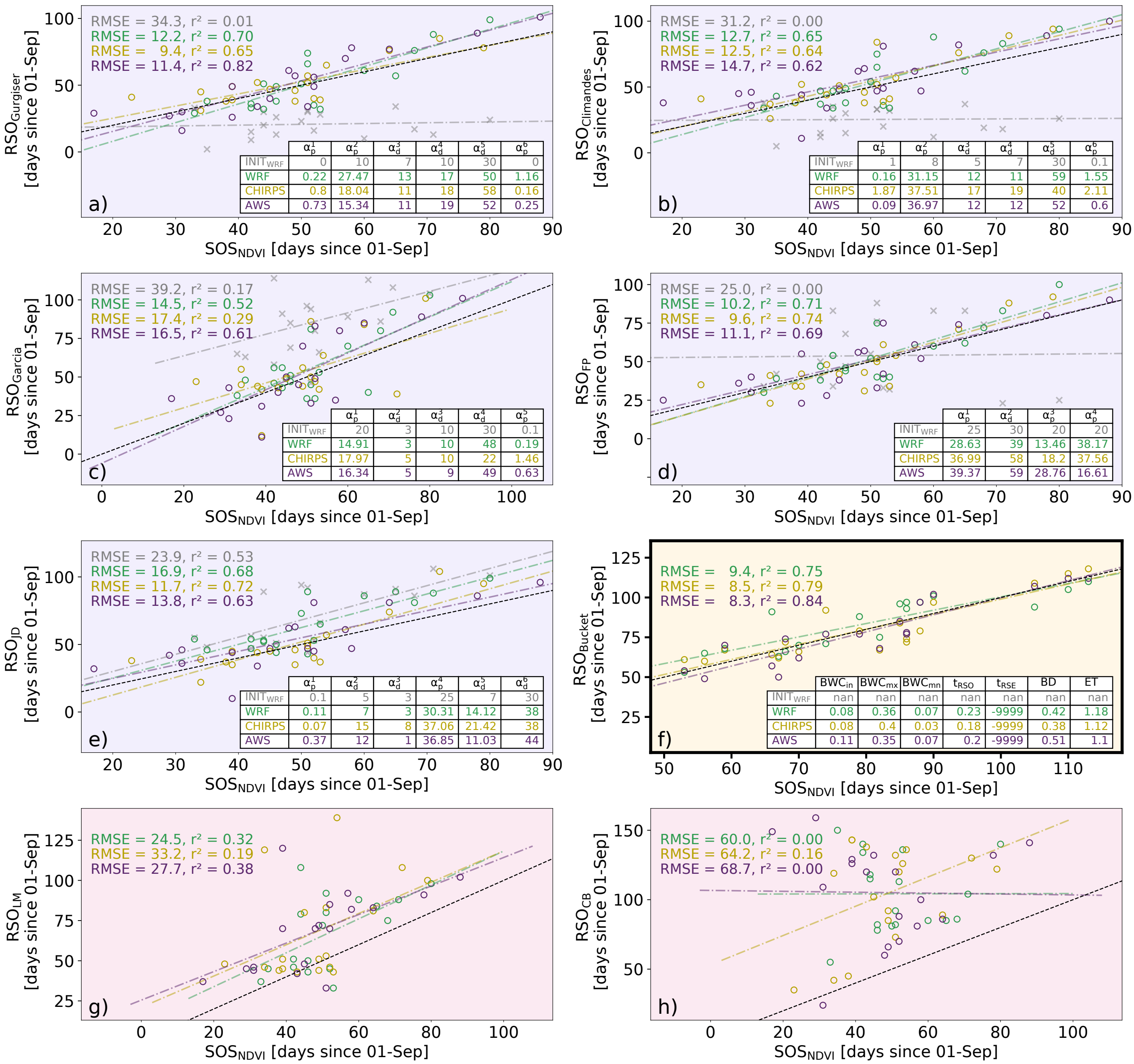

Figure 2. Rainy season onset metric outputs for WRF (teal), CHIRPS (orange) and AWS (purple) precipitation data. For the threshold-based metrics (purple background; panels a-e), results of the evaluation based on WRF but with uncalibrated thresholds as provided by the respective authors are also shown (gray). The novel bucket methodology is highlighted in yellow (panel f) and the objective methods in red (panels g-h). The black dashed line indicates the 1:1 relationship with the SOS_{NDVI}, while the colored lines correspond to the regression of the parameters. Annotated are Root Mean Squared Error (RMSE) and the coefficient of determination (r^2). The tables correspond to the parameters as outlined in Table 1 and the equation in Section 2.4 after calibration for each of the precipitation data. The LM and CB method have no calibration and therefore no table is shown.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

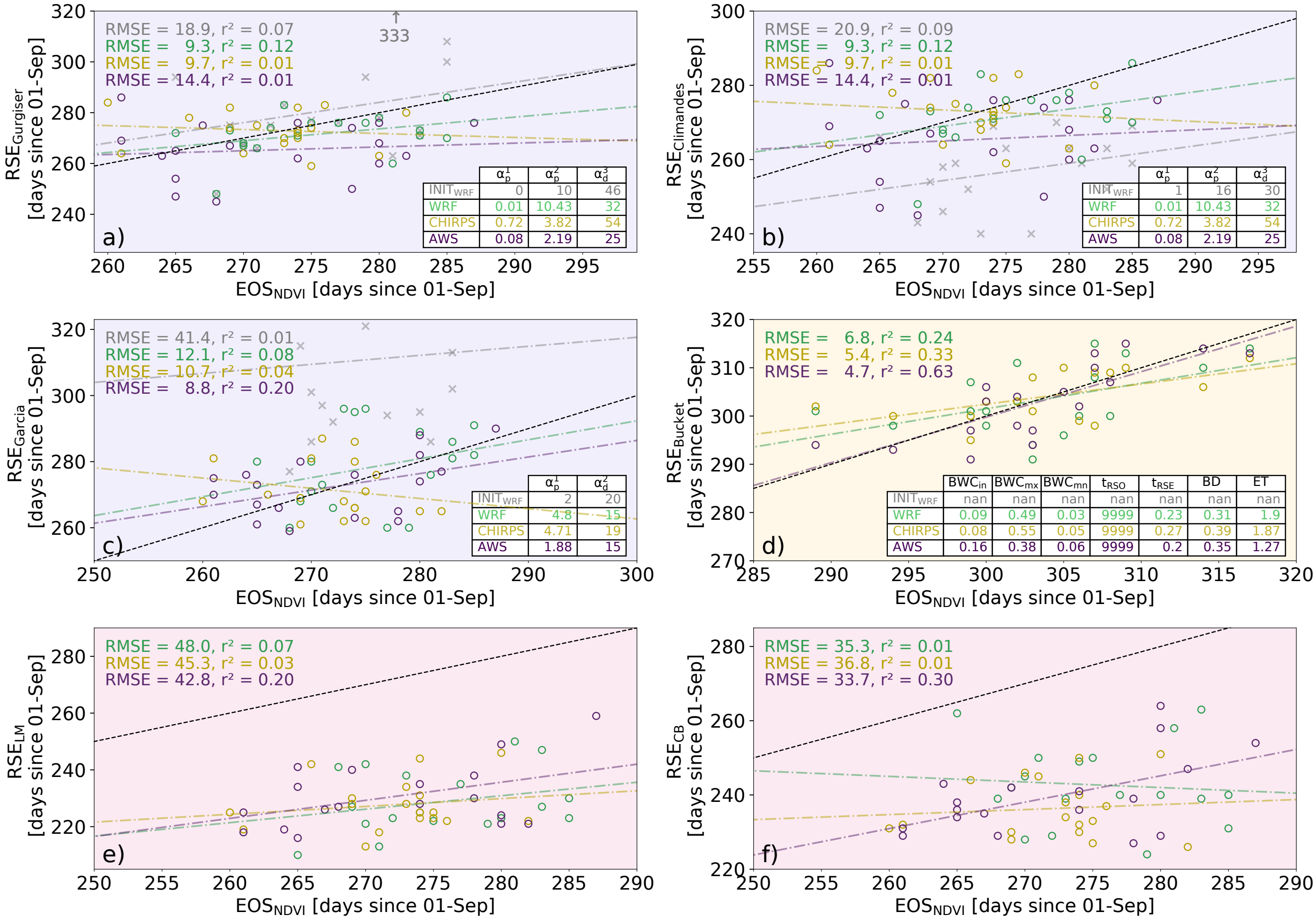

Figure 3. Same as 2 but for Rainy Season End and End of Season.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

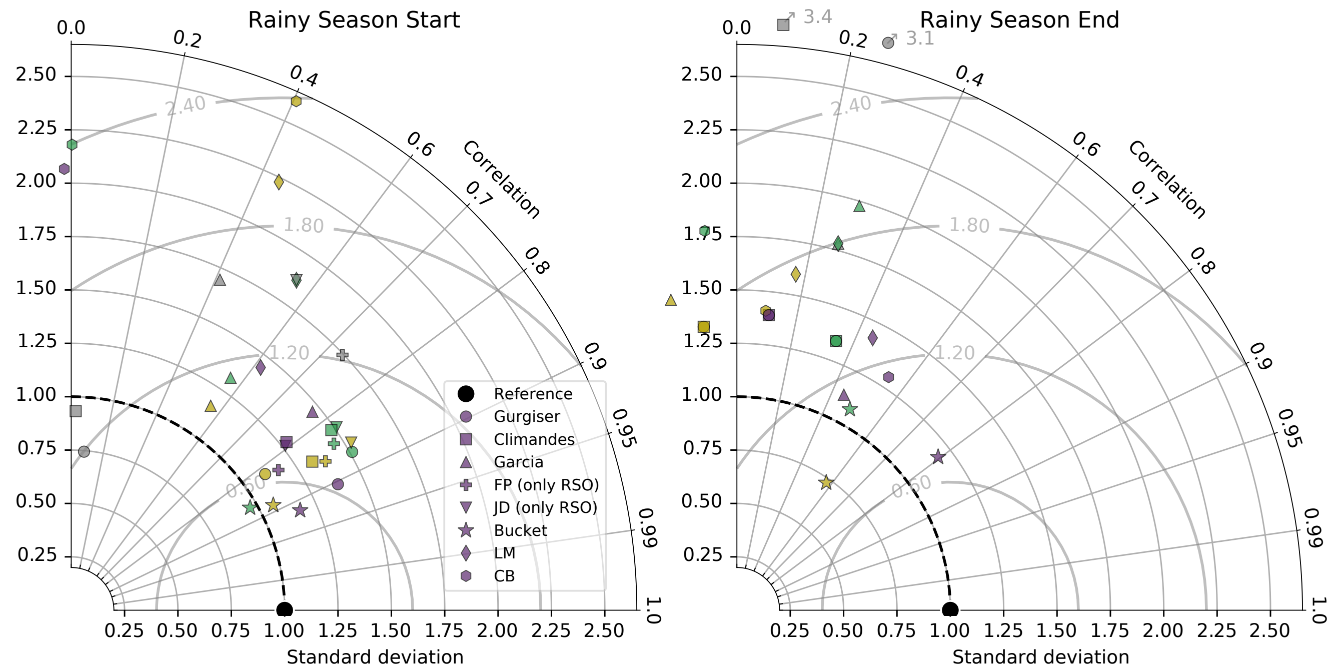

Figure 4. Taylor diagrams for normalized Rainy Season Onset (left) and End (right) for the calibration period 2000-2018. Each data point represents a metric depicted by the symbols, the colors represent the corresponding datasets (grey: WRF with initial metric parameters; teal: WRF calibrated; orange: CHIRPS; purple: AWS). The radial axes represent standard deviation, the azimuthal axis represents the correlation coefficient (r) and the circles the centered root mean squared difference. The black reference dot represents the normalized NDVI-derived SOS_{NDVI} and EOS_{NDVI} standard deviation.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

3.2 Metric sensitivity to rainfall changes

To understand how tuned threshold-based metrics and objective methods respond to hydro-climatological changes, we now utilize the large number of seasons provided by the CMIP5 models to assess sensitivities of the metrics in regard to potential and physically plausible changes of the rainy season. To account for a variety of scenarios, we correlated the rainy season metric outputs (RSO/RSE) calculated for all CMIP5 rainy seasons independent of model, year or scenario with both full hydrological year and sub-seasonal (SON, DJF, MAM, JJA) rainfall sums. The sub-seasonal rainfall sums are referring to the same rainy season RSO/RSE were derived on, meaning that, for example, JJA refers to the dry months after the RSO. The rationale for correlating the seasonal rainfall sums even beyond the period where the RSO typically occurs is to test whether some of the metrics show implausible sensitivities which reduce their usefulness from a practical perspective. Additionally, we used four Expert Team on Climate Change Detection and Indices (ETCCDI) climate indices (Zhang et al., 2011) by Potter et al. (2023). These are the number of dry and wet days (DD, WD), defined as days with precipitation less and greater than 1 mm; the Simple Precipitation Intensity Index (SDII), representing the average daily precipitation on wet days (WD) and the sum of precipitation above the 95th percentile relative to the historical (1980-2018) period (R95pTOT). Utilizing these rainfall sums and ETCCDI climate indices as independent variables we assess sensitivities by applying bin-weighted linear regression. To determine appropriate bin sizes for the regression in an objective fashion, we applied the Freedman-Diaconis rule (Freedman and Diaconis, 1981) to each of the nine independent variables. Figure 5 illustrates the summarized results of these regressions. Full regression plots can be found in Figs. A4 and A5.

In the context of the RSO (Fig. 5a, Fig. A4), all threshold-based metrics and the bucket metric show similar responses for both ETCCDI indices and precipitation sums, while LM and CB exhibit diverging sensitivities. Specifically, an increasing number of dry days (DD) results in a later season onset, while a higher number of wet days (WD) leads to an earlier onset, with weaker correlations for LM and CB. With increasing heavy precipitation (R95pTOT), represented by the sum of precipitation falling above the 95th percentile relative to the control period (1980-2018), the correlation is negative for all threshold-based and the bucket metric, indicating a correlation between earlier rainy season onsets and more heavy precipitation. However, for LM, this relationship is positive, and for CB, the resulting slope is not significant. Similarly, an increase in average precipitation on wet days, represented by the simple precipitation intensity index (SDII), results in an earlier season onset. LM again shows an opposite response, and CB shows just a weak correlation. All metrics except LM are sensitive to increased annual precipitation. With the exception of CB, which shows this correlation in DJF due its generally later onsets (c.f. Fig. 2), all metrics are strongly sensitive to SON precipitation. DJF, MAM, and partly JJA precipitation generally indicate this sensitivity as well, but this is most likely subject to autocorrelation. Notably, the bucket metric shows a stronger sensitivity to dry season (JJA) precipitation, as its design of the metric allows for transferring information regarding water availability between hydrological years. Both LM and CB show a distinct positive correlation to increased MAM precipitation, indicating later rainy season onset. This is problematic because, this correlation is a methodological artefact that does not reflect any physical process related to RSO water availability. This indicates limited metric robustness of the objective metrics to rainy season change. Note that the start of CB is based on the period of the first 250 days of the rainy season, meaning that the metric is based only on information of

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

the period of September 1st to May 8th (March 19th to August 31st for the end).

Similar as previously seen in the calibration results (c.f. Figs. 3, 4b)), the metrics for the end of the rainy season, RSE, show a weaker relationship with climate indices and precipitation sums, represented by lower r^2 values (see Fig. 5b and Fig. A5). For the number of dry (wet) days, all metrics suggest an earlier (later) season end, with the bucket model displaying the strongest sensitivity. For both R95pTOT and SDII, most regressions are insignificant or show weak correlations, with the exception of the bucket model, which suggests a moderate correlation towards later rainy season ends, and CB suggesting the opposite. All metrics except CB suggest a moderate sensitivity to seasonal and total rainfall sums, with Garcia, and LM suggesting a negative slope for SON precipitation (for LM, also DJF). Gurgiser and Climandes suggest a very strong sensitivity to JJA rainfall. The RSE calculated by CB appears to be relatively insensitive to altered rainfall sums, being only significantly correlated to JJA precipitation. Due to the lower overall correlation, interpreting these results is not as straightforward as for the RSO. However, the relatively high correlation of both the calibrated threshold-based and the bucket metric, along with revealing consistent correlations with our process understanding for all indices and precipitation sums, emphasizes their suitability for assessing potential changes in water availability in semi-arid areas such as the Rio Santa basin.

Taken together, the sensitivity analysis reveals that for the RSO, all threshold-based models and the bucket model appear to produce appropriate results, while LM and CB are subject to sensitivities that are likely to hinder a reliable interpretation regarding the temporal manifestation of the rainy season, in particular when rainy season characteristics are expected to change. While less clear for the RSE, the overall message is similar with the bucket metric along with the threshold-based metrics being most reliable.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

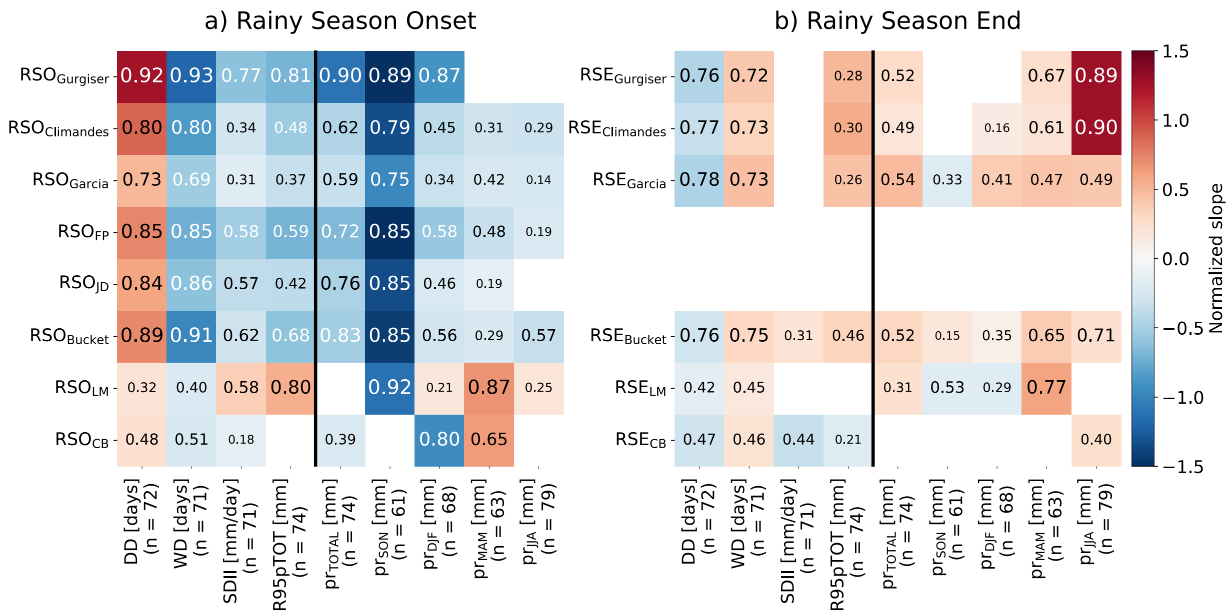

Figure 5. Heatmap of bin-weighted regression slopes with annotated r^2 values between ETCCDI indices and seasonal precipitation sums (independent variables) and rainy season metric derived onset and end (dependent variables). Corresponding bin sizes are noted on the x-axes labels as (n = x). Slope values are normalized and non-significant regressions ($p > 0.01$) are not shown. Full regressions including non-normalized slope values are displayed in Figs. A4 and A5.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

3.3 Past & Future Trends

Finally, we projected future metrics up to 2100 using a statistically-downscaled CMIP5 model ensemble based on historical WRF data, which comprises 30 individual models, and subsequently evaluated the trends for the historical and the future period. As depicted in Fig. 6, the substantial variability observed in the RSO from 2000 to 2018 (average IQR over all 8 metrics = 16.4) seems to have existed similarly, if not more pronounced, in the preceding decades before 2000 in both time series (IQR = 27.0). Regarding the historical RSE, the missing data points in 1989/1990 in three metric outputs (Fig. 7a-c) are due to a dry spell lasting about three months leading to non-fulfilment of metric criteria and thus no-data labelling. Interestingly, LM and CB do not show any anomaly for this event because these metrics do not have information about any form of climatology. Conversely, this is accounted for by the bucket and threshold-based metric as the calibrated parameters represent the average climate of 2000 – 2018, such that extreme cases exceeding the calibration period cannot be informatively processed. We believe this is a desirable feature as for a practitioner this can be more informative than an unrealistic result in such cases. None of the metric outputs suggests a trend for the past period, either for the rainy season onset or the end of the rainy season (c.f. Figs. 6 and 7).

In the following analysis, CMIP5 models were excluded if a rainy season metric produced five or more invalid values out of 81 seasons (2019 – 2100). Invalid values occur whenever the conditions for RSO or RSE are never met or were discarded if an unrealistically early end (before Feb 1st) occurred. We then calculate the CMIP5-ensemble mean and the ensemble standard deviation for each of the two RCP scenarios. This led to one RCP4.5 model simulation being removed from all analyses. Furthermore, the following number of models did not meet the required criteria for each metric: Regarding the RSO for the RCP4.5 (RCP8.5) scenario, 2 (1) models were discarded for Gurgiser and 1 (0) for Climandes. Regarding the RSE and RCP4.5 (RCP8.5) 3 (2) models were discarded for Gurgiser, 2 (2) for Climandes and 9 (6) for Garcia. In the light of the anticipated increase in future precipitation in the Rio Santa basin (by 5.8 % $\pm$ 6.3 %, RCP4.5 and 12.1 % $\pm$ 11.0 %, RCP8.5, Potter et al. (2023)), combined with the sensitivities of the metrics in the previous chapter, the the result of only JD and FP suggesting earlier RSOs (by approximately 0.5 days for the stabilization scenario (RCP4.5) (Fig. 6) and only the bucket metric suggesting a small delay (decadal slope of approximately 0.35 to 0.6 days for both scenarios) in the RSE (Fig. 7), appears surprising. However, as applying trend analysis for each month individually across the CMIP5 ensemble reveals that the months September and October do not show significant trends for either scenario and for RCP4.5 only January and April show significant precipitation increases (c.f. Fig.A6), the annual results seem consistent. While the early-season months are highly relevant for the determination of the RSO, changes in the peak rainy season months are generally outside of the periods used by the metrics to determine start and end. In absolute values, the trends in the dry months are very small (with 0.046 for May, 0.022 for June and 0.004 mm/day decadal slopes for July and August, c.f. Fig. A6) while the calibrated values for the dry day threshold to determine RSE (c.f. Fig. 3 & Table 1) are in the order of 2-10 mm. Therefore, the absolute changes are likely too small to significantly alter the outputs of the threshold-based metrics. The consistent trends for both scenarios derived from the bucket metric stem from the fact that higher peak rainy season rainfall will keep the BWC at a higher level (c.f. Fig. A2) and the decrease in water availability and thus the resulting rainy season end will be delayed. The between-model variability is high

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

in both absolute numbers and in temporal distribution of rainfall. This is represented by only 7 out of 30 RCP8.5 and 2 out of 29 RCP4.5 CMIP5 models suggesting a delay in the RSE in case of the bucket metric. An assessment of the distribution of significant model trends can be found in Figs. A7 and A8. Related to this, the CMIP models reflect observations and previous findings (e.g. Hänchen et al., 2022) regarding the larger variability in RSO compared to RSE as illustrated by the considerably smaller RSE standard deviation across all metrics (c.f. Figs. 6, 7).

At least for the continental scale, many CMIP5 models were previously reported to poorly represent the South American Monsoon System (Bombardi and Carvalho, 2008), which may be expected to be aggravated for the topographically complex Andes. Hence, our findings also contrast the results of Jones and Carvalho (2013), who used 6 CMIP5 models to predict future South American Monsoon System changes under an RCP 8.5 scenario on the continental scale and further suggested, by using the LM metric, earlier rainy season onsets and later retreats. However, this could be related to several key differences which are the larger model ensemble used here, a spatial mismatch between the Rio Santa basin and the greater region, resolution differences or the fact that the LM metric can be subject to inconsistent sensitivities to hydroclimatic change as we previously showed (c.f. Fig. 5). Additionally, future predictions are further complicated by the limited understanding of expected ENSO changes and its effects in the region. While Cai et al. (2023) recently suggested an increase in ENSO variability linked to anthropogenic climate change, reliable predictions about the potential alteration of the rainy season and general precipitation patterns in the Rio Santa basin specifically cannot confidently be made at this time. Nevertheless, our results created by using an unprecedented number of calibrated and sensitivity-tested rainy season metrics combined with a high-resolution, bias-corrected dataset of climate model precipitation, demonstrate that statements regarding future change should be interpreted with caution in terms of climate model ensemble robustness and critically reviewed towards the calibration of rainy season metrics.

Finally, as we are calibrating the metrics on a vegetation proxy, the increasing temperatures are expected to raise the evapotranspiration rate in the Rio Santa basin (see Potter et al., 2023). This is likely to affect actual plant water availability and introduce uncertainty of currently unknown magnitude in the region. While this does not affect the rationales of the metrics, it will likely alter the applicability from a practitioner's perspective. In future endeavors, the bucket metric could be modified to accommodate this by altering the evapotranspiration parameter over time, which for our demonstrative purposes was set to a constant value. We decided against pursuing this adjustment for the future projections presented here because the bucket metric is not intended to replace the tasks of sophisticated hydrological models, and realistically estimating actual evapotranspiration in a data-sparse environment is a complex task in itself.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

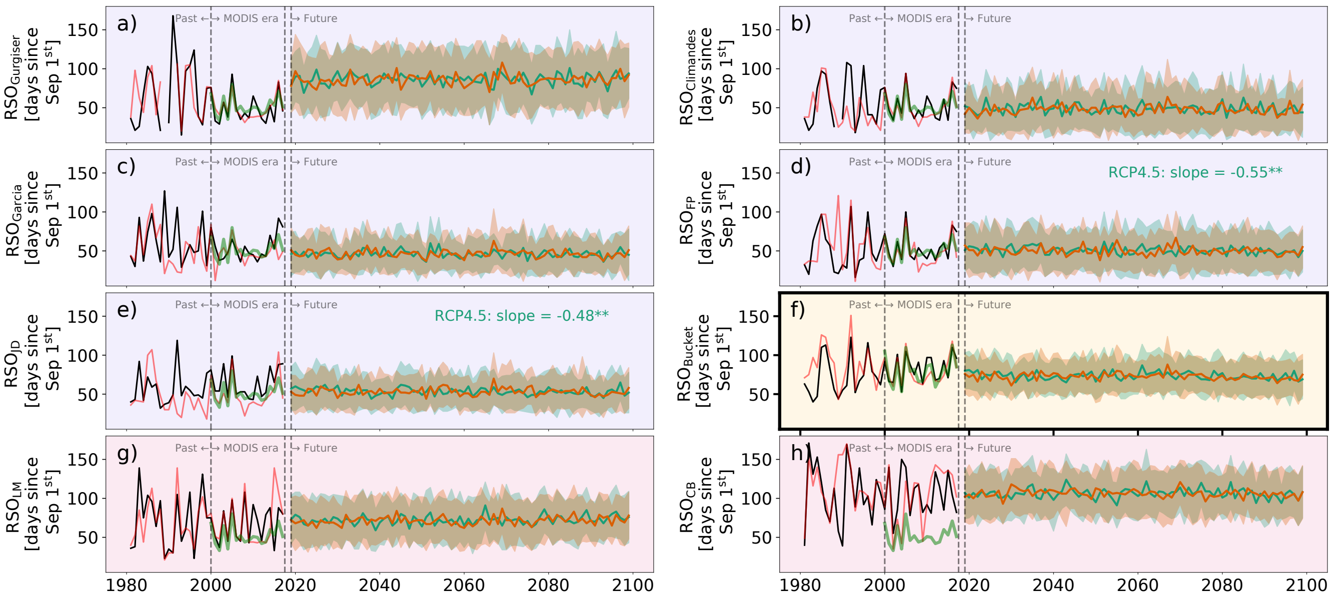

Figure 6. Rainy Season Onset (RSO) derived from 8 different metrics during the Past, Calibration, and Future periods, where threshold-based metrics are indicated by a purple, the bucket metric by a yellow and the objective metrics with a red background . Solid lines represent WRF (black) and CHIRPS (red) derived RSOs. The green line during the Calibration period indicates the SOS_{NDVI} used for metric calibration. Teal (RCP4.5) and orange (RCP8.5) lines represent statistically downscaled CMIP5-model ensemble averages. Shading around these lines indicates one standard deviation from the mean across the statistically downscaled CMIP5 models. For WRF, CHIRPS, and the two CMIP5 ensembles, trends (denoted as days per decade) were derived through linear regression. Significant trends are denoted by asterisks: *** for $p < 0.01$ and ** for $p < 0.05$ while not significant trends ($p > 0.05$) are not displayed.

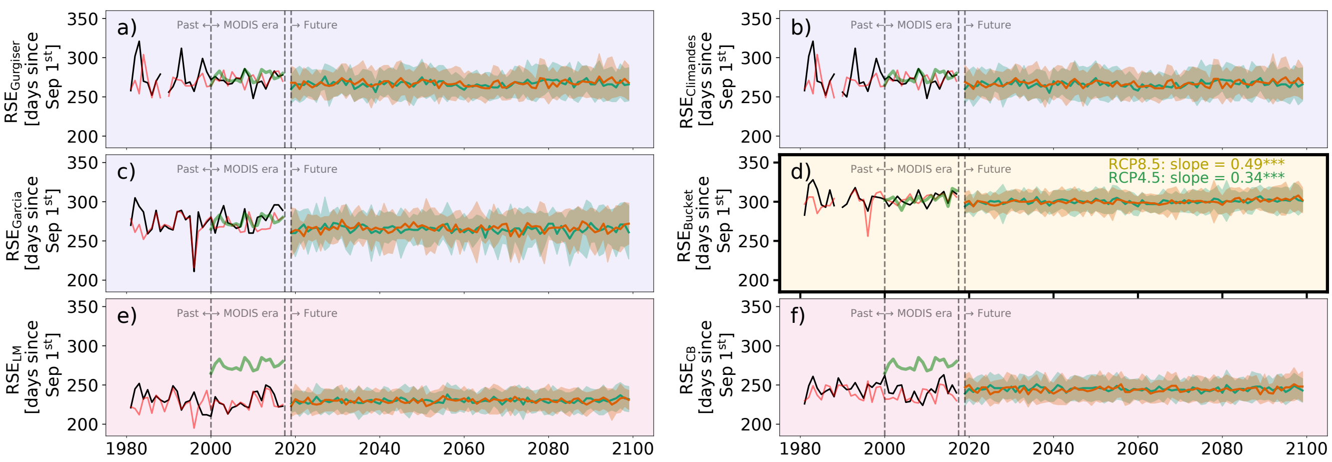

Figure 7. Same as Figure 6 but for Rainy Season End (RSE)

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

4 Conclusions

Based on several precipitation and remote-sensing derived land surface phenology data, we introduced a novel calibration strategy for rainy season metrics applied in semi-arid regions. For all three considered precipitation datasets, we find that the threshold-based rainy season metrics, once calibrated, are able to capture the interannual variability found in a vegetation greenness proxy in the Rio Santa basin, while more objective and flexible metrics have comparably low skill regarding this task. These objective metrics seem to exhibit implausible sensitivities that can potentially render them uninformative or even misleading under certain conditions of rainy season change.

Considering the numerous publications that highlight threshold-based metrics and propose a fixed parameter setup for specific regions, irrespective of the rainfall data source, we believe it is important to explore strategies for calibrating these metrics. This will enhance their practical application and effectiveness. Here, we demonstrated a framework for such an approach using remotely sensed data on vegetation greenness. In the specific case of the Rio Santa basin, the vegetation – rainfall correlation was proven reliable and, due to availability of NDVI data in relatively high spatial resolution, it is ideal to resolve the complex terrain, where gridded rainfall products are often subject to resolution biases. We however do believe that strategies for calibration different from using a proxy for vegetation greenness are also feasible as long as the variables are correlated with rainfall inputs into the hydrological system and available in sufficient quality. Examples could be, but are not limited to undisturbed runoff measurements or soil moisture data.

Moreover, our newly introduced bucket metric outperforms other metrics for both the onset and the end of the rainy season, shows physically consistent sensitivities and correcting for the vegetation – precipitation lag is optional, depending on the specific research question. The high skill and flexibility of the bucket metric allows for a wide range of applications in the context of hydroclimate in semi-arid areas. Additionally, it can likely be extended e.g. by making evapotranspiration dependent on energy- and/or water availability, while still remaining simplistic and efficient. The bucket metric is to our knowledge also the first attempt to take legacy effects of water availability into account; particularly relevant in regions such as the Rio Santa basin where large precipitation anomalies related to ENSO are common.

Based on an unprecedented number of future projections together with calibrated and sensitivity-tested rainy season metrics, we conclude that although precipitation is projected to increase, consistent trends for the rainy season onset cannot be derived and we find a comparably small delay of the rainy season end and consequently an increase of the rainy season length. In the light of high regional inter-annual variability, large intermodel spread of the CMIP5 projections and other factors currently poorly understood, such as the future impact of ENSO, reliable projections of climatic change in the tropical Andes remain challenging. While our novel framework allows crucial insights derived from rainfall time series, an adequate assessment of future water availability for practitioners' needs would benefit from more robust climate model forcings, eventually to be expected from the emergence of high-resolution convection-permitting model projections and hence information on evapotranspiration changes, most appropriately analyzed through a sophisticated hydrological model.

Code and data availability. Pre-processed data and python code to recreate the analysis and figures are available at https://github.com/lohae/RainySeasonMetrics and preserved at https://doi.org/10.5281/zenodo.13952139, allowing to apply and test the metrics for other regions or data. Full bias-corrected WRF data can be obtained at https://data.bas.ac.uk/full-record.php?id=GB/NERC/BAS/PDC/01728. The future precipitation from the statistically downscaled CMIP5 models are available at https://doi.org/10.5285/67CEB7C8-218C-46E1-9927-CFEF2DD95526, the future ETCCDI at https://doi.org/10.5285/B56D30E8-EDAA-4225-96D7-FCC689E930C7. Full CHIRPS data can be obtained through https://data.chc.ucsb.edu/products/CHIRPS-2.0/ while NDVI raw data can be acquired (for example) through https://appeears.earthdatacloud.nasa.gov/. The AWS data is publicly available at https://www.senamhi.gob.pe/?p=estaciones, we acquired it however through the METEODAT platform (available on request).

Author contributions. LH: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft preparation EP: Data curation, Formal analysis, Methodology, Writing – review & editing CK: Conceptualization, Data curation, Formal analysis, Software, Writing – review & editing PC: Conceptualization, Funding acquisition, Supervision, Writing – review & editing WG: Writing – review & editing FM: Data curation, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing GW: Conceptualization, Methodology, Resources, Software, Supervision, Writing – review & editing

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Parts of this study were conducted in the frame of the AgroClim Huaraz project, funded by the Earth System Sciences Program of the Austrian Academy of Sciences (OEAW) and the FFG project AustroSIF. EP acknowledges funding from a Leverhulme Trust Early Career Research Fellowship at the time of submission. CK also acknowledges funding from the NERC independent research fellowship COCOON (NE/X017419/1). We thank Mario Rohrer for providing access to the METEODAT platform where we acquired the weather station data. Special thanks to the developers of numpy (Harris et al., 2020), scipy (Virtanen et al., 2020), xarray (Hoyer and Hamman, 2017), pandas (McKinney), pwlf (Jekel and Venter, 2019), salem (Maussion et al., 2023) and their dependencies for making their code available on a free and open-source basis. Thanks to Yannick Copin for the taylor diagram code (available under https://gist.github.com/ycopin/3342888) and to Santiago Belda and the IPL team for the DaTimeS software (available under https://artmotoolbox.com/plugins-standalone/91-plugins-standalone/34-datimes.html).

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

References

Belda, S., Pipia, L., Morcillo-Pallares, P., Rivera-Caicedo, J. P., Amin, E., De Grave, C., and Verrelst, J.: DATimeS: A machine learning time series GUI toolbox for gap-filling and vegetation phenology trends detection, Environ Model Softw, 127, https://doi.org/10.1016/j.envsoft.2020.104666, 2020.

Bombardi, R. J. and Carvalho, L. M. V.: IPCC global coupled model simulations of the South America monsoon system, Climate Dynamics, 33, 893–916, https://doi.org/10.1007/s00382-008-0488-1, 2008.

Bombardi, R. J., Kinter, J. L., and Frauenfeld, O. W.: A Global Gridded Dataset of the Characteristics of the Rainy And Dry Seasons, Bulletin of the American Meteorological Society, 100, 1315–1328, https://doi.org/10.1175/bams-d-18-0177.1, 2019a.

Bombardi, R. J., Moron, V., and Goodnight, J. S.: Detection, variability, and predictability of monsoon onset and withdrawal dates: A review, International Journal of Climatology, 40, 641–667, https://doi.org/10.1002/joc.6264, 2019b.

Bury, J. T., Mark, B. G., McKenzie, J. M., French, A., Baraer, M., Huh, K. I., Zapata Luyo, M. A., and Gómez López, R. J.: Glacier recession and human vulnerability in the Yanamarey watershed of the Cordillera Blanca, Peru, Climatic Change, 105, 179–206, https://doi.org/10.1007/s10584-010-9870-1, 2010.

Cai, W., Ng, B., Geng, T., Jia, F., Wu, L., Wang, G., Liu, Y., Gan, B., Yang, K., Santoso, A., Lin, X., Li, Z., Liu, Y., Yang, Y., Jin, F.-F., Collins, M., and McPhaden, M. J.: Anthropogenic impacts on twentieth-century ENSO variability changes, Nature Reviews Earth Environment, 4, 407–418, https://doi.org/10.1038/s43017-023-00427-8, 2023.

Camberlin, P. and Diop, M.: Application of daily rainfall principal component analysis to the assessment of the rainy season characteristics in Senegal, Climate Research, 23, 159–169, https://doi.org/10.3354/cr023159, 2003.

Caparros-Santiago, J. A., Rodriguez-Galiano, V., and Dash, J.: Land surface phenology as indicator of global terrestrial ecosystem dynamics: A systematic review, ISPRS Journal of Photogrammetry and Remote Sensing, 171, 330–347, https://doi.org/10.1016/j.isprsjprs.2020.11.019, 2021.

Cook, B. I. and Buckley, B. M.: Objective determination of monsoon season onset, withdrawal, and length, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/10.1029/2009jd012795, 2009.

Dextre, R. M., Eschenhagen, M. L., Camacho Hernández, M., Rangecroft, S., Clason, C., Couldrick, L., and Morera, S.: Payment for ecosystem services in Peru: Assessing the socio-ecological dimension of water services in the upper Santa River basin, Ecosystem Services, 56, https://doi.org/10.1016/j.ecoser.2022.101454, 2022.

Didan, K.: https://doi.org/10.5067/MODIS/MOD13Q1.006, 2015a.

Didan, K.: https://doi.org/10.5067/MODIS/MYD13Q1.006, 2015b.

Drenkhan, F., Carey, M., Huggel, C., Seidel, J., and Oré, M. T.: The changing water cycle: climatic and socioeconomic drivers of water-related changes in the Andes of Peru, WIREs Water, 2, 715–733, https://doi.org/10.1002/wat2.1105, 2015.

Drenkhan, F., Buytaert, W., Mackay, J. D., Barrand, N. E., Hannah, D. M., and Huggel, C.: Looking beyond glaciers to understand mountain water security, Nature Sustainability, 6, 130–138, https://doi.org/10.1038/s41893-022-00996-4, 2022.

Dunning, C. M., Black, E. C. L., and Allan, R. P.: The onset and cessation of seasonal rainfall over Africa, Journal of Geophysical Research: Atmospheres, 121, https://doi.org/10.1002/2016jd025428, 2016.

Espinoza, J. C., Garreaud, R., Poveda, G., Arias, P. A., Molina-Carpio, J., Masiokas, M., Viale, M., and Scaff, L.: Hydroclimate of the Andes Part I: Main Climatic Features, Frontiers in Earth Science, 8, https://doi.org/10.3389/feart.2020.00064, 2020.

Ferijal, T., Batelaan, O., Shanafield, M., and Alfahmi, F.: Determination of rainy season onset and cessation based on a flexible driest period, Theoretical and Applied Climatology, 148, 91–104, https://doi.org/10.1007/s00704-021-03917-1, 2022.

Fitzpatrick, R. G. J., Bain, C. L., Knippertz, P., Marsham, J. H., and Parker, D. J.: The West African Monsoon Onset: A Concise Comparison of Definitions, Journal of Climate, 28, 8673–8694, https://doi.org/10.1175/jcli-d-15-0265.1, 2015.

Freedman, D. and Diaconis, P.: On the histogram as a density estimator: L 2 theory, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 57, 453–476, 1981.

Frere, M. and Popov, G.: Early agrometeorological crop yield assessment, FAO plant production and protection paper, 73, 144, 1986.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J.: The climate hazards infrared precipitation with stations–a new environmental record for monitoring extremes, Sci Data, 2, 150 066, https://doi.org/10.1038/sdata.2015.66, 2015.

Fyffe, C. L., Potter, E., Fugger, S., Orr, A., Fatichi, S., Loarte, E., Medina, K., Hellström, R. , Bernat, M., Aubry-Wake, C., Gurgiser, W., Perry, L. B., Suarez, W., Quincey, D. J., and Pellicciotti, F.: The Energy and Mass Balance of Peruvian Glaciers, Journal of Geophysical Research: Atmospheres, 126, https://doi.org/10.1029/2021jd034911, 2021.

Garcia, M., Raes, D., Jacobsen, S. E., and Michel, T.: Agroclimatic constraints for rainfed agriculture in the Bolivian Altiplano, Journal of Arid Environments, 71, 109–121, https://doi.org/10.1016/j.jaridenv.2007.02.005, 2007.

Garreaud, R. D.: The Andes climate and weather, Advances in Geosciences, 22, 3–11, https://doi.org/10.5194/adgeo-22-3-2009, 2009.

Giráldez, L., Silva, Y., Zubieta, R., and Sulca, J.: Change of the Rainfall Seasonality Over Central Peruvian Andes: Onset, End, Duration and Its Relationship With Large-Scale Atmospheric Circulation, Climate, 8, https://doi.org/10.3390/cli8020023, 2020.

Gurgiser, W., Juen, I., Singer, K., Neuburger, M., Schauwecker, S., Hofer, M., and Kaser, G.: Comparing peasants' perceptions of precipitation change with precipitation records in the tropical Callejón de Huaylas, Peru, Earth System Dynamics, 7, 499–515, https://doi.org/10.5194/esd-7-499-2016, 2016.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Del Rio, J. F., Wiebe, M., Peterson, P., Gerard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357–362, https://doi.org/10.1038/s41586-020-2649-2, 2020.

Hoyer, S. and Hamman, J.: xarray: N-D labeled Arrays and Datasets in Python, Journal of Open Research Software, 5, https://doi.org/10.5334/jors.148, 2017.

Hunziker, S., Gubler, S., Calle, J., Moreno, I., Andrade, M., Velarde, F., Ticona, L., Carrasco, G., Castellón, Y., and Oria, C.: Identifying, attributing, and overcoming common data quality issues of manned station observations, International journal of climatology, 37, 4131–4145, 2017.

Hänchen, L., Klein, C., Maussion, F., Gurgiser, W., Calanca, P., and Wohlfahrt, G.: Widespread greening suggests increased dry-season plant water availability in the Rio Santa valley, Peruvian Andes, Earth System Dynamics, 13, 595–611, https://doi.org/10.5194/esd-13-595-2022, 2022.

Jekel, C. F. and Venter, G.: https://github.com/cjekel/piecewise_linear_fit_py, 2019.

Jolliffe, I. T. and Sarria-Dodd, D. E.: Early detection of the start of the wet season in tropical climates, International Journal of Climatology, 14, 71–76, https://doi.org/10.1002/joc.3370140106, 1994.

Jones, C. and Carvalho, L. M. V.: Climate Change in the South American Monsoon System: Present Climate and CMIP5 Projections, Journal of Climate, 26, 6660–6678, https://doi.org/10.1175/jcli-d-12-00412.1, 2013.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D. B.: So, how much of the Earth's surface is covered by rain gauges?, Bull Am Meteorol Soc, 98, 69–78, https://doi.org/10.1175/BAMS-D-14-00283.1, 2017.

Klein, C., Hänchen, L., Potter, E. R., Junquas, C., Harris, B. L., and Maussion, F.: Untangling the importance of dynamic and thermodynamic drivers for wet and dry spells across the Tropical Andes, Environmental Research Letters, 18, https://doi.org/10.1088/1748-9326/acb72b, 2023a.

Klein, C., Potter, E. R., Zauner, C., Gurgiser, W., Cruz Encarnación, R., Cochachín Rapre, A., and Maussion, F.: Farmers' first rain: investigating dry season rainfall characteristics in the Peruvian Andes, Environmental Research Communications, 5, https://doi.org/10.1088/2515-7620/ace516, 2023b.

Laux, P., Kunstmann, H., and Bárdossy, A.: Predicting the regional onset of the rainy season in West Africa, International Journal of Climatology, 28, 329–342, https://doi.org/10.1002/joc.1542, 2008.

Liebmann, B. and Marengo, J.: Interannual Variability of the Rainy Season and Rainfall in the Brazilian Amazon Basin, Journal of Climate, 14, 4308–4318, https://doi.org/10.1175/1520-0442(2001)014<4308:Ivotrs>2.0.Co;2, 2001.

Liebmann, B., Camargo, S. J., Seth, A., Marengo, J. A., Carvalho, L. M., Allured, D., Fu, R., and Vera, C. S.: Onset and end of the rainy season in South America in observations and the ECHAM 4.5 atmospheric general circulation model, Journal of Climate, 20, 2037–2050, 2007.

MacLeod, D.: Seasonal predictability of onset and cessation of the east African rains, Weather and Climate Extremes, 21, 27–35, https://doi.org/10.1016/j.wace.2018.05.003, 2018.

Marengo, J. A., Liebmann, B., Kousky, V. E., Filizola, N. P., and Wainer, I. C.: Onset and End of the Rainy Season in the Brazilian Amazon Basin, Journal of Climate, 14, 833–852, https://doi.org/10.1175/1520-0442(2001)014<0833:Oaeotr>2.0.Co;2, 2001.

Mateo, E. I., Mark, B. G., Hellström, R. , Baraer, M., McKenzie, J. M., Condom, T., Rapre, A. C., Gonzales, G., Gómez, J. Q., and Encarnación, R. C. C.: High-temporal-resolution hydrometeorological data collected in the tropical Cordillera Blanca, Peru (2004–2020), Earth System Science Data, 14, 2865–2882, https://doi.org/10.5194/essd-14-2865-2022, 2022.

Maussion, F., Gurgiser, W., Großhauser, M., Kaser, G., and Marzeion, B.: ENSO influence on surface energy and mass balance at Shallap Glacier, Cordillera Blanca, Peru, The Cryosphere, 9, 1663–1683, https://doi.org/10.5194/tc-9-1663-2015, 2015.

Maussion, F., Siller, M., and Rothenberg, D.: fmaussion/salem v0.3.9, Zenodo, https://doi.org/https://doi.org/10.5281/zenodo.269646, 2023.

McKinney, W.: Data structures for statistical computing in python, in: Proceedings of the 9th Python in Science Conference, vol. 445, pp. 51–56, Austin, TX.

Pollock, M. D., O'Donnell, G., Quinn, P., Dutton, M., Black, A., Wilkinson, M. E., Colli, M., Stagnaro, M., Lanza, L. G., Lewis, E., Kilsby, C. G., and O'Connell, P. E.: Quantifying and Mitigating Wind-Induced Undercatch in Rainfall Measurements, Water Resources Research, 54, 3863–3875, https://doi.org/10.1029/2017wr022421, 2018.

Potter, E. R., Fyffe, C. L., Orr, A., Quincey, D. J., Ross, A. N., Rangecroft, S., Medina, K., Burns, H., Llacza, A., Jacome, G., Hellström, R. Castro, J., Cochachin, A., Montoya, N., Loarte, E., and Pellicciotti, F.: A future of extreme precipitation and droughts in the Peruvian Andes, npj Climate and Atmospheric Science, 6, https://doi.org/10.1038/s41612-023-00409-z, 2023.

Quiroz, R., Yarlequé, C., Posadas, A., Mares, V., and Immerzeel, W. W.: Improving daily rainfall estimation from NDVI using a wavelet transform, Environmental Modelling Software, 26, 201–209, https://doi.org/10.1016/j.envsoft.2010.07.006, 2011.

Rangecroft, S., Dextre, R. M., Richter, I., Grados Bueno, C. V., Kelly, C., Turin, C., Fuentealba, B., Hernandez, M. C., Morera, S., Martin, J., Guy, A., and Clason, C.: Unravelling and understanding local perceptions of water quality in the Santa basin, Peru, Journal of Hydrology, 625, https://doi.org/10.1016/j.jhydrol.2023.129949, 2023.

Sedlmeier, K., Imfeld, N., Gubler, S., Spirig, C., Caiña, K. Q., Escajadillo, Y., Rohrer, M., and Schwierz, C.: The rainy season in the Southern Peruvian Andes: A climatological analysis based on the new Climandes index, International Journal of Climatology, 43, 3005–3022, https://doi.org/10.1002/joc.8013, 2023.

Seregina, L. S., Fink, A. H., van der Linden, R., Elagib, N. A., and Pinto, J. G.: A new and flexible rainy season definition: Validation for the Greater Horn of Africa and application to rainfall trends, International Journal of Climatology, 39, 989–1012, https://doi.org/10.1002/joc.5856, 2018.

Silva, Y., Takahashi, K., and Chávez, R.: Dry and wet rainy seasons in the Mantaro river basin (Central Peruvian Andes), Advances in Geosciences, 14, 261–264, https://doi.org/10.5194/adgeo-14-261-2008, 2008.

Stern, R., Dennett, M., and Garbutt, D.: The start of the rains in West Africa, Journal of Climatology, 1, 59–68, 1981.

Storn, R. and Price, K.: Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces, Journal of Global Optimization, 11, 341–359, https://doi.org/10.1023/a:1008202821328, 1997.

USGS EROS Archive: Digital Elevation – Shuttle Radar Topography Mission (SRTM) 1 Arc-second Global, https://www.usgs.gov/centers/eros, uSGS EROS Archive [data set], last access: January 2021, 2021.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy, C.: SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Vuille, M., Kaser, G., and Juen, I.: Glacier mass balance variability in the Cordillera Blanca, Peru and its relationship with climate and the large-scale circulation, Global and Planetary Change, 62, 14–28, 2008.

Warner, K., Afifi, T., Henry, K., Rawe, T., Smith, C., and De Sherbinin, A.: Where the Rain Falls : Climate Change, food and Livelihood Security, and Migration, Global Policy Report of the Where the Rain Falls Project. Bonn: CARE France and UNU-EHS, 2012.

Yarleque, C., Vuille, M., Hardy, D. R., Posadas, A., and Quiroz, R.: Multiscale assessment of spatial precipitation variability over complex mountain terrain using a high-resolution spatiotemporal wavelet reconstruction method, Journal of Geophysical Research-Atmospheres, 121, 12 198–12 216, https://doi.org/10.1002/2016jd025647, 2016.

Zampieri, M., Toreti, A., Meroni, M., Bojovic, D., Octenjak, S., Marcos-Matamoros, R., Materia, S., Chang'a, L., Merchades, M., and Montero, M. d. M. C.: Seasonal forecasts of the rainy season onset over Africa: preliminary results from the FOCUS-Africa project, Climate Services, 32, 100 417, 2023.

Zhang, X. B., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., Trewin, B., and Zwiers, F. W.: Indices for monitoring changes in extremes based on daily temperature and precipitation data, Wiley Interdisciplinary Reviews-Climate Change, 2, 851–870, https://doi.org/10.1002/wcc.147, 2011.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

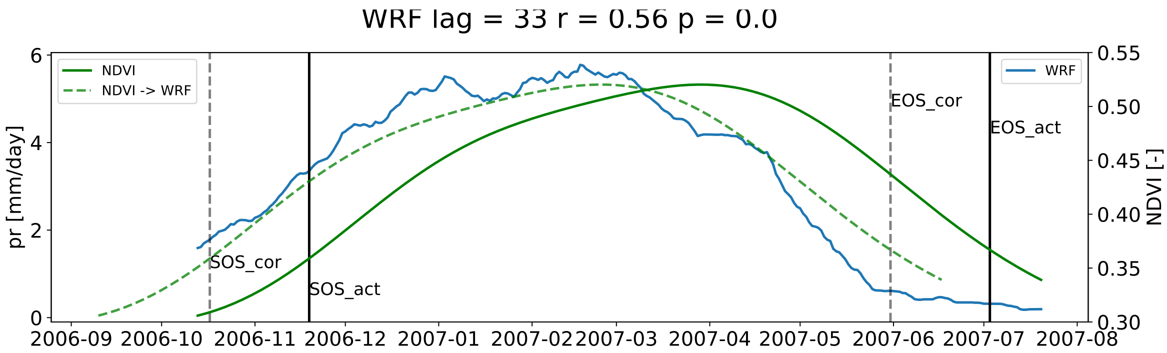

Figure A1. Visual example of the cross correlation function for lag correction of one hydrological year. The lag was determined based on WRF data smoothed by a 12-week rolling average and the processed NDVI data

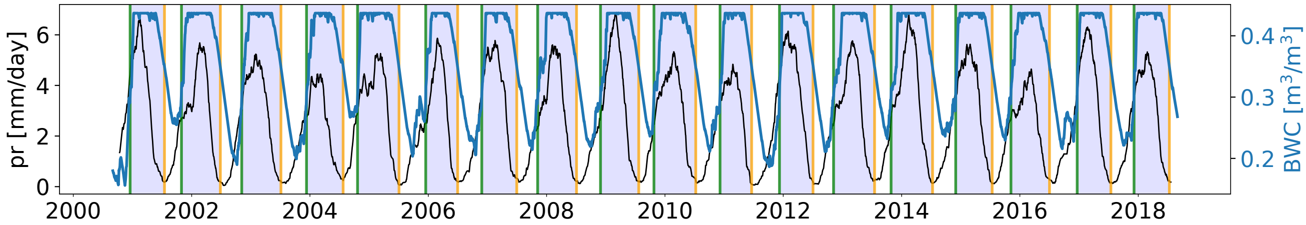

Figure A2. 12-week rolling window WRF time series (black) and BWC, modeled from daily (non-smoothed) precipitation from bucket metric (blue) for the calibration period 2000-2018. Green (orange) vertical lines indicate RSO (RSE) dates derived by the bucket metric. Blue shading indicates the resulting rainy season period

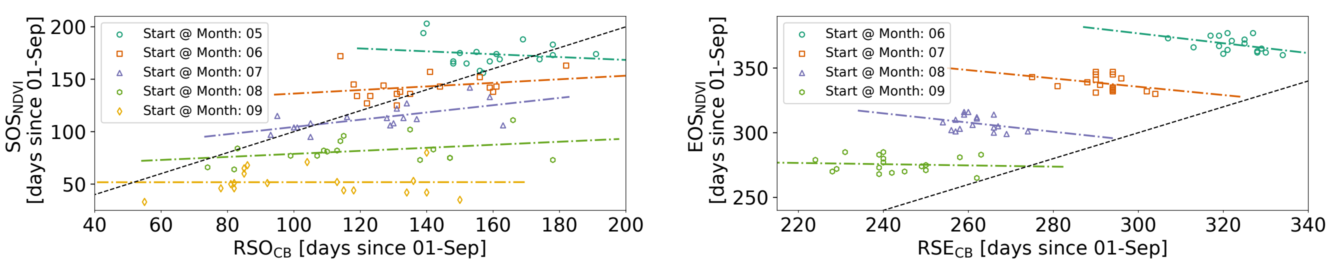

Figure A3. Sensitivity of Two-Phase linear Regression method to hydrological year definition by Cook and Buckley (2009).

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

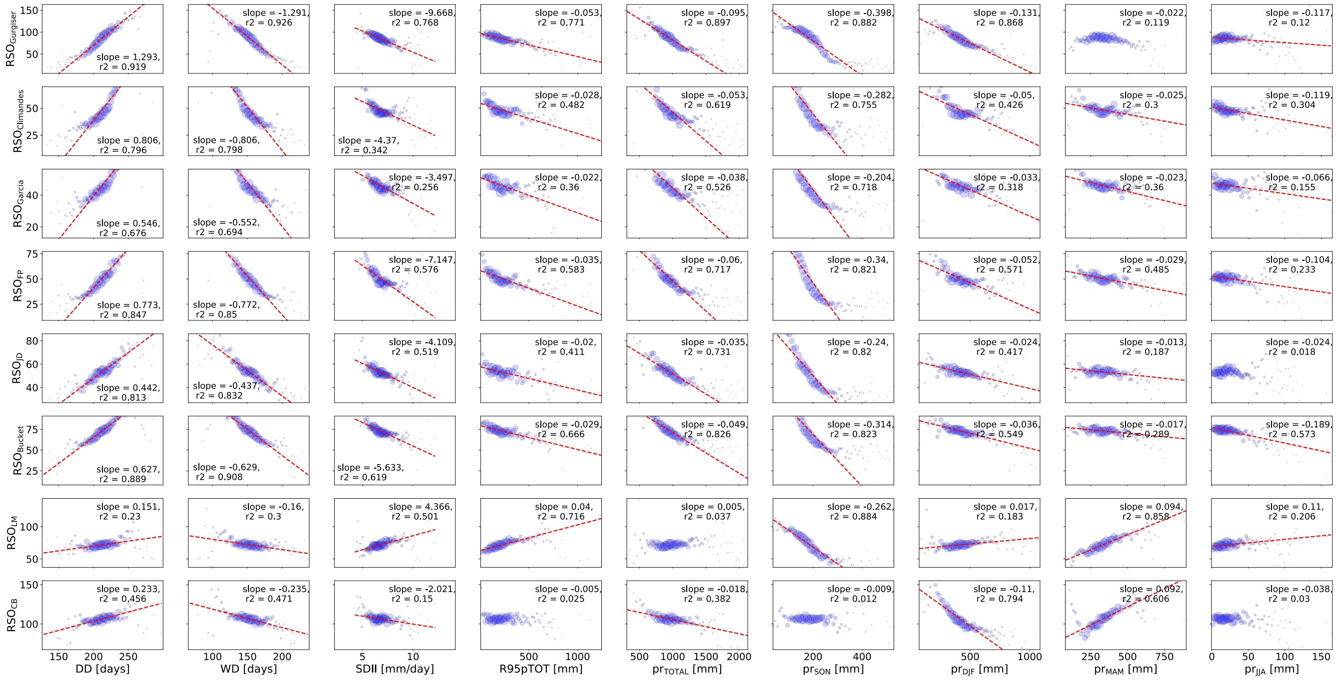

Figure A4. Bin-weighted regressions for RSO as summarized in 5a. Red regression lines are only shown for significant regressions (p < 0.01).

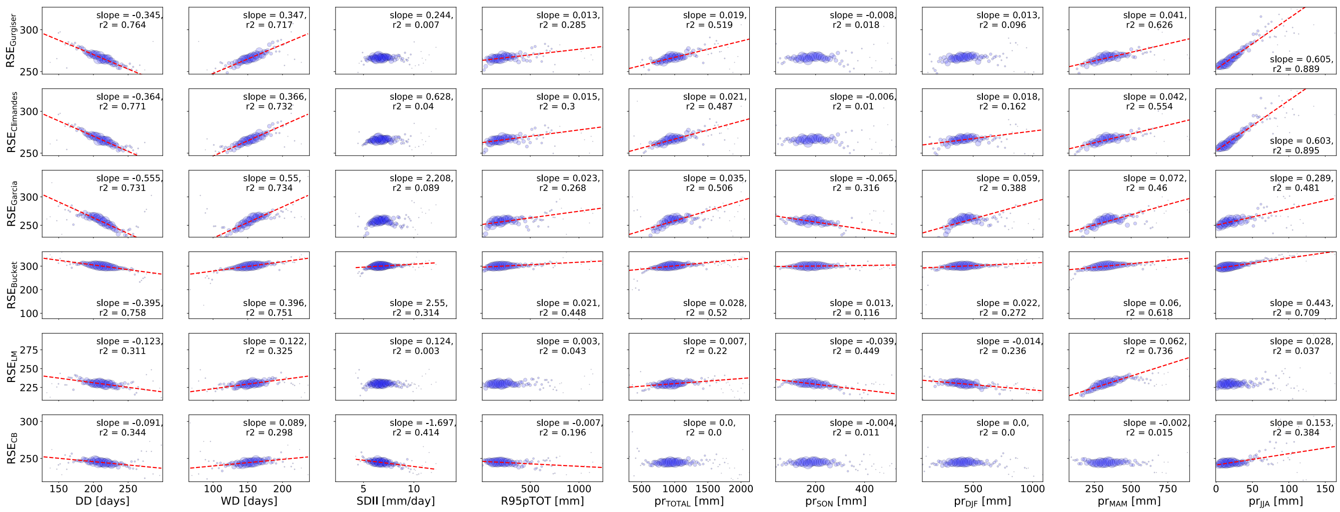

Figure A5. Bin-weighted regressions for RSE as summarized in 5b. Red regression lines are only shown for significant regressions (p < 0.01).

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

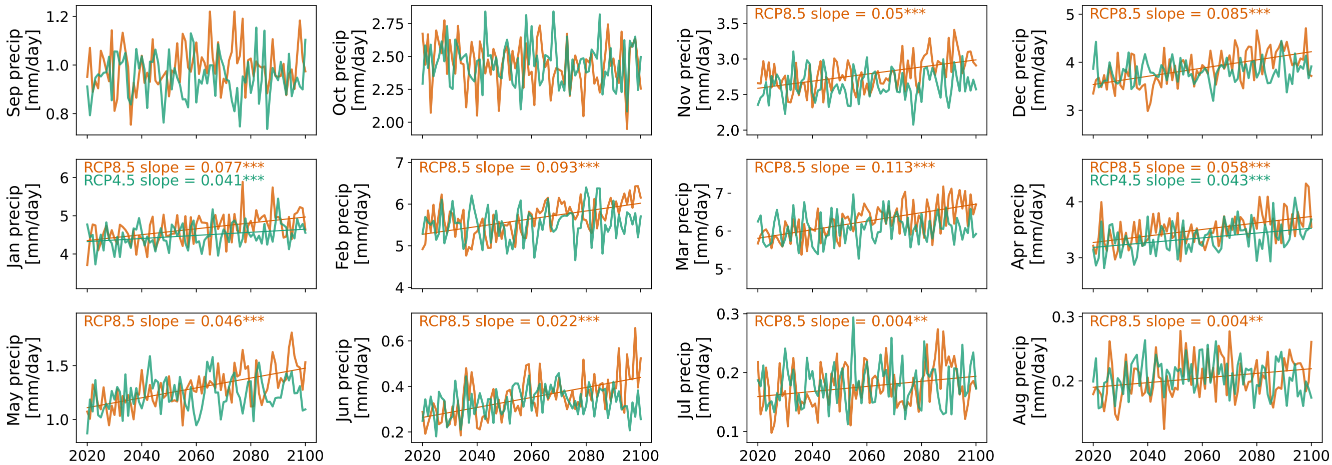

Figure A6. Monthly trends for the CMIP5 ensemble for both RCP4.5 (teal) and RCP8.5 (brown) scenarios. Decadal trends were derived through linear regression. Significant trends are denoted by asterisks: *** for p < 0.01 and ** for p < 0.05 while regression lines for non-significant trends (p > 0.05) are not displayed.

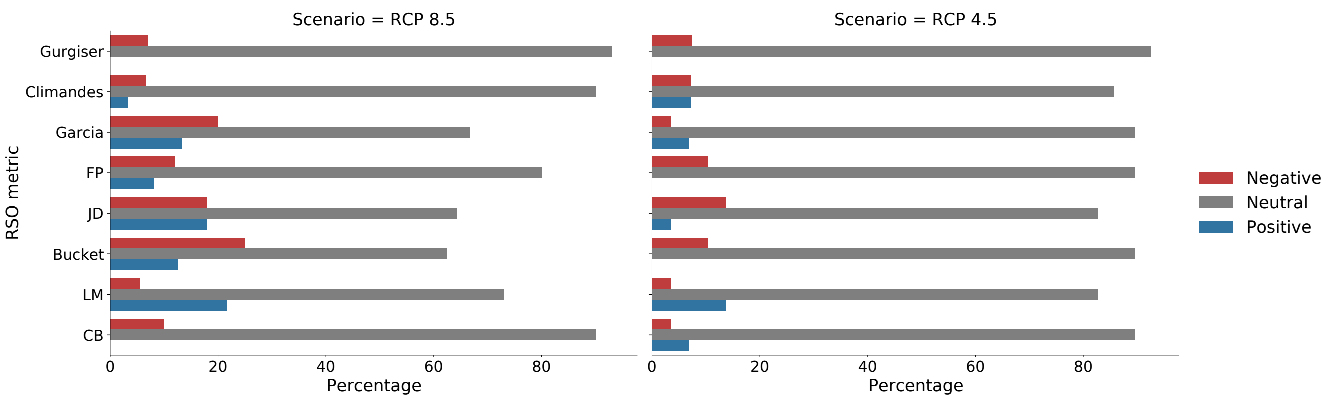

Figure A7. Relative Distribution of significant and non-significant CMIP5 model timeseries (p < 0.05) and their sign for the derived rainy season onsets for the timeperiod 2019-2100 for each rainy season metric and both RCP scenarios. A negative trend refers to an earlier season start and a positive trend to a later start.

https://doi.org/10.5194/egusphere-2024-3263
Preprint. Discussion started: 20 November 2024

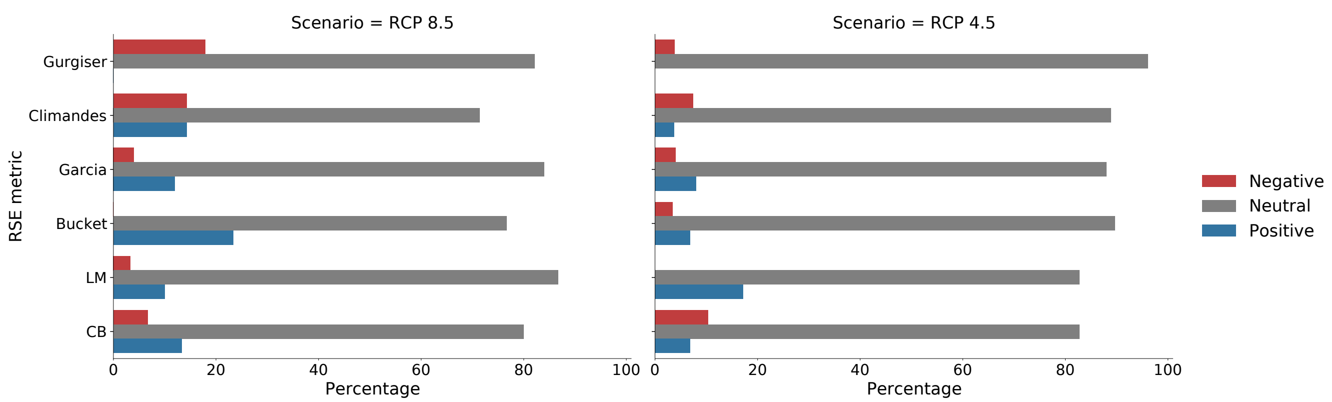

Figure A8. Same as Figure S8 but for RSE.