Hydrology and
Earth System
Sciences

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, School of Geography and Planning, Sheffield, UK
[3]UK Centre for Ecology and Hydrology, Wallingford, UK
[4]Climate and Agriculture, Agroscope Reckenholz, Zurich, Switzerland
[5]University of Bristol, School of Geographical Sciences, Bristol, 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 determines 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 seasonal 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. Furthermore, a metric's ability to capture observed vegetation variability can indicate its applicability over larger spatial or temporal scales. This study introduces a new bucket-type metric that incorporates a simplified water balance, accounts for both accumulation and storage, and also takes interannual legacy effects into account. We evaluate its performance against seven commonly used rainy season metrics, both calibrated and uncalibrated, using 18 years of the satellite-derived Normalized Difference Vegetation Index (NDVI) from the semi-arid Rio Santa basin in the Peruvian Andes. Our results demonstrate that calibrating metrics using vegetation data significantly enhances their ability to capture rainy season dynamics, with the bucket metric outperforming others in both accuracy and robustness. Furthermore, we examine the sensitivity of all metrics to variations in rainfall intensity and frequency under future climate scenarios, using a previously published high-resolution dataset specifically designed for the Rio Santa basin which provides historical (1981–2018) rainfall data and future projections (2019–2100) based on 30 statistically downscaled CMIP5 models for the Representative Concentration Pathway (RCP) 4.5 and 8.5 scenarios, respectively. While most rainy season metrics exhibit expected correlations in response to climatic changes, some established metrics display physically inconsistent behavior due to methodological artifacts, highlighting their limitations in assessing hydroclimatic changes. In addition to the sensitivity analysis, we evaluate long-term trends in rainy season characteristics. Statistically downscaled CMIP5 ensemble projections for the future period suggest only a slight delay in the rainy season end, with no consistent trends in onset timing. Instead, interannual variability and ensemble spread remain the dominant influences. Our findings emphasize the need for careful calibration of metrics across diverse climate scenarios and different locations to ensure their reliability for agricultural planning, policymaking, and climate adaptation strategies. By providing a novel framework for evaluating rainfall metrics, this study offers a scalable approach that can be readily applied to other semi-arid regions.

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 (e.g., Bombardi et al., 2019b; Fitzpatrick et al., 2015; Sedlmeier et al., 2023). Broadly, these metrics consist of threshold-based approaches which must be configured for each region (Sedlmeier et al., 2023) 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) in mind or for regions with high spatiotemporal variability in rainfall by employing data manipulation approaches such as Principal Component Analysis (Camberlin and Diop, 2003), 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, the importance of determining rainy season characteristics for either agricultural planning, monitoring of ecosystems, assessments of temporal water availability in light of a changing climate, or water management topics in general is emphasized (e.g., Bombardi et al., 2019b; Fitzpatrick et al., 2015). However, observational and gridded precipitation time series are typically subject to significant uncertainties, such as spatial representativeness issues, measurement errors (such as undercatch in windy conditions), temporal inconsistencies, and biases in retrieval algorithms (e.g., Kidd et al., 2017; Pollock et al., 2018). These uncertainties can lead water users and managers to make improper assumptions or take misguided actions. These uncertainties are particularly problematic in regions where decisions about planting, irrigation scheduling, or reservoir management rely heavily on short- or mid-term rainfall predictions. To the best of our knowledge, strategies to validate the outputs of rainy season metrics against independent observations are currently lacking. This raises concerns about whether such met-

rics accurately capture conditions on the ground and highlights the need for validation frameworks that ensure their relevance and reliability for practical applications and allow us to reliably deduce climatic changes. Furthermore, other aspects, such as legacy effects beyond 1 hydrological year or the sensitivity of rainy season metrics to the alteration of the hydrological cycle (which is to be expected under climate change), have so far not been assessed.

This raises the challenge of designing an independent validation approach. While variables directly linked to the hydrological cycle, such as soil moisture measurements, would represent the ideal choice, their availability at (near-)climatological timescales is limited. In semi-arid regions, vegetation dynamics provide a useful alternative, as they exhibit a strong correlation with the seasonal precipitation cycle, albeit with a characteristic time lag (Hänchen et al., 2022). Remotely sensed proxies for vegetation development offer high spatiotemporal resolution and have been successfully used to study vegetation development for more than 50 years (starting with Rouse et al., 1974). We therefore 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 specific data and locations, can be assessed for their applicability in different regions. Spectral vegetation indices, which serve as proxies for land surface greenness, are a promising candidate for calibrating rainy season metrics in semi-arid regions due to their high spatiotemporal resolution and availability from satellite data.

In this study, we develop and demonstrate a novel approach to calibrate rainy season metrics using vegetation dynamics, focusing on the upper Rio Santa basin (also Callejón de Huaylas) in the tropical Peruvian Andes. This region is characterized by high seasonal variability in precipitation, with the majority of annual precipitation occurring between September and April and little annual variability in temperature (see Fig. 1 for the geographic location and a climograph). The region 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 the El Niño–Southern Oscillation (ENSO) (e.g., Hänchen et al., 2022; Maussion et al., 2015). In this region, a thorough understanding of the dynamics of the rainy season is crucial for regional water resources and agriculture, as the seasonal rain provides water for irrigation, energy production, and the 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
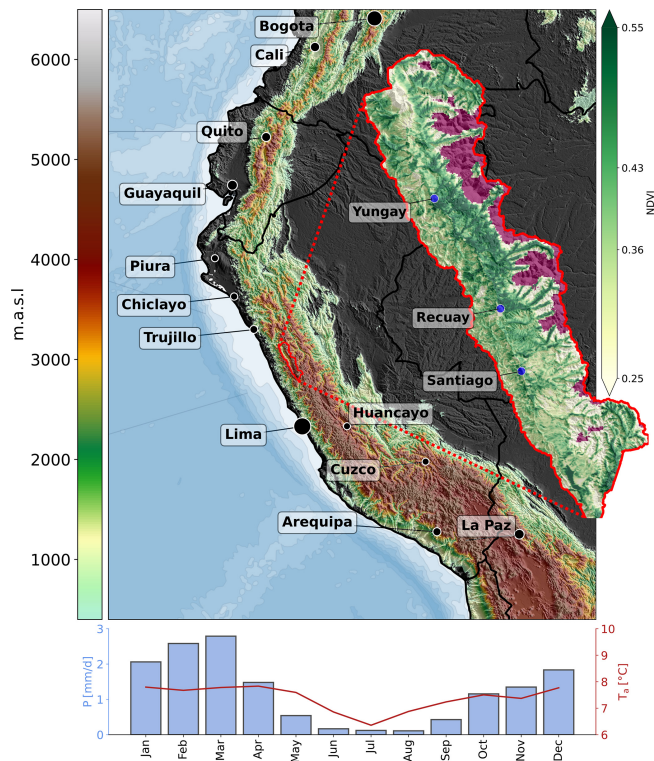
Figure 1. Overview of the study area. The large map shows the topography of the Andes, with elevations below 500 m shown in black (based on SRTM data; USGS EROS Archive, 2021) and with administrative borders and larger towns in the greater region. The enlarged area of the Rio Santa basin shows the long-term (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. The three blue dots indicate the locations of the local weather stations used in this study. 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., 2023a) for the Rio Santa basin.

glacial melt (e.g., Bury et al., 2010; Drenkhan et al., 2015; Fyffe et al., 2021). Small-scale farmers, however, often have no access 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., 2023a). Future climate scenarios indicate an overall increase in annual precipitation (Potter et al., 2023a).

Potential shifts in the timing of the rainy season in the region, despite their profound implications for both societal and ecological systems, have only recently been assessed. Notably, De la Cruz et al. (2025) used an objective metric to derive rainfall sums and rainy season onset and end

for a Peru-wide network of meteorological stations based on statistically downscaled CMIP6 projections to derive future changes. They found an increase in future annual precipitation and, similarly to other studies, show that past rainy season dynamics in the broader Andean region reveal high interannual variability in rainy season onset, with generally non-significant or weak longer-term trends (Garcia et al., 2007; Giràldez et al., 2020; Gurgiser et al., 2016; Sedlmeier et al., 2023). Across those studies, the end of the rainy season is notably less variable than the start, while showing no or small 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 the 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). This response contrasts with other basins in proximity, such as the Mantaro River basin, where the opposite pattern has been suggested (Giràldez et al., 2020), thus highlighting the spatial heterogeneity of hydroclimatic responses within the Andes.

To account for these difficulties, we employ a multifaceted approach capitalizing stem on previous studies: we combine several precipitation datasets with remote sensing data on temporal vegetation development. Specifically, we calculate the rainy season metrics based on convection-permitting, bias-corrected Weather Research and Forecasting (WRF) precipitation data and statistically downscaled CMIP5 projections (Potter et al., 2023a) and use gridded Climate Hazards InfraRed Precipitation with Station data (CHIRPS) (Funk et al., 2015) and data from three local weather stations for comparison. For validation and calibration, we utilize land surface phenology (LSP) data for the period 2000–2018 and the spatial extent of the Rio Santa basin, derived from the temporal development of the remotely sensed Normalized Difference Vegetation Index (NDVI) provided by Hänchen et al. (2022). Their research demonstrated that NDVI – an indicator of vegetation greenness available at high spatiotemporal resolution – captures variability and changes in water availability in the semi-arid Rio Santa basin, where water availability is the primary limiting factor for plant growth. This high spatial resolution is shown in Fig. 1, which shows the 2000–2018 average NDVI for the Rio Santa basin, illustrating both longitudinal and altitudinal gradients. Similarly, 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).

The principal objective of this study is to showcase a novel framework for characterizing the rainy season, emphasizing the importance of employing a calibration strategy for inferred rainy season onsets and ends. In addition, we test the sensitivity of rainy season metrics to plausible changes in rainfall intensity and frequency, as might occur due to global warming. By capturing shifts in seasonal rainfall dynamics, our approach provides a foundation for identifying and understanding hydrological changes that may inform future adaptation strategies. The proposed framework is designed to improve our understanding of variations in water availability within semi-arid regions, offering insights that extend beyond the Rio Santa basin and can be applied to similar climates. Regarding the Rio Santa basin, we aim to provide insights into past and future changes. We achieve this as follows:

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

2. We introduce 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. We test the response and sensitivity of each metric to physically plausible changes in the rainy season.

4. We analyze changes in 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 in the onset and end of the rainy season based on CMIP5 models for the region.

2 Material and methods

2.1 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 and MYD13Q1 (Didan, 2015a, b) NDVI time series in 250 m spatial resolution (cf. 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), and (iii) masked based on unimodal seasonal vegetation development and land cover data to exclude pixels which are evidently decoupled from the rainy season. LSP was assessed by applying a relative threshold to the average seasonal cycle of vegetation greenness in the Rio Santa basin (Caparros-Santiago et al., 2021) to obtain the start and end of the growing season, based on which we subsequently calibrated all rainy season

metrics. Specifically, the start (hereafter SOS_{NDVI}) and end (hereafter EOS_{NDVI}) of the growing season were derived as the day where the processed NDVI data reached 30 % of their seasonal amplitude (Hänchen et al., 2022). The resulting LSP data were averaged to the extent of the Rio Santa basin.

In our analysis, we utilize multiple precipitation datasets. A key component is the WRF bias-corrected regional climate model data published by Potter et al. (2023a), which provide consistent precipitation estimates at 4 km grid spacing from 1981 to 2018. As a second gridded precipitation dataset for the recent past, 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 (AWS) operated by the National Meteorological and Hydrological Service of Peru (SENAHMI) which were sufficiently gap-free.

These three stations, located at Yungay (9.14° S, 77.75° W), Recuay (9.73° S, 77.45° W), and Santiago (9.52° S, 77.52° W), are all located along the valley floor of the Rio Santa basin (Hunziker et al., 2017). In addition, Potter et al. (2023a) produced statistically downscaled projections based on a 30-member CMIP5 ensemble from 2019 to 2100 using quantile delta mapping for the Representative Concentration Pathway (RCP) 4.5 and 8.5 scenarios. These scenarios represent different greenhouse gas concentration trajectories, where the number indicates the associated radiative forcing in 2100 (in $W\,m^{-2}$). RCP4.5 is a stabilization scenario with moderate mitigation efforts, while RCP8.5 represents a high-emission, business-as-usual trajectory. These data preserve CMIP5 model trends while adjusting precipitation magnitude and the number of rainy days based on the bias-corrected WRF data, and they are available in the same 4 km grid spacing from 2019 to 2100. The two RCP scenarios allow us to assess multiple trajectories of future changes in the rainy season in the Rio Santa basin and provide a large dataset for metric sensitivity analysis. We do not evaluate metrics for raw, coarse-scale CMIP data in this study because, at their native resolution, they are known to inadequately represent orographic processes and interannual variability (e.g., Gutierrez et al., 2024).

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 were then spatially averaged (i.e., no spatial dimension). We excluded leap-year days (29 February) and performed the analysis based on a hydrological year definition suitable for the Rio Santa basin starting from 1 September and ending on 31 August 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 the use of 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 (see Fig. A1).

2.2 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-odd, 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, 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 (DD) over a certain length within a defined period after the onset. Gurgiser uses conditions 1, 2, and 3; Climandes uses conditions 1, 2, and 4; Garcia uses conditions 2 and 4; and JD uses conditions 2, 3, and 4. FP uses a different approach by dividing a 30 d 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 are focused exclusively on the onset of the rainy season, the three remaining threshold-based metrics provide a more comprehensive approach by also addressing 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 changepoint detection method, fitting a two-phase linear regression iteratively over (i) the first 250 and (ii) the last 200 d 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., Giràldez et al., 2020; Marengo et al., 2001), are beyond the scope of this study and thus not included.

2.3 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, additionally constrained by a minimum and maximum bucket water content, ensuring realistic water balance limits.

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

BWC ($m^3\,m^{-3}$) represents the bucket water content at time t, BD is the bucket depth (m), ρ is the water density (constant here as $1000\,\text{kg}\,\text{m}^{-3}$), P ($\text{mm}\,\text{d}^{-1}$) is the precipitation input, and ET ($\text{mm}\,\text{d}^{-1}$) is the daily output.

Note that ET ($\text{mm}\,\text{d}^{-1}$) 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 and partly integrates other hydrological components, such as runoff; thus within- and between-seasonal variation is not directly accounted for. The metric starts at day $d = 0$ at an initial BWC (BWC$_{\text{ini}}$). The model is constrained, as no further evaporation occurs as soon as a minimum value (BWC$_{\text{mn}}$) is reached. Similarly, a maximum value (BWC$_{\text{mx}}$) is defined where no more water is accumulated – any surplus conceptually runs off or drains from the bucket. The parameters BWC$_{\text{in}}$, BWC$_{\text{mx}}$, BWC$_{\text{mn}}$, 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} and t_{RSE} in Figs. 2 and 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 us 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.

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 $\sum P_{d:d+6}$, for example, is the sum of precipitation on each day from the onset to 6 d after the onset. $N[P_{d:d+30} > 0\,\text{mm}]$ 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., $N_c[P_{d:d+30} < 0.1\,\text{mm}] < 7$ is the condition that no dry spells of more than 7 d occur in the 30 d 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.

Name (reference)	RSO (original)	RSO (adapted)	RSE (original)	RSE (adapted)
Gurgiser (Gurgiser et al., 2016)	$P_d > 0\,\text{mm}$ $\sum P_{d:d+7} > 10\,\text{mm}$ $N[P_{d:d+30} > 0\,\text{mm}] \geq 11$	$P_d > \alpha_p^1$ $\sum P_{d:d+\alpha_d^3} > \alpha_p^2$ $N[P_{d:d+\alpha_d^5} > \alpha_p^6] \geq \alpha_d^4$	$P_d = 0\,\text{mm}$ $\sum P_{d:d+46} < 10\,\text{mm}$	$P_d \leq \alpha_p^1$ $\sum P_{d:d+\alpha_d^3} < \alpha_p^2$
Climandes (Sedlmeier et al., 2023)	$P_d > 1\,\text{mm}$ $\sum P_{d:d+5} \geq 8\,\text{mm}$ $N_c[P_{d:d+30} < 0.1\,\text{mm}] \leq 6$	$P_d > \alpha_p^1$ $\sum P_{d:d+\alpha_d^3} \geq \alpha_p^2$ $N_c[P_{d:d+\alpha_d^5} < \alpha_p^6] \leq \alpha_d^4$	$P_d \leq 1\,\text{mm}$ $\sum P_{d:d+30} < 16\,\text{mm}$	$P_d \leq \alpha_p^1$ $\sum P_{d:d+\alpha_d^3} < \alpha_p^2$
Garcia (Garcia et al., 2007)	$\sum P_{d:d+3} > 20\,\text{mm}$ $N_c[P_{d:d+30} < 0.1\,\text{mm}] \leq 9$	$\sum P_{d:d+\alpha_d^2} > \alpha_p^1$ $N_c[P_{d:d+\alpha_d^4} < \alpha_p^5] \leq \alpha_d^3$	$P_{d:d+20} \leq 2^*\,\text{mm}$	$P_{d:d+\alpha_d^2} \leq \alpha_{pr}^1$
FP (Frere and Popov, 1986)	$\sum P_{d:d+10} \geq 25\,\text{mm}$ $\sum P_{d+10:d+20} \geq 20\,\text{mm}$ $\sum P_{d+20:d+30} \geq 20\,\text{mm}$	$\sum P_{d:d+\frac{1}{3}\alpha_d^2} \geq \alpha_p^1$ $\sum P_{d+\frac{1}{3}\alpha_d^2:d+\frac{2}{3}\alpha_d^2} \geq \alpha_p^3$ $\sum P_{d+\frac{2}{3}\alpha_d^2:d+\alpha_d^2} \geq \alpha_p^4$	Onset only	Onset only
JD (Jolliffe and Sarria-odd, 1994; originally based on Stern et al., 1981)	$N[P_{d:d+5} > 0.1\,\text{mm} \geq 3$ $\sum P_{d:d+5} \geq 25\,\text{mm}$ $N_c[P_{d:d+30} < 0.1\,\text{mm}] \leq 7$	$N[P_{d:d+\alpha_d^2} > \alpha_p^1] \geq \alpha_d^3$ $\sum Pr_{d:d+\alpha_d^2} > \alpha_p^4$ $N_c[P_{d:d+\alpha_d^6} < \alpha_p^1] \leq \alpha_d^5$	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 changepoint detection to determine when daily precipitation changes from decreasing each day to increasing each day (onset) and from increasing to decreasing (end).			

* For RSE$_{Garcia}$, which is published as $P_{d:d+20} = 0\,\text{mm}$, we used a value of 2 mm instead, as 20 consecutive days with zero precipitation are not present in any of the datasets.

2.4 Calibration of threshold-based metrics

Using the NDVI-derived SOS$_{NDVI}$ and EOS$_{NDVI}$ as targets, we firstly tested the initial parameters provided by the corresponding authors. We then calibrated each threshold-based metric, along with our novel metric, by adjusting their parameters (see Table 1) for each of the three precipitation time series. This was done using 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 us to obtain robust correlation and was therefore omitted. To allow the robust and efficient processing of a large number of time series regarding the threshold-based metrics, we generally started the iterative search for the RSE starting from the previously derived RSO. Addition-

ally, when an RSE was found within the 90 subsequent days following the RSO, the iterative search was continued to account for erroneous detection of dry spells in the early rainy season. For all metrics, RSEs were discarded if an unrealistically early end (before 1 February) occurred.

2.5 Sensitivity analysis

To understand how tuned threshold-based metrics and objective methods respond to hydro-climatological changes, we utilize the large number of rainy seasons (~ 5000) provided by the future projections of Potter et al. (2023a) to assess sensitivities of the metrics with regard to potential and physically plausible changes in the rainy season. To account for a variety of scenarios, we correlate the rainy season metric outputs (RSO/RSE) calculated for all rainy seasons, independently of model, year, or scenario, with both
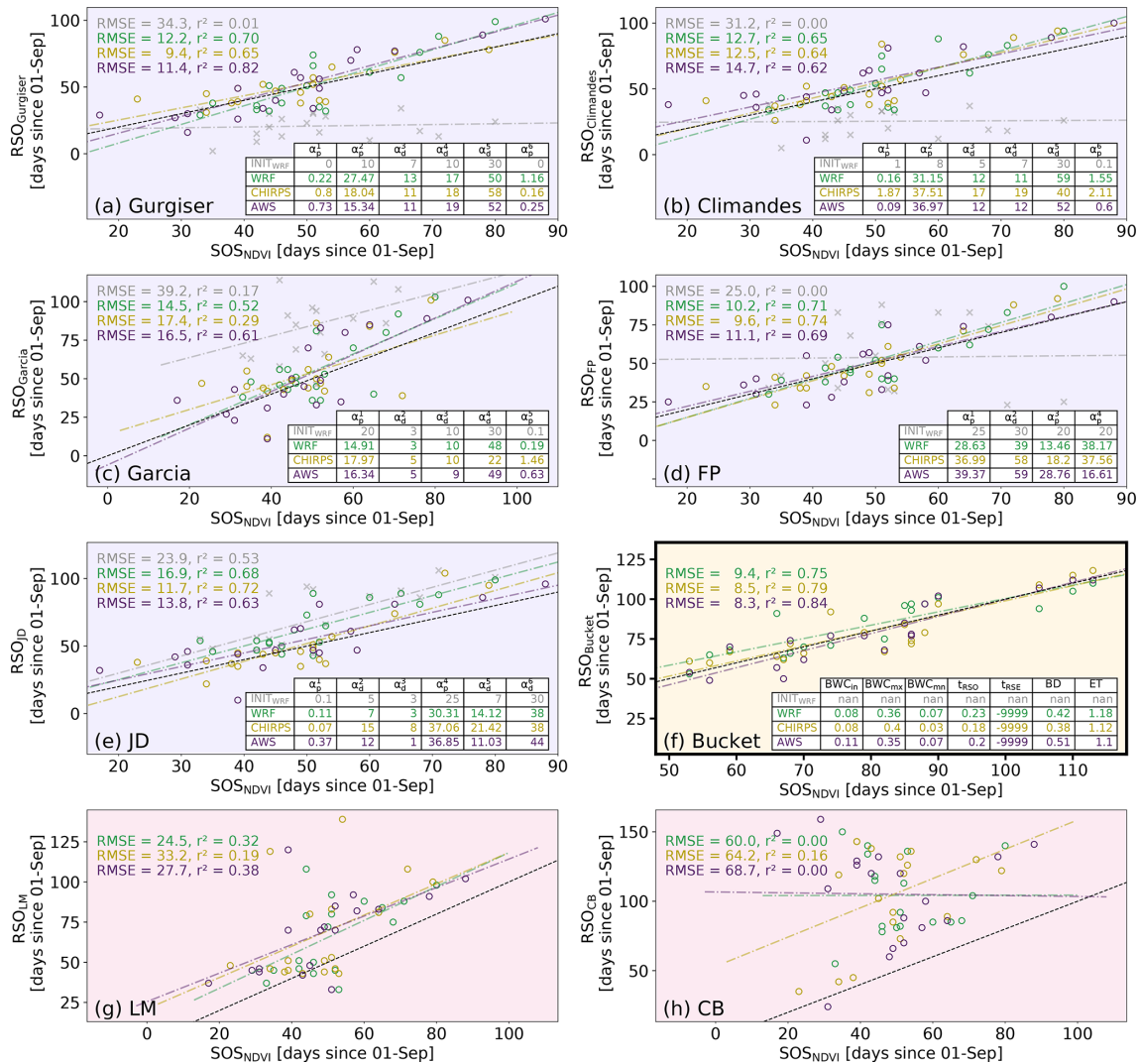
Figure 2. Rainy season onset metric outputs for WRF (teal), CHIRPS (orange), and AWS (purple) precipitation data. For the threshold-based metrics (purple background; **a–e**), results of the evaluation based on WRF but with uncalibrated thresholds as provided by the respective authors are also shown (gray) and are denoted as $\mathrm{INIT_{WRF}}$. The novel bucket methodology is highlighted in yellow (**f**), and the objective methods are highlighted in red (**g–h**). The dashed black line indicates the 1 : 1 relationship with the $\mathrm{SOS_{NDVI}}$, while the colored lines correspond to the regression of the parameters. Annotated are the root-mean-square error (RMSE) and the coefficient of determination (r^2). The tables correspond to the parameters as outlined in Table 1 and in the equation in Sect. 2.3 after calibration for all of the precipitation data. The LM and CB methods have no calibration; therefore no table is shown.

full-hydrological-year and sub-seasonal (SON, DJF, MAM, JJA) rainfall sums. The sub-seasonal rainfall sums refer to the same rainy season upon which RSO/RSE values were derived, 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) based on the WRF and statistically downscaled CMIP5 data created by Potter et al. (2023a).

These are the number of dry and wet days (DD and WD), defined as days with precipitation less than and greater than 1 mm; the simple precipitation intensity index (SDII) represents 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). Using LSP data as dependent variables and rainfall sums as independent variables, we assess sensitivities by applying bin-weighted linear regression for each variable independently. Bin sizes for the regressions are determined in an objective fashion by applying the Freedman–Diaconis rule (Freedman and Diaconis, 1981) to each of the nine independent variables.
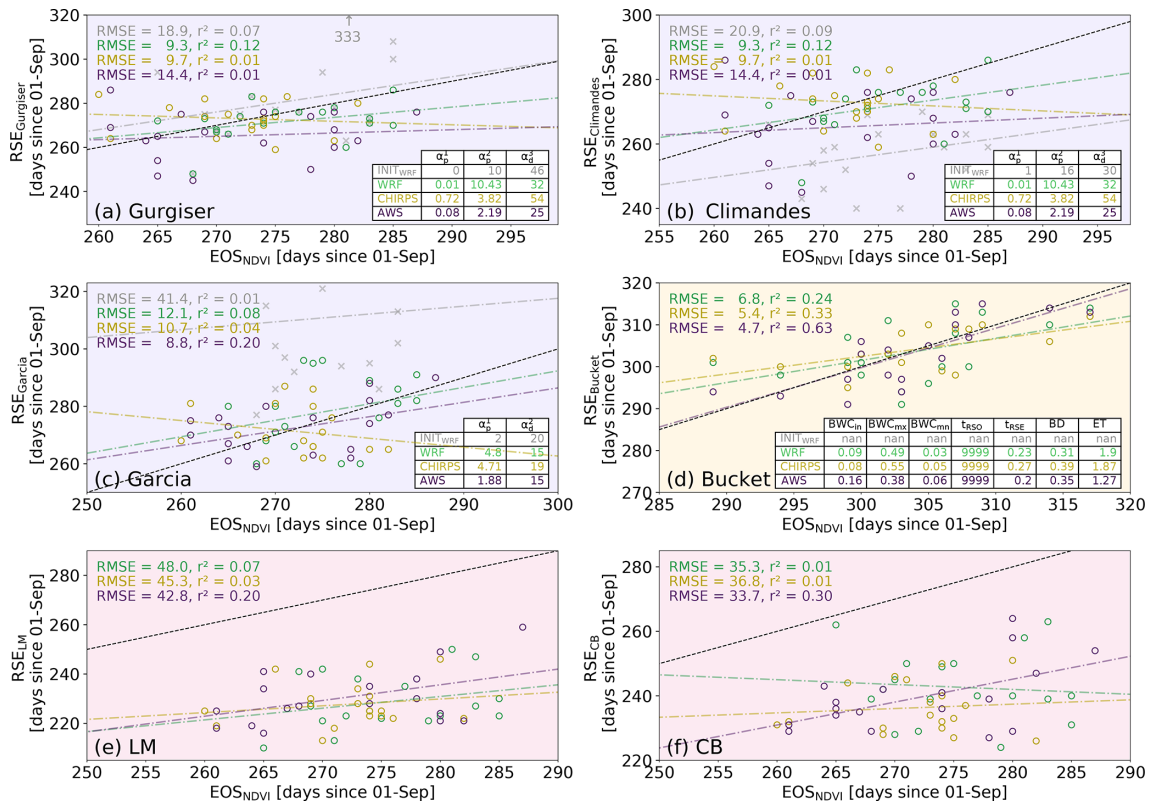
(a) Gurgiser — RMSE = 18.9, r^2 = 0.07; RMSE = 9.3, r^2 = 0.12; RMSE = 9.7, r^2 = 0.01; RMSE = 14.4, r^2 = 0.01

	α_p^1	α_p^2	α_d^3
INIT$_{WRF}$	0	10	46
WRF	0.01	10.43	32
CHIRPS	0.72	3.82	54
AWS	0.08	2.19	25

(b) Climandes — RMSE = 20.9, r^2 = 0.09; RMSE = 9.3, r^2 = 0.12; RMSE = 9.7, r^2 = 0.01; RMSE = 14.4, r^2 = 0.01

	α_p^1	α_p^2	α_d^3
INIT$_{WRF}$	1	16	30
WRF	0.01	10.43	32
CHIRPS	0.72	3.82	54
AWS	0.08	2.19	25

(c) Garcia — RMSE = 41.4, r^2 = 0.01; RMSE = 12.1, r^2 = 0.08; RMSE = 10.7, r^2 = 0.04; RMSE = 8.8, r^2 = 0.20

	α_p^1	α_p^2
INIT$_{WRF}$	2	20
WRF	4.8	-15
CHIRPS	4.71	19
AWS	1.88	15

(d) Bucket — RMSE = 6.8, r^2 = 0.24; RMSE = 5.4, r^2 = 0.33; RMSE = 4.7, r^2 = 0.63

	BWC$_{in}$	BWC$_{mx}$	BWC$_{mn}$	t_{RSO}	t_{RSE}	BD	ET
INIT$_{WRF}$	nan	nan	nan	nan	nan	nan	nan
WRF	0.09	0.49	0.03	9999	0.23	0.31	1.9
CHIRPS	0.08	0.55	0.05	9999	0.27	0.39	1.87
AWS	0.16	0.38	0.06	9999	0.2	0.35	1.27

(e) LM — RMSE = 48.0, r^2 = 0.07; RMSE = 45.3, r^2 = 0.03; RMSE = 42.8, r^2 = 0.20

(f) CB — RMSE = 35.3, r^2 = 0.01; RMSE = 36.8, r^2 = 0.01; RMSE = 33.7, r^2 = 0.30

Axes: RSE [days since 01-Sep] vs EOS$_{NDVI}$ [days since 01-Sep]

Figure 3. Same as Fig. 2 but for rainy season end (RSE) and end of season (EOS).

2.6 Future projections

To reliably determine trends in future CMIP5 projections, we firstly excluded individual models for each rainy season metric if they produced five or more invalid values out of 81 seasons (2019–2100). Invalid values occurred when the conditions for RSO or RSE were not met within a given hydrological year. For the remaining data, we calculated linear trends separately for the historical WRF and CHIRPS datasets, and for both RCP scenarios of the CMIP5 ensemble, using linear regression. Trend significance was assessed using a Wald test, with the null hypothesis that the slope is zero.

3 Results and discussion

3.1 Evaluation of rainy season metrics

We firstly 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 illustrate 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, INIT$_{WRF}$) setup of the metric, which, for all threshold-based metrics except JD, lacks correlation and showcases higher root-mean-square er-

ror (RMSE) values. The bucket metric stands out by exhibiting low errors (average RMSE = 8.7 d) and a robust correlation ($r^2 = 0.79$, on average) across all three input datasets (see Figs. 2 and 4a). This is likely related to the fact that the bucket metric was 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 time series 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 (see 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 d), albeit lacking correlation (maximum $r^2 = 0.25$), whereas the bucket metric (Fig. 3d) shows an even lower RMSE (5.6 d 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 Sect. 2.3). Gurgiser and Climandes share the same criteria for RSE; thus the resulting calibration is identical (see Table 1 and Fig. 3). The LM and CB metrics show
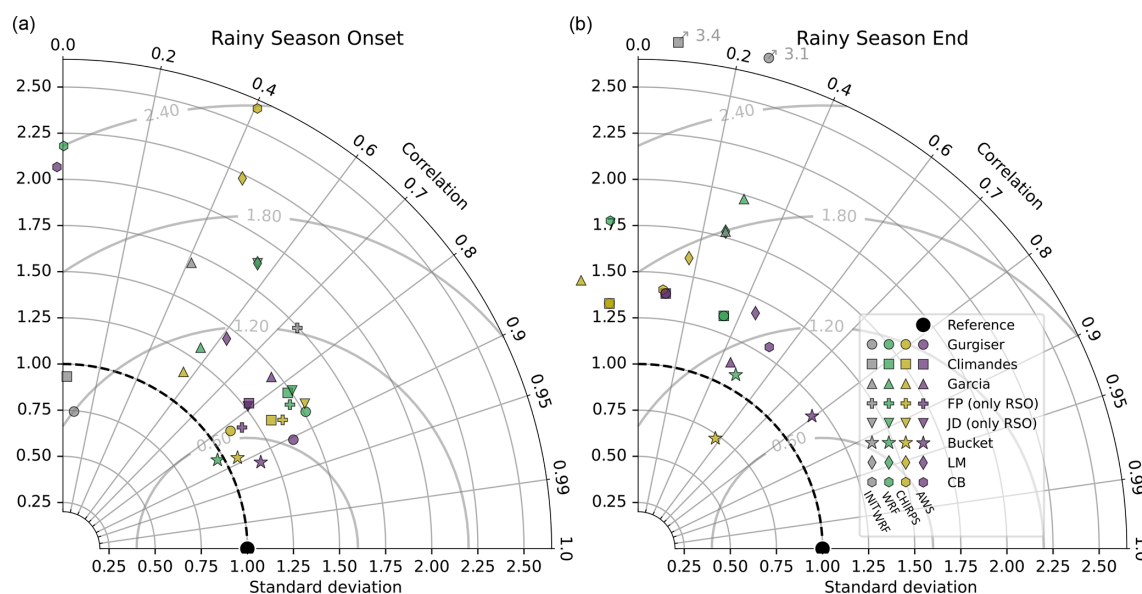
Figure 4. Taylor diagrams for normalized rainy season onset **(a)** and end **(b)** for the calibration period 2000–2018. Each data point represents a metric depicted by the symbols, and the colors represent the corresponding datasets (gray: WRF with initial metric parameters; teal: WRF-calibrated; orange: CHIRPS; purple: AWS). The radial axes represent standard deviation, the azimuthal axis represents the correlation coefficient (r), and the circles represent the centered root-mean-square difference. The black reference dot represents the normalized NDVI-derived SOS_{NDVI} and EOS_{NDVI} standard deviation.

an overall low skill in predicting the lag-corrected EOS_{NDVI} with high errors, with LM showing a weak correlation for two of three datasets (see Fig. 3f). The overall discrepancy across the metrics between skill in predicting SOS_{NDVI} and EOS_{NDVI} (see Fig. 4) may be linked to EOS_{NDVI} displaying low variability (standard deviation $\sigma = 6.63$ d), unlike the higher variability in SOS_{NDVI} ($\sigma = 17.61$ d). 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 (see tables in Figs. 2a–f and 3a–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. Our generally skillful results after calibration also illustrate

that existing concerns (e.g., MacLeod, 2018) regarding the sensitivity of threshold-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 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 the resulting values. Here, we followed the same approach as the original authors, using the first 250 d and the last 200 d 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 2 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 in 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 (Ferijal et al., 2022) when exploring this method.

3.2　Sensitivity analysis of rainy season metrics

To assess the sensitivity of rainy season metrics (RSO/RSE) to hydro-climatological changes, we correlated them with full-hydrological-year and sub-seasonal (SON, DJF, MAM, JJA) rainfall sums, as well as four ETCCDI climate indices as explained in Sect. 2.5. The results of these regressions are summarized in Fig. 5, with detailed plots provided in Figs. A4 and A5. In the context of the RSO (Figs. 5a, 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 metrics 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 to its generally later onsets (see 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 the transfer of 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 artifact that does not reflect any physical process related to RSO water availability. This indicates limited metric robustness of the objective metrics to changes in the rainy season. Note that the start of CB is based on the period of the first 250 d of the rainy season, meaning that the metric is based only on in-

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

Similarly to what was previously seen in the calibration results (see Figs. 3 and 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 Figs. 5b and 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, which suggests 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 metrics 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, particularly when rainy season characteristics are expected to change. While less clear for the RSE, the overall message is similar, with the bucket metric and the threshold-based metrics being the most reliable.

3.3　Past and future trends

Finally, we calculated past metrics based on WRF data from 1981 to 2018 and projected future metrics up to 2100 using the statistically downscaled CMIP5 model ensemble, which comprises 30 individual models (29 for RCP4.5), and subsequently evaluated the trends for the historical and future periods. 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, or was even 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 3 months leading to non-fulfillment of the metric criteria and thus resulting in NaN values. Interestingly, LM and CB do not show any anomaly for this event because these metrics do not have information about any form of climatology. Con-
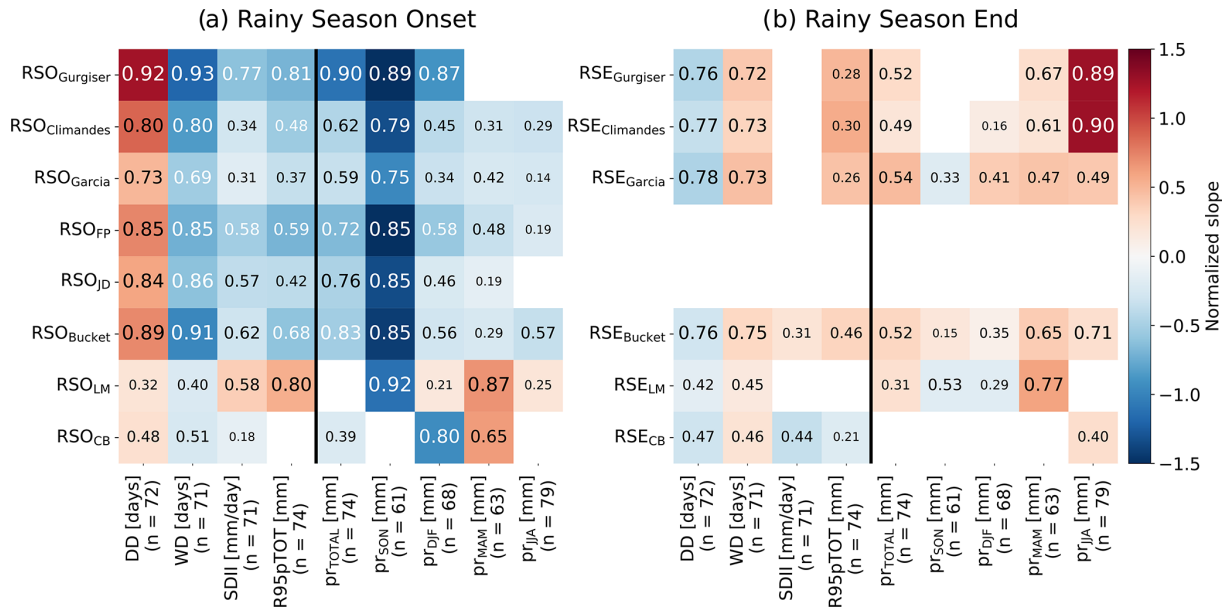
Figure 5. Heatmap of bin-weighted regression slopes with annotated r^2 values between ETCCDI indices and seasonal precipitation sums (independent variables) and rainy-season-metric derived onset and end (dependent variables). Corresponding bin sizes are noted on the x-axis labels as ($n = x$). Slope 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.

versely, 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 suggest a trend for the past period, either for the rainy season onset or for the end of the rainy season (see Figs. 6 and 7).

After establishing variability and trends for the historical period, we now explore the projected changes of rainy season metrics for the ensemble mean and standard deviation for each of the two RCP scenarios. Most of the metrics do not suggest a change in either the onset or the end of the rainy season until the end of the century (Figs. 6 and 7). Only the JD and FP metrics suggest earlier rainy season onsets (approximately 0.5 d per decade earlier for the stabilization scenario, RCP4.5; Fig. 6). Meanwhile, only the bucket metric suggests a small delay in rainy season ends, with a decadal slope of approximately 0.35 to 0.6 d for both scenarios (Fig. 7). In light of the anticipated increase in future precipitation in the Rio Santa basin (5.8 % ± 6.3 % for RCP4.5 and 12.1 % ± 11.0 % for RCP8.5; Potter et al., 2023a), combined with the sensitivities of the metrics discussed in Sect. 3.2, the results appear surprising. To investigate the apparent contradiction between increasing future annual precipitation trends and little change in the onset or end of the rainy season, we apply a trend analysis across monthly precipitation sums for each month of the year in the future

CMIP5 ensemble (Fig. A6). As this shows that the months September and October do not show significant trends for either scenario and that, for RCP4.5, only January and April show significant precipitation increases (see 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 most of the metrics to determine start and end. In absolute values, these trends in the dry months are very small (with decadal slopes of 0.046 for May, 0.022 for June, and 0.004 mm d^{-1} for July and August; see Fig. A6), while the calibrated values for the dry day threshold to determine RSE (see Fig. 3 and 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 (see Fig. A2) and that the decrease in water availability and thus the resulting rainy season end will be delayed.

There is between-model variability in future predictions of both rainy season start and end for all metrics (Figs. A7 and A8), making the resulting trends debatable. This is represented by only 7 out of 30 RCP8.5 and 2 out of 29 RCP4.5 CMIP5 models, suggesting a significant delay individually in the case of the bucket metric, and with 1 model even suggesting an earlier RSE under RCP4.5. An assessment of the distribution of significance of model trends for each metric and scenario can be found in Figs. A7 and A8. These results re-
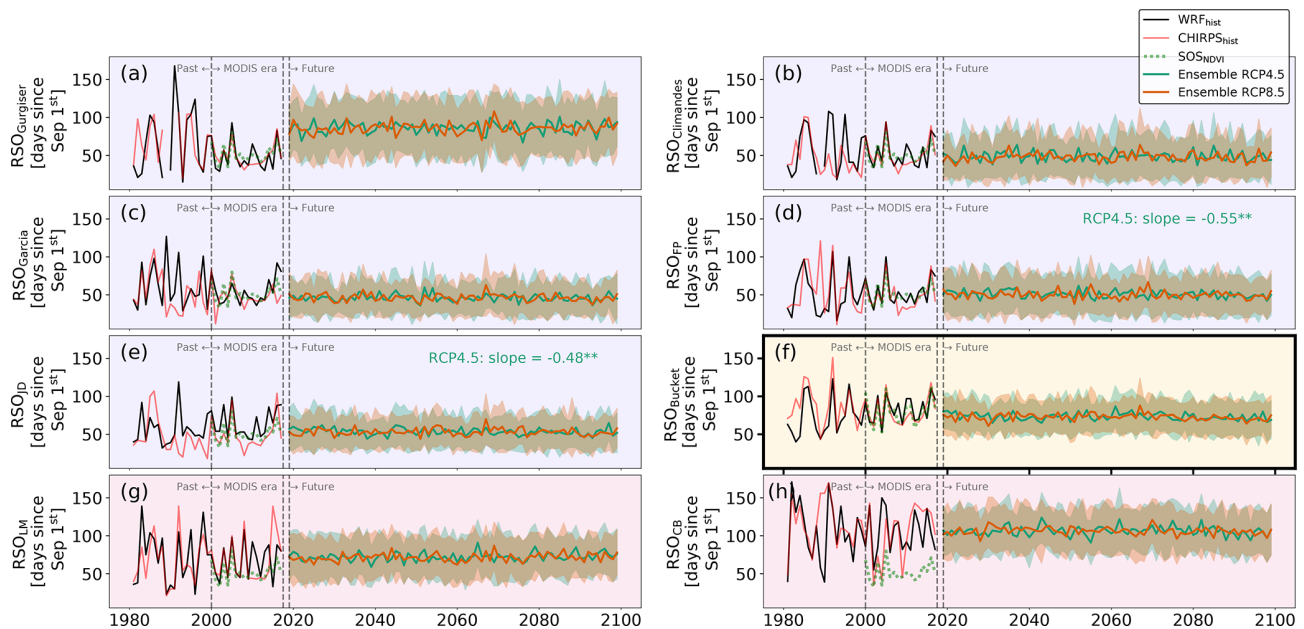
Figure 6. Rainy season onset (RSO) derived from eight different metrics during the past, calibration (MODIS era), and future periods, where threshold-based metrics are indicated by a purple background, the bucket metric is indicated by a yellow background, and the objective metrics are indicated by a red background. Solid lines represent WRF-derived (black) and CHIRPS-derived (red) 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 1 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 insignificant trends ($p > 0.05$) are not displayed.
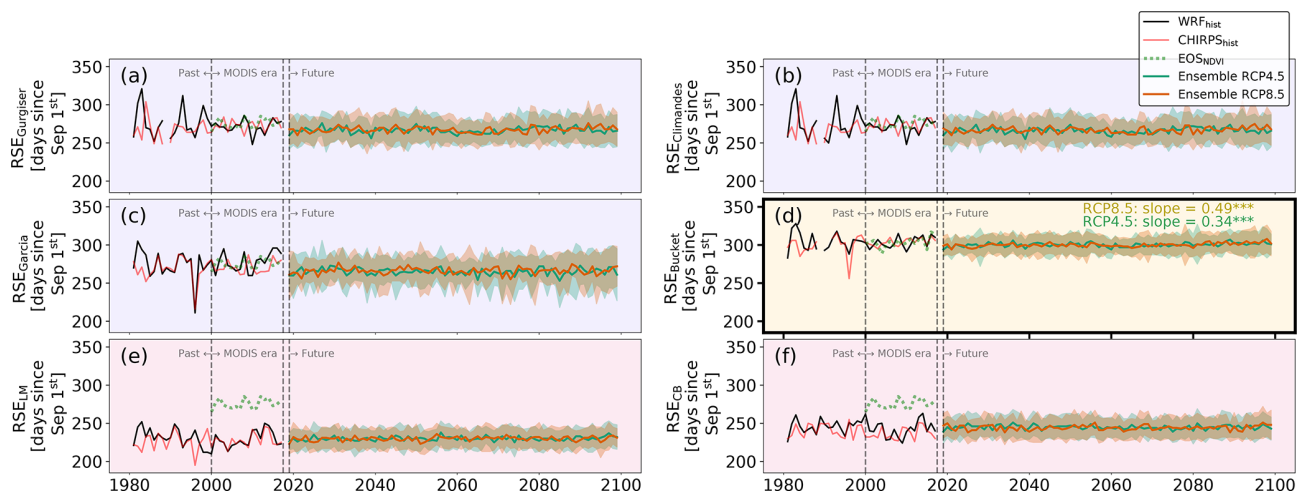
Figure 7. Same as Fig. 6 but for rainy season end (RSE)

flect 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 (as shown in Figs. 6 and 7).

The projections by Potter et al. (2023a) we use are based on statistical downscaling of CMIP5 models. At the continental scale, many CMIP5 models were previously reported

to poorly represent the South American Monsoon System (SAMS) (Bombardi and Carvalho, 2008), a challenge that is particularly pronounced in the topographically complex Andes. We compare our results to those of De la Cruz et al. (2025), who performed statistical downscaling based on meteorological stations in Peru using CMIP6 data and analyzed changes through the LM metric. De la Cruz et al. (2025) also

project an increase in total precipitation, consistent with the findings of Potter et al. (2023a), whose data informed our study. De la Cruz et al. (2025) also find no significant future changes in rainfall seasonality using the LM metric for the domain in which the Rio Santa basin is located. Furthermore, they highlight that GCMs have limited skill in simulating the interannual variability in rainy season onset and end, noting that many CMIP6 simulations still struggle to adequately represent the SAMS (see also Olmo et al., 2022). This suggests that the results from downscaled CMIP6 models and the downscaled CMIP5 models used in this study are consistent, at least based on the LM metric. Our findings 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 suggest, using the LM metric, earlier rainy season onsets and later retreats. This could be related to several key differences, which are the larger CMIP5 model ensemble used here, a spatial mismatch between the Rio Santa basin and the greater region, resolution differences, and the fact that the LM metric can be subject to inconsistent sensitivities to hydroclimatic change as we previously showed (Fig. 5).

Future predictions are further complicated by the limited understanding of expected ENSO changes and their effects in the region. While Cai et al. (2023) recently suggested an increase in ENSO variability linked to anthropogenic climate change, reliable ENSO-related predictions about the potential alteration of the rainy season and general precipitation patterns in the Rio Santa basin specifically cannot be made confidently at this time. Our results incorporate a large number of calibrated and sensitivity-tested rainy season metrics, combined with a high-resolution, bias-corrected large ensemble of future precipitation datasets. As such, we suggest that studies suggesting future change in rainy season timing should be interpreted with caution in terms of climate model ensemble robustness and, as our results indicate, critically reviewed towards the calibration of rainy season metrics.

Finally, as we are calibrating the metrics on a vegetation proxy, the effects of future increasing temperatures on evapotranspiration should also be considered, as these are expected increase in the Rio Santa basin with rising temperatures (see Potter et al., 2023a). 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. Meanwhile, it is therefore crucial to consider that, when metrics like these are applied with water users in mind, factors beyond precipitation change (i.e., rising temperatures, wet-/dry-spell frequency) must also be taken into consideration to ensure their practical relevance.

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 precipitation datasets considered, 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 and exhibit sensible sensitivities to potential hydroclimatic changes. More objective and flexible metrics, on the other hand, 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. We therefore recommend that the usage of such methods should at least be critically reviewed on a case-by-case basis to ensure that no false conclusions are drawn or misleading practical recommendations are made.

Considering the numerous publications that highlight threshold-based metrics and propose a fixed-parameter setup to be suitable 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 do, however, 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 are available in sufficient quality. Examples could be, but are not limited to, (undisturbed) runoff measurements or soil moisture data.

Motivated by limitations in existing metrics, we designed a novel bucket metric, which outperforms other metrics for both the onset and end of the rainy season, shows physically consistent sensitivities, and corrects for the vegetation–precipitation lag. The high skill and flexibility of the bucket metric allow a wide range of applications in the context of hydroclimate in semi-arid areas. Additionally, this metric can likely be extended, e.g., by making evapotranspiration dependent on energy and/or water availability or by altering parameters over time to simulate changes 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, which is particularly relevant in regions such as the Rio Santa basin, where large interannual precipitation anomalies, for example, related to ENSO, are common. Future attempts to address questions regarding the rainy season across semi-arid regions can readily use or adapt the bucket metric to suit a wide range of requirements.

Using the bucket metric together with other calibrated and sensitivity-tested rainy season metrics and a large number of future projections, 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 in the rainy season end and consequently an increase in the rainy season length. Considering high regional interannual 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, which will allow better representation of local precipitation. In addition, evapotranspiration changes should be further investigated, most appropriately analyzed through a sophisticated eco-hydrological model. Until then, both practitioners and researchers can profit from more robust predictions of water availability building on our novel framework.

Appendix A
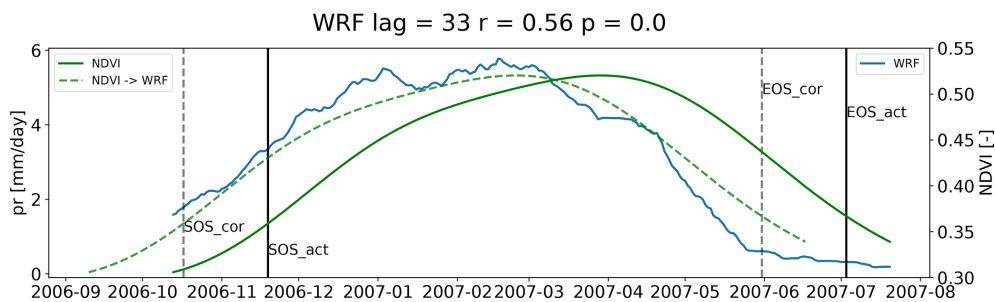
Figure A1. Visual example of the cross-correlation function for lag correction of 1 hydrological year. The lag was determined based on WRF data smoothed by a 12-week rolling average and the processed NDVI data
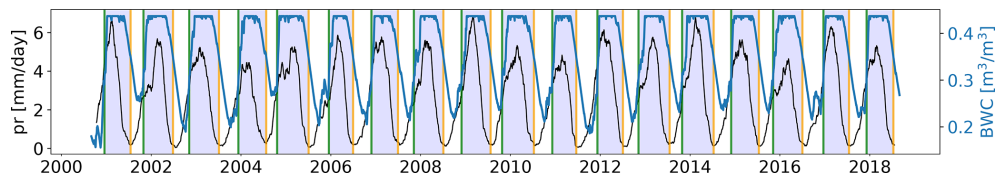
Figure A2. A 12-week rolling window WRF time series (black) and BWC, modeled from daily (non-smoothed) precipitation from the bucket metric (blue) for the calibration period 2000–2018. Green (orange) vertical lines indicate RSO (RSE) dates derived by the bucket metric. Blue shading indicates the resulting rainy season.
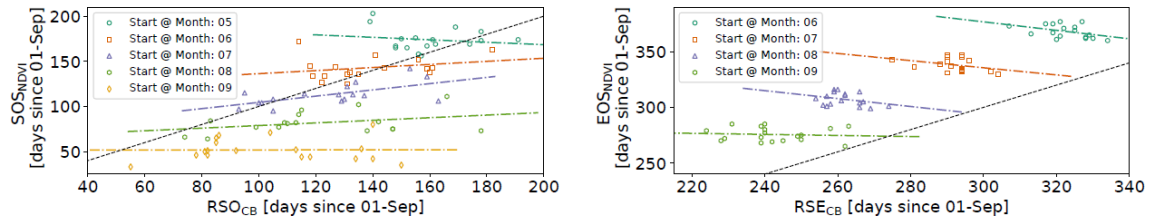
Figure A3. Sensitivity of the two-phase linear regression method to the hydrological year definition by Cook and Buckley (2009).
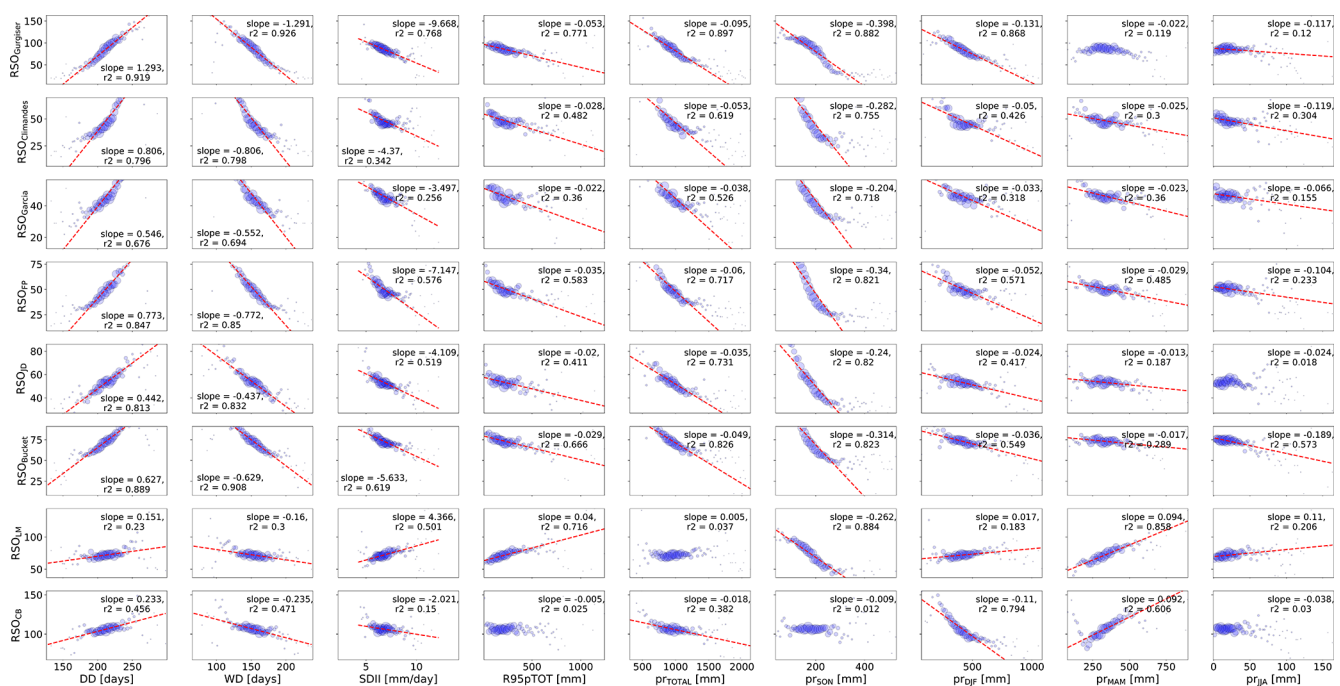

Figure A4. Bin-weighted regressions for RSO as summarized in Fig. 5a. Red regression lines are only shown for significant regressions ($p < 0.01$).
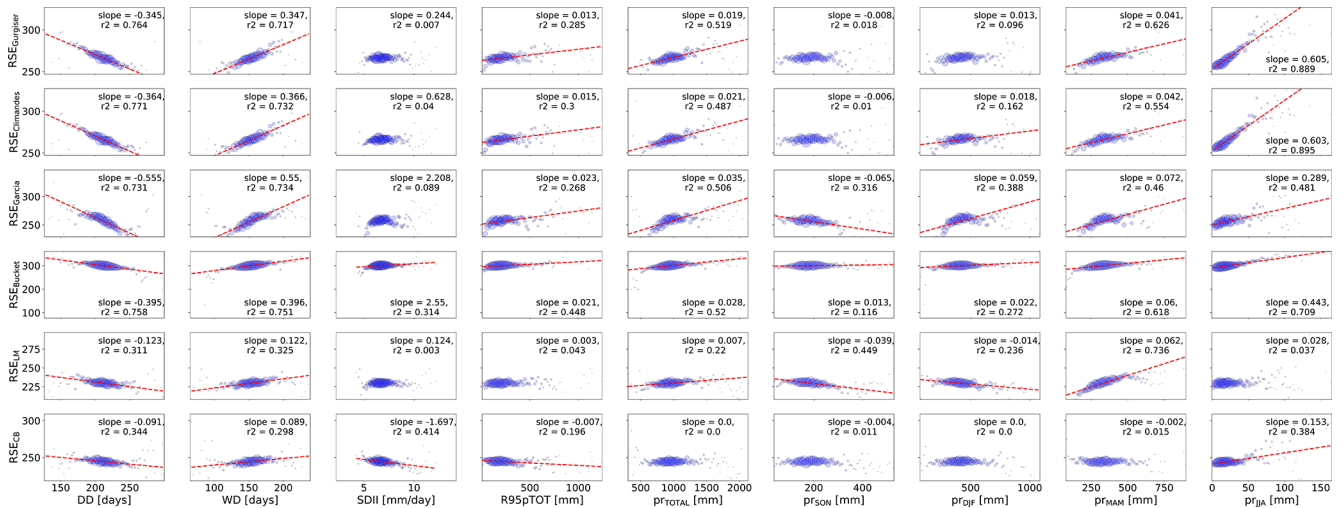
Figure A5. Bin-weighted regressions for RSE as summarized in Fig. 5b. Red regression lines are only shown for significant regressions ($p < 0.01$).
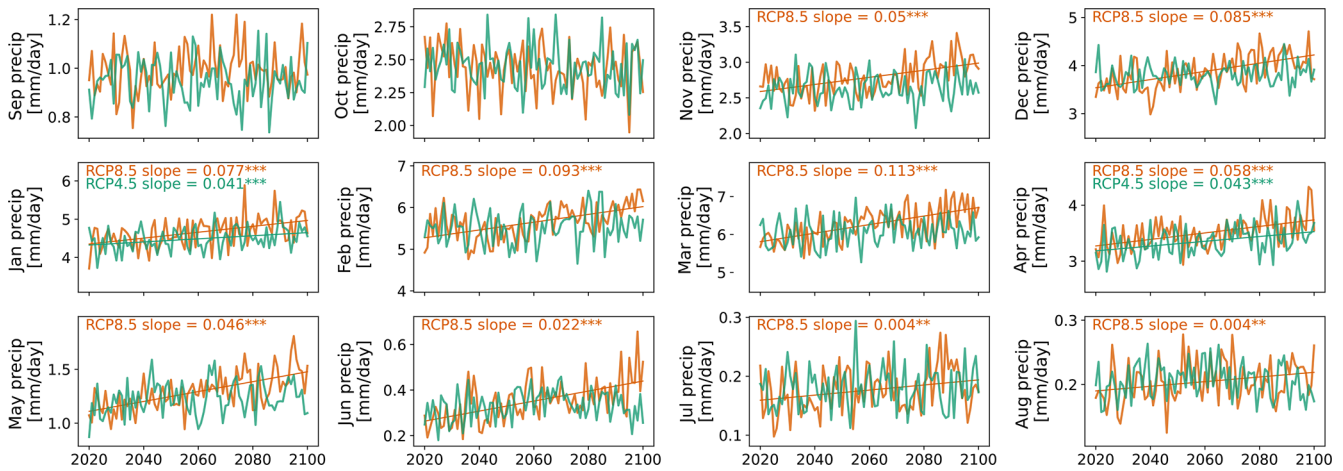
Figure A6. Monthly trends for the CMIP5 ensemble for both the RCP4.5 (teal) and RCP8.5 (brown) scenarios. Decadal trends were derived through linear regression. Significant trends are denoted by asterisks, *** for $p < 0.01$ and ** for $p < 0.05$, while regression lines for non-significant trends ($p > 0.05$) are not displayed.
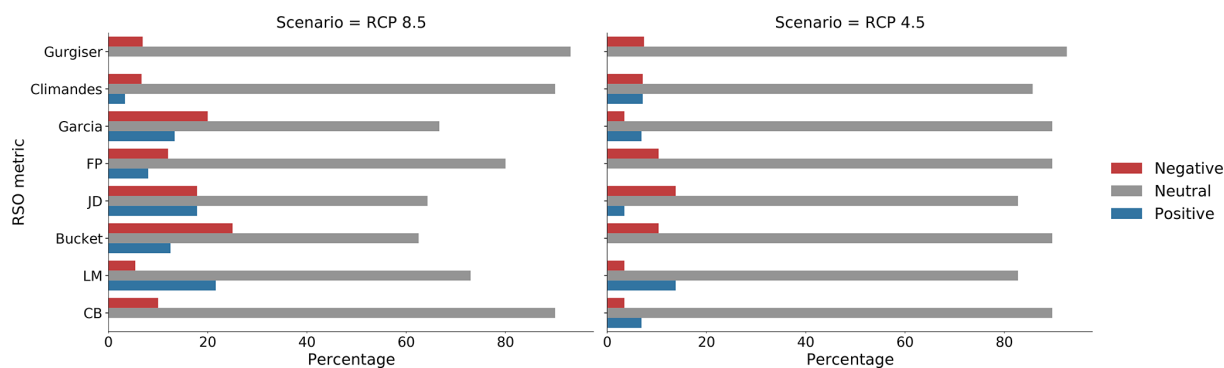
Figure A7. Relative distribution of significant and non-significant CMIP5 model time series ($p < 0.05$) and their sign for the derived rainy season onsets for the time period 2019–2100 for each rainy season metric and both RCP scenarios. A negative trend refers to an earlier season start, and a positive trend refers to a later start.
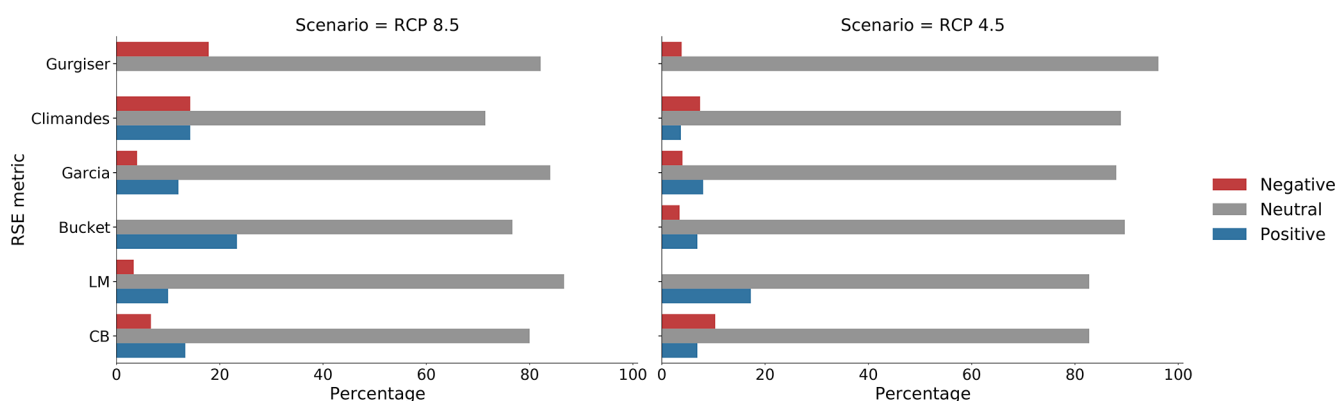
Figure A8. Same as Fig. A7 but for RSE.

Code and data availability. Pre-processed data and Python code to recreate the analysis and figures are available at https://github.com/lohae/RainySeasonMetrics (last access: 18 October 2024) and are preserved at https://doi.org/10.5281/zenodo.13952139 (Hänchen et al., 2024), allowing the application and testing of metrics for other regions or data. Full, bias-corrected WRF data can be obtained at https://doi.org/10.5285/2cf25580-9b79-440f-8505-6230dd377877 (Potter et al., 2023b). The future precipitation from the statistically downscaled CMIP5 models is available at https://doi.org/10.5285/67CEB7C8-218C-46E1-9927-CFEF2DD95526 (Potter et al., 2023c), and that of the ETCCDI is available at https://doi.org/10.5285/B56D30E8-EDAA-4225-96D7-FCC689E930C7 (Potter et al., 2023d). Full CHIRPS data can be obtained at https://data.chc.ucsb.edu/products/CHIRPS-2.0/ (Funk et al., 2015), while NDVI raw data can be acquired at (for example) https://doi.org/10.5067/MODIS/MOD13Q1.006 (Didan, 2015a) and https://doi.org/10.5067/MODIS/MYD13Q1.006 (Didan, 2015b). The AWS data are publicly available at https://www.senamhi.gob.pe/?p=estaciones (SENAMHI, 2025); however, we acquired them 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 and editing). CK: conceptualization, data curation, formal analysis, software, writing (review and editing). PC: conceptualization, funding acquisition, supervision, writing (review and editing). WG: writing (review and editing). FM: data curation, funding acquisition, project administration, resources, supervision, writing (review and editing). GW: conceptualization, methodology, resources, software, supervision, writing (review and editing).

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes ev-

ery effort to include appropriate place names, the final responsibility lies with the authors.

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, 104666, 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, Clim. Dynam., 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, B. Am. Meteorol. Soc., 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, Int. J. Climatol., 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, Clim. Res., 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 J. Photogramm., 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, J. Geophys. Res.-Atmos., 114, D23109, https://doi.org/10.1029/2009jd012795, 2009.

De la Cruz, G., Huerta, A., Espinoza, J.-C., and Lavado-Casimiro, W.: Present Variability and Future Change in Onset and Cessation of the Rainy Season Over Peru, Int. J. Climatol., 45, e8700, https://doi.org/10.1002/joc.8700, 2025.

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, Ecosyst. Serv., 56, 101454, https://doi.org/10.1016/j.ecoser.2022.101454, 2022.

Didan, K.: MOD13Q1 MODIS/Terra vegetation indices 16-day L3 global 250m SIN grid V006, NASA Land Processes Distributed Active Archive Center (LP DAAC) [data set], https://doi.org/10.5067/MODIS/MOD13Q1.006, 2015a.

Didan, K.: MYD13Q1 MODIS/Terra vegetation indices 16-day L3 global 250m SIN grid V006, NASA Land Processes Distributed Active Archive Center (LP DAAC) [data set], 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, J. Geophys. Res.-Atmos., 121, 11405–11424, 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, Front. Earth Sci., 8, 64, 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, Theor. Appl. Climatol., 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, J. 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: L2 theory, Z. Wahrscheinlichkeit., 57, 453–476, https://doi.org/10.1007/BF01025868, 1981.

Frere, M. and Popov, G.: Early agrometeorological Crop Yield Forecasting, FAO Plant Production and Protection Paper, 73, 144, ISBN 13 9789251024195, ISBN 10 9251024197, 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, 150066, https://doi.org/10.1038/sdata.2015.66, 2015 (data available at: https://data.chc.ucsb.edu/products/CHIRPS-2.0/, last access: 16 November 2020).

Fyffe, C. L., Potter, E., Fugger, S., Orr, A., Fatichi, S., Loarte, E., Medina, K., Hellström, R. Ã., 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, J. Geophys. Res.-Atmos., 126, e2021JD034911, 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, J. Arid Environ., 71, 109–121, https://doi.org/10.1016/j.jaridenv.2007.02.005, 2007.

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

Giràldez, 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, 23, 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 Syst. Dynam., 7, 499–515, https://doi.org/10.5194/esd-7-499-2016, 2016.

Gutierrez, R. A., Junquas, C., Armijos, E., Sörensson, A. A., and Espinoza, J.-C.: Performance of Regional Climate Model Precipitation Simulations Over the Terrain-Complex Andes-Amazon Transition Region, J. Geophys. Res.-Atmos., 129, e2023JD038618, https://doi.org/10.1029/2023JD038618, 2024.

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, 10, 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., Oria, C., Croci-Maspoli, M., Konzelmann, T., Rohrer, M., and Brönnimann, S.: Identifying, attributing, and overcoming common data quality issues of manned station observations, Int. J. Climatol., 37, 4131–4145, https://doi.org/10.1002/joc.5037, 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 Syst. Dynam., 13, 595–611, https://doi.org/10.5194/esd-13-595-2022, 2022.

Hänchen L., Potter, E., Klein, C., Calanca, P., Maussion, F., Gurgiser, W., and Wohlfahrt, G.: Code to recreate the analysis and figures of: "A Novel Framework for Analyzing Rainy Season Dynamics in semi-arid environments: A case study for the Peruvian Rio Santa Basin" (v1.0), Zenodo [code], https://doi.org/10.5281/zenodo.13952139, 2024.

Jekel, C. F. and Venter, G.: piecewise_linear_fit_py: fit piecewise linear data for a specified number of line segments, GitHub [code], https://github.com/cjekel/piecewise_linear_fit_py (last access: 2 November 2023), 2019.

Jolliffe, I. T. and Sarria-dodd, D. E.: Early detection of the start of the wet season in tropical climates, Int. J. Climatol., 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, J. Climate, 26, 6660–6678, https://doi.org/10.1175/jcli-d-12-00412.1, 2013.

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?, B. 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, Environ. Res. Lett., 18, 034002, 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, 071004, 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, Int. J. Climatol., 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, J. 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. V., 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, J. Climate, 20, 2037–2050, https://doi.org/10.1175/JCLI4122.1, 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, J. 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. Å., 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 Syst. Sci. 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., Rothenpieler, T., Bell, R., Li, F., Landmann, J., Dusch, M., Sun, T., hannah, paolodeidda, and tbridel: fmaussion/salem: v0.3.9 (v0.3.9). Zenodo [code], https://doi.org/10.5281/zenodo.7554820, 2023.

McKinney, W.: Data structures for statistical computing in python, in: Proceedings of the 9th Python in Science Conference, Austin, TX, 445, 51–56, https://doi.org/10.25080/Majora-92bf1922-00a, 2010.

Olmo, M. E., Espinoza, J.-C., Bettolli, M. L., Sierra, J. P., Junquas, C., Arias, P., Moron, V., and Balmaceda-Huarte, R.: Circulation patterns and associated rainfall over south tropical South America: GCMs evaluation during the dry-to-wet transition season, J. Geophys. Res.-Atmos., 127, e2022JD036468, https://doi.org/10.1029/2022JD036468, 2022.

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 Resour. Res., 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. Å., 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, 96, https://doi.org/10.1038/s41612-023-00409-z, 2023a.

Potter, E., Fyffe, C., Orr, A., Quincey, D., Ross, A., Rangecroft, S., Medina, K., Burns, H., Llacza, A., Jacome, G., Hellstrom, R., Castro, J., Cochachin, A., Montoya, N., Loarte, E., and Pellicciotti, F.: Bias-corrected temperature and precipitation data from the WRF regional climate model output, Cordillera Blanca and Vilcanota-Urubamba regions, Peru, from 1980 to 2018 (Version 1.0), NERC EDS UK Polar Data Centre [data set], https://doi.org/10.5285/2cf25580-9b79-440f-8505-6230dd377877, 2023b.

Potter, E., Fyffe, C., Orr, A., Quincey, D., Ross, A., Rangecroft, S., Medina, K., Burns, H., Llacza, A., Jacome, G., Hellstrom, R., Castro, J., Cochachin, A., Montoya, N., Loarte, E., and Pellicciotti, F.: Precipitation and temperature data from statistically downscaled CMIP5 models, Cordillera Blanca and Vilcanota-Urubamba regions, Peru, from 2019 to 2100 (Version 1.0), NERC EDS UK Polar Data Centre [data set], https://doi.org/10.5285/67ceb7c8-218c-46e1-9927-cfef2dd95526, 2023c.

Potter, E., Fyffe, C., Orr, A., Quincey, D., Ross, A., Rangecroft, S., Medina, K., Burns, H., Llacza, A., Jacome, G., Hellstrom, R., Castro, J., Cochachin, A., Montoya, N., Loarte, E., and Pel-

licciotti, F.: Precipitation and temperature climate change indices calculated from WRF data and statistically downscaled CMIP5 models, Cordillera Blanca and Vilcanota-Urubamba regions, Peru, from 1980 to 2100 (Version 1.0), NERC EDS UK Polar Data Centre [data set], https://doi.org/10.5285/b56d30e8-edaa-4225-96d7-fcc689e930c7, 2023d.

Quiroz, R., Yarlequé, C., Posadas, A., Mares, V., and Immerzeel, W. W.: Improving daily rainfall estimation from NDVI using a wavelet transform, Environ. Modell. Softw., 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, J. Hydrol., 625, 129949, https://doi.org/10.1016/j.jhydrol.2023.129949, 2023.

Rouse Jr., J. W., Haas, R. H., Schell, J., and Deering, D.: Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation, Tech. rep., NASA-CR-132982, 1974.

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, Int. J. Climatol., 43, 3005–3022, https://doi.org/10.1002/joc.8013, 2023.

Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI): Plataforma de Datos de Estaciones Meteorológicas, https://www.senamhi.gob.pe/?p=estaciones, last access: 25 June 2025.

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, Int. J. Climatol., 39, 989–1012, https://doi.org/10.1002/joc.5856, 2018.

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

Storn, R. and Price, K.: Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces, J. Global Optim., 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, uSGS EROS Archive [data set], https://www.usgs.gov/centers/eros (last access: 15 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 Planet. Change, 62, 14–28, https://doi.org/10.1016/j.gloplacha.2007.11.003, 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, CARE France and UNU-EHS, Bonn, ISBN 978-3-939923-88-6, eISBN 978-3-939923-89-3, 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, J. Geophys. Res.-Atmos., 121, 12198–12216, 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., del Mar Chaves Montero, M., Rembold, F., Troccoli, A., Roy, I., and Hoteit, I.: Seasonal forecasts of the rainy season onset over Africa: Preliminary results from the FOCUS-Africa project, Climate Services, 32, 100417, https://doi.org/10.1016/j.cliser.2023.100417, 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.