A new cost-effective approach to increase accuracy and lead time of flash flood forecast: Ecuador’s proof of concept

Fatima M. Pillosu1,2, Agathe Bucherie3, Carolynne Hultquist6, Andrew Kruczkiewicz3,4,5, Calum Baugh2, Humberto Vergara7,8, Florian Pappenberger2, Elisabeth Stephens1,4,9, Christel Prudhomme2,10,11, Hannah L. Cloke1,9,12,13

1 Department of Geography and Environmental Science, University of Reading, Reading, UK

2 Forecast Department, European Centre for Medium-range Weather Forecasts, Reading, UK

3 International Research Institute for Climate and Society (IRI), Columbia Climate School, New York, USA

4 Red Cross Red Crescent Climate Centre, The Hague, The Netherlands

5 Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede, The Netherlands

6 Center for International Earth Science Information Network (CIESIN), Columbia Climate School, New York, USA

7 Cooperative Institute for Mesoscale Meteorological Studies (CIMMS), The University of Oklahoma, Norman, OK

8 NOAA National Severe Storms Laboratory (NSSL), Norman, OK

9 Department of Meteorology, University of Reading, Reading, UK

10 Department of Geography and Environment, University of Loughborough, Loughborough, UK

11 UK Centre for Ecology and Hydrology, Wallingford, United Kingdom

12 Department of Earth Sciences, Air, Water and Landscape Science, Uppsala University, Sweden

13 Centre of Natural Hazards and Disaster Science, CNDS, Sweden

**Correspondence:** Fatima M. Pillosu(fatima.pillosu@ecmwf.int)

**Abstract.** Globally, flash floods are one of the costliest natural hazards for property damage and loss of life. The low accuracy of flash flood forecasts beyond a few hours limits their use in early warning systems. This study evaluates the performance in predicting flash flood risk based on global post-processed “ecPoint-Rainfall” forecasts compared with the ECMWF ensemble forecasts (ENS), up to 10 days lead time. A one-year verification for 2020 is carried out in Ecuador, whose varied climate and presence of a comprehensive flood report database made it an attractive site for verification. This study identifies the scenarios in which ecPoint provides better flash flood guidance than other forecasting systems, enabling more informed decision-making around flash flood events. Data suggests that ecPoint outperforms ENS in areas where rainfall originates from small-scale convective systems. Where rainfall originates from large-scale convective systems, ecPoint and ENS performances are comparable, except for correcting errors in the rainfall’s diurnal cycle. This study also contributes a new methodology to define the typical rainfall event that can cause flash floods when no point rainfall observations (e.g., rain gauges) are available. The method uses short-term ecPoint-Rainfall forecasts as a proxy for point rainfall observations, allowing us to verify flash flood forecasts and develop warning systems in regions with poor rainfall observation coverage.

**Keywords.** Ensemble rainfall prediction, ecPoint,flash flood forecasting, flash flood observations, Ecuador, flood risk, early warning systems, anticipatory action, forecast-based financing.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Introduction

Flash floods have significant societal, economic, and environmental consequences worldwide (Jonkman and Vrijling 2008; Dordevic et al. 2020). In Latin America, rapid and unregulated urbanisation of floodplains, human-induced degradation of catchments, absence of preparedness plans, persistent poverty levels, ineffective public policies, and inadequate infrastructure have amplified the impacts of flash floods (Pinos and Quesada-Román 2022). In Ecuador, flash floods have the highest mortality rate compared to other types of floods and can cause short- and long-term impacts, including infrastructure damage, agricultural losses, disruptions of businesses and education, disruptions of health services, and outbreaks of waterborne diseases (Galarza-Villamar et al. 2018). Kruczkiewicz et al. (2021a) estimate that approximately 60% of all floods are flash floods, and their frequency is expected to increase due to climate change (Hirabayashi et al. 2021).

Forecast-triggered mitigation strategies for flood risk reduction, such as early warning systems (Šakić Trogrlić et al. 2022; Coughlan de Perez et al. 2022) and forecast-based financing protocols (De Perez et al. 2016; Bischofites et al. 2019), have been shown to improve resilience, decrease mortality rates, and lower recovery costs. To allow prompt preparedness and action, accurate predictions of areas at risk of flash floods, with sufficient lead time, are crucial. In lower-income countries, accurate forecasts with even longer lead times might be required to set cost-effective mitigation strategies, such as the “ready-set-go” approach, in which various inexpensive actions are implemented at long lead times, and more specific or costly actions are later activated based on more accurate short-range forecasts (Bazo et al. 2019; Kiptum et al. 2023).

Flash floods are among the most challenging types of flood to predict due to the high uncertainty in their overall forecasting process (Zanchetta and Coulibaly 2020; Speight et al. 2021). Flash flood forecasting systems, either local/regional (Speight et al. 2018; Corral et al. 2019; Ibarreche et al. 2020; Ramos Filho et al. 2021; Shuvo et al. 2021), national (Javelle et al. 2016; Liu et al. 2018; Georgakakos et al. 2021) and continental scale (Raynaud et al. 2015; Gourley et al. 2017), share weaknesses that limit the predictability of flash floods: chronic lack of historical data on flash flood occurrence and impact (Lowrie et al. 2022), inaccurate predictions of extreme localised rainfall (Zeman et al. 2021), and challenging representation of detailed hydrological processes dependent on topography, soil conditions, and terrain coverage that modulate flash floods’ occurrence and severity (Xing et al. 2019).

Rainfall and soil moisture are critical indicators of flash flood occurrence (Georgakakos 2006). Their use in low-complexity index-based systems has proven effective in predicting areas at risk of flash floods from local to continental scale (Hurford et al. 2012; Raynaud et al. 2015; Ma et al. 2021; Zanchetta et al. 2022; Schroeder et al. 2016; Luong et al. 2021). Rainfall-wise, developers prefer the use of radar-derived rainfall totals (Javelle et al. 2010) or km-scale rainfall forecasts (Davolio et al. 2017; Song et al. 2019) to predict the extreme (localised) rainfall events that typically trigger flash floods. These two options are not viable for developing a system to predict areas at risk of flash floods in a continuous global domain with longer lead times. Radar-derived rainfall provides predictions of mere hours ahead (Imhoff et al. 2022). Km-scale NWP models can overestimate and misplace extreme rainfall even if they represent the distribution of rainfall totals better than their lower resolution counterparts (Cafaro et al. 2021; Thomassen et al. 2021; Nielsen and Schumacher 2016), and their skill decreases significantly after day two forecasts (Barrett et al. 2019; Schwartz 2019). Furthermore, their spatial coverage s is patchy.

Rainfall and soil moisture are standard outputs from global numerical weather prediction (NWP) models such as the Integrated Forecasting System (IFS) from the European Centre for Medium-range Weather Forecasts (ECMWF). However, the coarse spatial resolution and limited ability to capture localised rainfall events have hindered the use of global NWP models in flash flood prediction. Verification has shown that global NWP models have consistently extended the accuracy of rainfall forecasts up to medium-range lead times (Lavers et al. 2021; Haiden et al. 2023). Furthermore, statistical post-processing can downscale rainfall forecasts to resolutions more amenable for flash flood prediction while global NWP models move towards km-scales (Vannitsem et al. 2021). For example, ecPoint is a statistical post-processing technique that transforms global, grid-based forecasts into probabilistic point-scale predictions (Hewson and Pillosu 2021). It has improved the reliability and discrimination ability of the ECMWF ensemble (ENS) rainfall forecasts up to day 10, especially for extremes. Therefore, there has been a growing interest in using global NWP models to extend the lead time of flash flood predictions and improve the accuracy of forecasts (Bucherie et al. 2022b).

This study aims to establish whether rainfall forecasts from global NWP models can provide guidance in predicting areas at risk of flash floods. To the author's knowledge, no study in the literature has assessed this objectively. This analysis is imperative because such forecasts could be adopted in developing a low-complexity, rainfall-based index forecasting system and represent a cost-effective way to develop a flash flood forecasting system over a continuous global domain with reasonable skill up to medium-range lead times. This study analyses the accuracy of global ENS and ecPoint rainfall forecasts to identify areas at risk of flash floods. When assessing the performance of rainfall forecasts for flash flood prediction, objective verification is typically run against rainfall observations, and the suitability of the forecasts for flash flood prediction is then assumed (Gascón et al. 2023). Although reasonable, this assumption ignores that the relationship between triggering rainfall events and flash floods is not entirely linear. This study will instead perform a verification analysis against flash flood reports. The recent development of a well-documented flash flood database in Ecuador (Kruczkiewicz et al. 2021b) and the country’s high susceptibility to flash flooding make it an exceptional test bed for this study.

Two user categories have been defined in this study: the first one is interested only in whether or not there will be a flash flood event (the answer will be binary, i.e., yes or no), while the second one is also interested in verifying the amount of the rainfall event that generated such flash flood.

This study is organised into eight sections. Section 2 presents background information on Ecuador's geography and rainfall/flood climatology. Sections 3 and 4 present the data and methods used in the verification analysis. Sections 6 and 7, respectively, present and discuss the verification results. The conclusions of this study are finally drawn in Section 8.

# Background: geography, rainfall climatology and flooding in Ecuador

Located in north-western South America, Ecuador includes continental Ecuador and the Galápagos Islands in the Pacific Ocean, 1000 km from the mainland (insert in **Figure 1a**). The Andes run north to south through Ecuador (**Figure 1a**) and split it into three main regions (Vuille et al. 2000): “La Costa”, which comprises the Andes’ western slopes and the coastal plains along the Pacific Ocean; “El Oriente”, which covers a plateau containing 2% of the Amazon basin and the eastern slopes; and “La Sierra”, which contains the inter-Andean region between the western and eastern slopes of the Andes. This study considers only the continental landmass, hereafter referred to as “Ecuador”. The domain of interest, including the borders with Colombia, Peru and the Pacific Ocean, contains a total of 1090 grid boxes. This general domain is then split in the three main regions using Ecuador’s topography as represented in ENS (see **Figure 1**b). Grid boxes below 600 m above sea level and to the west and the east of the longitude 78.2 °W belong, respectively, to “La Costa” (i.e., 321 grid boxes) and “El Oriente” (i.e., 299 grid boxes). Grid boxes above 600 m were assigned to “La Sierra” region (i.e., 470 grid boxes).

The rainy season in “La Costa” spans from December to May. Broad-scale atmospheric and oceanic phenomena modulate rainfall events' intensity and spatial variability. The extreme phases of El Niño Southern Oscillation, known as El Niño (i.e., above-average sea surface temperature in the Pacific Ocean) and La Niña (i.e., below-average), enhance and decrease, respectively, the average rainfall during the rainy season (Recalde-Coronel et al. 2014; Tobar and Wyseure 2018). In addition, certain phases of the Madden-Julian Oscillation, 1 and 8 (i.e., when a convection centre is over the Western Hemisphere and Africa) and 4 and 5 (i.e., when a convection centre is over the Maritime continent), are associated, respectively, with an enhancement and decrease of precipitation (Recalde-Coronel et al. 2020). “La Sierra” has two main rainy seasons (i.e., February-May and October-November). Precipitation spatial patterns in the inter-Andean valleys are more complex than in “La Costa” because rainfall is typically generated by smaller-scale convective systems (Vuille et al. 2000). Additionally, as air masses lose much of their humidity on both flanks of the Andes, precipitation amounts in “la Sierra” are relatively lower than in the other two regions, varying between 800 and 1500 mm/year (Vuille et al. 2000). It rains throughout the year in “El Oriente”, with the wettest (driest) months being April-July (September-October). Rainfall climatology in “El Oriente” is primarily influenced by the strong convective activity across the Amazon Forest and the water vapour variations from the sea surface temperature of the tropical Atlantic Ocean (Vuille et al. 2000).

Floods can cause considerable material loss and deaths in “La Costa” and “La Sierra” because they are heavily populated and contain Ecuador’s two most important industrial areas, Guayaquil in Guayas and Quito in Pichincha (**Figure 1c**). Prolonged rainfall events in “La Costa” can generate extensive, severe surface runoff far from rivers and make rivers flood vast plain areas (Galarza-Villamar et al. 2018). Intense, shorter-lived rainfall events (i.e., less than one day) can also cause severe, sudden surface runoff (Galarza-Villamar et al. 2018). The rivers in “La Sierra” are susceptible to extreme localised rainfall events and, consequently, are prone to flash flooding (Laraque et al. 2009; Pinos and Timbe 2020). The river flows in “El Oriente” show a much stronger response to seasonal rainfall than a single rainfall event due to the size and length of Amazonian rivers, floodplain storage, and rivers’ shallow beds (Trigg et al. 2009).

# Data

## Rainfall forecasts: ECMWF ENS and ecPoint

ECMWF ENS consists of one control run started from the best possible representation of unperturbed initial conditions, and fifty perturbed members started from perturbed initial conditions (using singular vectors and a data assimilation ensemble) and stochastic model uncertainties (Buizza 2019). Up to day 15, ENS forecasts are saved in the native octahedral reduced-Gaussian with a resolution of ~18 km at the equator (Owens & Hewson, 2018). Over the period used to compute the climatology of rainfall events associated with flash flood events (1st January to 31st December 2019) and the verification period (1st January to 31st December 2020), three different model versions run operationally at ECMWF: 45r1[[1]](#footnote-2) (for forecasts from 1st January to 10th June 2019), 46r1[[2]](#footnote-3) (from 11th June 2019 to 12th July 2020) and 47r1[[3]](#footnote-4) (from 13th July to 31st December 2020). The mismatch of model versions over the periods considered in this study is unlikely to adversely affect the verification results because no significant changes were made in the physics of the rain generation mechanisms.

ecPoint is a decision-tree-based statistical post-processing technique that transforms global, grid-based forecasts into probabilistic point-scale forecasts (Hewson and Pillosu 2021). It helps to address the two main factors affecting the performance of global NWP model output against point verification: systematic biases (Lavers *et al.*, 2021) and lack of information on forecast sub-grid variability (Göber *et al.*, 2008). A one-year point verification shows that ecPoint-Rainfall (i.e., the family of ecPoint-related products that post-processes rainfall forecasts) provides forecasts with better reliability and discrimination ability for localised rainfall than ENS, especially for extremes, up to medium-range lead times (i.e., day ten forecast). ecPoint-Rainfall generates a distribution of 100 point-rainfall values for each raw ENS member, producing 5100 post-processed members per grid box. The operational product distils them in percentiles from 1st to 99th, providing them in the same native grid of ENS forecasts. Forecasts are provided in four overlapping 12-hourly accumulation periods ending at 0, 6, 12 and 18 UTC, up to day 10.

## Flash flood reports

The disaggregation by flood types and specific documentation about historical flash flood events and their impacts is rare in many regions of the world due to primarily a lack of commonly accepted flash flood definition (Kruczkiewicz et al. 2021a; Bucherie et al. 2022a). Kruczkiewicz *et al.* (2021b) developed a [method to assign an “Enhanced Flash Flood Confidence Index](https://www.mdpi.com/2072-4292/13/14/2764) (EFFCI)” for flood events in historical flood datasets based on text mining of disaster reports and a flash flood susceptibility index extracted from the geophysical properties of the location of the events. The EFFCI is an estimate of the likelihood of a flood event to be a flash flood, ranging from 1 (not very likely) to 10 (extremely likely). The flash flood database in Ecuador was mainly compiled from two datasets, DesInventar (UNDRR 2021) and the Ecuadorian Secretariat for Disaster Management (SNGRE), and it counts 4967 flood events from 2007 to 2020. In addition to the EFFCI index, most entries in the flash flood database contain information about the location (with latitude and longitude coordinates) and the day and time (in local time) of the flood occurrence. As a result of this method applied to Ecuador, a historical dataset of flood occurrences and impacts is available, with specific information on the likelihood of events being flash floods (Bucherie et al. 2021). Although this data set is the best attempt to address historical flash floods in Ecuador, it is essential to note that it is based on disaster reporting processes carried out on the ground, not systematically collected over time. As a result, it can present gaps, inconsistent descriptions of flood processes over time, and uncertainty in geolocation.

This study considered flood reports from 2019 to define the climatology of rainfall events associated with flash flood events. Events from 2020 were used to run the objective verification analysis. For flood reports in both years, three EFFCI thresholds were considered to evaluate the impact of uncertainty around a flood report being a flash flood event. EFFCI>=1 (i.e., all flood reports), EFFCI>=6 (i.e., flood reports that are likely to be flash floods), and EFFCI>=10 (i.e., flood reports that are highly likely to be flash floods). **Table 1** shows the total number of flood reports in 2019 and 2020 and the number of reports excluded because they did not include any reporting location (in lat/lon coordinates) and reporting time (with date and time). **Table 1** also shows the number of flood reports per region and the EFFCI threshold, while **Figure 2** shows their spatial distribution. To give the reader a feeling of how many flood reports are used per accumulation period in the objective verification analysis, **Figure 3a** shows the timeseries of the counts of flood reports with EFFCI>=6 for 2020, accumulated over the four overlapping 12-hourly accumulation periods at which the forecasts are provided. On average, around 90% of the days in 2020 had none reports of flood events. Most of the remaining 10% of the days (~ 30 days) contained only one flood report, with only one day (2020/20/28) with more than 5 reports within the whole domain (**Figure 3b**), corresponding to a domain coverage of only 0.6%. This correspond to a observational coverage that is several orders of magnitude lower to the one available for rainfall verification since there might be typically hundreds to thousands rainfall observations available at each accumulation period.

# Methods for the verification analysis

The two main attributes of any probabilistic forecasting system are reliability and discrimination ability, and together they determine the performance of the system (Jolliffe and Stephenson 2011). Reliability and discrimination ability are defined against a specific event being exceeded (e.g., 50 mm/12h), hereafter referred to as Verifying Rainfall Event (VRE). Reliability measures if an event (i.e., the VRE) is predicted with a probability that equals the average frequency at such an event is observed. Discrimination measures the ability of the forecasting system to distinguish situations leading to events (i.e., events exceeding the VRE) from those leading to non-events (i.e., events not exceeding the VRE).

Two main challenges were faced in this study. The magnitude (in mm) of rainfall events associated with flash floods in Ecuador defines the VREs to use in the objective verification analysis. Section 4.1 describes the methodology adopted in this study to define the VREs since they were not known to the authors. Since we are converting probabilistic forecast into binary, intrinsic reliability for such events in the usual probabilistic sense cannot be computed. However, the frequency bias can be calculated to determine the overall calibration of the system, indicating whether the system is on average under- or over forecasting the event. Section 4.2 describes the methodology used to estimate the frequency bias and the discrimination ability of ENS and ecPoint rainfall forecasts in identifying areas at risk of flash floods.

## Definition of the verifying rainfall events (VREs)

If not known a priori, VRE magnitudes can be defined in several ways (see **Figure 3**, green area “Definition of VREs”). If point rainfall observations are available (e.g., from rain gauges or radars), one can create the distribution of observed flash-flood-triggering rainfall totals. VREs would then correspond to specific percentiles of that distribution. The higher the percentile, the higher the magnitude of the VRE and the higher the severity level of flash flood events considered in the objective verification analysis. This approach requires high-density rainfall observations, in both space and time, to capture the extreme (localised) rainfall totals that triggered the flash floods (Haiden and Duffy 2016). In the absence of a suitable observational network (for example, in Ecuador, no high-density, in situ 12-hourly rainfall observations are available), the VREs could be defined only from gridded rainfall products such as reanalysis like ERA5 (Hersbach et al. 2020), reforecasts (Hamill et al. 2006), or blended gridded rainfall observations such as MSWEP (Beck et al. 2019) or GPCP (Adler et al. 2018). However, these datasets tend to underestimate rainfall extremes due to their coarse resolution (Tapiador et al. 2019).

This study proposes a methodology to create a synthetic distribution of flash-flood-triggering rainfall events using short-range ecPoint-Rainfall forecasts (**Figure 3**, green square). At short-range leads, the ecPoint-Rainfall realisations can be considered proxies for point rainfall observations within grid-boxes (Hewson and Pillosu 2021). Each flood report is associated with ecPoint-Rainfall forecasts from the nearest grid-box (**Figure 4a**). At each forecast run, two 12-hourly accumulation periods span each flood report’s reporting time (see the first and second column of **Table 4**), so a distribution of 396 ecPoint-Rainfall realisations (that is, 99 ecPoint-Rainfall values X 2 accumulation periods X 2 runs) can be built for each flood report (blue lines in **Figure 4b**). Due to this high number of forecast realizations per flood report, only one year is deemed necessary to define VREs. The rainfall value (tp) associated with an Xth percentile is extracted from each distribution (red dots in **Figure 4b**) to characterise different levels of severity of flash-flood-triggering rainfall events. The study tested high percentiles (e.g., 50th, 75th, 85th, 90th, 95th, 98th, and 99th) to exclude low rainfall totals from the analysis that were unlikely to be the drivers of any flash flood event. The distribution of flash-flood-triggering rainfall events is created (red line in **Figure 4c**). The VRE is then defined using the top 75% of the flooding events considered, which means selecting the rainfall values corresponding to the 25th percentile of the distribution (purple dot in **Figure 4c**). Separate VREs are calculated for “La Costa” and “La Sierra” to capture their different hydro-climatological regimes.

## Objective verification to assess the rainfall forecasts’ performance in the identification of areas at flash flood risk

### Assessment of forecasts’ calibration: Frequency Bias (FB)

The frequency bias is used here to evaluate the calibration of ENS and ecPoint rainfall forecasts in the prediction of areas at risk of flash floods. For each lead time, the frequency bias was determined by dividing the total number of ensemble members exceeding the considered VRE by the product of the number of ensemble members and the total number of instances a flash flood was observed. The frequency bias, for a specific lead time, is computed using the following formula:

|  |  |
| --- | --- |
|  |  |
| FB = | (1) |
|  |  |

A bias score of 1 indicates perfect calibration, with scores greater and smaller than 1 indicating that the forecasting system tends to over- and under-predict, respectively.

### Assessment of forecasts’ discrimination ability: ROC curves and Area Under the ROC curve (AROC)

This study uses the Relative Operating Characteristic (ROC) curve and the area under the ROC curve (AROC) to estimate and compare the discrimination ability of ENS and ecPoint rainfall forecasts in the prediction of areas at risk of flash floods (Jolliffe and Stephenson 2011). ROC curves are built using a 2x2 contingency table that quantifies the hits (H), misses (M), false alarms (FA) and correct negatives (CN) that realise when action is advised based on the VRE exceeding each sampled probability threshold (see **Table 3** for the definition of the constituting elements of the contingency table). Correspondent hit rates (HR) and false alarm rates (FAR) can be then computed, respectively, from equations (2) and (3):

|  |  |
| --- | --- |
| HR = H / (H+M) [values between 0 and 1] | (2) |
|  |  |
| FAR = FA / (FA+CN) [values between 0 and 1] | (3) |
|  |  |

For each sampled probability threshold, ROC curves map HRs on the Y-axis against FARs on the X-axis. The location of the ROC curve in the graph and the geometrical area under the ROC curve (AROC) determine the forecasting system’s discrimination ability. Perfect discrimination is obtained when only HRs grow while FARs always remain equal to zero. This is represented by a ROC curve that rises from the bottom left corner (0,0) along the Y-axis to the top left corner (0,1) and goes straight to the top right corner (1,1). AROC is in this case equal to 1. When the forecasting system has no discriminatory ability (i.e., it does not provide any additional information beyond climatological predictions), HRs and FARs grow at the same rate. Therefore, the ROC curve lies along the diagonal of the graph, and AROC equals to 0.5.

How ROC curves are built and AROCs are computed have a significant impact in the results’ interpretation. To ensure that ROC curves are as complete as possible, probability thresholds are determined using the full discretisation available in the ensemble rather than using fixed percentage bins (Ben Bouallègue and Richardson 2022). Hence, ROC curves for ENS and ecPoint are built, respectively, with 51 and 99 points. No curve fitting is used to build or complete the ROC curves Straight lines are drawn instead between consecutive points in the graph, as well as the last meaningful point of the ROC curve and the top-right corner (Ben Bouallègue and Richardson 2022). Moreover, AROCs are computed using the trapezoidal approximation, which simply sums the areas of single trapeziums formed by the straight lines between ROC’s consecutive points. As a result, ROC curves for high VREs might cluster on the bottom left corner of the graph, and if built with fewer points, they might look incomplete and AROCs might result smaller. However, this approach focuses the analysis on the “real” and not on the “potential” discrimination ability of the rainfall forecasts. To evaluate whether the differences between the AROC for ENS and ecPoint are significant, the percentile bootstrapping technique is applied (DiCiccio and Efron 1996). Sampling with replacement with 10,000 replicates and confidence intervals of 95% are considered.

Populating the contingency tables (**Table 2**) is the challenge of this verification analysis. Stationary observations (i.e., provided by instruments installed at a specific location, e.g., rain gauges) provide timeseries that record both yes- and non-events at the location where the instrument was installed. Thus, all four elements of the contingency table can be quantified. Non-stationary observations record only yes-events, at the location where the events happened. As a result, it is impossible to answer the question *“if there are no reports in an area, is it because an event happened, but nobody reported it, or because there was no event to report?”*. Some studies verify only yes-events with the caveat that only quadrants I (i.e., hits) and III (i.e., misses) of the contingency table can be populated (Robbins and Titley 2018). This study follows instead the method in Tsonevsky et al. (2018) which allows to fully populate the contingency table. This method assumes that a non-report corresponds to a non-event. Due to the care used to create the observational flood database, this assumption is considered valid also for this study. Observational fields are built assigning 1 to grid-boxes containing at least one flood report (i.e., observational yes-event); otherwise, the value 0 is assigned (i.e., observational non-event). Forecast fields are built by assigning the value 1 to those grid-boxes where the considered VRE is exceeded with a considered probability threshold (i.e., forecast yes-event); otherwise, the grid-boxes are assigned the value 0 (i.e., forecast non-event). The 2X2 contingency tables are built by examining overlapping grid boxes in correspondent observational and forecast fields: when both grid boxes are assigned the value 1 or 0, they count respectively as H and CN. When a grid box in the observational field is assigned the value 1, and the correspondent grid box in the forecast field is assigned the value 0, it counts as a M. It counts as a false alarm (FA) if it happens vice versa.

# Results

## Verifying rainfall events

**Figure 5** shows that, in both regions, the spread in the flash-flood-triggering rainfall totals increases with the EFFCI values. However, caution is required for flash floods with EFFCI>=10 (**Figure 5c** and **f**) due to the small sample size. Out of all the examined percentiles (i.e., 50th, 75th, 85th, 90th, 95th, 98th, and 99th), the verification analysis was conducted only for VREs that correspond to the 85th (VRE85) and 99th percentile (VRE99) as they were considered, respectively, representative of rainfall events associated with medium/severe and extremely severe rainfall events. Flood reports with EFFCI>=6 were considered in this study to maintain high the number of flood reports that are likely to be flash floods and, therefore, provide robust results in the objective verification analysis. The considered values in mm/12h for VRE85 and VRE99 are, respectively, indicated with a purple and orange dot in **Figure 5b** for “La Costa” and **Figure 5e** for “La Sierra”, and the rounded values are indicated in **Table 4**.

## Rainfall forecasts’ performance in the identification of areas at flash flood risk

### Frequency bias

Not added yet.

### Discrimination ability

AROCs denote the discrimination ability of a forecasting system for specific VREs and forecast lead times. **Figure 6** shows the evolution of AROC with lead time for “La Costa” and “La Sierra”, computed for flood reports with EFFCI>=6, and for VRE85 and VRE99 events. For both ENS and ecPoint, no degradation with lead time in the AROCs is observed in “La Costa” (**Figure 6a** and **b**). On the contrary, the AROCs in “La Sierra” tend to diminish with lead time (**Figure 6c** and **d**), at a similar rate for both ENS and ecPoint and regardless of VRE. Discrimination ability in ENS and ecPoint is similar in both regions for medium rainfall events as shown by overlapping AROC curves (**Figure 6a** and **c**), and it is generally higher for ecPoint when compared with ENS for extreme rainfall events (**Figure 6b** and **d**). In addition, the overall discrimination ability of both ENS and ecPoint forecasts is larger for medium rainfall events (**Figure 6a** and **c**) than for extreme ones (**Figure 6b** and **d**). A feature that stands out in all panels in **Figure 6** is the sinusoidal pattern shown by the AROC in correspondence of different accumulation periods throughout a day. **Figure 7** displays the rainfall’s annual mean in “La Costa” (continuous lines) and “La Sierra” (dashed lines) during different accumulation periods in a day, for both ENS (red lines) and ecPoint (blue lines), showing a strong rainfall diurnal cycle in the forecasts in Ecuador, and confirmed by Kikuchi and Wang’s study (2008) based on rainfall observations. In “La Costa” (**Figure 6a** and **b**), AROCs’ peaks are observed between 0000-1200 LT (i.e., lead time steps labelled in purple) while troughs are mostly observed between 1200-0000 LT (i.e., lead time steps labelled in fuchsia). These two accumulation periods correspond, respectively, to the ones with the second smallest and second biggest rainfall totals in the day (**Figure 7**, continuous red and blue lines). For both ENS and ecPoint, the amplitude between peaks and troughs increases with increasing VREs, although the amplitude is deeper for ENS as ecPoint improves (i.e., increase) the AROC values for the troughs for VREs = 99th percentile. This shows that, in “La Costa”, ecPoint adds value in the identification of areas at flash flood risk in those part of the day when higher rainfall totals are expected (i.e., during evening and night-time). A similar sinusoidal pattern is observed for the AROCs in “La Sierra” (**Figure 6c** and **d**), although, compared to “La Costa”, a 6- to 12-hour shift is observed between peaks and troughs. Unlike in “La Costa”, where ecPoint’s added value is observed at specific times of the day, ecPoint’s added value in “La Sierra” is observed in all accumulation periods. Similar sinusoidal patterns to those shown in **Figure 6** were observed also for flood reports with EFFCI>=1 and 10, and for VREs >= 50th, 75th, 90th, 95th and 98th percentiles (not shown). The only difference lies in noisier behaviours (i.e., deeper amplitudes between peaks and troughs) for increasing EFFCI thresholds and VREs. This is most likely due to the correspondent decrease in the number of flood reports associated with larger EFFCI thresholds as shown in **Table 1**. For increasing VREs, a decreasing number of events exceeding the VRE is most likely the reason why the trend with lead time of AROC appears noisier.

# Case Study: intense rainfall and flash floods on 8th March 2021

The following case study provide a practical example on how ENS and ecPoint rainfall forecasts can be used to predict areas at risk of flash floods. This case study assesses how the rainfall forecasts compare with rainfall observations. This is done because, in the authors’ experience receiving feedback from ecPoint users about the rainfall products, the co-verification of flash floods and rainfall events helps to increase users’ confidence in using ecPoint-Rainfall forecasts to forecast areas at flash flood risk.

March was one of the wettest months in 2021 in Ecuador. As a result of numerous heavy rainfall events, rivers such as Guayas, Los Ríos, Esmeraldas, and Manabí burst their banks, with landslides observed in many different regions[[4]](#footnote-5). The 8th of March was one of the wettest days (**Figure 9a**). Significant impacts resulted mainly in the highly populated city of Guayaquil, where very heavy rainfall was reported to occur in the afternoon after 4 pm (Local Time, LT)[[5]](#footnote-6), with rainfall totals exceeding 100 mm/24h in the city centre (zoomed red area in **Figure 9a**)[[6]](#footnote-7). Around 8th March, the MJO was reported by various centres to be in phase 8, which tends to be conducive to, or at least correlated with, onshore lower tropospheric westerly wind anomalies near the equatorial west-facing coasts of South America (Wheeler and Hendon 2004). In conjunction, analysts from NOAA highlighted the likelihood of enhanced convective activity in the region in routine bulletins[[7]](#footnote-8). From the dawn of 8th March, ECMWF’s numerical model sounding (**Figure 9b**) looked particularly conducive to flash-flood-triggering rainfall activity. For instance, the very high CAPE (Convective Available Potential Energy) shows that there is the potential for sufficiently high dewpoint depression insolation-based triggering that might not be impeded by thick cloud. It also shows the potential for very high-altitude convective cloud tops, very strong wind shear that favours prolonged convective cell life cycles (as down-draughts would not interfere with up-draughts), and relatively light steering winds (favouring the slow movement of the convective cells). This description is supported by SYNOP and METAR observations and satellite imagery (not shown), suggesting that the cause of this rainfall event was organised convective cells whose development was triggered by insolation.

**Figure 9c** shows ENS and ecPoint forecasts from the 00 UTC run for day 1 (first row), day 3 (second row), and day 7 (third row) lead times. The forecasts are valid for the 12-hourly accumulation period between 8th March 2021 at 12 am and 9th March 2021 at 0 am (LT), i.e., the fraction within the 24-hourly period of the observations reported in **Figure 9a** when most of the rainfall fell. The forecasts for the 50th (first and second column), 95th (third and fourth column), and 99th percentiles (fifth and sixth column) are shown. The median (i.e., the 50th percentile) should represent the dividing line for equi-probable observation categories. By comparing the rainfall observations (**Figure 9a**) and the forecast for the 50th percentile (first and second column in **Figure 9c**), one can see that, overall, ENS overestimates the mean rainfall. On the contrary, due primarily to its bias correction for rainfall overprediction at grid-scale, ecPoint’s 50th percentile is systematically smaller than in ENS, showing a better fit with the observations. At the same time, ecPoint’s 95th (third and fourth column in **Figure 9c**) and 99th percentiles (fifth and sixth column in **Figure 9c**) highlight a higher potential than ENS of having higher local rainfall totals in certain areas (e.g. Guayaquil). While far more observations would be needed to analyse the performance of ENS and ecPoint forecasts for such high percentile robustly, it appears that both ENS and ecPoint predicted well the local rainfall extremes in “La Costa”. For example, there is a signal of extreme rainfall in ENS for Guayaquil already in the 95th percentile). In contrast, it appears that ecPoint adds the most value in the prediction of extreme rainfall in “La Sierra”. For example, ecPoint’s 99th percentile shows a 1% chance of having up to 60 mm/12h at some location n “La Sierra”, and one location observed such amount.

# Discussion

## On ENS and ecPoint performance in predicting areas at flash flood risk

The stable/degraded performance of AROC with lead time of both ENS and ecPoint in “La Costa” and ”La Sierra” might be linked with how well rainfall-generation mechanisms are predicted in the two regions in the raw ENS forecasts. Rainfall in “La Costa” is generally driven by large-scale atmospheric and oceanic convective phenomena such as “El Niño” (Tobar and Wyseure 2018) and the “MJO” (Recalde-Coronel et al. 2020), which tend to be well predicted up to 6 weeks in advance in the ENS (Haiden et al. 2021). ecPoint corrections are observed only at specific accumulation periods, which it might be due to a bias correction operated by the ecPoint post-processing for night-time rainfall. In contrast, rainfall in “La Sierra” is usually linked to smaller-scale convective systems (Recalde-Coronel et al. 2014), whose predictability can diminish relatively quickly after only few days in ENS (Haiden et al. 2021). ecPoint can effectively address the limitations of the prediction of smaller-scale convective systems in the ENS, especially for very extreme rainfall events, improving the discrimination ability of the rainfall forecasts at all accumulation periods.

## On the importance to keep improving the collection of flash flood reports to better develop and verify flash flood forecasting systems

Due to its good spatio/temporal coverage and the reliability of its reports’ attributes (e.g., events’ location and reporting time) thanks to a close collaboration with local authorities, the flood database developed in Ecuador by Kruczkiewicz et al. (2021b) allowed the authors of this study to conduct an in-depth, long-term verification of the performance of ENS and ecPoint rainfall forecasts at predicting areas at flash flood risk. The verification analysis used ROC curves that were built for flood reports in a one-year period, allowing the authors to show the performance of the rainfall forecasts over a large territory with different climatological regions (“La Costa” versus “La Sierra”), and over different times of the day (morning versus afternoon). To the best authors’ knowledge, this is the first time that such in-depth, long-term verification analysis is possible for flash flood events, whilst before researchers and developers had to compromise adopting sub-optimal approaches. The literature shows that flash flood verification is performed primarily for case studies (c.f. Raynaud et al. (2015)) since detailed information is mostly available or more easily accessible for single flash flood events (Gaume et al. 2016). On the contrary, flood databases have typically poor spatial/time coverage and little or no information on the type of flood for each entry to allow an in-depth, long-term flash flood verification analysis (Gaume et al. 2009; Kruczkiewicz et al. 2021b). The problem with this approach is that the verification results are very much case-dependent, and generalisations in space and time might not be possible (e.g., would other regions different from the one in the case-study show the same performance? Would have the performance been the same if the event happened in the afternoon instead of in the morning?). To carry out more comprehensive, longer-term verification analysis, authors tend to alternatively use rainfall observations to infer the goodness of rainfall forecasts at predicting flash floods (c.f. Park et al. (2019) and Gascòn et al. (2022)). This approach is adopted because the quality and quantity of the spatial/temporal coverage is much better for rainfall than for flash flood observations (Haiden and Duffy 2016; Zhang et al. 2011). The problem with this approach is that, while there is a much more linear relationship between rainfall and flash flood events than for other types of flood, the results from the rainfall verification do not necessarily represent the performance of the rainfall forecasts at predicting areas at risk of flash floods.

In addition, the flood database’s good spatio/temporal coverage and the quality of its reports allowed us to build the climatology of rainfall events associated with flash flood events from short-range ecPoint-Rainfall forecasts. The uses of such methodology are twofold. It can be used to build the VREs to use in a verification analysis (as in this study), or to develop a flash flood warning system which uses such rainfall climatology to define the warning rainfall events (VREs) that, if exceeded, could lead to flash flood events. It is worth acknowledging that the use of short-range ecPoint-Rainfall forecasts to build such rainfall climatology can provide a larger distribution of rainfall totals than those defined by actual rainfall observations. However, the selection of a percentile (e.g., 50th, 75th, 85th or 99th) of the distribution of forecast rainfall values assigned to each flood report helps to mitigate the above-mentioned issue, as only such percentiles build the distribution that will be used to define the VREs. Therefore, the construction of good quality flash flood databases is essential to carry out in-depth, long-term flash flood verification analysis. Therefore, the development of accurate flash flood report databases with good spatio/temporal coverage should remain of a high priority for the scientific community.

# Conclusions

The study aims are twofold. First, it aims to assess the performance of ecPoint-Rainfall forecasts in predicting areas at flash flood risk in Ecuador, particularly in terms of forecast accuracy and lead-time extension to medium ranges. This is the first study that discusses ecPoint’s performance in flash flood forecasting using flash flood reports instead of rainfall observations. This contribution is important because, while there is a more linear correlation between rainfall and flash flood events compared to other types of floods, not all rainfall events might translate into flash floods. Second, the study aims at defining a methodology using short-range ecPoint-Rainfall forecasts to define warnings for areas at flash flood risk in regions with no or poor rain gauge coverage to define rainfall climatologies but with a good historical record of flash flood events. Since ecPoint-Rainfall forecasts are global, the ability to create flash flood warnings would not be hampered by the lack of rain gauge observations, provided historical records of flash flood events exist for the region of interest.

Two user categories have been defined in this study: the first one is interested only in whether there will be a flash flood event (the answer will be binary, i.e., yes or no), while the second one is also interested in verifying the amount of the rainfall event that generated such flash flood. Users in the first category can identify areas at flash flood risk using the ENS’s 85th percentile and obtain comparable results to the ecPoint’s 99th percentile. Users in the second category are likely to find that local rainfall extremes verify better in ecPoint’s 99th percentile than ENS’s 85th percentile. For extreme localised rainfall events (VREs = 99th percentile), the results suggest that ecPoint outperforms ENS in both regions, “La Costa” and “La Sierra”. The performance of ecPoint appears to be better in “la Sierra”, where rainfall is originated mainly from small-scale convective systems. In “La Costa”, where rainfall is mainly originated from large-scale convective systems, ecPoint and ENS performances are comparable except for certain times of the day, where for example, bias correction applied by ecPoint in the rainfall’s diurnal cycle increases the overall performance of the post-processed forecasts over the raw ENS. Another example of how ecPoint and ENS verify against point rainfall observations is provided in the case study.

This study has also provided a method that uses short-range (i.e., day 1) ecPoint-Rainfall forecasts as a proxy for point rainfall observations to create a synthetic rainfall climatologies that can be used to define VREs. This method could also be applied in the definition of thresholds to issue warnings for areas at risk of flash floods.

The authors identify two areas for further research. First, developing ecPoint-compatible model climatologies from the point of view of spatial resolution (i.e., both forecasts and climatologies refer to point-rainfall instead of average rainfall over a grid-box) contributes to the creation of flash flood warnings with a continuous global domain. Such a product could be used to provide flash flood forecasts with a continuous global domain, de facto providing flash flood forecasts to regions of the world that currently do not have any forecasting systems in place or extending to the medium ranges the flash flood forecasts in those countries where only shorter lead time forecasts are available. In addition, forecasters in global projects such as Aristotle[[8]](#footnote-9) or FbF could formulate their flash flood predictions without knowing the local rainfall climatology of the areas of interest. Finally, the authors believe that more resources should be spent to keep developing flood databases like the one presented in this study because they incorporate invaluable details on the type of flood that can be used to target verification and forecast development efforts.

# Tables

**Table 1** – The first column shows the total number of flood reports in the database for 2019 and 2020. The second column shows the number of excluded reports from the study because they did not contain any reporting location (in lat/lon coordinates) and/or reporting time (with date and time). The remaining columns show the distribution of flood reports per region and EFFCI threshold.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Year** | **N. of raw flood reports** | **N. of eliminated flood reports** | **N. of cleaned flood reports** | **N. of flood reports with EFFCI>=1** | | | **N. of flood reports with**  **EFFCI>=6** | | | **N. of flood reports**  **with EFFCI>=10** | | |
| Costa | Sierra | Oriente | Costa | Sierra | Oriente | Costa | Sierra | Oriente |
| 2019 | 302 | 3 | 299 | 176 | 116 | 7 | 93 | 102 | 0 | 17 | 34 | 0 |
| 2020 | 190 | 0 | 190 | 96 | 93 | 88 | 48 | 79 | 0 | 22 | 26 | 0 |

**Table 2** - Day 1 ecPoint-Rainfall forecasts used to define the verifying rainfall events (VREs).

|  |  |  |
| --- | --- | --- |
| **Flood reports’ reporting time**  **(in UTC for day X)** | **Valid 12-hourly periods (in UTC) containing the flood reports’ reporting time** | **Correspondent 12-hourly accumulation periods for day 1 ecPoint-Rainfall forecasts (forecast run date / forecast run time in UTC / lead time in hours)** |
| Between 0 and 5.:59 | 18 (on day X-1) to 5:59 | Day(X-1) / 00 UTC / (t+18,t+30) |
| Day(X-1) / 12 UTC / (t+6,t+18) |
| 0 to 11:59 | Day(X) / 00 UTC / (t+0,t+12) |
| Day(X-1) / 12 UTC / (t+12,t+24) |
| Between 6 and 11:59 | 0 to 11:59 | Day(X) / 00 UTC / (t+0,t+12) |
| Day(X-1) / 12 UTC / (t+12,t+24) |
| 6 to 17:59 | Day(X) / 00 UTC / (t+6,t+18) |
| Day(X-1) / 12 UTC / (t+18,t+30) |
| Between 12 and 17:59 | 6 to 17:59 | Day(X) / 00 UTC / (t+6,t+18) |
| Day(X-1) / 12 UTC / (t+18,t+30) |
| 12 to 23:59 | Day(X) / 00 UTC / (t+12,t+24) |
| Day(X) / 12 UTC / (t+0,t+12) |
| Between 18 and 23:59 | 12 to 23:59 | Day(X) / 00 UTC / (t+12,t+24) |
| Day(X) / 12 UTC / (t+0,t+12) |
| 18 to 5:59 (on day X+1) | Day(X) / 00 UTC / (t+18,t+30) |
| Day(X) / 12 UTC / (t+6,t+18) |

**Table 3** - Definition of the four quadrants in a contingency table.

|  |  |  |
| --- | --- | --- |
| **FORECASTS (COLUMNS) /**  **OBSERVATIONS (ROWS)** | **YES** | **NO** |
| **YES** | QUADRANT I  Hits (H)  The event *was observed* when it *was predicted*. | QUADRANT II  False Alarms (FA)  The event *was not observed* when it *was predicted*. |
| **NO** | QUADRANT III  Misses (M)  The event *was observed* when it *was not predicted*. | QUADRANT IV  Correct Negatives (CN)  The event *was not observed* when it *was not predicted*. |

**Table 4** - Verifying rainfall events (in mm/12h).

|  |  |  |
| --- | --- | --- |
| **Region** | **VRE85** | **VRE99** |
| La Costa | 9.865 mm/12h (rounded to 10 mm/12h) | 50.452 mm/12h (rounded to 50 mm/12h) |
| La Sierra | 5.885 mm/12h (rounded to 6 mm/12h) | 25.551 mm/12h (rounded to 26 mm/12h) |

# Figures

Immagine che contiene testo, mappa, atlante

Descrizione generata automaticamente

**Figure 1** - Panel (a) shows Ecuador’s orography, its political regions with white contours, and the location of Ecuador’s three main geographical regions: the coast (“La Costa”), the highlands (“La Sierra”), and the Amazon Forest (“EL Oriente”). The insert shows Ecuador’s location (in red) within South America. Ecuador’s domain within the ENS and ecPoint grid is shown in Panel (b). The location of the model’s grid-boxes is indicated with black dots, and the regions are indicated in yellow for “La Costa”, in brown for “La Sierra” and in green for “El Oriente”. Panel (c) shows the names of Ecuador’s political regions, and their estimated population density (in people/km2) from 2020’s census.

Immagine che contiene testo, mappa, diagramma, atlante

Descrizione generata automaticamente

**Figure 2** – Panels (a), (b), and (c) show, respectively, the spatial distribution of point flood reports for 2019 with an EFFCI>=1, EFFCI>=6, and EFFCI>=10. Panels (d), (e), and (f) are the same but for flood events occurred in 2020.

A screenshot of a map

Description automatically generated

**Figure 3** – Panel (a) shows the timeseries of the counts of flood reports, accumulated over four overlapping 12-hourly accumulation periods, starting at 0 (first row), 6 (second row), 12 (third row), and 18 UTC (fourth row). The flood reports are from 2020 with EFFCI>=6, whose spatial distribution was shown in **Figure 2e**. The bars that correspond to the counts of grid reports (in pink) are overlapped with the bars that correspond to the counts of point reports (in brown) for the same day, meaning that where brown bars are visible, more than one point flood report was assign to same grid-boxes. Panel (b) shows an example of the spatial distribution of (point and grid) flood reports for a specific 12-hourly accumulation period (starting on 2020/02/28 at 00 UTC).

,

Immagine che contiene testo, diagramma, Parallelo, Piano

Descrizione generata automaticamente

**Figure 4** – Flowchart containing the “Diagnostic methodological steps” followed in this study (in green) and the correspondent “Outputs” (in red) to answer the research question posed in the introduction (in yellow).

Map

Description automatically generated

**Figure 5** – Schematic representation of how the verifying rainfall events (VREs) are computed, for the N point flood events in “La Costa”. The figure is meant to be read from right to left. (a) Locations of point flood reports, with EFFCI>=6, in 2019. (b) Schematic representation of the distributions of short-range ecPoint-Rainfall realisations (blue) for a sample of the N point flood reports in “La Costa”. From such distributions, rainfall values tp(xth), in red, are extracted for the xth percentile. This percentile denotes a rainfall event presumed sufficiently extreme to generate flash floods. (c) Net distribution (red) of the N tp(xth) values. From this net distribution, the rainfall value tp(25th) is extracted (purple). Choice of the 25th percentile is subjective; it nominally retains 75% of the observed point flood events.

Chart, diagram

Description automatically generated

**Figure 6** – Amounts (in mm/12h) of typical rainfall events during flash flood events in “La Costa” (panels a, b and c) and “La Sierra” (panel d, e, and f). Distributions of short-range (i.e. day 1) ecPoint-Rainfall forecasts are built for each flood report (FRs) in 2019 as shown in **Figure 4**. Panels (a) and (d) are for FRs with EFFCI>=1, panels (b) and (e) are for FRs with EFFCI>=6, and panels (c) and (f) are for FRs with EFFCI>=10. How “extreme” is the typical rainfall event associated to each FR is defined by seven percentiles (i.e., 50th in yellow, 75th in green, 85th in purple, 90th in cyan, 95th in blue, 98th in fuchsia, and 99th in orange) extracted from the distribution of 99 ecPoint-Rainfall forecasts for each FR. The verifying rainfall events values (VREs) are then defined considering the values of the typical rainfall events for the top 75% FRs (highlighted by the grey line in the zoomed in figures in correspondence of the 25th percentile for the seven distributions). The VREs values for the distributions corresponding to the 85th and 99th percentiles are highlighted with grey circles.



**Figure 7** - Areas under the ROC curve (AROC) for flood reports with EFFCI>=6. Panels (a) and (b) show, respectively, the AROC for VRE85 ((tp>=10 mm/12) and VRE99 (tp>=50 mm/12) in “La Costa”. Panels (c) and (d) show, respectively, the AROC for VRE85 (tp>=6 mm/12) and VRE99 (tp>=26 mm/12) in “La Sierra”. The lines and the shaded areas represent, respectively, the values of the AROC and the confidence interval (CI) at 95% for ENS (in red) and ecPoint (in blue). The x-axis indicates the forecast lead times as steps at the end of the 12-hourly accumulation period expressed in hours from the 00 UTC run. The colours associated to each step indicate the correspondent valid 12-hourly accumulation periods in UTC and local time (LT): green for 0000-1200 UTC (or 1800-0600 LT), purple for 0600-1800 UTC (or 0000-1200 LT), cyan for 1200-0000 UTC (or 0600-1800 LT), and fuchsia for 1800-0600 UTC (or 1200-0000 LT). The x-axis indicates also the equivalent lead times in days (from 1 to 10).

Chart

Description automatically generated

**Figure 8** - ENS (red lines) and ecPoint (blue lines) diurnal cycle for 12-hourly rainfall annual mean in “La Costa” (solid lines) and “La Sierra” (dashed lines). The annual means are obtained from day 1 forecasts between 1st January to 31st December (only 00 UTC run were considered). Four overlapping 12-hourly accumulation periods in UTC (and the correspondent local time, LT) are shown, and are indicated with four different colours: green for 0000-1200 UTC (1800-0600 LT), purple for 0600-1800 UTC (or 0000-1200 LT), cyan for 1200-0000 UTC (or 0600-1800 LT), and fuchsia for 1800-0600 UTC (or 1200-0000 LT).

Chart, line chart

Description automatically generated

**Figure 9** – ROC curves for VRE99 and for the accumulation period between (t+60,t+72), i.e. for 0600 to 1800 local time (LT). The red and the blue line denote, respectively, the ROC curves for ENS and ecPoint. The continuous and the dashed lines correspond to the ROC curves for “La Costa” and “La Sierra”.

Diagram

Description automatically generated

**Figure 10** - Flash floods in Ecuador on 8th March 2021. Panel (a) shows 24-hourly synop rainfall observations between 8th March at 6 am and 9th March at 6 am (coloured dots), 24-hourly rainfall reports from INAMHI for Guayaquil between 8th March at 0 am and 9th March at 0 am (coloured triangles), and flash flood reports in different regions between 8th March at 0 am and 9th March at 0 am (black diamonds). Panel (b) shows the sounding for Guayaquil (lat:-2.2; lon:-79.9) valid for 8th March 2021 at 6 am. Panel (c) shows day 1, 3, and 7 forecasts from 00 UTC runs for ENS and ecPoint, valid for the accumulation period between 8th March at 12 am and 9th March at 0 am (when the rainfall event was at its peak). All reported times here are meant in LT.

# References

Adler, R. F., and Coauthors, 2018: The Global Precipitation Climatology Project (GPCP) monthly analysis (New Version 2.3) and a review of 2017 global precipitation. *Atmosphere (Basel)*, **9**, 138, https://doi.org/10.3390/atmos9040138.

Barrett, A. I., C. Wellmann, A. Seifert, C. Hoose, B. Vogel, and M. Kunz, 2019: One Step at a Time: How Model Time Step Significantly Affects Convection-Permitting Simulations. *J Adv Model Earth Syst*, **11**, 641–658, https://doi.org/10.1029/2018MS001418.

Bazo, J., R. Singh, M. Destrooper, and E. C. De Perez, 2019: Pilot experiences in using seamless forecasts for early action: The “ready-set-go!" approach in the red cross. *Sub-seasonal to Seasonal Prediction: The Gap Between Weather and Climate Forecasting*, 387–398.

Beck, H. E., E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. J. M. Van Dijk, T. R. McVicar, and R. F. Adler, 2019: MSWep v2 Global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bull Am Meteorol Soc*, **100**, 473–500, https://doi.org/10.1175/BAMS-D-17-0138.1.

Bischiniotis, K., B. van den Hurk, E. Coughlan de Perez, T. Veldkamp, G. G. Nobre, and J. Aerts, 2019: Assessing time, cost and quality trade-offs in forecast-based action for floods. *International Journal of Disaster Risk Reduction*, **40**, 101252, https://doi.org/10.1016/j.ijdrr.2019.101252.

Ben Bouallègue, Z., and D. S. Richardson, 2022: On the ROC area of ensemble forecasts for rare events. *Weather Forecast*, **37**, 787–796, https://doi.org/10.1175/waf-d-21-0195.1.

Bucherie, A., F. Ayala, and A. Kruczkiewicz, 2021: Ecuador historical flood occurrences and impacts dataset with Flash Flood Confidence Index (2007-2020). *Zenodo*, https://doi.org/10.5281/zenodo.4662886.

——, C. Hultquist, S. Adamo, C. Neely, F. Ayala, J. Bazo, and A. Kruczkiewicz, 2022a: A comparison of social vulnerability indices specific to flooding in Ecuador: principal component analysis (PCA) and expert knowledge. *International Journal of Disaster Risk Reduction*, **73**, 102897, https://doi.org/10.1016/j.ijdrr.2022.102897.

——, M. Werner, M. Van Den Homberg, and S. Tembo, 2022b: Flash flood warnings in context: Combining local knowledge and large-scale hydro-meteorological patterns. *Natural Hazards and Earth System Sciences*, **22**, 461–480, https://doi.org/10.5194/nhess-22-461-2022.

Buizza, R., 2019: Introduction to the special issue on “25 years of ensemble forecasting.” *Quarterly Journal of the Royal Meteorological Society*, **145**, 1–11, https://doi.org/10.1002/qj.3370.

Cafaro, C., and Coauthors, 2021: Do convection-permitting ensembles lead to more skillful short-range probabilistic rainfall forecasts over tropical east africa? *Weather Forecast*, **36**, 697–716, https://doi.org/10.1175/WAF-D-20-0172.1.

Corral, C., M. Berenguer, D. Sempere-Torres, L. Poletti, F. Silvestro, and N. Rebora, 2019: Comparison of two early warning systems for regional flash flood hazard forecasting. *J Hydrol (Amst)*, **572**, 603–619, https://doi.org/10.1016/j.jhydrol.2019.03.026.

Coughlan de Perez, E., and Coauthors, 2022: Adapting to climate change through anticipatory action: The potential use of weather-based early warnings. *Weather Clim Extrem*, **38**, 100508, https://doi.org/10.1016/j.wace.2022.100508.

Davolio, S., F. Silvestro, and T. Gastaldo, 2017: Impact of rainfall assimilation on high-resolution hydrometeorological forecasts over Liguria, Italy. *J Hydrometeorol*, **18**, 2659–2680, https://doi.org/10.1175/JHM-D-17-0073.1.

DiCiccio, T. J., and B. Efron, 1996: Bootstrap confidence intervals. *Statistical Science*, **11**, 189–212, https://doi.org/10.1214/ss/1032280214.

Dordevic, M., P. Mutic, and H. Kim, 2020: Flash Flood Guidance System: Response to one of the deadliest hazards. *WMO Bulletin*, **69**, 29–33.

Galarza-Villamar, J. A., C. Leeuwis, G. M. Pila-Quinga, F. Cecchi, and C. M. Párraga-Lema, 2018: Local understanding of disaster risk and livelihood resilience: The case of rice smallholders and floods in Ecuador. *International Journal of Disaster Risk Reduction*, **31**, 1107–1120, https://doi.org/10.1016/j.ijdrr.2018.08.009.

Gascón, E., A. Montani, and T. D. Hewson, 2023: Post-processing output from ensembles with and without parametrised convection, to create accurate, blended, high-fidelity rainfall forecasts. https://doi.org/10.48550/arxiv.2301.04485.

Gaume, E., and Coauthors, 2009: A compilation of data on European flash floods. *J Hydrol (Amst)*, **367**, 70–78, https://doi.org/10.1016/j.jhydrol.2008.12.028.

——, M. Borga, M. C. Llasat, S. Maouche, M. Lang, and M. Diakakis, 2016: Mediterranean extreme floods and flash floods. *The Mediterranean Region under Climate Change. A Scientific Update*, 133–144.

Georgakakos, K. P., 2006: Analytical results for operational flash flood guidance. *J Hydrol (Amst)*, **317**, 81–103, https://doi.org/10.1016/j.jhydrol.2005.05.009.

——, and Coauthors, 2021: The Flash Flood Guidance System Implementation Worldwide: A Successful Multidecadal Research-To-Operations Effort. *Bull Am Meteorol Soc*, **1**, 1–35, https://doi.org/10.1175/bams-d-20-0241.1.

Gneiting, T., and P. Vogel, 2021: Receiver operating characteristic (ROC) curves: equivalences, beta model, and minimum distance estimation. *Mach Learn*, 1–13, https://doi.org/10.1007/s10994-021-06115-2.

Gourley, J. J., and Coauthors, 2017: The FLASH Project - Improving the Tools for Flash Flood Monitoring and Prediction across the United States. *Bull Am Meteorol Soc*, **98**, 361–372, https://doi.org/https://doi.org/10.1175/BAMS-D-15-00247.1.

Haiden, T., and S. Duffy, 2016: Use of high-density observations in precipitation verification. *ECMWF Newsletter*, 20–25, https://doi.org/10.21957/hsacrdem.

Haiden, T., M. Jannousek, F. Vitart, Z. Ben-Bouallegue, L. Ferranti, C. Prates, and D. Richardson, 2021: Evaluation of ECMWF forecasts, including the 2020 upgrade. *ECMWF Technical Memoranda*, **880**, https://doi.org/10.21957/6njp8byz4.

Haiden, T., M. Janousek, F. Vitart, Z. Ben-Bouallegue, and F. Prates, 2023: Evaluation of ECMWF forecasts, including the 2023 upgrade. *ECMWF Technical Memoranda*, **911**, 1–60.

Hamill, T. M., J. S. Whitaker, and S. L. Mullen, 2006: Reforecasts: An important dataset for improving weather predictions. *Bull Am Meteorol Soc*, **87**, 33–46, https://doi.org/10.1175/BAMS-87-1-33.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hewson, T. D., and F. M. Pillosu, 2021: A new low-cost technique improves weather forecasts across the world. *Commun Earth Environ*, **2**, 132, https://doi.org/10.1038/s43247-021-00185-9.

Hirabayashi, Y., M. Tanoue, O. Sasaki, X. Zhou, and D. Yamazaki, 2021: Global exposure to flooding from the new CMIP6 climate model projections. *Sci Rep*, **11**, 1–7, https://doi.org/10.1038/s41598-021-83279-w.

Hurford, A. P., S. J. Priest, D. J. Parker, and D. M. Lumbroso, 2012: The effectiveness of extreme rainfall alerts in predicting surface water flooding in England and Wales. *International Journal of Climatology*, **32**, 1768–1774, https://doi.org/https://doi.org/10.1002/joc.2391.

Ibarreche, J., and Coauthors, 2020: Flash flood early warning system in Colima, Mexico. *Sensors (Switzerland)*, **20**, 1–26, https://doi.org/10.3390/s20185231.

Imhoff, R. O., C. C. Brauer, K. J. van Heeringen, R. Uijlenhoet, and A. H. Weerts, 2022: Large‐Sample Evaluation of Radar Rainfall Nowcasting for Flood Early Warning. *Water Resour Res*, **58**, e2021WR031591, https://doi.org/10.1029/2021WR031591.

Javelle, P., C. Fouchier, P. Arnaud, and J. Lavabre, 2010: Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations. *J Hydrol (Amst)*, **394**, 267–274, https://doi.org/https://doi.org/10.1016/j.jhydrol.2010.03.032.

——, D. Organde, J. Demargne, C. Saint-Martin, C. De Saint-Aubin, L. Garandeau, and B. Janet, 2016: Setting up a French national flash flood warning system for ungauged catchments based on the AIGA method. *E3S Web of Conferences*, Vol. 7 of, 718010–18010.

Jolliffe, I. T., and D. B. Stephenson, 2011: *Forecast Verification: A Practitioner’s Guide in Atmospheric Science*. 2nd Editio. John Wiley & Sons, Ltd,.

Jonkman, S. N., and J. K. Vrijling, 2008: Loss of life due to floods. *J Flood Risk Manag*, **1**, 43–56, https://doi.org/10.1111/j.1753-318x.2008.00006.x.

Kikuchi, K., and B. Wang, 2008: Diurnal precipitation regimes in the global tropics. *J Clim*, **21**, 2680–2696, https://doi.org/10.1175/2007JCLI2051.1.

Kiptum, A., and Coauthors, 2023: Advancing operational flood forecasting, early warning and risk management with new emerging science: Gaps, opportunities and barriers in Kenya. *Journal of Flood Risk Management*, John Wiley & Sons, Ltd, e12884.

Kruczkiewicz, A., A. Bucherie, F. Ayala, and J. Bazo, 2021a: Forecast-based financing for flash floods: a flash flood confidence index to improve flood reporting. *Anticipation Hub*,. https://www.anticipation-hub.org/news/forecast-based-financing-for-flash-floods-a-flash-flood-confidence-index.

——, ——, ——, C. Hultquist, H. Vergara, S. Mason, J. Bazo, and A. de Sherbinin, 2021b: Development of a flash flood confidence index from disaster reports and geophysical susceptibility. *Remote Sens (Basel)*, **13**, 2764, https://doi.org/10.3390/rs13142764.

Laraque, A., and Coauthors, 2009: Sediment budget of the Napo River, Amazon Basin, Ecuador and Peru. *Hydrol Process*, **23**, 3509–3524, https://doi.org/10.1002/hyp.7463.

Lavers, D. A., S. Harrigan, and C. Prudhomme, 2021: Precipitation Biases in the ECMWF Integrated Forecasting System. *J Hydrometeorol*, https://doi.org/10.1175/jhm-d-20-0308.1.

Liu, C., L. Guo, L. Ye, S. Zhang, Y. Zhao, and T. Song, 2018: A review of advances in China’s flash flood early-warning system. *Natural Hazards*, **92**, 619–634, https://doi.org/10.1007/s11069-018-3173-7.

Luong, T. T., J. Pöschmann, R. Kronenberg, and C. Bernhofer, 2021: Rainfall threshold for flash flood warning based on model output of soil moisture: Case study wernersbach, germany. *Water (Switzerland)*, **13**, 1–15, https://doi.org/10.3390/w13081061.

Ma, M., and Coauthors, 2021: Development of a new rainfall-triggering index of flash flood warning-case study in Yunnan province, China. *J Flood Risk Manag*, **14**, 1–9, https://doi.org/10.1111/jfr3.12676.

Nielsen, E. R., and R. S. Schumacher, 2016: Using convection-allowing ensembles to understand the predictability of an extreme rainfall event. *Mon Weather Rev*, **144**, 3651–3676, https://doi.org/10.1175/MWR-D-16-0083.1.

Park, S., M. Berenguer, and D. Sempere-Torres, 2019: Long-term analysis of gauge-adjusted radar rainfall accumulations at European scale. *J Hydrol (Amst)*, **573**, 768–777, https://doi.org/10.1016/j.jhydrol.2019.03.093.

De Perez, E. C., and Coauthors, 2016: Action-based flood forecasting for triggering humanitarian action. *Hydrol Earth Syst Sci*, **20**, 3549–3560, https://doi.org/https://doi.org/10.5194/hess-20-3549-2016.

Pinos, J., and L. Timbe, 2020: Mountain Riverine Floods in Ecuador: Issues, Challenges, and Opportunities. *Frontiers in Water*, **2**, 36, https://doi.org/10.3389/frwa.2020.545880.

——, and A. Quesada-Román, 2022: Flood risk-related research trends in Latin America and the Caribbean. *Water (Switzerland)*, **14**, 10, https://doi.org/10.3390/w14010010.

Ramos Filho, G. M., V. H. R. Coelho, E. da S. Freitas, Y. Xuan, and C. das N. Almeida, 2021: An improved rainfall-threshold approach for robust prediction and warning of flood and flash flood hazards. *Natural Hazards*, **105**, 2409–2429, https://doi.org/10.1007/s11069-020-04405-x.

Raynaud, D., J. Thielen, P. Salamon, P. Burek, S. Anquetin, and L. Alfieri, 2015: A dynamic runoff co-efficient to improve flash flood early warning in Europe: Evaluation on the 2013 central European floods in Germany. *Meteorological Applications*, **22**, 410–418, https://doi.org/10.1002/met.1469.

Recalde-Coronel, C. G., A. G. Barnston, and Á. G. Muñoz, 2014: Predictability of december-april rainfall in coastal and Andean Ecuador. *J Appl Meteorol Climatol*, **53**, 1471–1493, https://doi.org/10.1175/JAMC-D-13-0133.1.

Recalde-Coronel, G. C., B. Zaitchik, and W. K. Pan, 2020: Madden–Julian oscillation influence on sub-seasonal rainfall variability on the west of South America. *Clim Dyn*, **54**, 2167–2185, https://doi.org/10.1007/s00382-019-05107-2.

Robbins, J. C., and H. A. Titley, 2018: Evaluating high-impact precipitation forecasts from the Met Office Global Hazard Map (GHM) using a global impact database. *Meteorological Applications*, **25**, 548–560, https://doi.org/https://doi.org/10.1002/met.1720.

Šakić Trogrlić, R., M. van den Homberg, M. Budimir, C. McQuistan, A. Sneddon, and B. Golding, 2022: Early Warning Systems and Their Role in Disaster Risk Reduction. *Towards the “Perfect” Weather Warning*, 11–46.

Schroeder, A. J., and Coauthors, 2016: The development of a flash flood severity index. *J Hydrol (Amst)*, **541**, 523–532, https://doi.org/10.1016/j.jhydrol.2016.04.005.

Schwartz, C. S., 2019: Medium-range convection-allowing ensemble forecasts with a variable-resolution global model. *Mon Weather Rev*, **147**, 2997–3023, https://doi.org/10.1175/MWR-D-18-0452.1.

Shuvo, S. D., T. Rashid, S. K. Panda, S. Das, and D. A. Quadir, 2021: Forecasting of pre-monsoon flash flood events in the northeastern Bangladesh using coupled hydrometeorological NWP modelling system. *Meteorology and Atmospheric Physics*, 1–23, https://doi.org/10.1007/s00703-021-00831-z.

Song, H. J., B. Lim, and S. Joo, 2019: Evaluation of rainfall forecasts with heavy rain types in the high-resolution unified model over South Korea. *Weather Forecast*, **34**, 1277–1293, https://doi.org/10.1175/WAF-D-18-0140.1.

Speight, L., and Coauthors, 2018: Developing surface water flood forecasting capabilities in Scotland: an operational pilot for the 2014 Commonwealth Games in Glasgow. *J Flood Risk Manag*, **11**, S884–S901, https://doi.org/10.1111/jfr3.12281.

Speight, L. J., M. D. Cranston, C. J. White, and L. Kelly, 2021: Operational and emerging capabilities for surface water flood forecasting. *Wiley Interdisciplinary Reviews: Water*, **8**, https://doi.org/10.1002/wat2.1517.

Tapiador, F. J., R. Roca, A. Del Genio, B. Dewitt, W. Petersen, and F. Zhang, 2019: Is precipitation a good metric for model performance? *Bull Am Meteorol Soc*, **100**, 223–233, https://doi.org/10.1175/BAMS-D-17-0218.1.

Thomassen, E. D., E. J. Kendon, H. J. D. Sørup, S. C. Chan, P. L. Langen, O. B. Christensen, and K. Arnbjerg-Nielsen, 2021: Differences in representation of extreme precipitation events in two high resolution models. *Clim Dyn*, **57**, 3029–3043, https://doi.org/10.1007/s00382-021-05854-1.

Tobar, V., and G. Wyseure, 2018: Seasonal rainfall patterns classification, relationship to ENSO and rainfall trends in Ecuador. *International Journal of Climatology*, **38**, 1808–1819, https://doi.org/10.1002/joc.5297.

Trigg, M. A., M. D. Wilson, P. D. Bates, M. S. Horritt, D. E. Alsdorf, B. R. Forsberg, and M. C. Vega, 2009: Amazon flood wave hydraulics. *J Hydrol (Amst)*, **374**, 92–105, https://doi.org/10.1016/j.jhydrol.2009.06.004.

UNDRR, 2021: Desinventar Project. *Plataforma Desinventar Sendai*,. https://db.desinventar.org/.

Vannitsem, S., and Coauthors, 2021: Statistical postprocessing for weather forecasts review, challenges, and avenues in a big data world. *Bull Am Meteorol Soc*, **102**, E681–E699, https://doi.org/10.1175/BAMS-D-19-0308.1.

Vuille, M., R. S. Bradley, and F. Keimig, 2000: Climate variability in the Andes of Ecuador and its relation to tropical Pacific and Atlantic Sea Surface temperature anomalies. *J Clim*, **13**, 2520–2535, https://doi.org/10.1175/1520-0442(2000)013<2520:CVITAO>2.0.CO;2.

Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon Weather Rev*, **132**, 1917–1932, https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

Xing, Y., Q. Liang, G. Wang, X. Ming, and X. Xia, 2019: City-scale hydrodynamic modelling of urban flash floods: the issues of scale and resolution. *Natural Hazards*, **96**, 473–496, https://doi.org/10.1007/s11069-018-3553-z.

Zanchetta, A. D. L., and P. Coulibaly, 2020: Recent advances in real-time pluvial flash flood forecasting. *Water (Switzerland)*, **12**, 570, https://doi.org/10.3390/w12020570.

——, ——, and V. Fortin, 2022: Forecasting High-Flow Discharges in a Flashy Catchment Using Multiple Precipitation Estimates as Predictors in Machine Learning Models. *Hydrology*, **9**, 216, https://doi.org/10.3390/hydrology9120216.

Zeman, C., N. P. Wedi, P. D. Dueben, N. Ban, and C. Schär, 2021: Model intercomparison of COSMO 5.0 and IFS 45r1 at kilometer-scale grid spacing. *Geoscientific Model Development [preprint]*, **2021**, 1–35, https://doi.org/10.5194/gmd-2021-31.

Zhang, J., and Coauthors, 2011: National mosaic and multi-sensor QPE (NMQ) system description, results, and future plans. *Bull Am Meteorol Soc*, **92**, 1321–1338, https://doi.org/10.1175/2011BAMS-D-11-00047.1.

1. www.ecmwf.int/en/forecasts/documentation-and-support/evolution-ifs/cycles/summary-cycle-45r1 [↑](#footnote-ref-2)
2. www.ecmwf.int/en/forecasts/about-our-forecasts/evolution-ifs/cycles/summary-cycle-46r1 [↑](#footnote-ref-3)
3. www.ecmwf.int/en/forecasts/about-our-forecasts/evolution-ifs/cycles/summary-cycle-47r1 [↑](#footnote-ref-4)
4. https://www.pichinchacomunicaciones.com.ec/lluvias-causan-desborde-de-rios-e-inundaciones-en-varias-provincias/ [↑](#footnote-ref-5)
5. https://www.wunderground.com/history/daily/SEGU/date/2021-3-8 [↑](#footnote-ref-6)
6. https://www.eluniverso.com/guayaquil/comunidad/la-mayor-lluvia-del-2021-en-guayaquil-provoco-afectaciones-en-64-zonas-entre-inundaciones-arboles-caidos-canales-rebosados-y-otros-nota/ [↑](#footnote-ref-7)
7. https://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/ARCHIVE/PDF/mjo\_evol-status-fcsts-20210315.pdf [↑](#footnote-ref-8)
8. https://meetingorganizer.copernicus.org/EGU2020/EGU2020-20016.html [↑](#footnote-ref-9)