Assessing the ability of probabilistic point-scale rainfall forecasts in predicting flash floods: using historical flash flood reports to refine predictions

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](mailto:fatima.pillosu@ecmwf.int))

**Abstract.** Globally, flash floods are one of the most costly natural hazards for property damage and loss of life. The low prediction accuracy for forecast lead times beyond a few hours limits the use of flash flood forecasts in early warning systems. This study presents the verification results of a new flash flood forecasting system based on “ecPoint” rainfall forecasts using ECMWF ensemble forecasts (ENS) as a benchmark. A one-year verification for 2020 is carried out in Ecuador, whose varied climatology made it an attractive site for verification. The recent development of a detailed flash flood database in Ecuador enabled this verification analysis. The original contribution to knowledge is the identification of scenarios in which the use of 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 skills are comparable with the exception of corrections of errors in rainfall’s diurnal cycle. This study also contributes a new methodology to define the average rainfall event that can cause flash floods when no point rainfall observations (e.g., rain gauges) are available, by using short-term ecPoint forecasts as a proxy for point rainfall observations. This new method us to verify flash flood forecasts and develop warning systems that will help decision-makers mitigate the effects of flash floods in regions with poor rainfall observation coverage.

**Keywords.** Flash flood forecasting, historical flash flood reports, ecPoint.

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

# Introduction

Flash floods are one of the most concerning types of natural hazards in Ecuador as their unexpected, rapidly evolving, and difficult to predict features can cause severe distress in the affected areas. While there are no official statistics describing the occurrence of flash floods in Ecuador, a recent study analysing historical disaster reports has shown that about 60% of flood reports could represent flash flood events (Kruczkiewicz et al. 2021b). Furthermore, humanitarian and news reports from specialized websites such as FloodList[[1]](#footnote-2) and Reliefweb[[2]](#footnote-3) suggest that flash floods are one of the most recurrent and damaging type of floods in the country. Ecuador is one of many countries that face flash flood risk. Globally, flash floods account for ~85% of all types of floods. With more than 5000 fatalities every year (Dordevic et al. 2020), flash floods also have the highest mortality rate amongst other types of flood (Jonkman and Vrijling 2008). Flash floods also have severe social, economic, and environmental impacts due to the increased vulnerability of people living and having economic activities in flood-prone areas and cities (Dordevic et al. 2020). Such impacts are exacerbated in low-income countries that have fewer resources to recover from the impacts of extreme natural hazards (Winsemius et al. 2018). In Latin America, floods cause severe impacts driven by a variety of factors such as exponential, unregulated urbanization of floodplains, human-triggered catchment degradation, absence of preparedness plans, the persistence of poverty, inefficient public policies, and lack of sufficient infrastructure (Pinos and Quesada-Román 2022). In Ecuador, flash floods have caused both short term and long term impacts such as damage to infrastructure and agriculture, interruptions to business and education, disruption of healthcare services, and outbreaks of waterborne diseases (Galarza-Villamar et al. 2018). Future projections noting increases in flood risk, coupled with socioeconomic factors that are likely to persist, lead to a future whereby flash flood impact may increase in Ecuador in the coming decades (Hirabayashi et al. 2021).

Forecast-triggered strategies for flood risk reduction such as “early warning systems, EWSs” (Golnaraghi 2012) and “forecast-based financing, FbF” protocols (Coughlan De Perez et al. 2015) can increase resilience, reduce mortality, and reduce recovery costs (UNICEF and WFP 2015), especially in low-income countries with poor or no alternative solutions for flood protection (Golnaraghi 2012). The success of forecast-triggered strategies relies on forecasting systems able to produce accurate predictions with sufficient lead time to help flood risk managers act in a timely manner (UNICEF and WFP 2015), especially in low-income countries where longer preparation times might be needed (Bazo et al. 2018). Flash flood forecasting systems are developed at regional (Speight et al. 2018; Corral et al. 2019; Ibarreche et al. 2020; Ramos Filho et al. 2021; Shuvo et al. 2021; Georgakakos et al. 2021), national (Javelle et al. 2016; Liu et al. 2018), and continental scale (Raynaud et al. 2015; Park et al. 2019). Nonetheless, flash floods remain one of the most difficult types of flood to predict, with high levels of uncertainty in the overall forecasting process (Zanchetta and Coulibaly 2020) as well as a chronic lack of historical data of both occurrence and impact (Lowrie et al. 2022). Flash floods are rapid-onset events, typically occurring in small areas within minutes or few hours after a torrential triggering-rainfall event. Their limited predictability is linked to the challenge of predicting extreme localized rainfall accurately (Golding et al. 2016), and representing in detail hydrological factors such as topography, soil conditions, and terrain coverage that can modulate the occurrence and severity of flash floods (Xing et al. 2019; Kastridis and Stathis 2020). While improvements in flash flood forecasting are occurring, sufficient challenges remain. For example, km-scale numerical weather prediction (NWP) models have improved the skill of short-range forecasts for localized rainfall (up to 2 days ahead), issues related to extending the lead time of reliable predictions beyond day 3 are still not fully resolved. Furthermore, the need for high quantities of observational data and high computational costs can often be a limiting factor for the development of operational in-house flash flood forecasting systems for low-income and data-poor countries (Zanchetta and Coulibaly 2020).

ecPoint is a statistical post-processing technique that transforms global, grid-based forecasts into probabilistic point-scale forecasts (Hewson and Pillosu 2021). For rainfall, in particular for extremes, Hewson & Pillosu (2021) have shown that, against point verification, ecPoint provides more reliable and skilful forecasts than ECMWF’s ensemble (ENS) up to medium-range lead times (i.e., day 10). Furthermore, ecPoint provides global point-scale forecasts at a fraction of the cost of running a global km-scale NWP model. These features make ecPoint a good candidate for testing it as rainfall input in flash flood forecasting systems. Ecuador’s high susceptibility to flash flooding and its hydro-climatological diversity make the country a good test bed for the evaluation of ecPoint’s performance in the identification of areas at risk of flash flooding in diverse hydro-climatological scenarios. In addition, a well-documented flash flood database has been recently developed in Ecuador (Kruczkiewicz et al. 2021b) which allows for the analysis to be conducted. This article assesses the performance of ecPoint-Rainfall forecasts in the identification of areas at flash flood risk in Ecuador, in particular in terms of forecast accuracy and lead-time extension to medium-ranges. It also seeks to propose a methodology to define extreme rainfall events that can potentially generate flash floods using historical flash flood reports and short-range ecPoint-Rainfall forecasts when there are no adequate rainfall observations to define rainfall climatologies. Since Ecuador is working on developing a flash flood forecasting system with national level, this study could contribute at increasing the lead times of the flash flood predictions up to medium-ranges (i.e. up to day 5 to day 10), and at correcting biases due to rainfall’s diurnal cycles. Furthermore, Ecuador Red Cross could also benefit from improved flash flood forecasts for decision making and anticipatory actions.

# Background: geography, rainfall climatology and flooding in Ecuador

Continental Ecuador is located in north-western South America within 1.60 °N - 3.50 °S and 81.10°W - 75.28 °W (small box in **Figure 1a**). Ecuador’s Galápagos islands are located in the Pacific Ocean, 1000 km to the west of the mainland (insert in **Figure 1a**), but will not be included in this study. Thus, “continental Ecuador” will be referred to only as “Ecuador” from now on. The Andes run north to south through Ecuador (**Figure 1a**), and splits the country into three main geographic regions (Vuille et al. 2000): "La Costa” comprises the coastal plains along the Pacific Ocean (with hills that, on average, do not exceed 300 m above sea level), and the Andes’ western slopes; "El Oriente" covers a plateau containing 2% of the Amazon basin, and the Andes’ eastern slopes; "La Sierra", contains the inter-Andean region between the western and the eastern slopes of the Andes.

The rainy season in "La Costa" spans from December to May, and its intensity can be modulated by broader scale atmospheric and oceanic phenomena. The extreme phases of El Niño Southern Oscillation (ENSO) known as El Niño (i.e., above-average SST in the Pacific Ocean) and La Niña (i.e., below-average SST), respectively, enhance and decrease the average rainfall during the rainy season (Recalde-Coronel et al. 2014; Tobar and Wyseure 2018). In addition, the Madden-Julian Oscillation phases 1 and 8 (i.e., when the convection centre is found over the Western Hemisphere and Africa), and 4 and 5 (i.e., when the convection centre is found over the Maritime Continent) are associated, respectively, with an enhancement and decrease of precipitation (MJO, Recalde-Coronel *et al.*, 2020). “La Sierra” has a complex spatial precipitation pattern, with convective activity in the inter-Andean valleys being influenced by oceanic and continental air masses. “La Sierra” has two main rainy seasons (February to May and October-November), and as air masses lose much of their humidity on both flanks of the Andes, precipitation amounts in the inter-Andean region are rather low, 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). The rainfall climatology in “El Oriente” is considered to be 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.

Flooding in “La Costa”, especially during El Niño events, can cause considerable material loss and deaths. Prolonged rainfall events (from 1 to multiple days of continuous rainfall) can make rivers to burst their banks, and flood extensive plain areas, which also tend to be the most heavily populated (**Figure 1c**). Such rainfall events can also cause severe surface runoff in areas far from rivers due to the saturation of the ground (Galarza-Villamar et al. 2018). Rivers in “La Sierra” are highly sensitive to extreme localized rainfall events, and consequently are prone to flash flooding (Laraque et al. 2009; Pinos and Timbe 2020). “El Oriente”, as other parts of the Amazonian basin, displays a large lag between a single-rainfall-event-peak and river discharge peaks due to the size and length of Amazonian rivers, flood plain storage and relative interactions, and rivers generally having a shallow bed and topographical slopes, with relatively slow moving waters (Trigg et al. 2009). River flows show instead a much stronger response to seasonal rainfall patterns as opposed to single rainfall events (Trigg et al. 2009).

# Data

## Flash flood reports

​​ Like in many regions of the world, the disaggregation by flood types and specific documentation about historical flash flood events and their impacts is rare. This is mostly due to a lack of commonly accepted flash flood definition (Kruczkiewicz et al. 2021a) but also in part driven by differences in both definitions of flood types and in risk perception for various flood types that exist within a given area, which is the case in Ecuador (Bucherie et al. 2022). 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 estimation 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 compiled from two main datasets, DesInventar (UNDRR 2021) and the Ecuadorian Secretariat for Disaster Management (SNGRE), for a total of 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, an historical dataset of occurrences and impacts of floods, with specific information about the likelihood of events to be flash floods is available (Bucherie et al. 2021). One should notice that, while this dataset is the best attempt to address flash flood historical occurrence in Ecuador, it is important to note that it is based on disaster reporting processes made on the ground, not systematically collected through time. As a result, it can present gaps, inconsistent descriptions of the flood processes in time, as well as uncertainty in the geolocalization.

This study considered flood reports from 2019 to define the climatology of rainfall events associated with flash flood events, and 2020 to run the verification analysis. For flood reports in both 2019 and 2020, three EFFCI thresholds were considered to evaluate the impact of the uncertainty around a flood report being a flash flood event: EFFCI>=1 (i.e., all flood reports are considered, including not flash floods), EFFCI>=6 (i.e., only flood reports that are likely to be flash floods are considered), and EFFCI>=10 (i.e., only flood reports that are extremely likely to be flash floods are considered). **Table 2** shows the number of flood reports in 2019 and 2020, and how many reports were excluded because they did not have a reporting time. **Table 3** shows the number of flood reports divided per year and EFFCI threshold, while **Figure 2** shows their spatial distribution. From the analysis of **Table 3** and **Figure 2**, it was decided to not include “El Oriente” in the analysis as it does not have many flood reports.

## Rainfall forecasts: ECMWF ENS and ecPoint

ECMWF ENS consists of one control run which is started from the best possible representation of unperturbed initial conditions, and fifty perturbed members which are started from perturbed initial conditions (using singular vectors and a data assimilation ensemble) and stochastic model uncertainties (Buizza 2019). Up to day 15, ENS native resolution consists in an octahedral reduced-Gaussian grid with 640 latitude lines between the pole and the equator (O640), which corresponds to ~18 km spatial resolution at the equator (Owens & Hewson, 2018). The analysis period that in this study uses ENS forecasts goes from 1st January to 31st December 2020. Over this analysis, two different NWP versions (known as cycles) were running operationally at ECMWF: 46r1[[3]](#footnote-4) between 1st January and 6th July, and 47r1[[4]](#footnote-5) between 7th July and 31st December. While verification has shown that rainfall forecasts from cycle 47r1 are better than 46r1 (Haiden et al. 2021), such differences are not expected to condition the results in this study.

Hewson & Pillosu (2021) have developed a statistical post-processing technique, called ecPoint, that helps to address the two main factors that affect the performance of global NWP models forecasting extreme localized rainfall: systematic biases (Lavers *et al.*, 2021) and lack of information on forecasts’ sub-grid variability (Göber *et al.*, 2008). A one-year period of global verification shows that ecPoint-Rainfall (i.e., the family of ecPoint-related products that post-processes rainfall forecasts) can produce more reliable and skilful rainfall predictions than the raw ENS for point verification, especially in case of extreme events (Hewson and Pillosu 2021). Currently, ecPoint-Rainfall generates 100 new ensemble members for each raw ENS member, producing a total of new 5100 post-processed members per grid-box. They are then distilled in percentiles from 1st to 99th, and provided in the same native grid of ENS forecasts.

# Methods

## Verification analysis

The two main attributes of any probabilistic forecast are reliability and discrimination, and together these determine the usefulness of a probabilistic forecasting system (Jolliffe and Stephenson 2011). When verifying binary events (i.e., yes- and no-event) for which probabilities of occurrence cannot be computed, the only attribute of the forecasting system that can be verified is the discrimination. In this study, the Relative Operating Characteristic (ROC) curve and the Area Under the Roc Curve (AURC) are typically used as a summary measure of the discrimination attribute of a forecasting system (Jolliffe and Stephenson 2011). The calculation of the ROC curve involves sampling the probability distribution for a given rainfall threshold. For each probability threshold, a 2X2 contingency table can be defined (see **Table 1** for an example of a contingency table and the definition of its four elements). The ROC curve is then represented as a graph that maps hit rates (HR, in the Y-axis) against false alarm rates (FAR, in the X-axis), being them defined as follows:

|  |  |
| --- | --- |
| HR = H / (H+M) | (1) |
| FAR = FA / (FA+CN) | (2) |

Values of HRs and FARs lie between 0 and 1, so the ROC curve is defined within a unit square. The location of the ROC curve in the unit square is determined by the intrinsic discrimination capacity of the forecasting system. Perfect discrimination 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). This implies that HRs grow at a much faster rate than FARs. Such condition is represented by an AURC equal to 1. A ROC curve that lies on the diagonal of the unit square represents a condition in which the forecasts have zero skill because HRs and FARs grow at the same rate, and forecasts provide no added information. Such condition is represented by an AURC equal to 0.5.

The following choices were made to compute the ROC curves and the AURC. Following Hamill and Juras (2006), instead of using fixed rainfall values as rainfall thresholds, they are expressed in relative terms (e.g. quantiles of a climatology) for a consistent analysis across different climatic zones. The methodology used to create the rainfall thresholds is described in section 0. Following Bouallegue and Richardson (2021), in order to ensure ROC curves are as complete as possible, they are built using the maximum available discretization by computing the probability thresholds from each single member exceeding the rainfall threshold rather than use fixed percentage bins (Bouallegue and Richardson 2021). This means that ROC curves for ecPoint and ENS are built with 99 and 51 points, respectively. No curve fitting is adopted to close the ROC curves. A straight line between the last meaningful point of the ROC curve and the top-right corner is drawn to close the ROC curve and compute the AURC. With such approach, the AURC is underestimated compared to fitting a smooth curve over the ROC (Wilson 2000; Bouallegue and Richardson 2021). However, as this study focuses on the relative comparison of the AURC between two forecasting systems (i.e., ecPoint and ENS), such underestimation is not of a concern as it is the same for both systems. In addition, closing the ROC curve with a straight line allows to compare ecPoint and ENS considering their current operational configurations (Bouallegue and Richardson 2021), which will allow to evaluate whether having a bigger ensemble to forecast extreme rainfall, and consequently areas at risk of flash floods is important or not. Finally, the AURC is computed using the trapezoidal approximation (i.e., the area under the ROC curve is estimated considering straight lines between two consecutive points in the plot, and the area is equal to the sum of the areas of the single trapeziums). To test the significance of the differences between the AURC between ecPoint and ENS, a non-parametric bootstrapping technique with replacement and with 10000 replicates was adopted. A 95% confidence interval was considered.

The population of the 2X2 contingency table is the challenge of this verification analysis because we are not using stationary observations (e.g., gauges) that provide information about both yes- and non-events, allowing us to populate the four quadrants of the contingency table (**Table 1**). When observations are provided as reports, as they are in our study, they are typically received only at the events' location so reports provide information only about the yes-events. It is not possible to answer the questions: if there are no reports in an area, is it because an event happened, but nobody reported, or was there no event to report? Studies such as Robbins and Titley (2018) verify only the yes-events with the caveat that only quadrants I (i.e., hits) and III (i.e., misses) of the contingency table can be populated. Following the approach of Tsonevsky et al. (2018) for the verification of severe convection events using reports from the European Severe Weather Database (ESWD), it is possible to populate the four quadrants of the contingency table (**Table 1**), and provide an estimation of both HRs and FARs for the forecasting systems under evaluation. This aspect of the verification analysis is important because risk-averse users are as much interested in the estimation of false alarms as much as hits and misses.

In order to estimate the non-events, a no-report is considered as a non-event in this study. To allow a comparison between observations and forecast, a mask of the full domain of interest must be created. The mask used in this study (**Figure 1b**) is based on the ENS (and ecPoint) native grid. The mask covers all continental Ecuador, including the borders with Colombia and Peru and the Pacific coast, counting 1090 grid-boxes in total. “La Costa”, “La Sierra” and “El Oriente” are defined based on the topography of Ecuador as represented in the ENS. Grid-boxesto the left and to the right of the longitude 78.2 °W and below 600 m above sea level belong to “La Costa” (321 grid-boxes) and “El Oriente” (299 grid-boxes), respectively. Grid-boxes above 600 m belong to “La Sierra” (470 grid-boxes). The mask is then used to build observational fields that match the valid accumulation periods of the forecast fields i.e., 12-hourly accumulation periods ending at 0, 6, 12 and 18 UTC for each day in the verification analysis, which in this study goes from 1st January to 31st December 2020). In the observational field for a specific 12-hourly accumulation period, grid-boxes are assigned the value 1 if they contain at least one flood report: otherwise, they are assigned the value 0. Forecast fields are instead built by assigning to each grid-box the value 1 if a specific rainfall event is exceeded with a certain probability threshold; otherwise, it is assigned the value 0. The 2X2 contingency table is then built by examining the values of overlapping grid-boxes in the observational and forecast fields. When both grid-boxes are assigned the value 1 or 0, such instance is counted as a hit or correct negative, respectively. When a grid-box in the observational field has the value 1, and the correspondent grid-box in the forecast field has the value 0, such instance is counted as a missed. It is counted as a false alarm if it happens vice versa. From the population of the 2X2 contingency table is then possible to calculate the HRs and FARs, and build the ROC for a specific rainfall threshold.

## Defining verifying rainfall events

In this study, a “verifying rainfall event (VRE)” is a rainfall threshold that, if exceeded, is likely to generate flash floods. It is then used as a baseline to assess whether ecPoint can predict such rainfall events and identify areas at flash flood risk. There are two main ways to determine the value of VREs. The first way is to examine rainfall observations (e.g., rain gauges, radar) near flash flood reports to determine the average extreme rainfall values likely to lead to flash floods. This method benefits from defining VREs that represent the amounts of localised rainfall extremes that generate flash floods. In the absence of local rainfall observations associated with flash flood reports, it is also common practice to define VREs from rainfall climatologies obtained from model gridded rainfall products such as ERA5 reanalysis (Hersbach et al. 2020) and reforecasts (Hamill et al. 2006), or from gridded rainfall products obtained from the blending of different types of rainfall observations (e.g. gauges, radar, satellite) such as MSWEP (Beck et al. 2019), or GPCP (Adler et al. 2018). However, due to the coarse resolution of these rainfall products, local rainfall totals might be underestimated (Tapiador et al. 2019). The first way of computing VREs is preferred in this study because its output is more compatible with what ecPoint-Rainfall aims to predict, i.e., point rainfall totals.

To define VREs using the first method, high-density rainfall observations, in space and time, are required to have a greater chance to capture time and location of rainfall events that lead to flash floods (Haiden and Duffy, 2016). Due to the absence in Ecuador of high-density, in-situ 12-hourly rainfall observations, day 1 ecPoint-Rainfall forecasts are used as a proxy for rainfall observations to create a synthetic climatology of rainfall events associated with flash floods. At such short-range, the 99 rainfall values provided by ecPoint-Rainfall at each ENS grid-box are representative of rainfall sub-grid variability and can be thought of as point rainfall observations within each ENS grid-box (Hewson and Pillosu 2021). Both 0 and 12 UTC runs are considered to increase the sample size. ecPoint-Rainfall forecasts are valid for four 12-hourly overlapping accumulation periods ending at 0, 6, 12, and 18 UTC. To increase the sample size, all four overlapping accumulation periods were used in the analysis, which means that each flood report in the database was matched with two 12-hourly accumulation periods that best spanned the event (**Table 4**). Each flood report was assigned a distribution of 396 rainfall values from ecPoint, and 204 rainfall values from ENS. This methodological approach should not be seen as a disadvantage for the verification of the ENS forecasts as it is similar to what is generally done in rainfall verification, where rainfall forecasts from an NWP model are typically verified against in-situ observations instead of rainfall totals averaged over the NWP grid-box. Indeed, this approach allows for analysing how well the NWP model can represent local rainfall totals.

From the rainfall distribution assigned to each flood report, a representative extreme rainfall value, provided as an Xth percentile of such distribution, was extracted. This approach was adopted because, physically, the smallest rainfall values in the distribution are unlikely the driver of any flash flood event. The 50th, 75th, 85th, 90th, 95th, 98th, and 99th percentiles were considered in this study to represent the average rainfall that might lead to flash floods. Separate distributions were computed for "La Costa" and "La Sierra" to capture possible differences in the distribution of rainfall values associated with flash flood events in these two regions. A distribution of the average rainfall totals associated with the different Xth percentiles considered was created for all flood reports (**Figure 3**). The VRE was determined from such distribution based on how many flood reports one wants to retain in the analysis. Percentiles from such distribution depict the number of retained flood reports. This aspect is important because the higher the percentile, the more extreme the flash floods that defined the VRE are. The 25th percentile was considered in this study, meaning the top 75% of flood events were retained in the definition of VREs.

# Results

## Definition of verifying rainfall events

This study proposes a methodology to define the value of rainfall events that can potentially lead to flash floods and help to define verifying rainfall events for verification purposes. The method uses historical flash flood reports and short-range ecPoint-Rainfall forecasts as rainfall observation proxies to create a synthetic rainfall climatology. **Figure 3** presents the distribution of average rainfall events associated with flash floods in 2019 in “La Costa” and “La Sierra”. Overall, the distribution spread in “La Costa” (**Figure 3a**, **b, and c**) is bigger than in “La Sierra” (**Figure 3d**, **e, and f**), where the distribution is more vertical. This means that the rainfall events that can cause flash floods in “La Costa” are much more varied than in “La Sierra”. There are also hints that in both regions, “La Costa” and “La Sierra”, the spread of the distribution of rainfall totals that cause floods that are likely to be flash floods (i.e., EFFCI>=6 in **Figure 3b** and **e**) is slightly bigger than the distribution of rainfall totals that cause all types of floods (i.e., EFFCI>=1 in **Figure 3a** and **d**). Such difference appears to be even bigger for flood events that are extremely likely to be flash floods (i.e., EFFCI>=10 in **Figure 3a** and **f**). However, due to the small number of events in this category (**Table 3**), such difference cannot be considered significant as it might be due only to statistical sampling.

In both regions and for all EFFCI thresholds (see zoomed in panels in **Figure 3**), the distributions built from ecPoint forecasts tend to be to the left of the distribution built from ENS forecasts for small average rainfall events, to later cross to the right for bigger average rainfall events. In “La Costa”, the crossing point occurs for bigger average rainfall events (i.e., the 95th percentile, distributions in blue in **Figure 3a**, **b**, **c**) than in “La Sierra” (i.e., the 90th percentile, distributions in cyan in **Figure 3d**, **e**, **f**). The bigger crossing point for “La Costa” is indicative that the difference between ENS and ecPoint’s mean rainfall during flash floods is bigger than in “La Sierra”, meaning that bigger bias corrections are operated by ecPoint in “La Costa”. In addition, when ecPoint’s distribution is to the right of the ENS’s distribution, the distance between both lines is bigger in “La Sierra” than in “La Costa” indicating that there is bigger variability of rainfall totals between ENS and ecPoint in “La Sierra”. In the authors’ experience, this result reassures that ecPoint is providing a physically compatible result to what rainfall observations would typically provide. Rainfall tends to be caused by larger-scale convective phenomena such as “El Niño” (Tobar and Wyseure 2018) and MJO (Recalde-Coronel et al. 2020) in “La Costa”, while smaller-scale convective systems are the main cause of rainfall in “La Sierra” (Recalde-Coronel et al. 2014). Coarse resolution, global NWP models tend to represent better rainfall distributions caused by large-scale phenomena, and local peaks might not be well represented only if they are caused by the interaction of such large-scale systems with complex orography which are not well represented in coarse resolution NWP models. For this reason, the distribution of large average rainfall events for both ENS (continuous fuchsia and orange lines in **Figure 3a**, **b, and c**) and ecPoint (dashed fuchsia and orange lines in **Figure 3a**, **b, and c**) are not as distant from each other as in “La Sierra”. The ecPoint’s distribution for large average rainfall events (dashed fuchsia and orange lines in **Figure 3d**, **e, and f**) is well to the right of the ENS’s distribution because the smaller-scale convective systems in “La Sierra” can generate very high local peaks surrounded by very small totals. Such rainfall distribution is not well represented in coarse resolution NWP models, which tend to overestimate the smaller rainfall values and largely underestimate the local peaks. ecPoint corrects for this misrepresentation of local rainfall totals in coarse resolution NWP models.

The VREs used in this study are provided by the distribution of average rainfall events that correspond to the 85th and the 99th percentile of the ecPoint-Rainfall forecasts realizations for each flood report with EFFCI>=6 (purple and orange distributions in **Figure 3b** and **e**, respectively). The VRE values are obtained retaining the top 75% percent of all flood reports (i.e., the 25th percentile of the purple and orange distributions in **Figure 3b** and **e**, respectively, and indicated by the grey line in their correspondent zoomed in panels), and are shown in **Table 4**.

## ecPoint performance in the identification of areas at flash flood risk

AURCs denote the discrimination ability of a forecasting system for a specific VRE under consideration, and for a specific forecast lead time. **Figure 4** shows the trend with lead time of AURC for “La Costa” and “La Sierra”, computed for flood reports with EFFCI>=6, and for VRE = 85th and 99th percentile (in mm).

“La Costa” shows no degradation with lead time of both ENS and ecPoint’s AURCs (**Figure 4a** and **b**), while the AURCs in “La Sierra” tend to diminish with lead time (**Figure 4c** and **d**). In addition, ecPoint’s AURC appears to diminish with lead time with a similar rate of ENS’s AURC for both VREs. The stable/degraded performance with lead time of ENS, and consequently ecPoint, in “La Costa”/”La Sierra” might lie on how well those phenomena that generate rainfall in the two regions are predicted. Rainfall in “La Costa” tends to be generated 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 week 6 in ENS (Haiden et al. 2021). Therefore, no significant degradation in ENS and ecPoint forecasts should be expected to be observed up to day 10. Rainfall in “La Sierra” tends to be linked instead to smaller-scale convective systems (Recalde-Coronel et al. 2014), whose predictability diminishes relatively quickly after only few days in ENS forecasts (Haiden et al. 2021).

ENS and ecPoint’s AURC are almost identical for VREs = 85th percentile in “La Costa” (**Figure 4a**) and in “La Sierra” (**Figure 4c**), in particular when error bars are considered (i.e., error bars for ENS and ecPoint tend to fully overlap). The difference between ENS and ecPoint’s AURC increases for higher thresholds (i.e., VREs = 99th percentile, **Figure 4b** and **d**), with ecPoint showing more significantly bigger AURC than ENS overall in both regions. Therefore, when using small rainfall thresholds, a similar performance in the identification of areas at flash flood risk should be expected for the two forecasting systems. Instead, ecPoint shows a better discrimination ability than ENS in the identification of areas at flash flood risk when considering high rainfall thresholds.

Overall, ENS and ecPoint’s AURCs for VREs = 85th percentile (**Figure 4a** and **c**) are bigger than the corresponding AURCs for VREs = 99th percentile (**Figure 4b** and **d**). Such decrease does not necessarily imply a decrease in the intrinsic discrimination ability of the forecasting systems. It only means that the real configuration of the forecasting systems (i.e., 50 ensemble members for ENS and 99 for ecPoint) are not able to discriminate extreme rainfall events (i.e., VREs = 99th percentile) as well as less extreme rainfall events (i.e., VREs = 85th percentile). As a result of the lower number of events exceeding larger VREs, the points on the ROC curves move towards the bottom left of the unit square, and the correspondent AURC computed with the trapezoidal approximation diminishes (Gneiting and Vogel 2021; Bouallegue and Richardson 2021). According to Bouallegue and Richardson (2021), it is possible to identify the intrinsic discrimination ability of the forecasting systems for extreme events by adopting a bi-normal fitting model to compute the AURCs, but this is out of the scope of this study. Such intrinsic discrimination capability is theoretical, and the authors of this study are interested in identifying the discrimination abilities of the real configurations of the forecasting systems as those are the ones used by end-users.

It is interesting to notice that ENS’s AURC for VREs >= 85th percentile is bigger than ecPoint’s AURC for VREs >= 99th percentile in both regions. To examine in more detail the different behaviour between ENS and ecPoint, let us consider the ROC curves for the VRE = 85th percentile for “La Sierra” for ENS (**Figure 6a**, dashed red line) and for ecPoint (**Figure 6a**, dashed blue line). Although ecPoint’s AURC is overall smaller than ENS’s, all points in the curve contribute towards the AURC. In the ENS’s ROC curve for the VRE = 99th percentile (**Figure 6b**, dashed red line) only the points at the bottom left of the unit square contribute to the AURC; the points that correspond to the small probabilities of exceeding the VRE (i.e., <10%, that correspond to last five points closer to the top right corner of the unit square) are lined up in a straight line, not providing any contribution to the AURC. Instead, the bigger contribution to AURC for ecPoint (**Figure 6b**, dashed blue line) is provided by the small probabilities of exceeding the VRE (i.e., <10%, that correspond to last 10 points closer to the top right corner of the unit square). This means that, although ENS’s AURC for the VRE = 85th percentile is bigger than ecPoint’s AURC for the VRE >= 99th percentile, the skill in identifying areas at flash flood risk comes from different sources i.e., small probabilities of exceeding high rainfall totals for ecPoint and high probabilities of exceeding small rainfall totals for ENS.

A feature that stands out in all panels in **Figure 4** is the sinusoidal pattern shown by the AURC. This is likely due to the strong rainfall diurnal cycle present in Ecuador, shown by ENS and ecPoint behaviour of rainfall’s annual mean in “La Costa” and “La Sierra” during different accumulation periods (**Figure 7**), and confirmed by Kikuchi and Wang (2008), whose study is based on rainfall observations. For both VREs in “La Costa”, AURC peaks are observed for both ENS and ecPoint between 0000-1200 LT (i.e., lead time steps labelled in purple in **Figure 4a** and **Figure 4b**), which corresponds to the accumulation period with the second smallest rainfall totals in the day (i.e., 12-hourly accumulation period labelled in purple in **Figure 7**, continuous lines); troughs are mostly observed between 1200-0000 LT (i.e., lead time steps labelled in fuchsia in **Figure 4a** and **Figure 4b**), which corresponds to the accumulation period with the second highest rainfall totals in the day (i.e., 12-hourly accumulation period labelled in fuchsia in **Figure 7**, continuous lines). The amplitude between peaks and troughs increases with increasing VREs (**Figure 4a** and **Figure 4b**). However, such amplitude is deeper for ENS than ecPoint, which manages to improve (i.e., increase) the AURC 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). Substantial improvements are also observed for intermediate intervals. A similar behaviour with lead time for AURC is observed in “La Sierra”. Although a 6- to 12-hour shift is observed between peaks and troughs compared to “La Costa”. Peaks are observed in the accumulation period between 0600-1800 LT up to day 4 forecasts, and between 1200-0000 LT afterwards (i.e., lead time steps labelled in cyan and fuchsia, respectively, in **Figure 4c** and **Figure 4d**), which correspond to the accumulation periods with the highest rainfall totals during the day (i.e., 12-hourly accumulation period labelled in cyan and fuchsia, respectively, in **Figure 7**, dashed lines). Troughs are observed instead between 1800-0600 LT up to day 4, and between 0000-1200 LT afterwards (i.e., lead time steps labelled in green and purple, respectively, in **Figure 4c** and **Figure 4d**), which correspond to the accumulation periods with the smallest rainfall totals during the day (i.e., 12-hourly accumulation period labelled in green and purple, respectively, in **Figure 7**, dashed lines). Unlike in “La Costa”, where ecPoint’s added value is observed at a specific time of the day, ecPoint’s added value in “La Sierra” is observed in all accumulation periods.

The source of the correction is also different in the two regions as shown by the example ENS and ecPoint ROC curves in “La Costa” and “La Sierra” (**Figure 6b**). In “La Costa” (continuous lines), ENS and ecPoint’s ROC curves may belong to the same underlying curve as they are overlapping with the exception of a couple of points, so the underlying forecast information content in the ENS forecasts might be similar by ecPoint, which would mean that the added value provided by the post-processing lies in the tail (Bouallegue and Richardson 2021). Consequently, ENS will provide a very similar ability to ecPoint in identifying areas at risk of flash flood, with the advantage of using ecPoint being seen when comparing ENS and ecPoint rainfall forecasts with rainfall observations, as ecPoint is more likely to provide better guidance on the actual extreme observed rainfall totals than ENS. In “La Sierra” (dashed lines in **Figure 6b**), the points in the blue line (i.e., ecPoint ROC curve) are distinct from the points in the red line (i.e., ENS ROC curve), and are closer to the top left corner of the unit square. This is a typical result expected from an increase in forecast predictive skill of extreme rainfall events (Bouallegue and Richardson 2021).

The trend with lead time of AURC was computed also for flood reports with EFFCI>=1 and 10, and for VREs >= 50th, 75th, 90th, 95th and 98th percentiles (not shown). The trends for all these cases are similar to what shown in **Figure 4**. 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 3**. For increasing VREs, a decreasing number of events exceeding the VRE is most likely the reason why the trend with lead time of AURC appears noisier.

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

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[[5]](#footnote-6). March 8th was one of the wettest days (**Figure 8a**). Major impacts resulted mainly in the highly populated city of Guayaquil, where very heavy rainfall was reported to occur in the afternoon after 4 pm (LT)[[6]](#footnote-7), with rainfall totals exceeding 100 mm/24h in the city centre (zoomed red area in **Figure 8a**)[[7]](#footnote-8). Around March 8th, 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[[8]](#footnote-9). On the day, from dawn, a numerical model sounding from ECMWF (**Figure 8b**) looked particularly conducive to flash flood activity, denoting a very high CAPE (i.e., convective available potential energy) potential once insolation got to work, a sufficiently high dewpoint depression that suggests insolation-based triggering would not be impeded by thick cloud, a potential for very high altitude convective cloud tops, a wind shear that favours prolonged convective cell life-cycles (as down-draughts would not interfere with up-draughts), and relatively light steering winds favouring slow movement of those long lived convective cells. This is supported by SYNOP and METAR observations, together with satellite imagery, that suggest that the cause of the rainfall was organized convective cells whose development was triggered by insolation.

**Figure 8c** shows day 1 (first row), day 3 (second row), and day 5 (third row) forecasts for the 00 UTC runs of ENS and ecPoint. 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 8a** when the majority of the rainfall fell. The 50th (first and second column), 95th (third and fourth column), and 99th percentile (fifth and sixth column) are shown. Besides the wavy behaviour that can be observed in both forecasting systems, caused by the difficulty of representing steep slopes in spectral form, ecPoint provides a smoother view of the rainfall totals. This is expected in ecPoint as it intrinsically smooths out the percentile picture provided by the raw forecasts (Hewson and Pillosu 2021), providing a more useful and complete guidance for forecasters on the areas at risk of heavy rainfall. The median (i.e., the 50th percentile), which should be the dividing line for equi-probable observation categories, is systematically smaller in ecPoint than in ENS. This is due to ecPoint’s rainfall overprediction bias correction at the grid scale and also to its asymmetric correction for sub-grid variability (for the grid-box weather types characteristic of this case study). ecPoint highlights, in most areas, that there is the potential of having higher local totals than ENS (see the 95th and the 99th percentile in **Figure 8c**), while ENS overestimates the mean rainfall overall (see the 50th percentile in **Figure 8c**). The heavy rainfall events in “La Costa” appear better predicted than those in “La Sierra” by ENS. Although for luck, a point with rainfall between 50-60 mm/24h was observed in “La Sierra”. ENS does not reach such values even in day one (the maximum rainfall total expected is up to 30 mm/12h). On the other hand, ecPoint shows that reaching such total is possible, although with very small chance (i.e., 1%). In “La Costa”, both the 95th and the 99th percentiles from ENS and ecPoint identify similar areas at flash flood risk (i.e., Guayaquil and the north-west coast). While the 95th percentile for ENS and ecPoint appear very similar (confirming the result obtained in **Figure 3**), only ecPoint’s 99th percentile appears to provide a better guidance on the extreme rainfall totals. While there are not enough rainfall observations to conclude whether ecPoint’s 99th percentile (which should be observed on average in 1 observation over 100) overestimated the rainfall in “La Costa”, one can notice the benefits of having access to higher density observations in Guayaquil which allows for us to say something more meaningful about the behaviour of ecPoint. From 80 mm/24h in the SYNOP observations, the reports from INHAMI go up to 150 mm/24h (shown in both ENS and ecPoint). ecPoint’s 99th percentile reached values up to 300 mm/12h, which might have happened but were not observed.

# Discussions

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

## On the need to improve the collection of flash flood reports to better develop and verify flash flood forecasting systems

This study provides information on how a new rainfall product (ecPoint) performs in the identification of areas at flash flood risk using a verification method over a long period, i.e., one year instead of looking only at case studies (c.f.Raynaud et al. (2015)), and using flash flood reports as observational data, instead of rainfall observations (c.f. Park et al. (2019)). Developers and scientists were forced to adopt such sub-optimal approaches due to the lack of flash flood databases with good spatial/temporal coverage (Gaume et al. 2009; Kruczkiewicz et al. 2021b), and because more detailed information might be available for single flash flood events and spatial/time coverage of rainfall observations can be much better in quantity and quality (Haiden and Duffy 2016; Zhang et al. 2011). Flood databases tend to not contain enough information on flash floods to clearly show the value of flash flood forecasting systems. This study benefited from the use of a newly developed database in Ecuador (Kruczkiewicz et al. 2021b), which allowed us to draw interesting conclusions on the performance of ecPoint in the identification of areas at flash flood risk. Not only did this study benefit by the higher density of flash flood reports that one would typically get from other databases which, by their nature, tend to report only internationally reported large, widespread events (e.g., EM-DAT or FloodList), but it also benefited by the more reliable quality of the attributes of each flood report (e.g., events’ location and reporting time) thanks to the development of such database in close collaboration with local authorities which allowed us to verify such information. Such information allowed us to develop a verification study that shows the performance of a new forecasting system over a large territory and the difference between different climatological regions as well as different times of the day. It also allowed us to develop a methodology using short-range ecPoint-Rainfall forecasts to build rainfall climatologies that define the average rainfall events associated with flash flood events. Such methodology can be used for verification (as in this study) but also for the development of a warning system, which uses such average rainfall events to produce warnings when the forecasts exceed such totals. While the former application could be carried out by the development of a climatology that is compatible with ecPoint’s resolution (e.g., post-processing with the ecPoint methodology model climatologies built from ERA 5 or reforecasts), a continuous improvement of flash flood databases remains the only way to improve the verification of any flash flood forecasting system.

## On the benefit of extending flash flood forecasts for humanitarian applications

# Conclusions

This study aimed to assess the performance of ecPoint-Rainfall forecasts in the identification of areas at flash flood risk in Ecuador, in particular in terms of forecast accuracy and lead-time extension to medium-ranges. The results suggest that ecPoint outperforms ENS in areas where rainfall is originated mainly from small-scale convective systems. Where rainfall is originated from large-scale convective systems, ecPoint and ENS skills are comparable.

To the best knowledge of the authors, 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. Therefore, this study provides a more precise picture of whatecPoint’s performance for flash flood forecasting could be.

Some limitations are worth noting. In the definition of the average rainfall events that might lead to flash floods, the use of short-range ecPoint-Rainfall forecasts can provide a larger distribution of rainfall totals than those defined by actual rainfall observations. In addition, such methodology depends on having good quality, medium-to-high resolution flash flood reports available. While this study claims that

* Summarize the key findings of the study
* Answer the research questions and aims
* Outline the study contributions
* Discuss limitations/weaknesses of the study
* Make recommendations for future research

Restatement of the results. How I did it (talk about the original method). What are the original insights of this study.

Repeat what the forecasting systems has shown. This is a better forecasting system because …

Repeat what is the method, what is different. This is a better method because …

Why all this matters.

Why should we care about the results.

What is the next for this research.

* ecPoint is a better forecasting system because…
* This study shows that forecasters could provide guidance on the areas at risk of flash floods using only ENS forecasts: indeed, ENS forecasts with high probabilities of having small rainfall totals provide a similar guidance for flash flood forecasting than ecPoint forecasts with small probabilities of having high rainfall totals. However, ecPoint provides a better reference in the verification against the rainfall totals that generated the flash flood event. This aspect would instil in forecasters more confidence in forecasts provided by ecPoint than ENS.
* Developing a flash flood forecasting system with ecPoint forecasts has the challenge that it requires a reference that is compatible with its resolution. If it is not possible to provide this through point-based rainfall observations (e.g., rain gauges), the coarse rainfall climatologies available today would not provide an acceptable compromise (underestimation of localized extremes). There is the need to post-process also rainfall climatologies in order to make them compatible. If this is done, forecasts could be provided with a continuous domain for the whole globe.
* FURTHER RESEARCH (MENTION AT THE END). Also, flash flood observations would not be needed for the implementation of the system. However, they would be needed for its verification. Ergo, there is still needed to keep developing flash flood databases like the one used in this study.
* A system that is made available with a continuous global domain, could be used by forecasters in projects such as Aristotle or FbF that do not know the local climatology, but could provide flash flood forecasts in different regions of the world that currently do not have any systems in place, or extend to medium ranges the forecasts in those countries where shorter lead time forecast are available.

# Tables

**Table 1** - 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 2** - Number of all flood reports in the database in 2019 and 2020 (i.e., flood reports with EFFCI>=1). The second column of the table shows the number of reports eliminated from the database because the reports did not contain the location of the report in lat/lon coordinates, or the reports did not contain the date and time in which the flood event happened.

|  |  |  |  |
| --- | --- | --- | --- |
| **Year** | **Total n. of flood reports** | **N. of eliminated flood reports** | **N. of retained flood reports** |
| 2019 | 300 | 3 | 297 |
| 2020 | 190 | 0 | 190 |

**Table 3** - Number of flood reports for 2019 and 2020, in each region, and for EFFCI>=1 (first group), EFFCI>=6 (second group), and EFFCI>=10 (third group). The percentages within parenthesis for the reports with EFFCI>=6 and EFFCI>=10 represent the reduction in percent of the number of flood reports compared to the number of flood reports with EFFCI>=1, which correspond to the total number of flood reports in the correspondent year (see **Table 2**).

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **N. of flood reports with EFFCI>=1** | | | | **N. of flood reports with EFFCI>=6** | | | | **N. of flood reports with EFFCI>=10** | | | |
| Region/  Year | Costa | Sierra | Oriente | All | Costa | Sierra | Oriente | All | Costa | Sierra | Oriente | All |
| 2019 | 175 | 114 | 8 | 297 | 92  (-47%) | 101  (-12%) | 0  (-100%) | 193  (-35%) | 18  (-90 %) | 33  (-71 %) | 0  (-100 %) | 51  (-83 %) |
| 2020 | 91 | 88 | 11 | 190 | 48  (-47%) | 77  (-12 %) | 2  (-82%) | 127  (-33%) | 22  (-76 %) | 26  (-70 %) | 0  (-100 %) | 48  (-75 %) |

**Table 4** - Day 1 ecPoint and ENS forecasts used to define the verifying rainfall events.

|  |  |
| --- | --- |
| **Valid 12-hourly periods (in UTC for day X) containing the reporting time of the flood reports** | **Correspondent 12-hourly accumulation periods for day 1 rainfall forecasts (forecast base date / forecast base time in UTC / lead time in hours)** |
| 18 (on day X-1) to 6 | Day(X-1) / 00 UTC / (t+18,t+30) |
| Day(X-1) / 12 UTC / (t+6,t+18) |
| 0 to 12 | Day(X) / 00 UTC / (t+0,t+12) |
| Day(X-1) / 12 UTC / (t+12,t+24) |
| 0 to 12 | Day(X) / 00 UTC / (t+0,t+12) |
| Day(X-1) / 12 UTC / (t+12,t+24) |
| 6 to 18 | Day(X) / 00 UTC / (t+6,t+18) |
| Day(X-1) / 12 UTC / (t+18,t+30) |
| 6 to 18 | Day(X) / 00 UTC / (t+6,t+18) |
| Day(X-1) / 12 UTC / (t+18,t+30) |
| 12 to 0 (on day X+1) | Day(X) / 00 UTC / (t+12,t+24) |
| Day(X) / 12 UTC / (t+0,t+12) |
| 12 to 0 (on day X+1) UTC | Day(X) / 00 UTC / (t+12,t+24) |
| Day(X) / 12 UTC / (t+0,t+12) |
| 18 to 6 (on day X+1) UTC | Day(X) / 00 UTC / (t+18,t+30) |
| Day(X) / 12 UTC / (t+6,t+18) |

**Table 5** – Verifying rainfall events (in mm/12h).

|  |  |  |
| --- | --- | --- |
|  | **85th percentile** | **99th percentile** |
| **La Costa** | 9.865 mm/12h | 50.452 mm/12h |
| **La Sierra** | 5.885 mm/12h | 25.551 mm/12h |

# Figures

Map

Description automatically generated

**Figure 1** - Panel (a) shows Ecuador’s orography, its political regions, and the location of Ecuador’s three main geographical regions: the coast (“La Costa”), the highlands (“La Sierra”), and the Amazon (“EL Oriente”). The small box shows Ecuador’s location (in red) within South America. Panel (b) shows the definition of the Ecuador’s three main geographical regions using the ENS and ecPoint grid (“La Costa”, “La Sierra” and “El Oriente” are represented in yellow, brown, and green, respectively) and the location of the ENS and ecPoint grid-boxes (black dots) within Ecuador’s domain. Panel (c) shows the population density (in people/km2) for each region from the 2020 census (data source: <https://es.wikipedia.org/wiki/Provincias_de_Ecuador>). Panel (d) lists the names of Ecuador’s political regions following the numbers indicated in panel (c).

Map

Description automatically generated

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

Graphical user interface, chart, diagram

Description automatically generated

**Figure 3** - Distributions of average rainfall events for 2019’s flood reports (FRs). Continuous and dashed lines represent the distributions for ENS and ecPoint, respectively. Panels (a), (b), (c), and (d), (e), (f) represent, respectively, the distributions for “La Costa” and “La Sierra”, where distributions in panels (a) and (d) were built with FRs with EFFCI>=1, distributions in panels (b) and (e) were built with FRs with EFFCI>=6, and distributions in panels (c) and (f) were built with FRs with EFFCI>=10. How “extreme” is the average rainfall event associated with flood events is defined by seven percentiles computed from the distribution of rainfall forecasts associated to each FR: the 50th (yellow), 75th (green), 85th (purple), 90th (cyan), 95th (blue), 98th (fuchsia), and 99th (orange) percentile. In the zoomed figures, the 25th percentile of the distributions is indicated in grey to stress that the values of the VREs were selected considering the top 75% flood events recorded in the database.

Chart

Description automatically generated

**Figure 4** - Areas under the ROC curve (AURC) for flood reports with EFFCI>=6. Panels (a) and (b), and (c) and (d) show, respectively, the AURC for “La Costa” and “La Sierra”. AURCs for VREs greater than the 85th (tp>=9.865 mm/12 for “La Costa”, and tp>=5.885 mm/12h for “La Sierra”) and the 99th percentile (tp>=50.452 mm/12 for “La Costa”, and tp>=25.551 mm/12h for “La Sierra”) are shown, respectively, in panels (a) and (c), and (b) and (d). The red and blue lines represent, respectively, the AURC for ENS and ecPoint. The red and blue shaded areas represent, respectively, the 95% confidence interval (CI) for AURC for ENS and ecPoint. The forecast lead times are indicated in hours as the step at the end of the 12-hourly accumulation period. The correspondent valid 12-hourly accumulation periods in UTC (and local time, LT) are shown in four different colours: 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 equivalent lead times in days (from day 1 to 10) are also indicated.

Chart

Description automatically generated

Figure 5 – ROC curves for VREs = 85th percentile (in purple) and VRE=99th percentile (in orange). Panels (a) and (b) show ROC curves for “La Costa”, and panels (c) and (d) show ROC curves for “La Sierra”. Panels (a) and (C) show ROC curves for ENS, and panels (b) and (d) show ROC curves for ecPoint.

Chart, scatter chart

Description automatically generated

**Figure 6** – ROC curves for the lead time step at the end of the 12-hourly accumulation period t+72 (i.e., day 3 forecasts), whose valid time corresponds to the accumulation period between 0600-1800 LT. Panel (a) and (b) show, respectively, the ROC curve for the 85th and the 99th percentiles.Continues and dashed lines correspond, respectively, to “La Costa” and “La Sierra”. Read and blue lines correspond, respectively, to ENS and ecPoint.

Chart

Description automatically generated

**Figure 7** - 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), obtained from 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).

Diagram

Description automatically generated

**Figure 8** - Flash floods in Ecuador on March 8th, 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.

Bazo, J., R. Singh, M. Destrooper, and E. C. De Perez, 2018: 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.

Bouallegue, Z. Ben, and D. S. Richardson, 2021: On the ROC Area of Ensemble Forecasts for Rare Events. *Preprints*, https://doi.org/10.20944/PREPRINTS202111.0535.V1.

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, 2022: A comparison of social vulnerability indices specific to flooding in Ecuador: Principal component analysis (PCA) and expert knowledge. *Int. J. Disaster Risk Reduct.*, **73**, 102897, https://doi.org/10.1016/j.ijdrr.2022.102897.

Buizza, R., 2019: Introduction to the special issue on “25 years of ensemble forecasting.” *Q. J. R. Meteorol. Soc.*, **145**, 1–11, https://doi.org/10.1002/qj.3370.

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.*, **572**, 603–619, https://doi.org/10.1016/j.jhydrol.2019.03.026.

Coughlan De Perez, E., B. Van Den Hurk, M. K. Van Aalst, B. Jongman, T. Klose, and P. Suarez, 2015: Forecast-based financing: An approach for catalyzing humanitarian action based on extreme weather and climate forecasts. *Nat. Hazards Earth Syst. Sci.*, **15**, 895–904, https://doi.org/https://doi.org/10.5194/nhess-15-895-2015.

Dordevic, M., P. Mutic, and H. Kim, 2020: Flash Flood Guidance System: Response to one of the deadliest hazards. *WMO Bull.*, **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. *Int. J. Disaster Risk Reduct.*, **31**, 1107–1120, https://doi.org/10.1016/j.ijdrr.2018.08.009.

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

Georgakakos, K. P., 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.

Golding, B., N. Roberts, G. Leoncini, K. Mylne, and R. Swinbank, 2016: MOGREPS-UK convection-permitting ensemble products for surface water flood forecasting: Rationale and first results. *J. Hydrometeorol.*, **17**, 1383–1406, https://doi.org/10.1175/JHM-D-15-0083.1.

Golnaraghi, M., 2012: *Institutional partnerships in multi-hazard early warning systems: a compilation of seven national good practices and guiding principles*. Springer,.

Haiden, T., and S. Duffy, 2016: Use of high-density observations in precipitation verification. *ECMWF Newsl.*, 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 Tech. Memo.*, **880**, https://doi.org/10.21957/6njp8byz4.

Hamill, T. M., and J. Juras, 2006: Measuring forecast skill: Is it real skill or is it the varying climatology? *Q. J. R. Meteorol. Soc.*, **132**, 2905–2923, https://doi.org/10.1256/qj.06.25.

——, 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. *Q. J. R. Meteorol. Soc.*, **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.

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

Javelle, P., 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.

Kastridis, A., and D. Stathis, 2020: Evaluation of hydrological and hydraulic models applied in typical mediterranean ungauged watersheds using post-flash-flood measurements. *Hydrology*, **7**, 12, https://doi.org/10.3390/hydrology7010012.

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.

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. *Anticip. 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.*, **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.

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. *Nat. Hazards*, **92**, 619–634, https://doi.org/10.1007/s11069-018-3173-7.

Lowrie, C., A. Kruczkiewicz, S. N. McClain, M. Nielsen, and S. J. Mason, 2022: Evaluating the usefulness of VGI from Waze for the reporting of flash floods. *Sci. Reports 2022 121*, **12**, 1–13, https://doi.org/10.1038/s41598-022-08751-7.

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

Pinos, J., and L. Timbe, 2020: Mountain Riverine Floods in Ecuador: Issues, Challenges, and Opportunities. *Front. 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. *Nat. 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. *Meteorol. Appl.*, **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. *Meteorol. Appl.*, **25**, 548–560, https://doi.org/https://doi.org/10.1002/met.1720.

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. *Meteorol. Atmos. Phys.*, 1–23, https://doi.org/10.1007/s00703-021-00831-z.

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.

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.

Tobar, V., and G. Wyseure, 2018: Seasonal rainfall patterns classification, relationship to ENSO and rainfall trends in Ecuador. *Int. J. Climatol.*, **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.*, **374**, 92–105, https://doi.org/10.1016/j.jhydrol.2009.06.004.

Tsonevsky, I., C. A. Doswell, and H. E. Brooks, 2018: Early warnings of severe convection using the ECMWF extreme forecast index. *Weather Forecast.*, **33**, 857–871, https://doi.org/10.1175/WAF-D-18-0030.1.

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

UNICEF, and WFP, 2015: UNICEF/WFP Return on Investment for Emergency Preparedness Study. 34.

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.

Wilson, L., 2000: Comments on “Probabilistic predictions of precipitation using the ECMWF Ensemble Prediction System". *Weather Forecast.*, **15**, 361–364, https://doi.org/10.1175/1520-0434(2000)015<0361:COPPOP>2.0.CO;2.

Winsemius, H. C., B. Jongman, T. I. E. Veldkamp, S. Hallegatte, M. Bangalore, and P. J. Ward, 2018: Disaster risk, climate change, and poverty: Assessing the global exposure of poor people to floods and droughts. *Environ. Dev. Econ.*, **23**, 328–348, https://doi.org/10.1017/S1355770X17000444.

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. *Nat. 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.

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. <https://floodlist.com/> [↑](#footnote-ref-2)
2. <https://reliefweb.int/disasters> [↑](#footnote-ref-3)
3. <https://www.ecmwf.int/en/forecasts/about-our-forecasts/evolution-ifs/cycles/summary-cycle-46r1> [↑](#footnote-ref-4)
4. <https://www.ecmwf.int/en/forecasts/about-our-forecasts/evolution-ifs/cycles/summary-cycle-47r1> [↑](#footnote-ref-5)
5. https://www.pichinchacomunicaciones.com.ec/lluvias-causan-desborde-de-rios-e-inundaciones-en-varias-provincias/ [↑](#footnote-ref-6)
6. https://www.wunderground.com/history/daily/SEGU/date/2021-3-8 [↑](#footnote-ref-7)
7. 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-8)
8. https://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/ARCHIVE/PDF/mjo\_evol-status-fcsts-20210315.pdf [↑](#footnote-ref-9)