**RESEARCH ARTICLE**

Co-production of training guidelines for new post-processed “point rainfall” forecast products

Fatima M. Pillosu1,2, Boglárka Tóth3, Istvan Ihasz3, Roberto Vindas Morán4, Werner Stolz4, Tim Hewson2, Christel Prudhomme2,5,6, Elisabeth Stephens7,8, Hannah L. Cloke1,7,9,10

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 Hungarian Meteorological Service, Budapest, Hungary

4 National Meteorological Institute of Costa Rica, San José, Costa Rica

5 Department of Geography and Environment, University of Loughborough, Loughborough, UK}

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

7 Department of Meteorology, University of Reading, Reading, UK

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

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

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

**Correspondence:** Fatima M. Pillosu, Department of Geography and Environmental Science, University of Reading, Reading, UK & Forecast Department, European Centre for Medium-range Weather Forecasts, Reading, UK([fatima.pillosu@ecmwf.int](mailto:fatima.pillosu@ecmwf.int))

**Abstract.** The ecPoint statistical post-processing technique, provides improved rainfall forecasts for points across the world. Whilst verification shows that marked improvements are achieved, for extremes also, the post-processed forecasts add layers of complexity to traditional probabilistic forecasts which could render interpretation difficult for users. So, a collaborative pilot study with the national hydro-meteorological services of Costa Rica and Hungary was conducted to assess the extent to which the products would be correctly used, and fully exploited, in operational forecasting environments with different climatological characteristics. The key component regarding interpretation was the guidelines written for users. This study found that whilst correct utilisation of the new products was indeed initially challenging, by adopting an approach that fully caters for different levels of experience with probabilistic forecasts this could be achieved, and indeed decision-making in critical situations could be improved. This is encouraging for the wider adoption of this new forecast product and we use the findings to provide new guidelines which should cover a wide range of settings, including forecasting environments in which the product usage background is primarily deterministic.

**Keywords.** Forecasters training, ensemble-based probabilistic forecasts, extreme localized rainfall**,** statistical post-processing.

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

# Introduction

Worldwide, extreme localized rainfall, and associated hazards, such as flash floods and landslides, cost thousands of lives and billions of dollars in economic losses every year (UNDRR, 2020). Moreover, extreme localized rainfall is expected to become more frequent due to climate change (Hoegh-Guldberg *et al.*, 2019) and more damaging due to increased exposure/vulnerability of people and assets (Ward *et al.*, 2020). This has led the scientific community to develop numerical weather prediction (NWP) models to try to predict location, timing and magnitude of extreme localized rainfall events, possibly several days in advance to gauge the range of possible impacts (Doswell *et al.*, 1996). Km-scale NWP models improve the realism of extreme localized rainfall forecasts but they are not always necessarily more accurate than other coarser-resolution NWP models (Roberts, 2008; Roberts & Lean, 2008; Golding *et al.*, 2016), including global models. Moreover, km-scale NWP models often provide forecasts only for very short lead times (Yussouf *et al.*, 2020). Statistical post-processing can improve accuracy and increase lead times (Vannitsem *et al.*, 2021). However, the wide range of statistical post-processing techniques reviewed in Vannitsem *et al.*, (2021) shows that accurate prediction of extreme localized rainfall is a long-standing problem, mainly because the amount of required observational data tends to limit their development to only regions with good observational networks. A new statistical post-processing technique, called ecPoint (Hewson & Pillosu, 2021), has been successful in this sphere, with global verification showing that it delivers accurate rainfall forecasts for points across the world whilst also overcoming many of the calibration challenges. Whilst a focal point of this study is extreme rainfall, Hewson and Pillosu showed that ecPoint post-processing also markedly improves forecasts of dry weather, which for certain applications is also very beneficial.

Weather forecasters have the challenging task of translating the wealth of information in forecast products, with all its strengths and weaknesses, into meaningful guidance for end-users, without reducing trust, so that that information can be appropriately used for decision-making (Losee & Joslyn, 2018). However, the adoption in operational routines of new products can be slow because forecasters have the tendency to work with well-established products, whose strengths and weaknesses are well understood (Nielsen & Schumacher, 2016). Therefore, it is not only cutting-edge, high-quality products that must be supplied but also clear proof of the new-products’ added value in the context of forecasters’ decision-making processes (Murphy, 1993). Moreover, forecasters’ exhausting shifts and tight schedules tend to leave little time for regular updates and the study of new products (Sills, 2009). Accompanying training guidelines can accelerate forecasters’ adoption of a new products as long as they (1) contain extensive verification to prove new-product’s added value compared to others, (2) explain how to interpret them correctly given the strengths/weaknesses of the new approach, (3) explain the meteorological circumstances in which added value is likely to be most prominent, and (4) clearly indicate how to access the product efficiently under real-time operational constraints (Bullón & Viana, 2018). However, extensive research on the interaction between developers and forecasters (Morss *et al.*, 2016; Novak *et al.*, 2008; Morss *et al.*, 2008; Demeritt *et al.*, 2010, 2013; Joslyn & LeClerc, 2013; Evans *et al.*, 2014; LeClerc & Joslyn, 2015; Losee & Joslyn, 2018; Wilson *et al.*, 2019; Fundel *et al.*, 2019; Demuth *et al.*, 2020; Emerton *et al.*, 2020) has highlighted four main issues with most training guidelines: (1) content focuses too much on scientific background and not enough on practical applications; (2) strengths, weaknesses and special features of the forecast products are often omitted; (3) examples and language are not appropriate for user needs and experience; (4) information regarding product availability and accessibility is lacking. This evidence shows that the way training guidelines are provided still requires special attention from the scientific development community, not in isolation, but instead by working closely with the user community.

To promote the adoption of the ecPoint-derived rainfall products (ecPoint-Rainfall) within the forecaster community, training guidelines on the product’s strengths/weaknesses and how to use them were published at the early stages of development in the ECMWF Newsletter (Pillosu & Hewson, 2017) and when the product was released operationally in April 2019 in the ECMWF Forecast User Guide (Owens & Hewson, 2018). The severe flooding experienced by Peru in 2017 provided the first opportunity for ecPoint developers to hear about strengths and weaknesses of ecPoint-Rainfall forecasts for a specific region and from local developers, e.g. forecasters from the Peruvian national hydro-meteorological service SENAMHI (Pillosu *et al.*, 2017). ecPoint developers were interested to know whether SENAMHI forecasters would have perceived ecPoint-Rainfall forecasts useful if they had used the new product in the creation of forecasts/warnings for the flash flood events. For a forecaster, a new rainfall product is perceived as more “useful” than others if it can provide better pointers regarding the location, time, and intensity of the rainfall event, ideally with longer lead times (Nielsen & Schumacher, 2016). SENAMHI did not perceive ecPoint-Rainfall forecasts as more useful, in spite of developers’ analysis having demonstrated that ecPoint-Rainfall was more skilful than raw ECMWF ensemble (ENS) in detecting the extreme localized rainfall during the floods (Pillosu *et al.*, 2017) and globally (Pillosu & Hewson, 2017). Subsequent discussions with SENAMHI forecasters showed that guidelines provided did not help forecasters to exploit the new forecasts, and their usefulness was consequently underrated. Was this outcome related to forecasters’ previous experience with probabilistic forecasts? Or was it related to the fact that even within a classical framework of probabilistic forecasts ecPoint introduces new layers of complexity that render interpretation particularly challenging?

This article discusses the outcomes of a pilot study carried out to identify weaknesses in ecPoint-Rainfall guidelines, and to establish directions to improve them. ecPoint-Rainfall is a global probabilistic product that could be used by forecasters around the world, who may or may not have experience with probabilistic forecasts. Therefore, guidelines must reflect this. The user base of ecPoint-Rainfall is wide-ranging. For example, forecasters who use convection-resolving NWP models to forecast extreme localized rainfall could be one of the main users of ecPoint-Rainfall, e.g. to complement that output by providing less jumpy forecasts at the short lead times covered by the km-scale runs (typically 1-3 days), or to extend lead times beyond the 2-3 day limit typical of convection-resolving models (Hewson & Pillosu, 2021). Also, forecasters who do not have access to km-scale models for their region of interest could use the ecPoint-Rainfall output to produce forecasts for extreme localized rainfall events. The pilot study involved the collaboration between ecPoint developers and forecasters from the national hydro-meteorological services (NHMSs) of Hungary (Országos Meteorológiai Szolgálat, OMSZ) and Costa Rica (Instituto Meteorológico Nacional de Costa Rica, IMN). These services were invited to participate to represent different levels of experience with probabilistic forecasts (OMSZ has been providing probabilistic forecasts since 1990s, whilst IMN uses primarily deterministic NWP models to this day). This can ensure that any conclusions on ecPoint-Rainfall use are linked to guideline quality and not to forecasters’ prior experience with probabilistic forecasts. A second consideration was the different climatological backgrounds of Hungary and Costa Rica (respectively tropical and extra-tropical), which of itself could also help further worldwide adoption.

# Data

## The ecPoint methodology

ecPoint (Hewson & Pillosu, 2021) is a statistical post-processing technique that addresses the two main factors that affect the utility of global NWP model outputs, especially for guidance on extreme local events: lack of information on forecast sub-grid variability (Göber *et al.*, 2008) and systematic biases (Lavers *et al.*, 2021). Systematic biases are model errors which lead to forecast under-/overestimations at grid scale, whilst the lack of information on forecast sub-grid variability is a characteristic of NWP models. NWP models forecast only grid-box averages and do not enable forecasters to A picture containing diagram

Description automatically generatedassess when/where smaller/higher local values than the model average might occur. Statistical post-processing can improve the forecast accuracy and reliability by correcting biases and converting forecasts from grid-box to point scale (Buizza, 2018). A plethora of statistical post-processing techniques have been developed over the last 50 years (see Vannitsem *et al.* (2021) for a review). However, the operational creation of statistically post-processed forecasts remains problematic to this day, especially in the case of global NWP models and extreme events. Table 1 in Hewson & Pillosu (2021) presents a list of thirteen challenges faced by well-established and state-of-the-art statistical post-processing techniques, and how ecPoint has addressed them by applying a “global remote calibration” approach. The concept behind this approach is that sub-grid variability and systematic biases are weather-dependent, and such dependency is generally not location-related. Thus, observations from everywhere in the world are gathered to define a set of grid-box weather types (G\_WT, **Fig. 1a**) that are used to convert each single raw grid-box forecast into a distribution of equally probable bias-corrected point-scale realizations (**Fig. 1b**).

**Fig. 1** - Panels (a) and (b) show, respectively, the workflow for the calibration and the forecasts generation using the ecPoint methodology. Panel (c) shows an example of the two ecPoint-Rainfall products for 12-hourly rainfall, available on ecCharts (https://www.ecmwf.int/en/forecasts/ ecCharts), i.e., a map plot for the 99th percentile (first row) and a map for the probabilities of exceeding 10 mm/12h (second row).

## ecPoint and forecasters’ heuristics

Kahneman & Tversky (1973) showed that people tend to not make decisions by processing information using quantitative statistical methods. Instead, information flow is simplified through a set of subjective methods called heuristics. Especially in weather forecasting, the high volumes of information that needs to be processed within the finite time of forecast operations has made forecasters rely particularly on a heuristic approach (Doswell, 2004). The rise of automated objective guidance (e.g. NWP and machine learning) makes the weather forecasting community consider whether things should change (Stuart *et al.*, 2006; Novak *et al.*, 2008; Sills, 2009). Studies have shown that raw forecast accuracy can indeed be improved upon using heuristics, e.g. by understanding strengths and weaknesses in the model in different weather situations (Murphy & Winkler, 1977; Carroll and Hewson, 2008; Sills, 2009). However, it has also been shown that heuristics can be a source of bias, particularly in complex, less frequent scenarios (Tversky & Kahneman, 1974). In weather forecasting, anchoring and adjustment, availability, representativeness, overconfidence, and inconsistency are the main sources of these biases (Doswell, 2004). Anchoring and adjustment refers to the process by which quantitative uncertainty assessments begin with some anchoring value and thereby tend to not fully account for the available information. Availability refers to how readily relevant information comes to the fore when assessing uncertainty. Representativeness refers to whether the situation under consideration genuinely fits a pre-defined conceptual model. Overconfidence refers to the tendency for forecasters’ self-belief, through experience, to increasingly outstep what is actually the reality. Inconsistency refers to how the same input data tends to not provide the same output.

ecPoint adopts the well-founded aspects of the forecaster’s heuristic approach, whilst simultaneously addressing its innate limitations and biases (**Fig. 2**). Forecasters effectively perform subjective calibration by relating upcoming weather patterns to their past experience of similar scenarios, using local observations. This can lead to representivity and availability biases. Forecasters then intuitively generate qualitative probabilities. This can lead to inconsistency, overconfidence and anchoring and adjustment bias. Meanwhile, ecPoint uses expert elicitation and statistical tests to calibrate the forecasts using *global* data (**Fig. 1a**). It then applies an objective process, involving simple mathematical transformations that are fully informed by the past performance to generate objective probabilities (**Fig. 1b**). ecPoint then provides a dynamical calibration process that favours a quick and cost-effective adaptation to any new model versions, which in the heuristic approach would require a costly assignment of human resources to maintain the forecaster advantage. Diagram, timeline

Description automatically generated

**Fig. 2** - Correlation between weather forecasters’ heuristics approach to assess uncertainty in raw NWP guidance and ecPoint methodology.

## ecPoint-Rainfall

ecPoint-Rainfall is a product branch of the ecPoint family[[1]](#footnote-2) that post-processes raw NWP rainfall forecasts to mirror rain gauge measurements. 12-hourly rainfall forecasts from the 00 and 12 UTC run of the ECMWF ensemble (ENS) are currently post-processed operationally, up to day 10. The outputs (percentiles from 1st to 99th) are available for visualization on ecCharts[[2]](#footnote-3) in the form of map plots of percentiles (**Fig. 1c**, first row) or probabilities of exceeding a certain rainfall threshold (**Fig. 1c**, second row).

# Methods

Diagram

Description automatically generatedThe participants in the pilot study were operational weather forecasters at IMN and OMSZ with experience in forecasting extreme localized rainfall events. Two intermediaries were identified at each NHMS to ensure an efficient communication between operational forecasters and ecPoint developers. The intermediaries were required to be operational forecasters themselves or be familiar with operational procedures and preferences for forecasting extreme localized rainfall events.

**Fig. 3** - Experiment design. The white boxes describe the different steps carried out during the “real-time” and the “offline” phase of the experiment. The fuchsia, blue, green, and cyan frames encompass the steps done by the ecPoint developers, NMHS forecasters, the intermediaries, and the ecPoint developers and intermediaries together, respectively. The yellow rhombus contained the questions asked at key moments of the study to define the path to take in the experiment. The green boxes represent the actions taken based on the followed path.

The pilot study was structured in two consecutive phases (**Fig. 3**). The first phase (called “real-time”) gathered feedback from operational forecasters on aspects that can contribute to the adoption, under real-time constraints, of new forecasting products, i.e., performance, perceived usefulness, guidelines efficacy, and the compatibility with current operational systems. The aim of the second phase (called “offline”) was to encourage ecPoint developers and intermediaries to reflect on the received feedback and find more efficient and effective ways to convey ecPoint-Rainfall guidelines to users. An “informal discussion” approach was chosen for all the conversations carried out in the “offline” phase as it helps interviewers to put respondents at ease and does not inhibit comments or ideas (Harding, 2018).

The “real-time” phase was organized in two steps: forecast delivery setup and guidelines provision. Real-time forecasts were delivered via file transfer protocol (ftp) for 12 months from May 2018. **Table 1** shows the characteristics of the ecPoint-Rainfall forecasts provided to IMN and OMSZ. ecPoint-Rainfall guidelines were provided via email, and intermediaries were offered on-request e-mail support, e.g., on the guidelines provided or for suggestions on how to develop local ecPoint-Rainfall products. This “email-exchange” approach is consistent with how the bulk of User Support functions at ECMWF. The guidelines content is shown in **Table 2**. They were divided in two categories. The “scientific” guidelines illustrated the scientific basis of the ecPoint methodology and the forecasts calibration that are required to be fully understood for a correct use of the forecasts. The “technical” guidelines described the technical characteristics of the forecasts that are most relevant to introduce smoothly any new forecasts in an operational system, such as file dimensions and formats, times of delivery and data formats. At the end of “real-time” phase, the intermediaries were asked to provide a written report on the forecaster’s experience of using and testing ecPoint-Rainfall forecasts operationally. This approach is consistent with how ECMWF requests Member States to provide feedback on standard ECMWF products. Intermediaries were left free to organize the report as they deemed appropriate. However, ecPoint developers asked intermediaries to include in their reports comments on (1) whether products were derived from ecPoint-Rainfall, and if so, describe them; (2) what was the perceived usefulness of ecPoint-Rainfall; (3) what was the effectiveness of the guidelines provided; (4) show a case study or verification results for regions of particular interest in their countries.

**Table 1** - Characteristics of the ecPoint-Rainfall forecasts provided at the beginning of the “real-time” phase.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Forecast format** | **Model Runs** | **Accumulation** | **Accumulation periods ending at** | **Lead times** | **Files format** |
| 1st to 99th percentiles | 00, 12 UTC | 12 hours | 0, 6, 12,18 UTC | Up to (t+246) | Grib2 |
| **Domain (N/S/W/E) for IMN (Costa Rica):** 12°N / 7°N / 87°W / 82°W; **for OMSZ (Hungary):** 49°N / 45°N / 15°W / 24°W. | | | | | |

**Table 2** - ecPoint-Rainfall guidelines provided (via email) at the beginning of the “real-time” phase.

|  |  |
| --- | --- |
| **Scientific guidelines** | **Technical guidelines** |
| Links to pre-prepared resources (mainly ECMWF Newsletters):   * https://www.ecmwf.int/en/newsletter/153/news/new-point-rainfall-forecasts-flash-flood-prediction * https://www.ecmwf.int/en/newsletter/159/news/new-point-rainfall-products-eccharts (\*) * https://confluence.ecmwf.int/display/FUG/ (\*)   (\*) Guidance provided from April 2019 when ecPoint-Rainfall was released operationally. | “ecPoint-Rainfall forecasts are produced in gridded binary (grib) format twice a day (for a 00 and 12 UTC run) in four 12-hourly overlapping periods, with valid times finishing at 00, 06, 12, 18 UTC. The maximum lead time computed is t+246 (i.e. day 10). Each file contains 99 global fields, which correspond to percentiles from 1st to 99th. The rainfall values are already accumulated (over the 12h period). Due to the size of each file (400 MBs for a global field), only forecasts for the domain of interest (N/S/W/E coordinates) can be sent. Such domain will be extracted in a regular lat-lon grid. The file names indicate the date (in YYYYMMDD format) and the run (UTC time, in HH format) in which the forecasts were computed. They also contain the step (hours, in hhh format) at which the 12-hourly accumulation period ends (example: ecPointRain\_12h\_YYYYMMDD\_HH\_hhh.grib). The forecasts from the 00 and 12 UTC run will be sent out at 9 and 21 UTC, respectively.” |

The “offline” phase was organized in four steps. As a first step, ecPoint developers reviewed the written reports provided by IMN and OMSZ at the end of the “real-time” phase and highlighted possible issues in the interpretation or the use of ecPoint-Rainfall forecasts. The aspects evaluated in the reports are described in **Table 3**. As a second step, ecPoint developers discussed the outcomes of their analysis with the intermediaries at IMN and OMSZ separately in a three-hour videocall. The conversation was guided by a list of open-ended questions (see Appendix A) to stimulate the conversation without stopping intermediaries to naturally raise their points for discussion (Harding, 2018). The list of open-ended questions was sent out one week before the videocall, and it was structured in two main sections. The first section aimed at formalizing the ecPoint developers’ pre-existing understanding of the participants’ experience in areas such as probabilistic forecasting and statistical post-processing to understand the context under which forecasters at IMN and OMSZ formulate their predictions for extreme localized rainfall events. This background information is presented in section 4, ahead of the results section, to put the outcomes of the experiment in a more meaningful context. The second section focused on the ecPoint-Rainfall forecasts and aimed at clarifying the participants’ thoughts about its performance and perceived usefulness, and the effectiveness of the ecPoint-Rainfall guidelines. As a third step, ecPoint developers highlighted similarities and differences between IMN and OMSZ in the use of ecPoint-Rainfall to forecast extreme localized rainfall events to abstract a new set of products and guidelines that should better fit users with similar needs. The main similarities/differences between IMN and OMSZ are reported in **Table 4**. As a fourth step, ecPoint developers emailed to the intermediaries a new set of guidelines and mock-up products for the case studies provided in the written reports which were taken account of the feedback received during the informal conversations. The aim was to determine whether the new products and guidelines would fit better with the participants’ needs. The intermediaries were asked to email back a comment about the new guidelines and the mock-up products stating (1) whether the revised products/guidelines were more useful than those provided in the “real-time” phase, and (2) whether intermediaries would change their initial thoughts about ecPoint-Rainfall and use the new product operationally to forecast extreme localized rainfall events in their countries.

|  |  |
| --- | --- |
| **Aspects evaluated: To what extent did…** | **Why?** |
| …the participants appreciate the difference in scale between ecPoint-Rainfall and NWP rainfall forecasts? | NWP forecasts represent rainfall averages over the model grid-box, whilst ecPoint-Rainfall represent the forecast at a point within the grid-box, even though no information can be provided on where that point is within the grid-box. |
| …the participants appreciate the difference between the spread in ecPoint-Rainfall and NWP rainfall forecasts? | The spread provided by ecPoint-Rainfall refers to rainfall at a point, whilst the spread provided by NWP rainfall forecasts refers to the grid-box average. |
| …the participants focus on the high percentiles of ecPoint-Rainfall (i.e. > 90th percentile) to assess the location and the magnitude of extreme localized rainfall? | Lower percentiles might not forecast extreme localized rainfall events which, by definition, are very low-probability. |
| …the participants appreciate the difficulties on verifying ecPoint-Rainfall forecasts for extreme localized rainfall events in small regions with sparse observational network, for restricted periods of time, (e.g. < 1 year)? | For example, the 99th percentile represents rainfall events with a 1 in 100 chance of happening. Such a rare event might not be observed; however, it does not mean the forecasts were wrong because the event could have happened but not be recorded. |

**Table 3** - Aspects evaluated in the written reports delivered at the end of the “real-time” phase.

**Table 4** - Examination of similarities/differences between IMN (Costa Rica) and OMSZ (Hungary) during the third step of the “offline” phase.

|  |  |
| --- | --- |
| **What were the similarities/differences…** | **Why?** |
| … in the forecasting remit of the two organizations? | To assess whether the forecasting remit between the two centres explain any differences that might come out from the discussions (e.g. one centre is mainly interested in short range forecasts instead or medium-range lead times.) |
| …in the quality/number of products developed from ecPoint-Rainfall? | To assess whether the approaches adopted when using ecPoint-Rainfall were due to the novelty of the product, the quality of the guidelines provided, or the previous forecasters’ experience with probabilistic or statistically post-processed forecasts, and better tailor future guidelines according to specific user’s needs. |
| …in the operational use of ecPoint-Rainfall to forecast extreme localized rainfall? |
| …on how ecPoint-Rainfall might be used in any decision-making process? |

# Pilot study context

This section summarizes information collected from the informal discussions during the “offline” phase of the study on the challenges in the prediction of extreme localized rainfall events in Costa Rica and Hungary, and on the forecasters’ experience with probabilistic NWP models and statistical post-processing at IMN and OMSZ. This information is presented ahead of the results section to help the user to put the outcomes of the real-time and the offline phase in a more meaningful context.

## Instituto Meteorológico Nacional de Costa Rica (IMN)

Rainfall in Costa Rica is typically generated by cold fronts, hurricanes, and low-pressure systems. Their interaction with prevailing winds and Costa Rica’s complex orography (**Fig. 4a**) make the Pacific coast typically wetter than the Caribbean coast (**Fig. 4c**). They can also enhance rainfall amounts or produce severe localized storms which can generate severe flash floods. Extreme localized rainfall predictions at IMN rely substantially on real-time observations of rainfall, river discharge and soil moisture, which help forecasters to estimate the impacts of a rainfall event:

“Rainfall forecasts at IMN typically rely 60% on NWP guidance and 40% on human expertise, which is built by experience and research. The human expertise can indeed sometimes provide more detailed information on the nature of the phenomenon affecting us (e.g., a cold front does not produce the same impact of a tropical wave) and on the state of the regions that might be affected (e.g. if it has been raining in previous days and catchments are saturated).”

Global NWP models are typically not used at IMN to forecast rainfall in Costa Rica:

“Costa Rica is a small country with complex orography. Global models have a too coarse spatial resolution to represent the details of the spatial variation of rainfall events, even more in Costa Rica. Furthermore, they parametrize convection and the rainfall values might not be correct. At occasions we will look at the GFS [Global Forecast System, developed by the National Centers for Environmental Prediction in the USA], and now we also have access [from 2018] to the IFS [Integrated Forecasting System from ECMWF] but our **Map

Description automatically generated**forecasters tend to rely mainly on guidance from our high-resolution models.”

**Fig. 4** - Panel (a) and (b) show, respectively, the orography in Costa Rica and Hungary (source: www.maps-for-free.com). The black contours define the climatological regions in Costa Rica (named in black) and the administrative counties (not named) in Hungary. The red circle indicates the most impacted areas (named in red) in the rainfall case study provided, respectively, by IMN and Hungary. Panel (c) and (d) show, respectively the annual rainfall amounts in Costa Rica (www.imn.ac.cr/atlas-climatologico) and Hungary (www.met.hu/en/eghajlat/magyarorszag\_eghajlata/ altalanos\_eghajlati\_jellemzes/csapadek).

IMN developed instead a series of in-house models, based on the Weather Research and Forecasting (WRF) system, that have different spatial/temporal resolutions, domains, and model configurations to tailor their use for specific hazards. The model WRF-1.5, developed in 2018, produces rainfall forecasts at 1.5 km resolution, up to day 5, and aims to improve predictions for those systems that can produce localized storms.

“Even if we [forecasters at IMN] know that sometimes the deterministic model [WRF-1-5] can overestimate the rainfall amounts, their spatial distribution is much better than the global models. From experience we know that the rainfall values just need to be multiplied by a factor of 0.4.”

These models are, however, mainly deterministic:

“95% of predictions at IMN are created using deterministic guidance, and only 5% derives from probabilistic models.”

Although the main focus is on deterministic models, IMN forecaster are trained also to use and understand ensemble model outputs:

“50% of forecasters have attended the NOAA’s Weather Prediction Centre International Tropical Desk, training in weather and climate forecasting for the Americas. Other forecasters have attended other training centres over the years.”

## Országos Meteorológiai Szolgálat (OMSZ, Hungary)

Hungary’s climate is mainly warm continental. Rainfall is typically connected to large scale systems (e.g., cyclones, squall lines or cold fronts). However, their interaction with the orography (**Fig. 4b**) can enhance the rainfall amounts (**Fig. 4d**), making the southwest and northeast (mostly hilly) typically wetter than the southeast (mostly flat). Hungary’s orography is not particularly complex. However, small differences in the orography (e.g., 100-250 m) can be enough to trigger extreme localized storms which can generate severe flash floods. OMSZ has a long-standing experience in developing ensemble NWP models:

*“Since the 1990s, extreme rainfall predictions at OMSZ are mostly generated using ensembles. Forecasters typically look at a suite of models, including AROME* [2.5 km horizontal resolution]*, ALADIN* [8 km horizontal resolution]*, and ECMWF’s deterministic* [HRES, 9km horizontal *resolution*] *and ENS* [18 km horizontal resolution]*. Forecasters exploit their knowledge about the models’ strengths/weaknesses in different weather scenarios to arrive to a consensus forecast and provide the best possible guidance.”*

OMSZ has also an extensive experience verifying ensemble model outputs:

*“We regularly verify the models that we develop at OMSZ [AROME and ALADIN]. We also provide regular feedback to ECMWF about HRES and ENS performance. We have regular email communications with the User Support at ECMWF or the person on duty for the daily report*[[3]](#footnote-4)*. Like other ECMWF’s cooperating and member states, every two years we also send to ECMWF a summary of our verification results on HRES and ENS. They would normally be summarized and presented in the “Green Book”* [see example for 2019 in Hewson (2020)]*”.*

OMSZ forecasters also take part regularly in training programmes at ECMWF to keep updated on the developments in the ECMWF models, and improve the use of probabilistic products. OMSZ is experienced in rainfall post-processing, which is regularly used at OMSZ to add value to forecasts for small-scale, low-predictability phenomena like extreme localized rainfall (Matrai and Ihász 2017; Ihász et al. 2018).

# Results

The summary of activities undertaken by IMN and OMSZ during the “real-time” phase are presented in **Table 5**. Verification activities from the “real-time” phase comprised case-studies of high-impact events (see summary in **Table 6** and objective verification (see summary in **Table 7**). Even though both NMHSs expressed positive feedback regarding the general utility of ecPoint-Rainfall to forecast extreme localized rainfall (**Table 6**, Outcomes row), intermediaries also highlighted some issues regarding the underestimation of the highest observed rainfall amounts, the rainfall overestimation of the highest ecPoint-Rainfall percentiles and the definition of the wettest day in a multiple-day event (“Outcomes” row in **Table 6** and **Table 7**). The following sub-sections present the discussions carried out between ecPoint developers and intermediaries, the products developed to overcome the found issues, the guideline amendments, and preliminary forecasters’ impressions about the proposed solutions.

|  |  |  |
| --- | --- | --- |
| **Activities/Outcomes** | **IMN (Costa Rica)** | **OMSZ (Hungary)** |
| Forecasts used | Only forecasts from 12 UTC runs. They are the earliest ones available to forecasters for daytime work due to time zone differences (UTC-6). Lead times up to day 7. Max lead-time used at IMN for warnings. File format was converted to netCDF, which is more commonplace in the Americas. | Only forecasts from 00 UTC runs. They were considered sufficient to verify ecPoint-Rainfall. Lead times up to day 5. Max lead-time used at OMSZ for warnings. |
| Requested training follow-up | On what products could be created from ecPoint-Rainfall, shown in **Table 6**. | None. |
| Products created | Map plot of the ecPoint-Rainfall 85th percentile (**Fig.5a**). | Four-panel meteogram (**Fig.5b**) containing: ENS and ecPoint-Rainfall 12-hourly precipitation, ENS rate of convective precipitation ratio, ENS 700 hPa wind speed, and ENS CAPE. Map plot displaying the grid-box values of the ecPoint-Rainfall 90th, 75th, 50th, 25th, 10th percentiles (numerical values in **Fig.5c**) and the probabilities of exceeding certain rainfall thresholds (contour shades in **Fig.5c**). |
| Independent verification | Case study (see **Table 7**, “IMN (Costa Rica)” column for a summary). Results not peer-reviewed published. | Case study (see **Table 7**, “OMSZ (Hungary)” column for a summary) and objective verification (see  **Table 8** for a summary). Methodology and results have been published in Tóth and Ihász (2021). |

**Table 5** - Summary of activities carried out by IMN and OMSZ during the “real-time” phase and outcomes.

## On the selection of the 85th percentile for ecPoint-Rainfall forecasts to predict extreme localized rainfall

**Fig. 5** - Panel (a) shows the product developed by IMN, i.e. a map plot of the ecPoint-Rainfall’s 85th percentile over Costa Rica and parts of Central America. Panel (b) and (c) show the products developed by OMSZ. Panel (b) shows a meteogram displaying 12-hourly precipitation from the ECMWF ENS in blue and ecPoint-Rainfall in orange (first panel), ECMWF ENS rate of convective precipitation ratio (second), 700 hPa wind speed (third), and CAPE (fourth). Panel (c) shows a map plot of the probabilities of exceeding 10 mm/12h (contour shades) and the values (numbers, from top to bottom) of the 90th , 75th, 50th, 25th and 10th percentiles for each 0.5-degree resolution grid boxes. The grid box resolution was decreased to make numbers more visible. The contour shades correspond instead to the original ecPoint-Rainfall resolution of ~18km.

Both IMN and OMSZ decided to use the ecPoint-Rainfall 85th percentile to forecast extreme localized rainfall. In the case of IMN, the decision was taken intuitively, considering the 85th percentile as a reasonable choice to forecast extreme localized rainfall without missing some less extreme, but still impactful events. In the case of OMSZ, the decision was taken after an objective verification analysis of the post-processed forecasts for the summer-months of 2018.

**Table 6** – Guidelines provided via email to IMN on the development/interpretation of products derived from ecPoint-Rainfall.

|  |  |
| --- | --- |
| **Map plot Percentiles** | **Probabilities of not exceeding a rainfall threshold** |
| “ecPoint-Rainfall forecasts are provided in percentiles from 1st to 99th. A map plot of the 99th percentile displays the rainfall values (in mm/12h) that have 1% (or 99%) risk of (not) being exceeded. When looking at extreme localized rainfall events, it would be advised to use high percentiles, namely >90th percentile, but preferably the 98th or 99th, since extreme localized rainfall is by definition a very low-probability event. Nonetheless, the chosen percentile would depend on the level of risk your institution would act upon: if it is required that an event has at least 10% risk of exceedance to issue a warning, the 90th percentile would be considered. An important aspect to consider is that, although the rainfall value for a percentile is assigned to the whole grid-box, it refers to a point within the grid-box, and nothing can be said about its location within that grid-box.” | “To compute the probabilities of exceeding a rainfall threshold for a grid-box, consider its corresponding 99 percentiles values, assign 1 to those values that exceed the threshold, and 0 otherwise. Compute the mean and multiply it by 100 to obtain the probability in %. The rainfall threshold to choose relates to the event that can generate some impacts (e.g., flash floods) in the area of interest. Notice that extreme rainfall thresholds are likely to have very small probabilities of occurrence, namely <5%, as they are by definition a very low-probability event.An important aspect to consider is that, although the probability of oc**c**urrence is assigned to the whole grid-box, it refers to a rainfall value observed at a point within the grid-box, and nothing can be said about its location with that grid-box.” |

At IMN, forecasters know that events of 50 mm/12h can generate some impacts (e.g. flash floods or landslides) in the Pacific coast of Costa Rica and in the capital city San José. Forecasters might be therefore inclined to use the 50th, 75th or 85th percentiles in their predictions. Such percentiles would indeed provide forecasts for more frequent events (i.e. with a 1 in 2, 1 in 4 or 1 in 7 chance, on average, to be observed), whilst the 98th or 99th percentiles would instead provide forecasts for much more rare events (i.e. with a 1 in 20 or 1 in 100 chance to be observed).

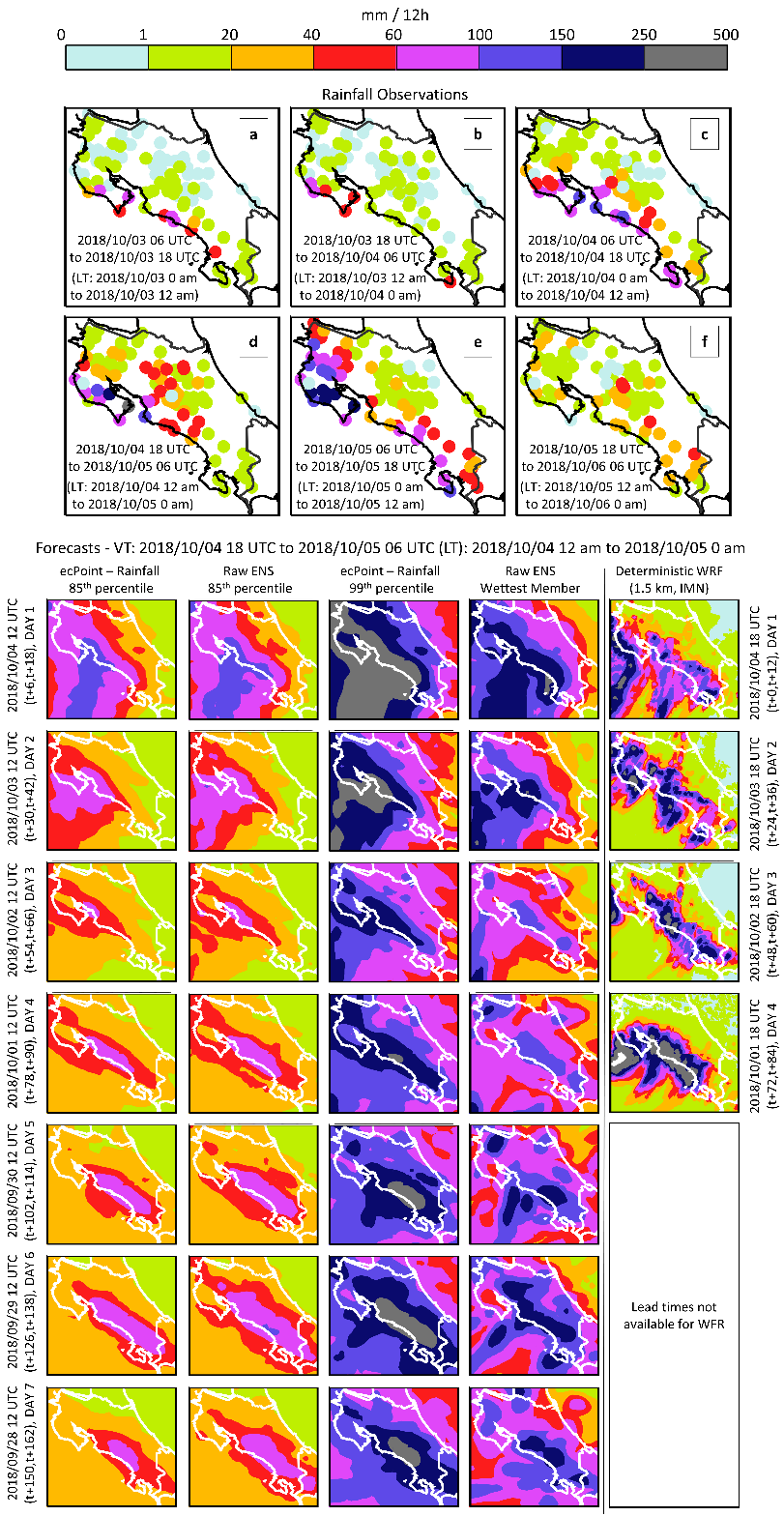
“Events of around 50 mm/12h can already cause some impacts on the Pacific coast of Costa Rica or in the capital city, San José. Those events are relatively frequent. If we only consider the 98th or 99th percentile [as suggested by the original guidelines presented in **Table 6**], we could miss many more frequent events that can already generate some impacts in the region. The 85th percentile was considered a reasonable choice.” (IMN)

**Table 7** - Summary of case studies analysed during the real-time phase

|  |  |  |
| --- | --- | --- |
| **Analysis by** | **IMN (Costa Rica)** | **OMSZ (Hungary)** |
| Event description | An extreme rainfall event occurred in the Pacific coast regions, between October 3rd and 5th, 2018 due to an almost stationary low pressure over the Caribbean Sea. The most intense rainfall fell on October 4th (~400 mm/24h in the south of Nicoya peninsula, ~200 mm/24h and ~90 mm/24h in Regin Pacífico Central and Sur, respectively). From October 6th Costa Rica was out of the influence of the low-pressure system. 12-hourly rainfall observations are shown in **Fig.6a-f**. The rainfall event caused severe flash floods and landslides. One person died, hundreds were moved to refuges, and thousands of people were affected by electricity/water service interruption or road closures. | An extreme rainfall event occurred in the Bükk mountains (**Fig.7a**) between 10th June and 11th June 2018. A slow-moving low-pressure system, coming from Turkey, interacted with the Northern Central Mountains in Hungary, and stayed over the Bükk mountains for hours generating very localized high rainfall event on 11th June, between 0 and 12 UTC. The highest rainfall amount (92 mm/24h) was recorded in Bükkszentlélek (~10km east from Szilvásvárad, purple circle with a cross in Fig. 8a). The rainfall event caused flash floods in the whole Bükk area, especially in Szilvásvárad. Such events occur rarely in Hungary, approximately every 10-20 years. |
| Forecasts used | IMN tested the developed product (**Table 5**, Products created), up to day 7. | OMSZ tested the developed products (**Table 5**, Products created), in particular the probabilities of exceeding 10 and 30 mm/12h at day 2 and 4. |
| Outcomes | *“ecPoint-Rainfall predicted well the start and the end of the rainfall event from September 27th. Forecasts from October 4th for the same day, pointed out that Región Pacífico Central and Sur would have been the most impacted areas and underestimated the rainfall amounts in the Nicoya peninsula. No ecPoint-Rainfall forecasts [based on the 85th percentile] reached the observed totals [****Fig.6****, forecasts,**first column]. Most runs predicted that October 5th would have been the wettest day, whilst it was October 4th."* | *“Based on ecPoint-Rainfall, high rainfall amounts were expected in the Bükk area. The map plot for the ecPoint-Rainfall probabilities of exceeding 10 and 30 mm/12h outlined better than the ENS the area where the highest local precipitation was observed (****Fig.7b-i****). Although the probabilities of exceeding 30 mm/12h are relatively low* [between 1 and 3%], *this information is very important for forecasters.”* |

|  |  |
| --- | --- |
| **Analysis by** | **OMSZ (Hungary)** |
| Forecasts used | 00 UTC runs for forecasts between 1st June and 31st August 2018.  *“This period was chosen because extreme localized rainfall events in Hungary are mainly generated by convection during summertime. Since convective rainfall is difficult to predict with NWP models, especially global models, we thought this could be a good test for ecPoint-Rainfall.”* |
| Observations used | 12-hourly rainfall observations from 310 rain gauges (**Fig.7j**), provided by the Hungarian Meteorological Service Unit of Informatics Applications. |
| Scores used | Talagrand diagram to test ecPoint-Rainfall reliability.  New score (see Tóth and Ihász (2021) for a detailed description) to test ecPoint-Rainfall’s ability to predict extreme localized rainfall. The aim was to define which percentile provides better guidance in the prediction of extreme localized rainfall by comparing forecasts (for a specific percentile) and observations and estimating the overall under-/overestimation for a specific rainfall event. OMSZ considered a rainfall event exceeding 15 mm/12h and tested the 85th and 95th percentiles. |
| Outcomes | Talagrand diagrams (**Fig.7k**):  *“The flatter Talagrand diagrams for ecPoint-Rainfall suggest that it provides more reliable forecasts than ECMWF ENS. The Talagrand diagram for ENS has a clear L shape, suggesting a consistently underestimation and overestimation of small and large rainfall values, respectively. However, ecPoint-Rainfall members show probabilities of occurrence between 2 and 4 %, suggesting low systematic failures in their reliability.”*  New score (**Fig.7l-m**):  *“The ecPoint-Rainfall 85th percentile provides the most accurate prediction of extreme localized rainfall events, whilst higher percentiles tend systematically overestimate the actual measured values. Indeed, it is worth highlighting that higher percentiles often tend to overestimate rainfall observations.”* |

**Table 8** - Summary of objective verification analysis from the real-time phase

By using percentiles lower than the 90th, there is a chance to observe, at a point, a much more extreme event than the one predicted. Such an event will be referred to as the “reasonable worst-case scenario” from now on. Using such low percentiles could result in missing a high impact event. Instead, products A (percentiles, **Table 9** - first row) and B (probabilities of exceeding a certain threshold, **Table 9** - second row) can be combined: product A shows the probabilities of exceeding the rainfall value that forecasters know can generate some impacts in the region of interest, and product B shows what could be the “reasonable worst-case scenario” (i.e., rainfall values associate with the ecPoint-Rainfall 99th percentile). The threshold of 50 mm/12h was chosen to create the prototype product A in **Table 9**, and the 99th percentile was chosen to create the prototype product B.

**Fig. 6** – Rainfall observations and forecasts for IMN case study. The top half of the figure shows panels from (a) to (f) that contain the 12-hourly rainfall observations for the extreme rainfall event occurred between October 3rd panel and 5th, 2018. The bottom half of the figure shows the forecasts’ evolution (from day 7 to day 1) for the rainfall event occurred between October 4th at 12am and October 5th at 0am (local time, LT). From the left, the first column shows the 85th percentile for ecPoint-Rainfall, the second column shows the 85th percentile for the raw ENS, the third column shows the 99th percentile for ecPoint-Rainfall, the fourth column shows the wettest member of the raw ENS, and the fifth column shows the deterministic forecast from WRF-1.5 (spatial resolution of 1.5 km) typically used by IMN in the forecasts of extreme localized rainfall events. ECMWF ENS and ecPoint-Rainfall forecasts correspond to runs at 12 UTC (i.e. first available run from Europe to IMN forecasters in the morning due to time difference between Europe and America). WRF-1.5 forecasts correspond to runs at 18 UTC (first run available to IMN forecasters in the morning). The forecasts’ colour scheme has been modified compared to the original IMN products (**Fig. 5a**), to standardize all the plots in this figure, and make easier the forecasts. The times to which observations and forecasts refer to are indicated in UTC time and local time (LT).

“This product [combination of prototype product A and B] could help us to forecast an event that can already generate an impact in the Pacific coast but at the same time, have an idea what could be the local maxima. For example, the Nicoya peninsula was affected by much higher rainfall, but we did not see it in the 85th percentile.” (IMN)

Immagine che contiene mappa

Descrizione generata automaticamenteForecasters at OMSZ developed a verification methodology to identify which percentile could be used to provide guidelines on extreme localized rainfall events (Tóth & Ihász, 2021). OMSZ estimated that an event above 15 mm/12h could be considered extreme. OMSZ found that the 85th percentile provided the best performance as there is no under or overestimation of the rainfall events exceeding 15 mm/12h (see **Fig. 7l**).

**Fig. 7** - Panels (a) to (i) show the results of the case study carried out by OMSZ. Panel (a) shows the 12 hourly observations for the intense rainfall event on June 11th, 2018 between 0 and 12 UTC time. The purple circle with a cross refers to a manually added record of 92 mm/24h in Bükkszentlélek between 12 UTC June 10th and 12 UTC June 11th (most of the rain fell on June 11th between 0 and 12 UTC). Panels (b), (c), (d), (e) show forecasts for day 2; (b) and (d) show the probabilities of not exceeding 10 mm/12h for ecPoint-Rainfall and ECMWF ENS, respectively; (c) and (d) show the probabilities of not exceeding 30 mm/12h for ecPoint-Rainfall and ECMWF ENS, respectively. The inserted box shows the CDF for ecPoint-Rainfall (in blue) and ECMWF ENS (in red) for day 2 rainfall forecasts for Bükkszentlélek. Panels (f), (g), (h), (i) are the same but for day 4 forecasts. Panels (j) to (m) show the results of the objective verification for ecPoint-Rainfall forecasts carried out by OMSZ. The verification period goes from June 1st to August 31st, 2018. Panel (j) shows the relative position between OMSZ rain gauges (coloured circles) and ecPoint-Rainfall and ECMWF ENS grid-boxes (grey squares). The colours associated to the rain gauges indicate their heigh in meters above sea level. Panel (k) shows the Talagrand diagrams for ecPoint-Rainfall (left column) and ECMWF ENS (right column). The first and the second row correspond, respectively, to forecasts for the accumulation period (t+0,t+12; i.e. day 1) and (t+96,t+108; i.e. day 4). Panel (l) shows the overall performance of ecPoint-Rainfall in the prediction of rainfall exceeding 15 mm/12h using the 85th percentile (top) and 95th percentile (bottom). The number of days with overall good performance for each grid-box (i.e. days with no overall rainfall under- or overestimation) are coloured in shades between white and blue. Black indicates grid-boxes with no associated observations. Grey indicates grid-boxes with overall rainfall over- or underestimation. Panel (m) shows the overall rainfall under- and overestimation of ecPoint-Rainfall in the prediction of rainfall exceeding 15 mm/12h using the 85th percentile (top) and 95th percentile (bottom). The number of days with overall under- and overestimation for each grid-box are coloured in shades of blue and orange, respectively. Black indicates grid-boxes with no associated observations. White indicates grid-boxes with no overall rainfall under- or overestimation.

The methodology developed at OMSZ employs the probabilistic information provided by ecPoint-Rainfall (or any other ensemble forecast) in a deterministic way as it defines a point in a two-dimensional space (given by the combination of a percentile and a rainfall threshold), allowing only for a yes or no answer to the question “Did the X percentile exceed Y rainfall threshold?”. Whilst this approach can be useful at the early warning stages to identify the locations that could be affected by extreme localized rainfall, it does not exploit all the information contained in the full ecPoint-Rainfall distribution which might suggest different actions based on the impacts of the different probability structures of exceeding a rainfall threshold. **Fig. 8** conceptualizes this point. The cumulative distribution function (CDF) (a) represents a typical convective rainfall event that would satisfy the criterion to issue a warning for extreme localized rainfall. Whilst the CDFs (b), (c), and (d) would also satisfy this criterion (i.e., 15 mm/12h are obtained at smaller percentiles than the 85th), they have different probability structures for rainfall values greater than 15 mm/12h. For example, (b) does not show any probabilities of having a rainfall event much higher than 15 mm/12h , while (d), although is very similar to (b) around the 85th percentile, it shows some (although small) probabilities of having a much more extreme localized rainfall event. This is important information for a forecaster because the impacts of the rainfall event in (d) can be much higher than those for the rainfall event represented by (b), even if the probabilities are very small (i.e., around 1 or 2%). The CDF (c) represents another possible rainfall event, this time with high probabilities of having greater rainfall on average at the grid-box scale than (b) and (d), although the tail is not as big as the tail in (d). Finally, the CDF (e) represents an event that would have not triggered any warning. However, it shows some chance of a very extreme localized rainfall event, and therefore its impacts could be severe. All this information would be lost if considering only one percentile and one rainfall threshold. If a meteorological service does not have a requirement for this type of information (e.g. no action can be taken on very small probabilities), there would be scope to explore with their users how to exploit this additional low probability but potentially Chart

Description automatically generatedhigh-impact information.

**Fig. 8** - Conceptual CDFs for possible ecPoint-Rainfall output scenarios (indicated by the colour shades). The CDF (A) corresponds to a typical convective rainfall event. Assuming the criterion for issuing warnings for extreme localized rainfall is having at least 15% probability of exceeding 15 mm/12h, the red point at the interception of the 85th percentile and 15 mm/12h provides the lower limit. The solid-line CDFs (i.e., A, B, C, and D) correspond to those rainfall events that would satisfy the criterion (i.e., 15 mm/12h is obtained at lower percentiles than the 85th); the dashed-line CDFs correspond to those rainfall events that would not (i.e., 15 mm/12h is obtained at higher percentiles than the 85th).

Moreover, previous case study analysis conducted at ECMWF suggest that the raw ENS and ecPoint-Rainfall’s CDFs tend to cross around the 85th percentile. This can be also seen in IMN’s case study by comparing the forecasts of the 85th percentile of ecPoint-Rainfall (forecasts, first column) and the 85th percentile of the Raw ENS (forecasts, second column) in **Fig. 6**. We expect that the crossing point for other ensemble systems would be the same because this reflects a general property of point rainfall observation distributions around the world (provided the grid-scale biases in those systems are similar).

## On the underestimation of localized rainfall extremes by ecPoint-Rainfall

The observed rainfall patterns for the rainfall on October 4th in Costa Rica (**Fig. 6d**) typify a low-probability high-impact event: high spatial variability of rainfall totals and very localized extremes (Legg & Mylne, 2004). As described by the intermediaries in section 4.1, forecasters at IMN do not generally look at global NWP models to forecast these type of rainfall events due to their coarse resolution. From the informal discussions, it transpired that, even if IMN forecasters tested ecPoint-Rainfall to forecast extreme localized rainfall events, the post-processed forecasts were considered as global NWP models and for this reason, forecasters were not surprised when the product underestimated the highest rainfall observations in the Nicoya peninsula. IMN intermediaries provided the forecasts from WRF-1.5, typically used at IMN to forecast extreme localized rainfall events to demonstrate that km-scale models (1.5 km spatial resolution) are better at forecasting these type of rainfall events than global models.

“ecPoint-Rainfall has an 18 km resolution. That is too coarse. Global models do not handle convection properly, and do not represent the real orography. To forecast localized rainfall in Costa Rica, which is a small country with complex orography, we need very high-resolution forecasts, otherwise the forecasts do not reach the observed values.” (IMN)

**Table 9** - Prototype name, guidelines and example of products that could be derived from ecPoint-Rainfall

|  |  |
| --- | --- |
| **Prototype product name and accompanying guidelines** | **Product example** |
| 1. Map plot of the probabilities of exceeding a rainfall threshold Y [in mm/accumulation period]   It provides users with an idea of how likely it is to observe the event of interest. The rainfall threshold can be provided (1) by the experience of the local forecaster who knows what rainfall event might generate some impacts in the region of interest, or (2) from a (observational or model) rainfall climatology in case the forecaster doesn’t know what rainfall event might generate some impacts. This product does not provide any information on whether larger rainfall events are likely to be observed, and with which likelihood. | Immagine che contiene mappa  Descrizione generata automaticamente |
| 1. Map plot for an Xth percentile   ecPoint aims to anticipate sub-grid variability and provide a wider distribution of possible rainfall values. The tail of ecPoint-Rainfall’s distribution should capture those low-probability, extreme localized rainfall events (referred to as the “reasonable worst-case scenario”). Therefore, it is recommended to plot high percentiles, ideally the 98th or 99th (i.e., 1 in 20 or 100 event, respectively). Lower percentiles would not catch the “reasonable worst-case scenario”. Furthermore, case study analysis has shown that ecPoint-Rainfall’s 85th percentile tends to be similar to ENS. Particular attention should be paid to the fact that although the raw forecasts and ecPoint-Rainfall are provided with the same grid resolution, they do not refer to the same spatial scale. The first one provides rainfall averages over the grid-box, whilst ecPoint-Rainfall provides rainfall forecasts at a point within the grid-box, which can be much higher than the grid-box average in case of high rainfall sub-grid variability. However, nothing can be said on where that point is within the grid-box. | Immagine che contiene mappa  Descrizione generata automaticamente |
| 1. Cumulative Distribution Function (CDF) of all percentiles   The product shows the CDF for raw ENS (in red) and ecPoint-Rainfall (in blue). The spread of the CDF allows one to identify the range of rainfall values that could be observed. Second, the shape of the CDF allows the user to identify what rainfall values are more likely to be observed. For example, a more vertical CDF would typically indicate a more confident forecast. | Diagram  Description automatically generated |

Even though the post-processed forecasts are provided in a grid with the same spatial resolution as the raw forecasts, ecPoint-Rainfall provides forecasts at point-scale, and not average at grid scale. This is not clear from the guidelines and should be stated more strongly. The reason why ecPoint-Rainfall did not show forecast values of the same magnitude of those observed in the Nicoya peninsula is because the product developed by IMN was based on the ecPoint-Rainfall 85th percentile (**a**).

The ecPoint-Rainfall 99th percentile maps (**Fig. 6**, forecasts, third column) were compared with the raw ECMWF ENS (**Fig. 6**, forecasts, fourth column) and the forecasts from WRF-1.5 (**Fig. 6**, forecasts, last column), and showed that from day 4, ecPoint-Rainfall provides a more consistent guidance on the location of the areas (i.e. Pacific cost) at higher risk of localized extreme rainfall; from day 2, ecPoint-Rainfall also provides a better signal (i.e. better location and magnitude of the extreme rainfall event) than the raw ENS and WRF-1.5 which still appear noisier and jumpier. Moreover, the raw ENS underestimated the rainfall observations by ~50% also at day 1, whilst WRF-1.5 significantly overestimated them. IMN intermediaries reported indeed that, from experience, forecasters know the must reduce the raw WF-1.5 forecasts by a factor of 0.4. ecPoint-Rainfall provides useful guidance even for localise heavy rainfall, but this is only captured by the highest percentile maps.

“The map plots for the 99th percentile provide forecasts that are very similar to the rainfall amounts shown on our WRF-1.5, and it is easy to interpret the two outputs in the same way. However, we realize now they are not. The documentation should tackle better the issue regarding the difference in resolution between ecPoint-Rainfall and raw NWP models to avoid the misinterpretation of the forecasts.” (IMN)

## On the misplacement of the “wettest day” in a multi-day event

In the following section, “wettest day” means the day when the highest local peak was observed. IMN reported that the wettest day was October 4th because the observations recorded the rainfall peak (309.2 mm/12h) on that day. It was also reported that the 85th percentile of ecPoint-Rainfall indicated that October 5th would have been the wettest day instead, i.e., the day with the highest rainfall observation.

The rainfall totals are more uniformly distributed over the Nicoya peninsula on October 5th (**Fig. 6e**) than on October 4th when there was much more variability in the rainfall totals (there are very small rainfall totals, although the rainfall peak (309.2 mm) was recorded on October 4th (**Fig. 6d**)). Therefore, the average rainfall is higher on October 5th than October 4th. Since IMN used the 85th percentile for ecPoint-Rainfall and the rainfall distributions of ecPoint-Rainfall and ENS tend to cross at such a percentile (as highlighted in section 5.1), IMN saw what they would have seen by inspecting the ECMWF ENS, i.e., that, on average, the Nicoya peninsula would have been wetter on October 5th than on October 4th because the average rainfall was higher on the first day compared to the second. Therefore, IMN arrived at the right conclusion (i.e. predicting that October 5th would have been wetter - on average) but from wrong premises (i.e. expecting to see the highest rainfall peak on October 5th). Thus, no information could have been provided on when the highest local peak could have been observed.

To provide guidance on when the highest local rainfall event might be observed, product (C) in **Table 9** is most appropriate. The prototype product represents the CDFs for the ecPoint-Rainfall and raw ENS distributions for a location in the south coast and inland parts of the Nicoya peninsula on October 4th and 5th. The conclusions that can be drawn from the comparison of the CDFs are twofold and are related to the spatial scale that ecPoint-Rainfall and the raw ENS refer to. First, the raw ENS CDFs suggest that a big range of (grid-scale) rainfall totals (from 0 to 200 mm/12h) could be observed over the Nicoya peninsula on both days. However, a bigger area to the left of the raw ENS CDF (which is proportional to the ensemble mean), for both coast and inland parts of the peninsula, suggests that bigger rainfall averages at grid-scale should be expected on October 5th rather than on October 4th. This is supported in the observations (**Fig. 6d-e**). Second, ecPoint-Rainfall CDFs suggest that the rainfall variability at point-scale can be much higher (from 0 to 450 mm/12h) than the one observed at grid-scale. So, although with low probabilities (e.g., 1% or 2%), severe localized events could be observed in the Nicoya peninsula, either on October 4th or 5th. However, ecPoint-Rainfall CDFs suggest that the probabilities of having an extreme localized event is very similar on both days, therefore, on the basis of the CDFs no conclusion could be drawn on which day would be the wettest.

The fact that the highest rainfall peak was observed on October 4th instead than on October 5th could be only a matter of chance. Only part of the sub-grid scale variability of the precipitation field can be known from observations, so the verification of extreme localized rainfall events in small regions is difficult. Denser observational networks or reliable radar-derived rainfall totals would provide a better representation of the rainfall sub-grid variability in small regions (Haiden & Duffy, 2016), so they are vital for the evaluation of ecPoint-Rainfall in the prediction of extreme localized rainfall events in small regions. Costa Rica has a good observational network. However, **Fig. 6a-f** show that there are some uncovered spots, where a higher peak on October 5th could have happened but would not be recorded. Therefore, ecPoint-Rainfall guidelines should tackle the fact that no deterministic conclusions can be drawn about ecPoint-Rainfall correctly placing the wettest day in a multiple-day event, especially in the verification of small regions. Whether ecPoint-Rainfall provides good guidance on when the highest local rainfall peak is going to happen is information that should be seen either over a much longer period of time and over a much bigger region, so the effects of the random measurement of rainfall extremes (and the fact that ecPoint-Rainfall output is probabilistic) can be counterbalanced by the bigger number of cases examined.

## On the verification of ecPoint-Rainfall

OMSZ developed a methodology to verify ecPoint-Rainfall based on assigning only one observation to the nearest model grid-box. It was assumed that the closer the observation to the model grid-box, the more representative the observation is of the forecast provided for that grid-box. This assumption might be relevant for traditional NWP model outputs, where the model grid-box forecast represents a rainfall average over the grid-box. However, in the case of ecPoint-Rainfall, the verification is much simpler. By construction, ecPoint-Rainfall mirrors what is provided by rain gauges, i.e., point rainfall observations. Therefore, verification statistics must be calculated by using point rainfall observations and the forecasts from the nearest model grid-box, even if the forecast from the same model grid-box gets assigned to more than one observation. Second, the verification methodology merges adjacent model grid-boxes to assess whether the forecasts provided a good guidance compared to observations. This is another very common technique used for the verification of traditional NWP model outputs to compensate for the lack of ensemble members. The merging of adjacent model grid-boxes is not necessary when verifying ecPoint-Rainfall which, by construction, already expands the number of ensemble members for each raw member (see **Fig. 1b**).

## On using parallels between ecPoint and forecasters heuristics to improve guideline efficacy

Whilst the parallels between the ecPoint methodology and forecasters heuristics did not directly change the way ecPoint-Rainfall guidelines were articulated, explaining to forecasters such parallels proved to be a turning point in the way guidelines were understood. The importance of such parallel came out during the “informal discussions” and were recalled when the mock-up products were sent for a second evaluation by the forecasters. Drawing a direct link between ecPoint methodology and the way forecasters provide Bayesian probabilities for an event on daily basis removed the complexity factor and opened up the forecasters to a wider understanding of how ecPoint post-processes raw NWP model outputs which translated into a more effective use of the products presented. This discovery is an important point to include in the ecPoint-Rainfall guidelines. Indeed, every forecaster could relate to forecasters heuristics, independent of their previous probabilistic knowledge, because whether they work in a deterministic or a probabilistic (ensemble) framework, forecasters always provide some sort of adjustments to the raw model guidance. Typically forecasters working with deterministic models have to provide a probabilistic distribution around the single value forecast, which is what an NWP ensemble would provide in a numerical way (Buizza, 2019); and forecasters working with NWP ensembles have to adjust such distributions with some sort of calibration in mind as these ensembles are typically biased or under-spread (Gneiting & Katzfuss, 2014).

# Discussions and concluding remarks

This article discusses the outcomes of a pilot study carried out with IMN and OMSZ to identify weaknesses in ecPoint-Rainfall training guidelines to help forecasters make a more efficient use of the complex but intrinsically skilful post-processed product. Specifically, this study aimed to establish how the guidelines should be tailored to favour widespread and scientifically correct adoption of the new product, across a diverse range of users (in particular NHMSs that issue warnings for extreme localized rainfall events), with the ultimate goal of improving the quality and effectiveness of rainfall warnings. Findings are drawn from the participation of two NHMSs in the pilot study, and so may not be generalizable across all potential user communities. However, the two participants were significantly different in several aspects. IMN and OMSZ are located in different geographical regions, with tropical (Costa Rica) and continental (Hungary) climates, respectively. The two services also have a different level of operational experience producing forecasts and warnings with NWP ensembles: IMN still works mainly with deterministic models, whilst OMSZ could be fairly described as an “early adopter” of the ensemble approach (since the 2000s). Consequently, we consider that conclusions from this pilot study should have some worldwide validity.

ecPoint-Rainfall was found to provide skilful forecasts for extreme localized rainfall events in diverse geographical regions. This is naturally a pre-requisite for further activities, and complements the more general verification results for the globe, and for tropical and extratropical sub-regions, discussed by Hewson & Pillosu (2021). However, further work on focussed regional verification would be beneficial.

The guideline documentation about ecPoint and ecPoint-Rainfall provided to users must be markedly improved because the potential for misinterpretation of the forecast products, by practising forecasters, was still quite high. Guidelines can be improved by drawing a direct link between the ecPoint methodology and forecasters heuristic approach to providing Bayesian probabilities for an event. Such an analogy can remove the complexity factor around the ecPoint methodology and open forecasters to a wider understanding of how ecPoint post-processes raw NWP model outputs. Indeed, every forecaster can relate to forecaster heuristics, independently of whether they are used to working with deterministic or probabilistic NWP guidance, or whether they have previous probabilistic knowledge, because forecasters ordinarily impart some sort of adjustments to raw model output based on various factors such as experience. It is part of their job. ecPoint developers will thus include in the guidelines an infographic like **Fig.2** that draws such parallels. Such presentations can help to improve the processing of complex information since human brains manage visual information better than text, and infographics are indeed a great way to land links (Yarbrough, 2019).

The ecPoint guidelines will now stress that whilst ecPoint-Rainfall output is provided on the same grid as the raw forecast, it provides forecasts for a point within the grid-box whilst raw NWP models provide values averaged to the grid-box resolution. Then we will highlight that to provide better guidance for low-probability, high-impact events, called “a reasonable worst-case scenario”, users should consider using high percentiles (i.e., 98th or 99th). Whilst such events are by definition in the tail of the distribution (Legg & Mylne, 2004) by not using the highest percentiles the most impactful events will likely be under-predicted; users did not fully grasp this aspect so more emphasis is clearly needed. Indeed, the high ecPoint percentiles do not ordinarily overestimate the rainfall since, when sub-grid variability is high, local maxima, which they tend to represent, will anyway be much higher than average rainfall over the grid-box (predicted by raw NWP models). It was clear from section 5.1 that both NHMSs used lower percentiles (i.e., 85th) to provide guidance on extreme localized rainfall. The reality is that rainfall values thus delivered will be too low. A key point is that on a given day it is much more likely that a high percentile ecPoint value will be exceeded *somewhere* in the grid-box than it is that a high percentile grid-box average forecast will be exceeded over the grid-box as a whole (i.e., observed grid-box average). This has not been highlighted before and seems to not be recognised. ecPoint developers will include attention-getters, like info-boxes, to highlight such issues in the revised guidelines.

ecPoint-Rainfall guidelines will be further improved by adding more example use-cases (in video or written format) and by creating an online Q&A forum. The first option allows for real-world practical applications to be highlighted, and the scientific background prominence to be relegated, which recommendations in other studies have shown should increase uptake. Whilst ecPoint developers did not try specifically video user-cases in this study, the direct verbal contact with IMN and OMSZ intermediaries proved to be fundamental for explaining multiple aspects and led to the conclusion that pre-recorded video-format use-case discussions can be an efficient communication tool. Forecasters are ordinarily under considerable time pressures and need efficient information dissemination (Novak *et al.*, 2008; Bullón & Viana, 2018; Demuth *et al.*, 2020) and given that more direct verbal interaction (as used in this study) cannot be implemented widely, video-formats probably offer the best solution. Q&A forums can also be an extremely helpful, scalable resource for more direct contact with users, especially those with less relevant probabilistic experience who likely require closer interaction to be convinced of the benefits of a new approach (Bullón & Viana, 2018).

Although not so directly critical for forecaster interpretation, in revised user guidelines we will also briefly discuss how ecPoint forecasts can be correctly and usefully verified (pursuant to Section 5.4). We are also investigating provision of real-time CDF-style products (Table 9, row C) to exploit their comprehensive information content for sites.

In summary, this study has been pivotal in highlighting how fundamental it is to properly educate each and every class of user when introducing new concepts and products into operational forecasting. There is a danger of misusing or misinterpreting the forecasts which potentially has adverse impacts downstream of the users of the forecasts, e.g., inappropriate warning issues or warnings missed. For ecPoint-Rainfall output, the potential pitfalls of incomplete training were in some respects similar for user groups with more or less experience of probabilistic forecasting, but in other respects were different. For experienced users the danger of inappropriately applying previous knowledge may be paramount, whilst for inexperienced users the main difficulty is potentially the leap in understanding required. We conclude this study with insightful feedback from our two user groups, that highlights how training guideline upgrades are vital for the widespread uptake of ecPoint forecasts:

“If all these changes are applied to the way the new ecPoint-Rainfall forecasts are presented to us, forecasters, the new product would be extremely useful to forecast extreme localized rainfall events.” (IMN)

“ecPoint-Rainfall could be a consequential tool for the early detection of such localized extreme rainfall events, but guidelines need to be clearer about what the product exactly means.” (OMSZ)

As short-term follow-ups from this study, the ECMWF Forecast User Guide (Owens and Hewson, 2018) and the face-to-face forecaster training offered by ECMWF will be updated in accordance with the findings of this study whilst also setting up a channel in the ECMWF Forecast User Forum for more direct information exchange with users. As a long-term follow-up, it would be beneficial to assess the extent of any positive benefits in the training material changes.

# Acknowledgements

The authors would like to thank David Richardson for their time providing constructive comments and feedback that have led to a much more improved article. The research leading to these results has received funding from the Copernicus program of the European Commission ([https://www.copernicus.eu](https://www.copernicus.eu/)). The authors declare also that they have no conflict of interest.

# Orcid

***Fatima M. Pillosu*** <https://orcid.org/0000-0001-8127-0990>

***Tim Hewson*** <https://orcid.org/0000-0002-3266-8828>

***Christel Prudhomme*** <https://orcid.org/0000-0003-1722-2497>

***Elisabeth Stephens*** <https://orcid.org/0000-0002-5439-7563>

***Hannah L. Cloke*** <https://orcid.org/0000-0002-1472-868X>

# References

Buizza R. 2018. Ensemble Forecasting and the Need for Calibration. In: *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, 15–48.

Buizza R. 2019. Introduction to the special issue on “25 years of ensemble forecasting.” *Q. J. R. Meteorol. Soc.* **145**:1–11.

Bullón JC., Viana S. 2018. La predicción operativa y el papel del predictor. In: *Física del caos en la predicción meteorológica*. Madrid: Agencia Estatal de Meteorología, 29–46.

Demeritt D., Nobert Ś., Cloke H., Pappenberg F. 2010. Challenges in communicating and using ensembles in operational flood forecasting. *Meteorol. Appl.* **17**:209–222.

Demeritt D., Nobert S., Cloke HL., Pappenberger F. 2013. The European Flood Alert System and the communication, perception, and use of ensemble predictions for operational flood risk management. *Hydrol. Process.* **27**:147–157.

Demuth JL., Morss RE., Jankov I., Alcott TI., Alexander CR., Nietfeld D., Jensen TL., Novak DR., Benjamin SG. 2020. Recommendations for developing useful and usable convection-allowing model ensemble information for NWS forecasters. *Weather Forecast.*

Doswell CA. 2004. Weather forecasting by humans - Heuristics and decision making. *Weather Forecast.* **19**:01115–01126.

Doswell CA., Brooks HE., Maddox RA. 1996. Flash flood forecasting: An ingredients-based methodology. *Weather Forecast.* **11**:560–581.

Emerton R., Cloke H., Ficchi A., Hawker L., de Wit S., Speight L., Prudhomme C., Rundell P., West R., Neal J., Cuna J., Harrigan S., Titley H., Magnusson L., Pappenberger F., Klingaman N., Stephens E. 2020. Emergency flood bulletins for Cyclones Idai and Kenneth: A critical evaluation of the use of global flood forecasts for international humanitarian preparedness and response. *Int. J. Disaster Risk Reduct.* **50**:101811.

Evans C., Van dyke DF., Lericos T. 2014. How do forecasters utilize output from a convection-permitting ensemble forecast system? Case study of a high-impact precipitation event. *Weather Forecast.* **29**:466–486.

Fundel VJ., Fleischhut N., Herzog SM., Göber M., Hagedorn R. 2019. Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers and end-users. *Q. J. R. Meteorol. Soc.* **145**:210–231.

Gneiting T., Katzfuss M. 2014. Probabilistic Forecasting. *Annu. Rev. Stat. Its Appl.* **1**:125–151.

Göber M., Zsótér E., Richardson DS. 2008. Could a perfect model ever satisfy a naïve forecaster? On grid box mean versus point verification. *Meteorol. Appl.* **15**:359–365.

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

Haiden T., Duffy S. 2016. Use of high-density observations in precipitation verification. *ECMWF Newsl.*:20–25.

Harding J. 2018. *Qualitative data analysis: From start to finish*. Sage Publications.

Hewson T. 2020. *Use and Verification of ECMWF products in Member and Co-operating States (2019)*.

Hewson TD., Pillosu FM. 2021. A new low-cost technique improves weather forecasts across the world. *Commun. Earth Environ.* **2**:132.

Hoegh-Guldberg O., Jacob D., Taylor M., Guillén Bolaños T., Bindi M., Brown S., Camilloni IA., Diedhiou A., Djalante R., Ebi K., Engelbrecht F., Guiot J., Hijioka Y., Mehrotra S., Hope CW., Payne AJ., Pörtner HO., Seneviratne SI., Thomas A., Warren R., Zhou G. 2019. The human imperative of stabilizing global climate change at 1.5°C. *Science (80-. ).* **365**.

Joslyn S., LeClerc J. 2013. Decisions With Uncertainty: The Glass Half Full. *Curr. Dir. Psychol. Sci.* **22**:308–315.

Kahneman D., Tversky A. 1973. On the psychology of prediction. *Psychol. Rev.* **80**:237–251.

Lavers DA., Harrigan S., Prudhomme C. 2021. Precipitation Biases in the ECMWF Integrated Forecasting System. *J. Hydrometeorol.*

LeClerc J., Joslyn S. 2015. The cry wolf effect and weather-related decision making. *Risk Anal.* **35**:385–395.

Legg TP., Mylne KR. 2004. Early warnings of severe weather from ensemble forecast information. *Weather Forecast.* **19**:891–906.

Losee JE., Joslyn S. 2018. The need to trust: How features of the forecasted weather influence forecast trust. *Int. J. Disaster Risk Reduct.* **30**:95–104.

Morss RE., Demuth JL., Lazo JK. 2008. Communicating Uncertainty in Weather Forecasts: A Survey of the U.S. Public. *Weather Forecast.* **23**:974–991.

Morss RE., Mulder KJ., Lazo JK., Demuth JL. 2016. How do people perceive, understand, and anticipate responding to flash flood risks and warnings? Results from a public survey in Boulder, Colorado, USA. *J. Hydrol.* **541**:649–664.

Murphy AH. 1993. What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather Forecast.* **8**:281–293.

Murphy AH., Winkler RL. 1977. Can weather forecasters formulate reliable probability forecasts of precipitation and temperatures? *Natl. Weather Dig.* **2**:2–9.

Nielsen ER., Schumacher RS. 2016. Using convection-allowing ensembles to understand the predictability of an extreme rainfall event. *Mon. Weather Rev.* **144**:3651–3676.

Novak DR., Bright DR., Brennan MJ. 2008. Operational forecaster uncertainty needs and future roles. *Weather Forecast.* **23**:1069–1084.

Owens RG., Hewson T. 2018. *ECMWF Forecast User Guide*.

Pillosu F., Hewson T. 2017. New point-rainfall forecasts for flash flood prediction. *ECMWF Newsl.* **153**.

Pillosu F., Modigliani U., Magnusson L., Calvelo MB., Sterponi L., Ramos MH., Valderrama P. 2017. ECMWF supports flood disaster response in Peru. *ECMWF Newsl.* **152**.

Roberts N. 2008. Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. In: *Meteorological Applications*. 163–169.

Roberts NM., Lean HW. 2008. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Mon. Weather Rev.* **136**:78–97.

Sills DML. 2009. On the MSC forecasters forums and the future role of the human forecaster. *Bull. Am. Meteorol. Soc.* **90**:619–627.

Stuart NA., Market PS., Telfeyan B., Lackmann GM., Carey K., Brooks HE., Nietfeld D., Motta BC., Reeves K. 2006. The future of humans in an increasingly automated forecast process. *Bull. Am. Meteorol. Soc.* **87**:1497–1502.

Tóth B., Ihász I. 2021. Validation of subgrid scale ensemble precipitation forecasts based on ECMWF’s ecPoint Rainfall project. *IDŐJÁRÁS (pre-print)*.

Tversky A., Kahneman D. 1974. Judgment under uncertainty: Heuristics and biases. *Science (80-. ).* **185**:1124–1131.

UNDRR. 2020. *The human cost of disasters: an overview of the last 20 years 2000-2019*.

Vannitsem S., Bremnes JB., Demaeyer J., Evans GR., Flowerdew J., Hemri S., Lerch S., Roberts N., Theis S., Atencia A., Bouallègue Z Ben., Bhend J., Dabernig M., de Cruz L., Hieta L., Mestre O., Moret L., Plenković IO., Schmeits M., Taillardat M., van den Bergh J., van Schaeybroeck B., Whan K., Ylhaisi J. 2021. Statistical postprocessing for weather forecasts review, challenges, and avenues in a big data world. *Bull. Am. Meteorol. Soc.* **102**:E681–E699.

Ward PJ., Blauhut V., Bloemendaal N., Daniell EJ., De Ruiter CM., Duncan JM., Emberson R., Jenkins FS., Kirschbaum D., Kunz M., Mohr S., Muis S., Riddell AG., Schäfer A., Stanley T., Veldkamp IET., Hessel WC. 2020. Review article: Natural hazard risk assessments at the global scale. *Nat. Hazards Earth Syst. Sci.* **20**:1069–1096.

Wilson KA., Heinselman PL., Skinner PS., Choate JJ., Klockow-McClain KE. 2019. Meteorologists’ interpretations of storm-scale ensemble-based forecast guidance. *Weather. Clim. Soc.* **11**:337–354.

Yussouf N., Jones TA., Skinner PS. 2020. Probabilistic high-impact rainfall forecasts from landfalling tropical cyclones using Warn-on-Forecast system. *Q. J. R. Meteorol. Soc.*:qj.3779.

# Appendix A – Guide questions for the informal discussions during the “offline” phase

**BACKGROUND QUESTIONS (PREGUNTAS DE CONTEXTO)**

1. Has the Met-Service any experience with ensemble forecasts? If so, which ensemble forecasts are mainly used?

¿El servicio meteorológico tiene experiencia con pronósticos de conjunto? Si es así, ¿qué pronósticos de conjunto utilizan principalmente?

1. If there is some access to ensemble forecasts in the Met-Service, how are they used? Are they used as the primary source to issue alerts and create products for end-users and inform them also about the uncertainty on the forecast? Or are they used as a background knowledge to complement the information provided by a deterministic model, and if so, why?

Si tiene acceso a pronósticos de conjunto, ¿cómo los usan? ¿Los usan como fuente primaria para emitir alertas y crear productos para usuarios finales e informarlos también sobre la incertidumbre en los pronósticos? ¿O los usan para complementar la información proporcionada por un modelo determinista? Si es así, ¿por qué?

1. What is the general impression of ensemble forecasts and their use in operational environments? Is there any internal disagreement on the practical value of ensemble forecasts (e.g., due to issues in the communication of probabilistic forecast or their reception by end-users)? Is there any discomfort surrounding how to deal with probabilistic forecasts in an operational environment?

¿Cuál es la impresión general sobre las predicciones de conjunto y su uso en entornos operativos? ¿Existe algún desacuerdo interno sobre el valor práctico de los pronósticos de conjunto (por ejemplo, debido a problemas en la comunicación del pronóstico probabilístico o su recepción por parte de los usuarios finales)? ¿Hay alguna molestia sobre cómo lidiar con pronósticos probabilísticos en el entorno operativo?

1. Has the Met-Service any experience of post-processing or calibrating rainfall forecasts? If so, for what purpose (e.g., improving quality of operational forecasts for end-users, making the forecasts more suitable for downstream applications such as hydrological forecasts)?

¿El servicio meteorológico tiene experiencia con el postproceso o calibración de pronósticos de lluvia? Si es así, ¿con qué propósito se hacen (por ejemplo, para mejorar la calidad de los pronósticos operativos para los usuarios finales, hacer que los pronósticos sean más adecuados para segundas aplicaciones como pronósticos hidrológicos)?

1. Who has been receiving ecPoint-Rainfall forecasts? What is their background (e.g., operational, research)?

¿Quién recibió los pronósticos de ecPoint-Rainfall? ¿Cuáles son sus antecedentes (por ejemplo, operativos, investigación)?

1. Do those particular people have general experience working with ensemble forecasts? Do they have experience working with post-processed forecasts? If not, do they have much time to devote to learning?

¿Las personas que recibieron las predicciones de ecPoint-Rainfall tienen experiencia de trabajo con predicciones de conjunto? ¿Tienen experiencia trabajando con predicciones calibradas? Si no es así, ¿tienen tiempo para dedicar al aprendizaje?

**QUESTIONS ON ECPOINT-RAINFALL (PREGUNTAS SOBRE ECPOINT-RAINFALL)**

1. Was ecPoint-Rainfall used operationally, experimentally, or for research?

¿Se usó ecPoint-Rainfall de manera operacional, experimental o para investigación?

1. Was it difficult to become accustomed to the meaning/structure of ecPoint-Rainfall forecasts? Did the fact that ecPoint-Rainfall products you received did not provide grid box forecasts create any issues?

¿Fue difícil acostumbrarse al significado o a la estructura de ecPoint-Rainfall? ¿El hecho que los productos de ecPoint-Rainfall que recibieron no proporcionan predicciones a escala de celda creó problemas?

1. If ecPoint-Rainfall was used operationally, were there any technical issues to integrate the forecasts in your operational workflows? Evaluation would include configuration of ecPoint-Rainfall, data volumes, run times, displaying the forecasts, etc.

Si se utilizó ecPoint-Rainfall operacionalmente, ¿hubieron problemas técnicos para integrar las predicciones en sus sistemas operativos? La evaluación incluiría configuración de ecPoint-Rainfall, volúmenes de datos, tiempos de ejecución, o representar gráficamente productos, etc.

1. Did you develop products from ecPoint-Rainfall? Did you use percentiles? Which percentiles? Why? Did you use probabilities? Which probabilities? Why?

¿Desarrollaron productos basados en ecPoint-Rainfall? ¿Usaron percentiles? ¿Cuáles percentiles? ¿Por qué? ¿Usaron probabilidades? ¿Qué probabilidades? ¿Por qué?

1. Where (e.g., over mountainous, coastal, flat areas), in which weather situations or for which type of events (e.g., deep convection, flash floods, etc.) do you think you could get most benefit from ecPoint-Rainfall? Why?

¿Dónde (por ejemplo, en zonas montañosas, costeras y planas), en qué situaciones climáticas, o para qué tipo de eventos (por ejemplo, convección, inundaciones, etc.) creen que podría obtener mayor beneficio de ecPoint-Rainfall? ¿Por qué?

1. Was ecPoint-Rainfall found useful? Do you think it added value to raw ECMWF ensemble and/or the model used in-house?

¿Se encontró ecPoint-Rainfall útil? ¿Te parece que añade valor a las predicciones de ECMWF y/o al modelo que utilizan comúnmente?

1. Do you think that ecPoint-Rainfall could change the way that alerts are issued for localized extreme rainfall, flash floods, etc? Perhaps increasing the lead-time at which alerts are issued (e.g., up to medium ranges)? Why?

¿Cree que ecPoint-Rainfall podría cambiar la forma en que se emiten las alertas para lluvias extremas localizadas, inundaciones repentinas, etc.? ¿Quizás podrían aumentar el plazo con el que se emiten las alertas (por ejemplo, hasta un plazo medio)? ¿Por qué?

1. Do you think ecPoint-Rainfall is useful information to have? If so, in which way, as preliminary information to raise internal awareness to prompt increase preparedness within the forecasting centre? Or would it also be used to trigger early actions to mitigate or manage high risk events?

¿Cree que ecPoint-Rainfall proporciona información útil? Si es así, ¿de qué manera? ¿Cómo información preliminar usada internamente para aumentar rápidamente la preparación el centro de predicción? ¿O también se usaría para activar acciones tempranas con el objetivo de mitigar o gestionar eventos de alto riesgo?

1. If you think ecPoint-Rainfall has improved raw model rainfall forecasts, based on your experience, what aspects stand out as being better (e.g., less false alarm rates, better representation of point rainfall values, etc)?

Si cree que, en base a su experiencia, ecPoint-Rainfall ha mejorado las predicciones de lluvia, ¿cuáles son lo que aspectos se destacan por ser mejores (por ejemplo, la menor frecuencia de falsas alarmas, una mejor representación de los valores puntuales de lluvia, etc.)?

1. Can you think of other useful applications for ecPoint-Rainfall (e.g., predicting dry weather)?

¿Tiene sugerencias para otra útil aplicaciones para ecPoint-Rainfall (por ejemplo, la predicción de no-lluvia)?

1. Currently the maximum percentile available is 99th (1 in 100 chance). We could in principle deliver up to percentile 99.98th (1 in 5000 chance), or, let’s say, we could even restrict the maximum available percentile to 95th (1 in 20 chance). What level do you think we should use as the maximum?

Actualmente el percentil máximo disponible es el 99° (1 entre 100 posibilidades). En principio, podríamos computar hasta el percentil 99,98° (1 en 5000 posibilidades) o podríamos restringir el percentil máximo al 95° (1 en 20 posibilidades). ¿Qué nivel cree que deberíamos usar como máximo?

1. Have you used/verified ecPoint-Rainfall for a particular event or case study? Briefly describe the geographical region and the weather conditions for which the tests were conducted.

¿Ha utilizado/verificado ecPoint-Rainfall para un evento o un caso de estudio en particular? Describa brevemente la región geográfica y las condiciones climáticas para las cuales se realizaron las pruebas.

1. Why was this particular case study or event chosen (e.g., the forecasts for the region or the particular synoptic situation are usually not very good)?

¿Por qué fue elegido ese particular caso de estudio o evento (por ejemplo, los pronósticos para la región o para la situación sinóptica considerada no son muy buenos por lo general)?

1. With hindsight, did ecPoint-Rainfall provide useful guidance for this particular situation? How would you rate its performance?

En retrospectiva, ¿ecPoint-Rainfall proporcionó información útil para esta situación en particular? ¿Como evaluaría su rendimiento?

1. Are there any improvements that you would like to see in ecPoint-Rainfall (e.g., 12-hourly rainfall accumulations were ok for your needs, or would you like to see other durations, etc.)?

¿Hay alguna mejora que le gustaría ver en ecPoint-Rainfall (por ejemplo, las acumulaciones de lluvia de 12 horas fueron adecuadas para sus necesidades o le gustaría ver otras acumulaciones, etc.)?

1. ecPoint family is a generic term covering post-processing applications for different variables using the ecPoint approach. Currently, ecPoint-Temperature is in undergoing development within the Highlander project (<https://highlanderproject.eu/wp-content/uploads/2021/07/D4.3_Downscaled_subseasonal_forecast_update20210716.pdf>) and ecPoint-Rainfall products are available on ecCharts and on the Mistral portal (<https://meteohub.mistralportal.it/app/maps/flashflood>) [↑](#footnote-ref-2)
2. ecCharts (<https://www.ecmwf.int/en/forecasts/eccharts>) is an interactive web-based service developed at ECMWF to explore and visualize tailored ECMWF graphical forecast products. [↑](#footnote-ref-3)
3. The “daily report” is a report generated on daily basis at ECMWF only for internal verification purposes. Main considered topics are about general or case-specific performance of ECMWF models. Whenever data is available, comparisons with other NWP models (global or regional) are also explored. [↑](#footnote-ref-4)