Evaluation of the performance of seasonal forecast models applied to the monitoring of water reserves: case study of a watershed in Northern Peru

Héloïse Lafargue[1] and Martin Espitalié[2]

[1]Department Big Data & AI, Imaging and multimedia, ENSEEIHT Engineering school
[2]Department Data Science and Meteorology, WeatherForce

Abstract

This study focuses on optimizing water resources thanks to meteorologic forecast, in northern Peru, an agriculturally vulnerable region prone to climatic extremes. The research evaluates seasonal forecasting models' performance to monitor water reserves, specifically implementing ECMWF's SEAS5 model for mid and long-term precipitation forecast and river flow forecast from GloFAS. Precipitation and flow rivers forecast trends are obtained for up to 5 months with good accuracy, and are easily accessible through anomaly maps. By integrating precipitation and river flow forecasts, the study makes it possible to enhance water stock management in northern Peru through improved forecasting and accessible visualization methods, contributing to more informed decision-making in the agricultural and water resource management sectors.

Keywords

Meteorology, Climatology, Hydrology, Anomaly, Seasonal models, Water Stock Managment, Peru, ECMWF Copernicus C3S, GloFAS.

1 Introduction

Peru's agricultural areas, which are climatologically arid (red area in fig.1), are highly susceptible to drought and flooding. Rainfall is therefore the meteorological element that requires the most monitoring, particularly in terms of irrigation management, especially in the context of global climate change and a region subject to fluctuations in the El Niño Southern Oscillation. The use of seasonal rainfall and river flow forecasts, along with their anomalies, provides a better long-term view of water reserve management. As a result, the WeatherForce team worked on a project in collaboration with ESA Space Solutions. The primary task was to evaluate the performance of seasonal forecasting models applied to monitoring water reserves, using watersheds of northern Peru as a specific case study. To address this objective, the overarching question was:

"How can meteorological insights be leveraged to enhance agricultural water stock management in Northern Peru?"

Methods based on the ECMWF SEAS5 seasonal precipitation forecasting model [1, 2, 3] demonstrate its efficacy within the diverse climates of Latin America, even though under the influence of the El Niño–Southern Oscillation [4, 5]. However, the specific case of Peru remains unexplored. Therefore, the objective is to conduct a more targeted study in a specific region: Northern Peru. Additionally, understanding river flow is crucial for water resource management, facilitating communication between farmers and reservoir authorities in the studied area.

While the utility of GloFAS [6, 7] data has been demonstrated in other regions [8, 9, 10], it remains unused in Latin America. Consequently, our proposal seeks to combine both approaches to enhance prior efforts. This entails evaluating the ECMWF C3S and GloFAS seasonal forecasting models in the Northern Peru region, with a specific focus on Cajamarca and the Jequetepeque River.

Additionally, research specific to Peru's climate diversity and investments in watershed services for regions like Cajamarca [11] and the Jequetepeque river [12, 13, 14] (fig.2) shed light on the region's unique challenges and opportunities. The approach aligns with the standard methods recommended by the World Meteorological Organization [15, 16, 17].

As a result, we arrive at the following question:

"How do the complementary seasonal models of ECMWF Copernicus C3S precipitation and GloFAS river discharge assist in improving decision-making for water resource management in Northern Peru?"

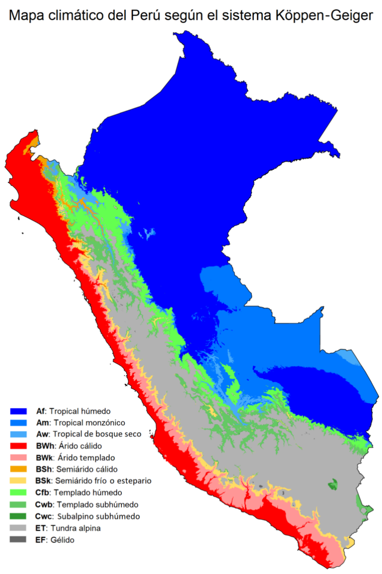


Figure 1: Heterogeneous climate of Peru [18]

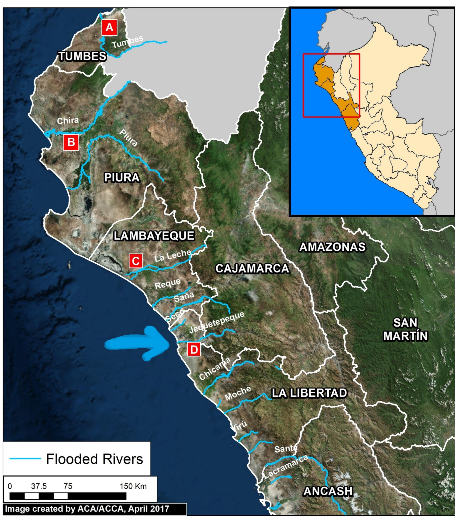


Figure 2: Main rivers of Northern Peru and Jequetepeque river [19]

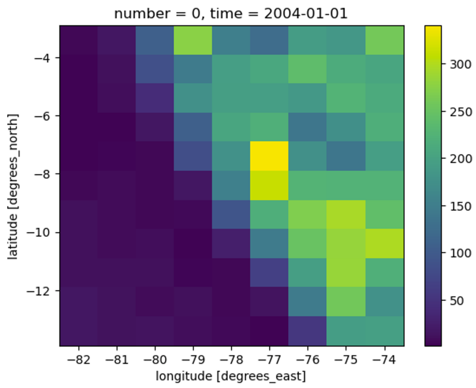


Figure 3: CHIRPS data before downscaling, 100km

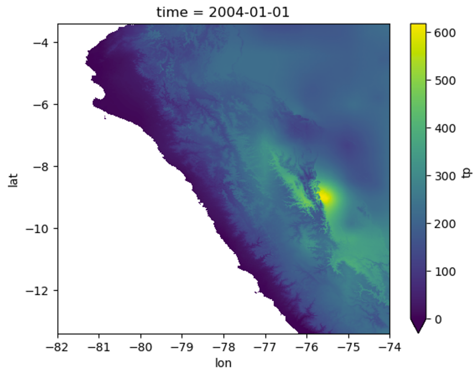


Figure 4: CHIRPS data after downscaling, 1km

2 Materials & Methods

This section will consist of two parts relating to precipitation, river flow models.

2.1 Seasonal precipitation model

For the seasonal forecast precipitation model, the **ECMWF SEAS5 Copernicus C3S** (European Centre for Medium-Range Weather Forecasts, Copernicus Climate Change Service) model was used. It is global seasonal model of the Earth's System that simulates interactions among the atmosphere, ocean, land, and other components. It considers variability phenomena such as El Niño, a very important weather pattern in South America.

Then, a statistical **downscaling** method relying on the WorldClim climatology was used [20, 21], allowing for a significant improvement in spatial resolution, from 100 km to 1km (fig.3 and fig.4). The WorldClim interpolation method involves a dataset of monthly climate data, which is relevant for a seasonal study.

After that, the **climatology** has been computed spanning over a 24-year period (1993-2016), almost in accordance with the timeframe of 30-year recommended by the World Meteorological Organization (WMO) [15, 17]. Climatology will then serve as a reference for calculating forecast anomalies.

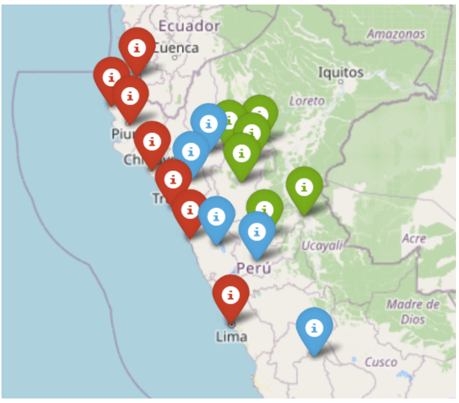


Figure 5: Clustered observation stations

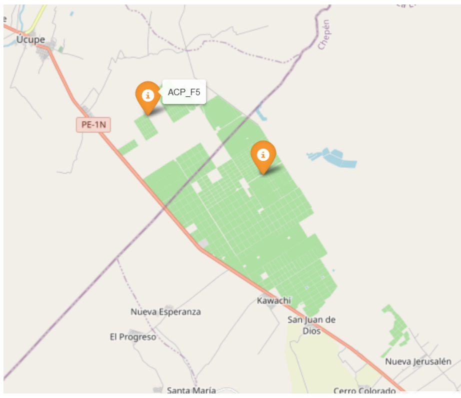


Figure 6: Field stations

The calculation of **anomalies** enables understanding the difference from the expected average, which is helpful in communicating with local farmers. The absolute anomaly, measured in the unit of the manipulated quantity, provides farmers with precise information such as the number of millimeters more or less that should be expected, directly relevant to them as they measure it in the fields. Additionally, relative anomaly calculations allow to disregard the unit, enabling comparison of various parameters like precipitation rates and water flow.

$$anom_{abs} = data - reference$$

$$anom_{rel} = \frac{data - reference}{reference} * 100$$

The **evaluation** of the model relied on evaluation methods recommended by the World Meteorological Organization (WMO) [16], encompassing the following indicators: bias, contingency tables, mean absolute error (MAE), root mean square error (RMSE), accuracy, Pearson correlation, skill score, etc.

The evaluation of the precipitation model was done first with the observation data from GSOD stations and field stations (fig.6) (fig.5 stations clustered in "selva" (green), "sierra" (blue), "costa" (red) based on altitude crossed to mean precipitation rate) and secondly with the re-analysis downscaled data from CHIRPS.

An approach based on determinism was favored, by averaging all the members provided by the model (25 members for the reforecast and 51 members for the forecast), to balance computational weight and time constraints during a 3-month study. This allowed us to work with metrics such as MAE, RMSE, and bias.

2.2 Seasonal river discharge model

For the hydrological forecasting model, the Global Flood Awareness System model, **GloFAS**, was used, with a resolution of 0.1° (≈ 10 km). This hydrological model is providing forecasts

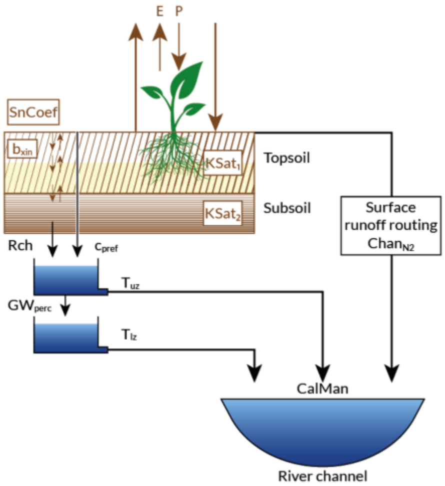


Figure 7: GloFAS, a distributed hydrological rainfall-runoff model [22]

up to D+210 (≈ 7 months). It performs a comprehensive water balance calculation for each grid cell within its domain, considering various factors: precipitation, temperature, potential evapotranspiration, soil evaporation rates (fig.7).

The **climatology** has been computed spanning over a 39-year period (1993-2020), in accordance with the timeframe recommended by the WMO.

The **evaluation** of the hydrological forecasting model was done using the historical data from GloFAS with the method of bias comparison and comparing with the in situ data provided by the Peruvian government [23].

$$bias = forecast_{step} - hist_{step}$$

3 Results and Discussion

3.1 Results of the precipitation model evaluation

The initial evaluation, using GSOD station data (1991-2020) as a reference, is based on a "real" and more comprehensive reference (a 30-year period), and reveals an overestimation by the C3S model. Furthermore, this approach, using local data, does not yield gridded anomalies. Significant biases in the C3S model are observed, justifying an approach based on relative anomalies. Due to the lack of spatial data and the presence of missing data in GSOD, CHIRPS will serve as a valuable complement.

A comprehensive analysis was conducted at both the station and cluster levels, categorizing them as "selva," "costa," and "sierra." Model evaluation was performed by comparing observations using relative anomalies, allowing for the removal of biases. This involved comparing the relative anomalies of the model to those of observations for each station, month, and year.

For each station, the process included:

- Generating cumulative monthly mean precipitation plots and model error plots (fig.8).
- Saving key metrics such as mean absolute error (MAE), root mean square error (RMSE), Pearson correlation, skill score, and normalized bias.
- Evaluating the model's forecasting performance by determining if its predictions fell within a specified percentage range around the observed values for each time step.

At the cluster level, the following steps were taken:

- Plotting the normalized bias for each station within the cluster.
- Calculating metrics as the means of station values grouped by cluster.
- Creating contingency tables, which sum the confusion matrices of each station in the cluster (fig.9).
- Plotting the mean bias for each cluster (fig.10).

The second approach utilizes the hindcast period of ECMWF C3S (1993-2026), downscaled at 1km with WorldClim interpolation, as the reference. This evaluation adheres to the official method, employing improved 1km data compared to the previous 1° data. This allows for the generation of gridded anomalies in the area. It is important to note that the overestimation by the baseline model results in a loss of precision.

The evaluation was conducted over the period 1994-2020, using reforecast and forecast data, with a count of correctly detected anomalies, whether they are positive or negative (fig.11). For example, "5/9" means that 5 out of 9 instances where the anomaly is positive, the C3S model correctly

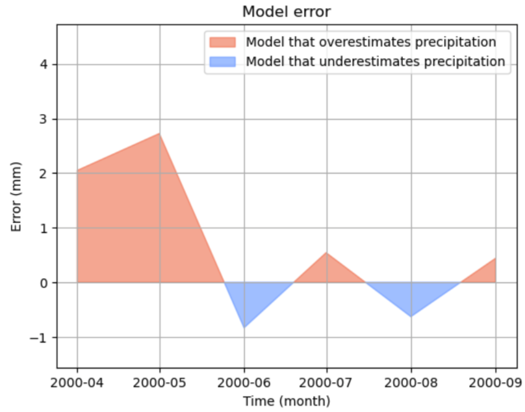


Figure 8: Bias evaluated for GSOD station Capitan FAP Guillermo Concha Iberico

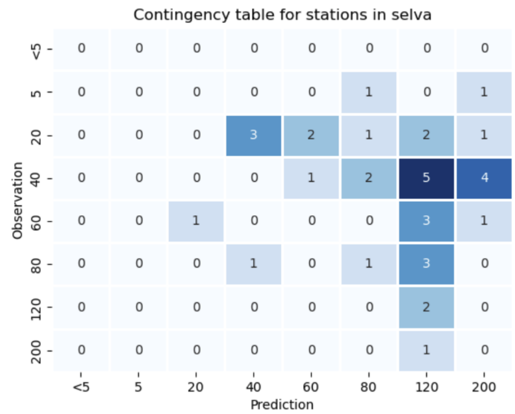


Figure 9: Contingency table for precipitation at runtime 2000-04

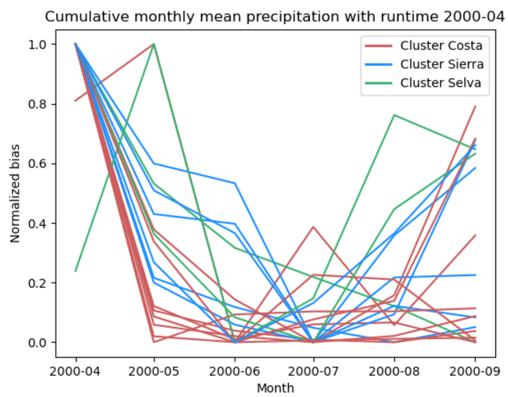


Figure 10: Evaluation of C3S with the GSOD stations data, bias by cluster

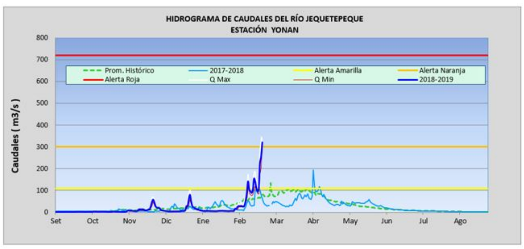
	GSOD	CHIRPS (threshold 0)	CHIRPS (threshold +/- 10%)	CHIRPS (threshold +/- 20%)
Positive anomaly detected	85/173	39/81	33/65	26/52
(%)	49.13%	48.15%	50.77%	50.00%
Good sign prevision until lead	+5	+5	+5	+5
Negative anomaly detected	151/248	71/108	61/93	45/70
(%)	60.89%	65.74%	65.59%	64.29%
Good sign prevision until lead	+5	+5	+5	+5
Total anomalies <0 or >0	421	189/325		
No anomaly (+/- 10%)			35/166 = 28.57%	87/202 = 28.57%

Figure 11: Evaluation of ECMWF C3S with the CHIRPS forecats data

detected it 5 times, at least one month in advance. The threshold corresponds to an allowed range around the anomaly value (0 if only the sign of the anomaly is considered, 10% around the value, etc.). The "Good sign prevision until lead" in the model evaluation through CHIRPS could be considered as an indicator of reliability.

3.2 Results of the river discharge model evaluation

The same method as for C3S anomalies was used (reforecast as the reference), along with a comparison with historical GloFAS data to validate the model. The normalized bias was also validated. However, we encounter a shortage of in-situ data, although the limited data available from government sources allow for validation of the order of magnitude.



Estación	Río	Caudal Promedio 01:00 hrs (m^3/s)	Caudal Promedio Histórico (m^3/s)	Anomalía (%)	Caudal Máximo (m^3/s)	Caudal Crítico (m^3/s)
Yonán	Jequetepeque	320.6	80.55	298.1	320.6	720

Figure 12: Jequetepeque official data [23]

3.3 Anomalies maps

Anomalies maps covering the entire Northern Peru region, as well as centered on the Cajamarca region and specifically the area of interest for farmers and the Galito Ciego reservoir, have been

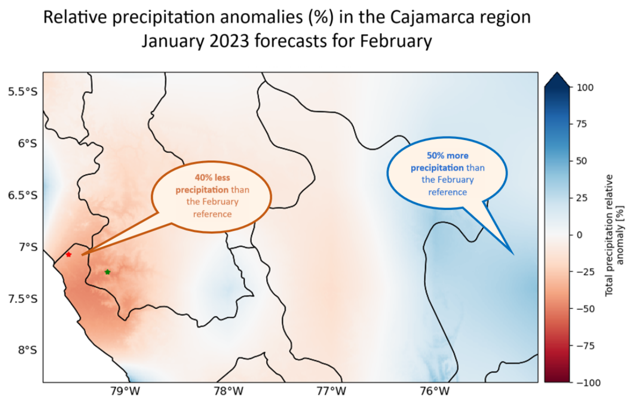


Figure 13: Precipitation relative anomaly map with notices

generated. These forecasts extend up to month + 6, enabling the derivation of precise recommendations for farmers.

For precipitation, both relative and absolute (fig.14) anomaly maps are available. However, for river discharge, we only have absolute anomalies (fig.15) due to difficulties in reading relative anomalies. In the case of river discharge, GLOFAS detects all watercourses, so to have the relative anomaly data we have to extract it from specific points of interest such as the dam location.

3.4 Relationship between precipitation and river discharge

In the end, a parameter comparison was conducted and discussions with an expert in hydrology were engaged.

GloFAS model primarily considers precipitation as a parameter, implying that the evolution of precipitation and river discharge should be correlated. However, in reality, there are other specific regional parameters that can account for differences, such as groundwater, evapotranspiration and runoff.

This is what the expert presented : river discharge is directly affected by precipitation. If this is not what we observe, it indicates that either it is due to reservoir management or groundwater recharge. Upstream areas are more reliant on rainfall. If, during a dry period, there is a relatively constant non-zero discharge, it implies that it originates from groundwater sources. During rainy periods, the river is primarily fed by surface runoff.

Therefore, river discharge is directly and predominantly influenced by precipitation, which explains why it should be expected similar trends between precipitation and river discharge, with some variations due to the additional parameters at play in this phenomenon (fig.16, precipitation in blue and river discharge in orange).

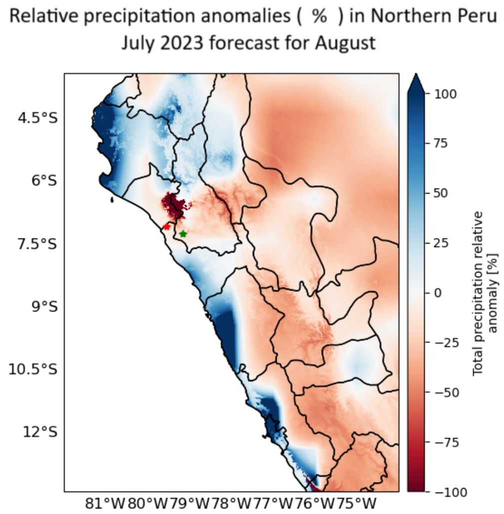


Figure 14: Precipitation relative anomaly

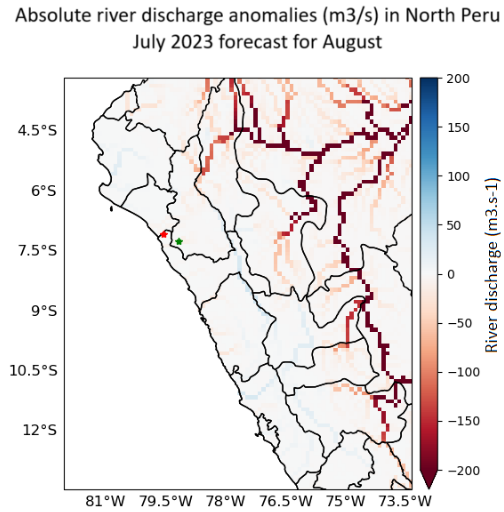


Figure 15: River discharge absolute anomaly

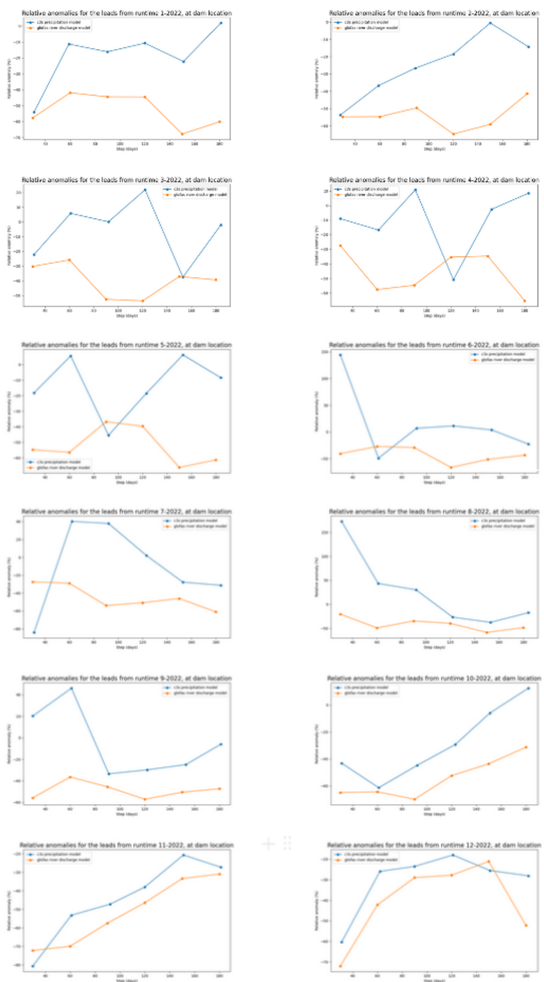


Figure 16: Comparison of relative anomalies in precipitation and river flow for each month in 2022 at the dam location

Furthermore, by linking the two anomaly forecasting models, it becomes feasible to provide specific recommendations to farmers, aiding them in preparing for water storage ahead of drought periods (fig.17).

3.5 Data-visualization and customer presentation

The last step was to present the work to the Peruvian pilot clients, in Spanish, to get their feedback. A prototype to integrate into the existing dashboard was created (fig.18), showcasing the data visualizations obtained. This enabled to have a discussion about what was most understandable and useful for the customer.

4 Conclusions

This study aimed to enhance decision-making in agricultural water resource management by evaluating the effectiveness of seasonal precipitation and river discharge forecasting models from Copernicus C3S and GLoFAS in Northern Peru. To achieve this, the hindcasts and forecasts were meticulously validated.

The study's successful forecasting for up to 5 months offers valuable insights for decision-makers, enabling them to make informed choices for long-term planning. Furthermore, the models' potential for reliable predictions beyond 5 months holds promise for the development of proactive strategies for climate-related challenges. The incorporation of anomaly maps for both precipitation and river discharge, extending up to 5 months, empowers informed decision-making.

In essence, the research suggests that meteorological forecasts can enhance water resource management in northern Peru, with their

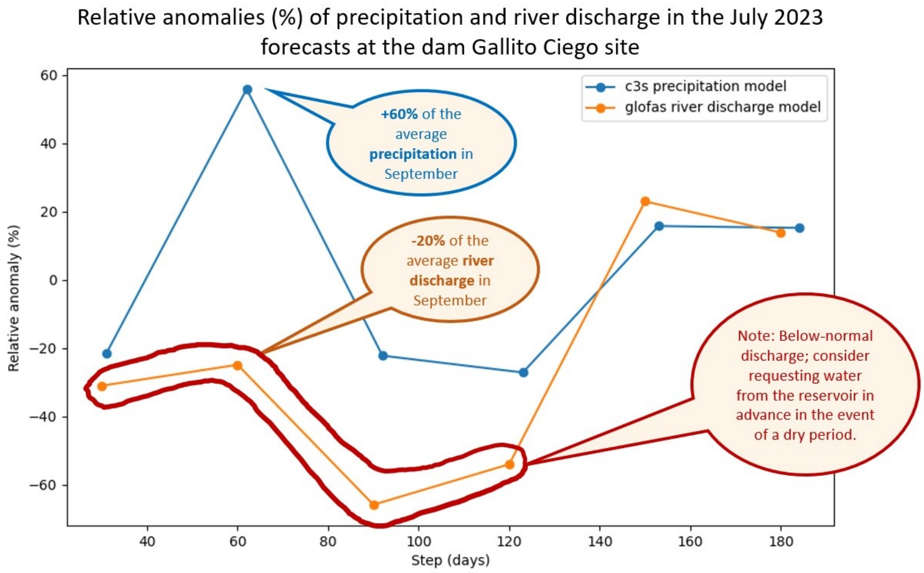


Figure 17: Comparison of relative anomalies in precipitation and river flow

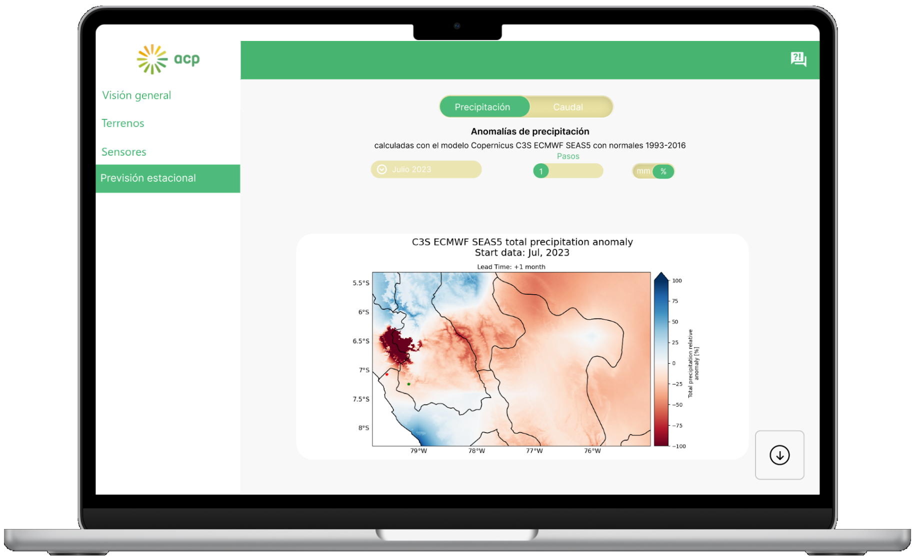


Figure 18: Customer prototype [24]

reliability extending well beyond the immediate future. These forecasts empower the region to adopt sustainable practices and adapt to climate variations, fostering resilience in the face of environmental uncertainties.

5 Acknowledgements

The authors thank the internship tutor Martin Espitalié for providing guidance, support and expertise throughout this project about meteorology and climate.

The authors would also like to acknowledge all the direction of WeatherForce, M.Pascal Venzac and Ms. Christine David, for this great opportunity and for their confidence.

Thanks to all the WeatherForce team, especially Kelly Grassi, Iñigo Reizabal Arregui, Arthur Albaz and Emilie Soursou for their help and advice.

Thanks to Ludovic Cassan, professor-researcher at ENSEEIHT for his external hydrological expertise intervention.

References

[1] ECMWF [Text];. Available from: https://www.ecmwf.int/.

[2] Description of SEAS5-v20171101 C3S contribution - Copernicus Knowledge Base - ECMWF Confluence Wiki;. Available from: https://confluence.ecmwf.int/display/CKB/Description+of+SEAS5-v20171101+C3S+contribution.

[3] Description of the C3S seasonal multi-system - Copernicus Knowledge Base - ECMWF Confluence Wiki;. Available from: https://confluence.ecmwf.int/display/CKB/Description+of+the+C3S+seasonal+multi-system.

[4] Gubler S, Sedlmeier K, Bhend J, Avalos G, Coelho CAS, Escajadillo Y, et al. Assessment of ECMWF SEAS5 Seasonal Forecast Performance over South America;35(2):561-84. Available from: https://journals.ametsoc.org/doi/10.1175/WAF-D-19-0106.1.

[5] Ferreira GWS, Reboita MS, Drumond A. Evaluation of ECMWF-SEAS5 Seasonal Temperature and Precipitation Predictions over South America;10(9):128. Available from: https://www.mdpi.com/2225-1154/10/9/128.

[6] Emerton R, Zsoter E, Arnal L, Cloke HL, Muraro D, Prudhomme C, et al. Developing a global operational seasonal hydro-meteorological forecasting system: GloFAS-Seasonal v1.0;11(8):3327-46. Available from: https://gmd.copernicus.org/articles/11/3327/2018/.

[7] Global Flood Awareness System - Copernicus Emergency Management Service - CEMS - ECMWF Confluence Wiki;. Available from: https://confluence.ecmwf.int/display/CEMS/Global+Flood+Awareness+System.

[8] Imaran M, Chakraborty A, Tripathy S. Skill assessment of GloFAS-ERA5 operational river discharge for the major Indian River Catchments [preprint];. Available from: https://essopenarchive.org/doi/full/10.1002/essoar.10508045.1.

[9] Emerton RE, Stephens EM, Cloke HL. What is the most useful approach for forecasting hydrological extremes during El Niño?;1(3):031002. Available from: https://dx.doi.org/10.1088/2515-7620/ab114e.

[10] Chen K, Guo S, Wang J, Qin P, He S, Sun S, et al. Evaluation of GloFAS-Seasonal Forecasts for Cascade Reservoir Impoundment Operation in the Upper Yangtze River;11(12):2539. Available from: https://www.mdpi.com/2073-4441/11/12/2539.

[11] Stern M, Echavarria M. Investments in watershed services services for the Jequetepeque watershed in Northwestern Peru, Departments of Cajamarca and la Libertad;. Available from: https://www.forest-trends.org/wp-content/uploads/imported/jequetepeque-iws_-11-5-13-pdf.pdf.

[12] Hídricos ANdADdCyPdR, Zarumilla ANdAAAdAJ. Evaluación de recursos hídricos en la cuenca del río Jequetepeque. Available from: https://repositorio.ana.gob.pe/handle/20.500.12543/1932.

[13] Raben K. Description of Jequetepeque watershed and the research areas;. Available from: https://www.jstor.org/stable/resrep15634.6.

[14] Hídricos ANdADdCyPdR, Zarumilla ANdAAAdAJ. Evaluación de recursos hídricos en la cuenca del río Jequetepeque. Available from: https://repositorio.ana.gob.pe/handle/20.500.12543/1932.

[15] Arguez A, Vose RS. The Definition of the Standard WMO Climate Normal: The Key to Deriving Alternative Climate Normals;92(6):699-704. Available from: https://journals.ametsoc.org/doi/10.1175/2010BAMS2955.1.

[16] Organization (WMO) WM, World Meteorological Organization (WMO). Guidance on Verification of Operational Seasonal Climate Forecasts. WMO. WMO;.

[17] Organization (WMO) WM, World Meteorological Organization (WMO). Guidance on Operational Practices for Objective Seasonal Forecasting (WMO-No. 1246). 2020th ed. WMO. WMO;. Available from: https://library.wmo.int/doc_num.php?explnum_id=11073.

[18] Climate of Peru;. Available from: https://en.wikipedia.org/w/index.php?title=Climate_of_Peru&oldid=1176667980.

[19] MAAP: What satellites show us about Peru's flooding;. Available from: https://www.maaproject.org/2017/peru_flood_summary/.

[20] Moreno A, Hasenauer H. Spatial downscaling of European climate data: SPATIAL DOWNSCALING OF EUROPEAN CLIMATE DATA;36(3):1444-58. Available from: https://onlinelibrary.wiley.com/doi/10.1002/joc.4436.

[21] Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas;37(12):4302-15. Available from: https://onlinelibrary.wiley.com/doi/10.1002/joc.5086.

[22] LISFLOOD - Copernicus Emergency Management Service - CEMS - ECMWF Confluence Wiki;. Available from: https://confluence.ecmwf.int/display/CEMS/LISFLOOD.

[23] SENAMHI - Perú;. Available from: https://www.senamhi.gob.pe/?p=aviso-hidrologico-detalle&a=2019&b=085&c=026&d=SENA.

[24] Figma;. Available from: https://www.figma.com/proto/BGAMGEK5ToQPQ5Pv9KaKLE/acp-dashboard?type=design&node-id=135-240&t=y6bFxQsep4hA4Chb-1&scaling=scale-down&page-id=0%3A1&starting-point-node-id=1%3A2&mode=design.