Journal of Natural Science and Exploration ISSN: 3068-9457

PRIME OPEN ACCESS

Volume 1, Issue 2

Research Article

Date of Submission: 20 Jul, 2025
Date of Acceptance: 27 Aug, 2025
Date of Publication: 23 Sep, 2025

Development of a Quantile Regression Model for Replicating Extreme Precipitation in Different Climatic Zones of Morocco

Oumechtaq Ismail[1*], Oulidi Abderrahim[1] and Bahaj Tarik[2]

[1]Solidarity Fund Against Catastrophic Events, Department of Studies and Risk Management, Morocco

[2]University FS Rabat, Department of Geology, Morocco

***Corresponding Author:**
Oumechtaq Ismail, Solidarity Fund Against Catastrophic Events, Department of Studies and Risk Management, Morocco.

Citation: Ismail. O., Abderrahim. O., Tarik. B. (2025). Development of a Quantile Regression Model for Replicating Extreme Precipitation in Different Climatic Zones of Morocco. *J Nat Sci Explor*. *1*(1), 01-16.

Abstract
Given the dispersion and scarcity of meteorological stations across Morocco, as well as gaps in recorded data series, many researchers turn to satellite data to model flood risk. However, the direct use of these data can introduce biases, requiring corrections. Existing solutions primarily focus on the central values of the data, neglecting extreme events, which are actually responsible for flooding.

To date, no model has been developed to reliably reproduce extreme precipitation using satellite data while accounting for the specific climatic factors in each region that drive the generation of these extreme events. In this work, we aimed to fill this gap. To do so, we first evaluated the reliability of different satellite products. Then, we developed a quantile regression (QR) model tailored to each region. The Kullback-Leibler distance was used to identify the quantile that best reproduces ground-based precipitation.

The model was validated through a spatio-temporal approach, using extreme events recorded at stations not included in the model development, as well as considering events beyond the 2000-2018 period used to build the model.

The results show that the QR model, developed for each region, accurately estimates extreme precipitation. For example, the difference between the observed cumulative precipitation and that calculated by the model ranges from 1% to 13%. This addresses three major issues: the lack of information in areas without stations, the correction of biases associated with satellite precipitation, and the improvement in estimating extreme precipitation.

Keywords: Biotypes, Cryptic Species, Endosymbionts, Genetic Groups, Virus-Vector, Whitefly

Introduction
Morocco faces various medium-level climate risks, with floods being the most frequent and costly natural disasters, resulting in significant human and economic consequences. This situation prompted the Moroccan government to implement a national strategy for managing natural risks, leading to the adoption of Law 110-14 in 2016. This law establishes a mandatory natural disaster risk coverage regime and resulted in the effective creation in 2020 of the Solidarity Fund against Catastrophic Events (FSEC).

The FSEC's mission is to compensate victims in the event of a natural disaster, particularly floods, earthquakes, and tsunamis. To achieve this, the fund conducts modeling studies and develops CAT models based on the following components: hazard, exposure, vulnerability, and damage. These models enable the assessment of damages, the estimation of consequent financial losses, and the transfer of risks.

The hazard model allows for the physical characterization of an event's impact. In the case of floods, this may include evaluating the extent of affected areas, river overflows, and the height and speed of water. Precipitation remains the most important element of the hydrological cycle and serves as the primary input for hydrological modeling. These precipitation data are obtained from ground measurements or satellite measurements.

Our research problem consists of three aspects. The first aspect is the very low number of rain gauges; consequently, the precipitation measured by these stations is localized and may not be representative of the entire country. The second aspect involves the use of satellite precipitation data, leading us to ask the following questions: Which products best estimate precipitation? Is there a satellite product valid for the different climatic regions of Morocco? What is the best technique for correcting the bias between satellite precipitation and ground-based rainfall?

Several studies have evaluated the reliability of satellite products and proposed bias corrections [1-4]. According to the analysis of these studies, we observed that most of them proposed corrections focusing on the central part of the distribution (correction by the mean, relative bias, linear regression), leaving extreme values uncorrected.

The novelty we introduced in this work is the development of a Quantile Regression (QR) model that considers extreme values for each climatic region. These regions were defined based on the research of Mr. Chaqdid [5]. This new model aims to replace the currently applied models and techniques, which use low and moderate precipitation and rely on one or two satellite products for all of Morocco.

The division of the Moroccan territory into homogeneous climatic regions, based on the climatic factors responsible for extreme events, aims to develop a consistent model for each region. This model will be based on extreme events of the same nature, i.e., generated by the same climatic factors. This approach will help us avoid the heterogeneity of the baseline data and focus our analysis on each specific region, rather than constructing a single model for all of Morocco, as is often the case in other studies.

The main objective of this study is to construct a QR model for each climatic region of Morocco instead of a single model for the entire territory, as is the case in several studies. This model will be capable of accurately reproducing extreme precipitation responsible for floods using satellite precipitation data.

Study Area

In order to account for the spatio-temporal variation of precipitation prevailing from North to South of Morocco and to reproduce precipitation at any point based on satellite precipitation data, we used 38 rain gauge stations covering all climatic regions of Morocco. The study area corresponds to the Moroccan territory. Figure 1 illustrates the boundaries of the area as well as the geographical locations of the rain gauge stations:

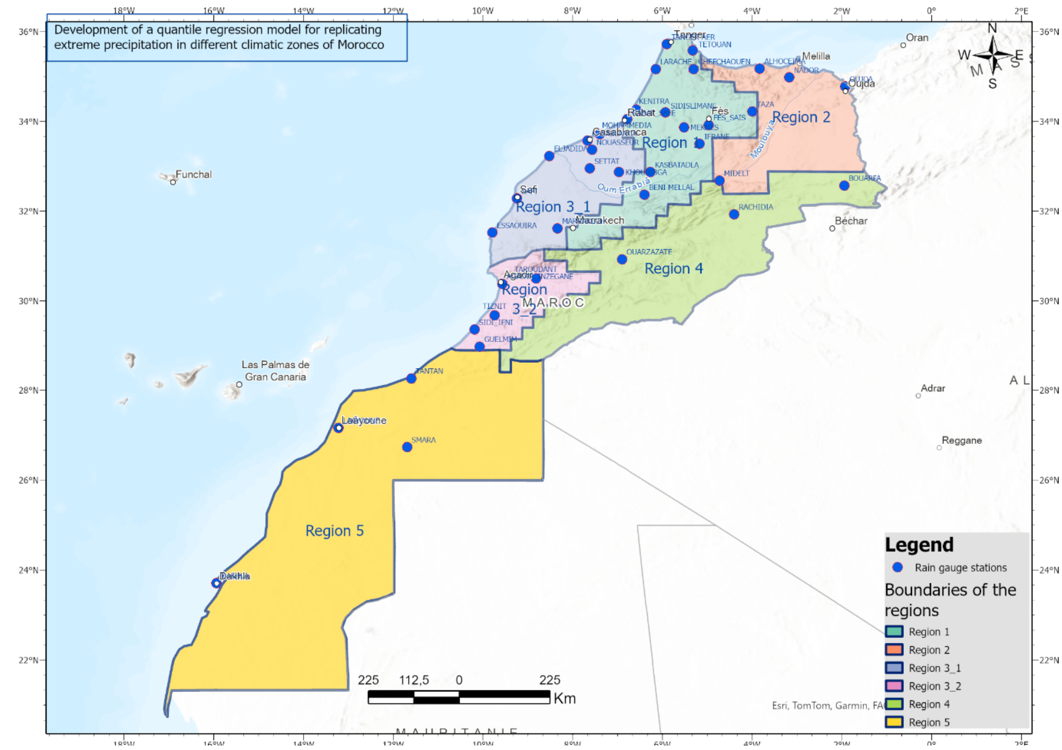


Figure 1: Study Area

Data, Materials, and Methods

Methodology

The methodology adopted in this work is based on five main axes:

- **Data Preparation**
  We began by preparing the precipitation data for each homogeneous climatic region defined in the article by [5]. Each region was analyzed separately to ensure the relevance of the results.

- **Downloading Satellite Precipitation Data**
  The satellite precipitation data were downloaded from the Climate Engine platform (https://app.climateengine.org/climateEngine). This platform provides access to high-resolution satellite data, which is essential for our analysis.

- **Development of Quantile Regression (QR) Models**
  For each region, we developed a Quantile Regression (QR) model for different quantiles (0.1, 0.2, 0.3, ..., 0.9). A rainfall threshold of 20 mm was used for all regions except regions 4 and 5, where the threshold was set to 10 mm. This threshold helps improve the performance of our model in reproducing extreme events and minimizes the impact of low rainfall values on model construction.

- **Calculation of KL Distance**
  We calculated the Kullback-Leibler (KL) distance between the developed quantile regression models and the observed rainfall data. The regression model with the smallest KL distance was selected as the most effective.

- **Model Validation**
  The selected model was validated using extreme events recorded in locations different from those used for model construction. This allowed us to verify the robustness and reliability of our model under real conditions.

The following flowchart summarizes the methodology we adopted to achieve our objectives:

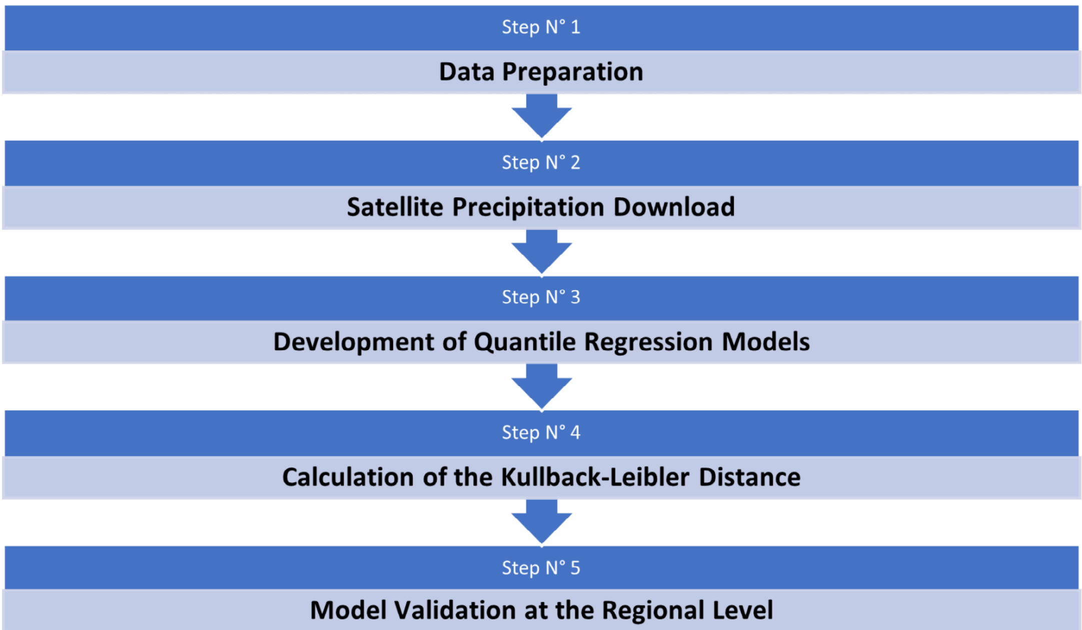


Figure 2: Validation The Model

Data

The data used in this study are grouped into three categories:
Ground-Measured Precipitation by DGM: These data are assumed to be accurate and serve as a reference.

Ground-Measured Precipitation by DRPE: Used to validate the obtained results.

Satellite Precipitation Data: ERA5, GPM, TRMM, and PERSIANN are the primary sources of satellite precipitation studied:

- **ERA5** is a climate reanalysis model developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). It utilizes various techniques including ground-based meteorological station measurements, satellite data to estimate real-time global precipitation, and numerical weather prediction models to produce precipitation estimates. The spatial resolution of ERA5 data is 11 km [6].

- **GPM** (Global Precipitation Measurement): GPM measures precipitation quantity using satellites equipped with microwave and radar sensors that measure precipitation through clouds and weather conditions. Raw data are then corrected for systematic errors such as calibration errors and biases. Validated data using ground-based measurements are used to produce a variety of precipitation products [7]. The spatial resolution of GPM data is 9.6 km.

- **TRMM** (Tropical Rainfall Measuring Mission): This satellite provides precipitation estimates based on microwave and radar sensor measurements. TRMM data are corrected for biases and validated with ground-based precipitation data, offering global coverage of tropical and subtropical precipitation with a spatial resolution of 0.25° x 0.25° [8].

- **PERSIANN** (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks): This product uses artificial neural networks to estimate precipitation from satellite data. Infrared data are combined with ground-based precipitation observations to improve estimation accuracy. PERSIANN has a spatial resolution of 0.25° x 0.25° [9].

Materials
The software R was employed for performing all necessary calculations, leveraging its extensive documentation, package availability, and performance in handling complex and repetitive operations [10].

Regarding satellite data, various platforms now offer direct download options for climate parameters such as temperature and precipitation. For this study, the Climate Engine platform (https://app.climateengine.com/climateEngine) was used to download daily precipitation data measured by the three satellites [11].

Theoretical Methods
To assess the accuracy of satellite-measured precipitation compared to ground-based measurements, commonly used performance criteria were applied. These criteria include the correlation coefficient (CC), root mean square error (RMSE), normalized root mean square error (NMSE), and relative bias (BR). These criteria have been employed in various studies, including the one conducted by Guo et al. (2015) [1].

a. **Quantile Regression (QR):** The objective is to estimate the parameters of a linear or nonlinear regression model that minimizes the sum of weighted absolute residuals, where the weights are determined by the quantile to be estimated. Quantile Regression (QR) provides a more comprehensive understanding of the response variable distribution at different quantiles. The equation is provided in the appendices. This method has been used by several authors in various locations. Recent studies in environmental science have utilized linear quantile regression for estimating conditional quantiles, for example. In 2017, Chiu and Yohann Moanahere (2017) used quantile regression models to study in detail health peaks and their relationship with weather in the cities of Montreal and Quebec (Canada). In another context, Eric Eboulet & Alina Matei (2013) applied quantile regression and its statistical tests to PISA data in the French-speaking part of Switzerland.

b. **Kullback-Leibler Distance:** The Kullback-Leibler (KL) distance is a measure of divergence between two probability distributions, used to quantify the difference between observed precipitation and that modeled by quantile regression. Quantile regression models the relationships between variables by estimating different quantiles of the response variable's distribution, which is useful for precipitation data that is often asymmetric and extreme [12]. This method allowed us to measure the divergence between the distributions of observed and modeled precipitation at different quantiles, identifying potential anomalies or biases in our model. The use of KL distance to evaluate probabilistic forecasts has been validated in previous studies [13].

For continuous distributions, as is our case, the KL distance is defined as follows:

$$Dkl(P||Q) = \int_{-\infty}^{\infty} p(x)\, log\frac{p(x)}{q(x)}\, dx$$

Where p(x) and q(x) are the probability density functions of P and Q, respectively.

Results and Discussions
Descriptive Statistics
To highlight that a single satellite product cannot be universally applied across the five climatic regions, it is essential to understand the varying accuracy of satellite products across these regions. Studies have shown that each satellite product performs differently across different zones [14]. For instance, satellite precipitation estimation performance can vary significantly between tropical and arid regions due to differing atmospheric dynamics and surface conditions.

In order to overcome these limitations, we developed a quantile regression model that integrates multiple satellite products for each climatic region. This model leverages the strengths of each product in different zones, thereby improving the accuracy of precipitation estimates.

Furthermore, given that available records mostly contain very low or zero precipitation values (less than 2 mm), we chose to compute boxplots only for precipitation values strictly exceeding 2 mm in each zone. This approach better captures the variability of significant precipitation events and provides a more representative analysis of actual climatic conditions.

We did not use RMSE, MSE, R^2, and BR metrics as we believe these metrics provide relevant values primarily for the central parts of distributions, which does not align with the objectives of this study.

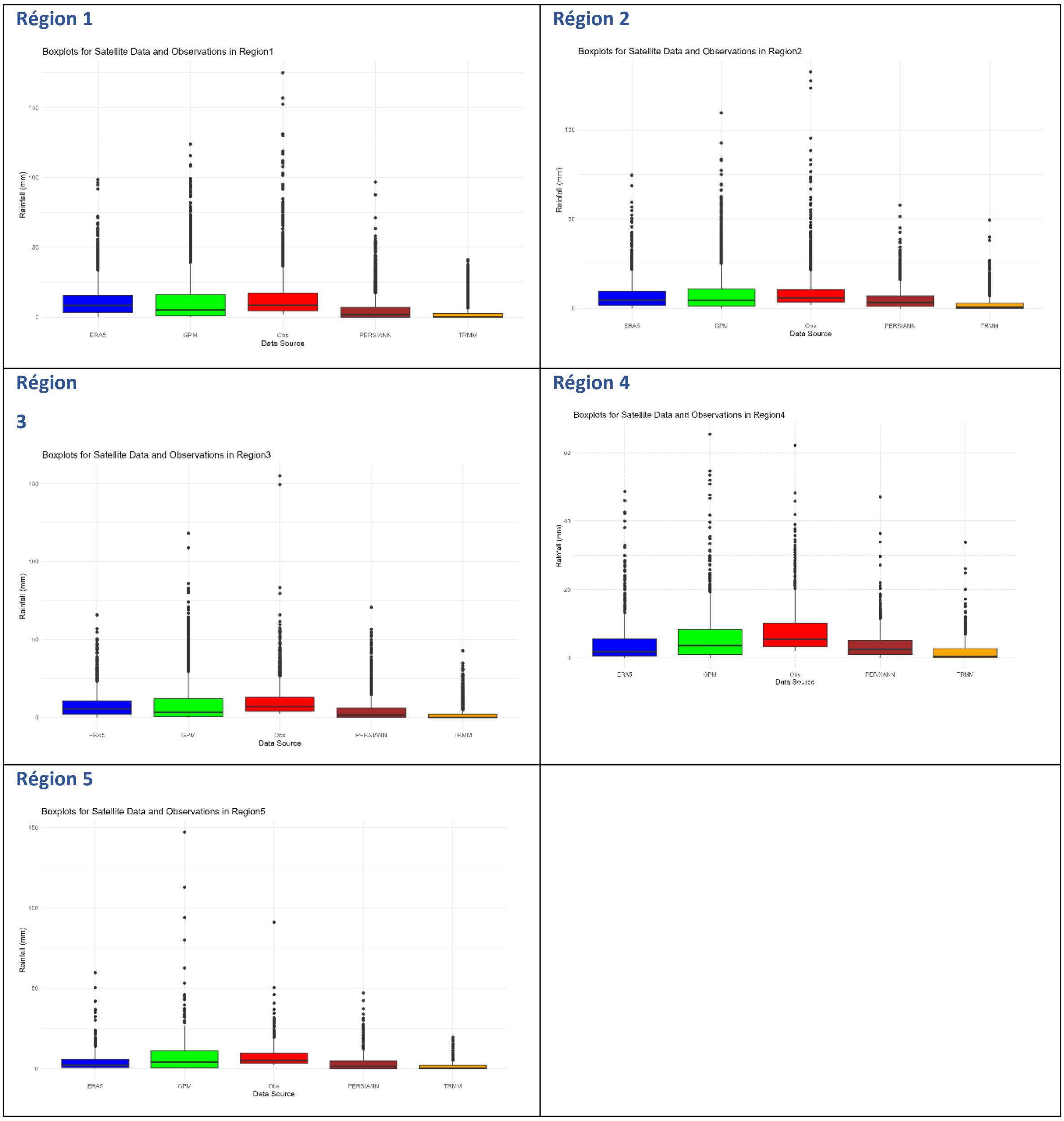


Table 1: Boxplot of Rainfall > 2mm in the Five Regions

We also quantitatively evaluated the accuracy of different satellite products compared to ground station measurements. This evaluation was conducted using 35 extreme events observed at various stations across different climatic regions. The results obtained, presented in Table No. 2, clearly demonstrate that a single specific satellite product cannot be used to replicate events on a national scale. This highlights the importance and novelty of our approach, which involves constructing a Quantile Regression (QR) model that considers all satellite products but favors (through the coefficient of the quantile regression equation) one or more products over others, depending on the rainfall regime of each region.

Station	Date	Obs	ERA5	TRMM	GPM	PERSIANN	Obs-ERA5	Obs-TRMM	Obs-GPM	Obs-PERSIANN
Fès	17/04/2007	15.40	0.34	1.27	5.11	13.58	98%	92%	67%	12%
Fès	19/04/2007	31.40	0.22	8.18	15.66	3.09	202%	151%	102%	184%
Fès	20/04/2007	70.30	29.36	23.30	37.38	10.14	266%	305%	214%	391%
Fès	17/05/2011	12.00	3.37	0.01	7.56	0.83	56%	78%	29%	73%
Fès	18/05/2011	55.30	28.59	0.00	24.44	0.00	173%	359%	200%	359%
Tanger	30/10/2018	82.50	24.49	17.84	66.70	24.29	377%	420%	103%	378%
Tanger	31/10/2018	18.40	36.16	13.03	27.56	11.73	-115%	35%	-59%	43%
Al Hociema	22/10/2008	27.20	6.86	0.07	7.13	4.14	132%	176%	130%	150%
Al Hociema	23/10/2008	39.50	39.66	11.93	55.47	12.62	-1%	179%	-104%	175%
Al Hociema	24/10/2008	11.30	8.26	0.00	4.22	4.82	20%	73%	46%	42%
Al Hociema	25/10/2008	35.80	23.84	0.07	1.19	0.00	78%	232%	225%	232%
Al Hociema	26/10/2008	132.50	48.22	49.40	92.55	23.93	547%	540%	259%	705%
Al Hociema	27/10/2008	13.70	5.86	2.80	0.78	2.38	51%	71%	84%	74%
Nador	24/10/2008	10.40	21.19	0.13	0.88	1.69	-70%	67%	62%	57%
Nador	25/10/2008	83.20	28.29	3.21	10.98	2.16	357%	519%	469%	526%
Nador	26/10/2008	123.40	74.48	40.03	47.51	27.46	318%	541%	493%	623%
Taroudant	30/10/2012	59.8	16.76	0.00	2.50	4.82	279%	388%	372%	357%
Taroudant	31/10/2012	80	62.37	9.17	10.41	3.69	114%	460%	452%	496%
Marrakech	28/11/2014	59.4	58.95	12.18	76.72	53.05	3%	307%	-112%	41%
Marrakech	30/11/2014	13.9	14.82	6.14	14.98	6.51	-6%	50%	-7%	48%
Amsoul	21/12/2009	35.7	6.22	17.30	47.39	9.81	191%	119%	-76%	168%
Amsoul	22/12/2009	10.1	41.57	12.69	47.90	25.88	-204%	-17%	-245%	-102%
Amsoul	23/12/2009	56.3	36.59	3.71	18.38	28.04	128%	342%	246%	184%
Amsoul	24/12/2009	33.4	36.66	5.40	36.19	13.54	-21%	182%	-18%	129%
Amsoul	25/12/2009	23.7	31.95	8.40	5.42	4.64	-54%	99%	119%	124%
Ouarzazate	31/03/2002	48.20	1.40	9.96	24.37	33.88	304%	248%	155%	93%
Ouarzazate	01/04/2002	23.70	38.05	4.32	29.41	13.33	-93%	126%	-37%	67%
Errachidia	07/11/2006	37.40	14.58	8.38	17.21	3.06	148%	188%	131%	223%
Errachidia	08/11/2006	10.90	6.62	8.60	7.60	0.90	28%	15%	21%	65%
Bouarfa	27/11/2014	14.90	9.40	0.00	12.68	2.46	36%	97%	14%	81%
Bouarfa	28/11/2014	28.60	32.27	0.00	23.67	20.18	-24%	186%	32%	55%
Bouarfa	29/11/2014	32.00	28.21	0.00	41.71	6.23	25%	208%	-63%	167%
TanTan	12/08/2003	91.10	0.62	6.84	9.95	42.33	588%	547%	527%	317%
TanTan	01/10/2002	46.00	1.01	2.82	7.72	14.63	292%	280%	249%	204%
DAKHLA	22/10/2018	40.7	36.48	13.63	43.85	13.72	27%	176%	-20%	175%

Table 2: Quantitative Evaluation in % of the Difference Between Satellite Product Measurements and Ground-Based Measurements

Development of RQ and RL Regression Models

The objective of this stage is to construct a quantile regression model and a linear regression model for the five regions, with a dual purpose: firstly, to accurately compute the coefficients of the quantile regression equation, and secondly, to model especially extreme precipitation events. We will use precipitation data measured by satellites GPM, ERA5, TRMM, and PERSIANN, which have been identified through various analyses as closest to ground-based data.

To enhance the performance of our model in reproducing extreme events, we have built our model using precipitation exceeding 20 mm for all regions except region 4 and region 5, where we lowered the threshold from 20 mm to 10 mm. This adjustment is due to the different precipitation regime in these two regions compared to others, where observed maximums are lower and events are intense but short-lived.

We also calculated the linear regression model for each region with the aim of demonstrating that our developed model is the best.

Results of the RQ Model

We built our model for each region by selecting only precipitation events (that could cause floods) above 20 mm for all regions except region 4, where we selected precipitation above 10 mm due to lower precipitation levels compared to other regions. Subsequently, when calculating the Kullback-Leibler distance between the precipitation estimated by quantile regression and the ground-measured precipitation, we chose the quantile that minimized this distance. The coefficients and quantiles obtained by region are provided in Table 3:

During the validation of the quantile regression (RQ) model developed for region 3, stretching from Mohammadia to Sidi Ifni, we observed two sub-groups of discrepancies between the RQ model and ground observations. The first group covers the area between Mohammadia and Souira, while the second extends from the Agadir region to the city of Sidi Ifni. This led us to subdivide this area into two sub-regions: region 3_1 and region 3_2. We recalculated the coefficients of the RQ model for each sub-region, resulting in improved accuracy and a significant reduction in the gap between observations and the RQ model.

This subdivision is geographically justified by the influence of the High Atlas range, which plays a crucial role in distinguishing between the precipitation regimes of sub-regions 3_1 and 3_2.

Region	Region 1	Region 2	Region 3_1	Region 3_2	Region 4	Region 5
Tau	0,6	0,7	0,6	0,6	0,4	0,9
ERA5	1,14	0,98	0,88	0,69	-0,06	1,23
GPM	0,21	0,23	0,17	0,02	0,90	-1,20
TRMM	0,42	0,69	0,12	0,71	0,74	0,84
PERSIANN	0,06	0,77	0,21	0,39	0,48	2,13

Table 3: Quantiles and Coefficients of the RQ Model by Region

Results of LR Model

The linear regression model was developed to demonstrate that this model does not yield good results in reproducing extreme precipitation events. The results obtained for each region are given in Table 4.

Region	Region 1	Region 2	Region 3_1	Region 3_2	Region 4	Region 5
ERA5	0,97	0,85	0,88	0,85	0,12	0,12
GPM	0,27	0,20	0,46	-0,04	0,93	0,46
TRMM	0,20	0,88	0,23	0,91	0,61	0,24
PERSIANN	0,02	0,08	0,07	0,40	0,57	0,80

Table 4: Coefficients of the RL Model by Region

Validation of the Developed Model

We validated the performance of our developed RQ model through different approaches. First, we performed validation using extreme events recorded at the stations that were used to construct the model. Second, we used extreme events recorded at stations managed by the DGM that were not used in constructing the RQ model. Finally, we conducted a final validation using extreme events observed at stations outside the DGM network, specifically at stations managed by the DRPE.

Another validation of the RQ model was carried out by testing the model's performance on events that occurred outside the [2000-2018] timeframe, specifically the events of January 2021 at the El Maleh dam station, located about 25 km upstream from the city of Mohammadia.

Validation of RQ Model in Region 1

The model we developed in region 1 allows the calculation of precipitation using the following equation:

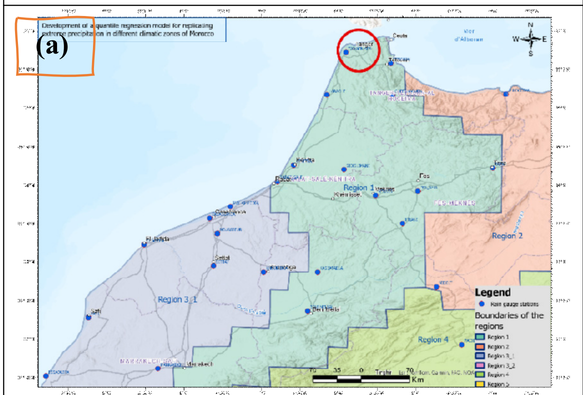
$$P_{RQ} = 1.14 * P_{(ERA5)} + 0.21 * P_{(GPM)} + 0.42 * P_{(TRMM)} + 0.06 * P_{(PERSIANN)}$$

Station: Tanger City

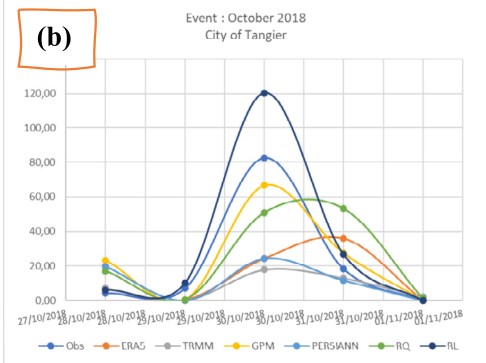
The recent event from October 2018 recorded at the Tangier station, located in the north of region 1, was used to validate our model. We observed that the difference between the observed cumulative rainfall and that calculated by the QR model is -9%, as shown in figure 3 (C).

Station : Tangier	Lat = 35.72 ; Long = -5.9

(a)



(b)



(c)

Date	Obs	ERA5	TRMM	GPM	PERSIANN	RQ	RL
28/10/2018	4,20	7,07	7,07	23,40	19,78	17,12	6,13
29/10/2018	7,00	0,30	0,00	0,25	0,00	0,40	10,21
30/10/2018	82,50	24,49	17,84	66,70	24,29	50,92	120,38
31/10/2018	18,40	36,16	13,03	27,56	11,73	53,27	26,85
01/11/2018	0,00	0,10	2,07	0,00	0,00	0,98	0,00
Cumul evnement	112,10	68,12	40,01	117,91	55,80	122,69	163,57
Différence/sol		39%	64%	-5%	50%	-9%	-46%

(d)

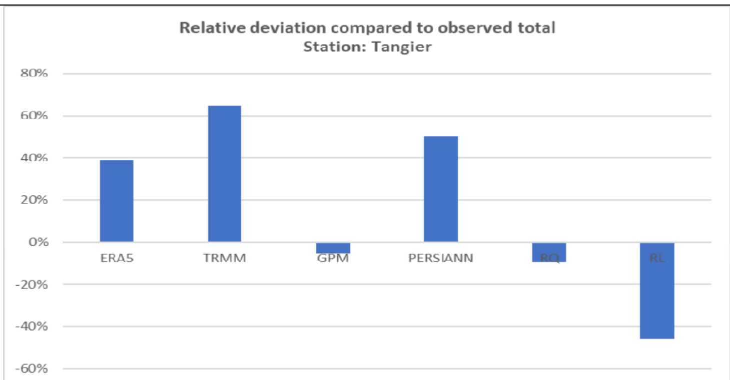
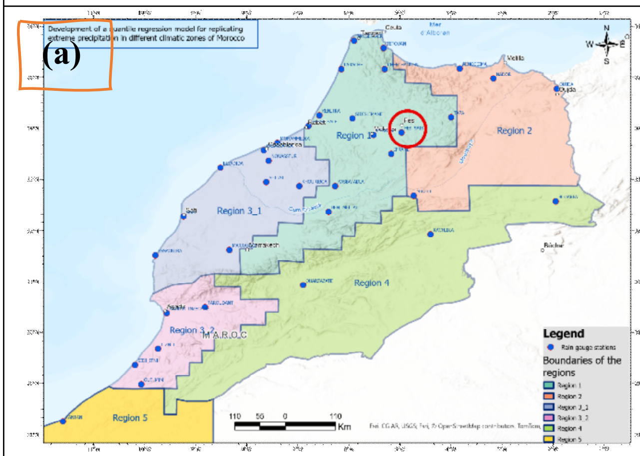


Figure 3: (a) Location of the Tangier Station, (b) Observed Rainfall, Rainfall Calculated by Different Satellites, Rainfall Calculated by the RQ Model and RL Model, (c) Model, (d) Difference Compared to the Observed Cumulative Rainfall

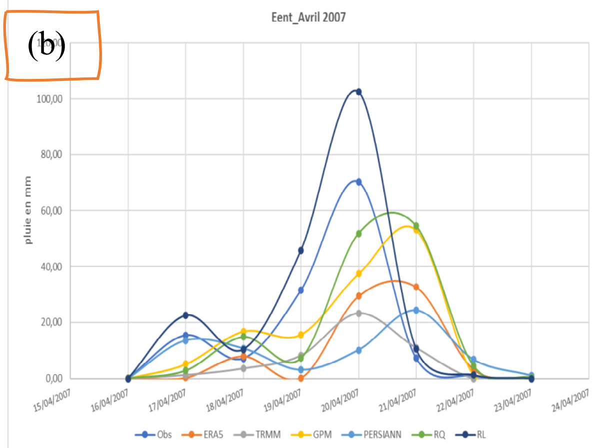
Station: City of Fez We also validated the developed model in region 1 using the event from April 2007 observed at the Fès station. We found that the difference between the observed cumulative rainfall and that calculated by the QR model is -3%, as shown in figure 4 (c).

Station : Fès	Lat = 33.92 ; Long = -4.97

(a)



(b)



(c)

Date	Obs	ERA5	TRMM	GPM	PERSIANN	RQ	RL
16/04/2007	0,00	0,04	0,00	0,00	0,00	0,05	0,00
17/04/2007	15,40	0,34	1,27	5,11	13,58	2,79	22,47
18/04/2007	7,10	7,99	3,59	16,79	10,74	14,80	10,36
19/04/2007	31,40	0,22	8,18	15,66	3,09	7,16	45,82
20/04/2007	70,30	29,36	23,30	37,38	10,14	51,79	102,58
21/04/2007	7,20	32,67	10,98	53,13	24,28	54,53	10,51
22/04/2007	0,80	3,28	0,00	1,23	6,70	4,40	1,17
23/04/2007	0,00	0,19	0,00	0,34	1,07	0,35	0,00
Cumul evnement	**132,20**	**74,09**	**47,34**	**129,63**	**69,60**	**135,86**	**192,90**
Différence/sol		**44%**	**64%**	**2%**	**47%**	**-3%**	**-46%**

(d)

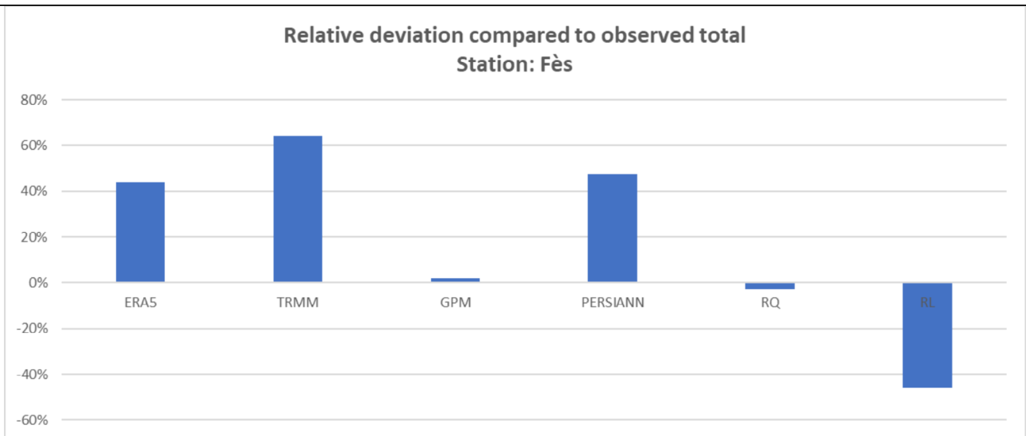


Figure 4: (a) Location of the Fez Station, (b) Observed Rainfall, Rainfall Calculated by Different Satellites, Rainfall Calculated by the RQ Model and RL Model, (c) Model, (d) Difference Compared to the Observed Cumulative Rainfall

Validation of RQ Model in Region 2

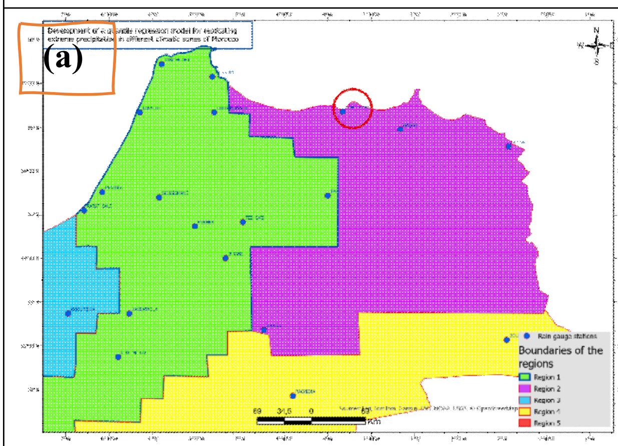
The model we developed in region 2 allows the calculation of precipitation using the following equation:

$$P_{RQ} = 0.98 * P_{(ERA5)} + 0.23 * P_{(GPM)} + 0.69 * P_{(TRMM)} + 0.77 * P_{(PERSIANN)}$$

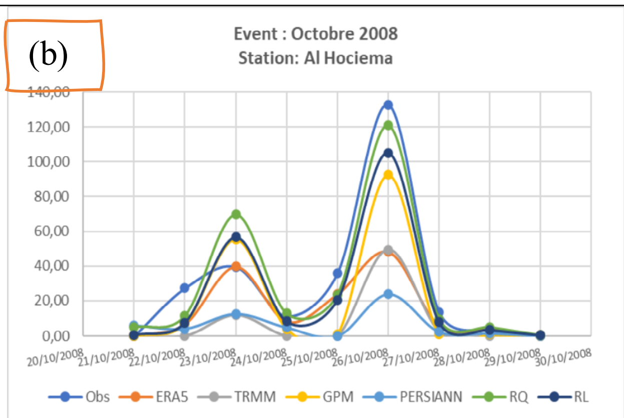
The validation of our QR model for region 2 was carried out using the event from October 2008, which occurred from 10/21/2008 to 10/29/2008, recorded at the Al Hoceima station, as shown in figure 5 (a). The recorded cumulative rainfall was 263.20 mm. In general, all satellite products significantly underestimated this event, as shown in figure 5 (c). The linear regression (LR) model performed better than all satellite products, but not as well as the QR model, which estimated a cumulative rainfall of 258.78 mm, resulting in a difference of 2%.

Station : Al Hociema	Lat = 35.18 ; Long = -3.83

(a)



(b)



(c)

Date	Obs	ERA5	TRMM	GPM	PERSIANN	RQ	RL
21/10/2008	0,00	0,29	0,00	0,00	6,10	4,97	0,73
22/10/2008	27,20	6,86	0,07	7,13	4,14	11,59	7,69
23/10/2008	39,50	39,66	11,93	55,47	12,62	69,51	56,58
24/10/2008	11,30	8,26	0,00	4,22	4,82	12,77	8,28
25/10/2008	35,80	23,84	0,07	1,19	0,00	23,67	20,62
26/10/2008	132,50	48,22	49,40	92,55	23,93	120,84	105,20
27/10/2008	13,70	5,86	2,80	0,78	2,38	9,66	7,79
28/10/2008	3,20	3,96	0,00	0,77	1,48	5,19	3,65
29/10/2008	0,00	0,61	0,00	0,00	0,00	0,60	0,52
Cumul_Ev	**263,20**	**137,56**	**64,27**	**162,11**	**55,46**	**258,78**	**211,07**
Diff par rapport Obs		**48%**	**76%**	**38%**	**79%**	**2%**	**20%**

(d)

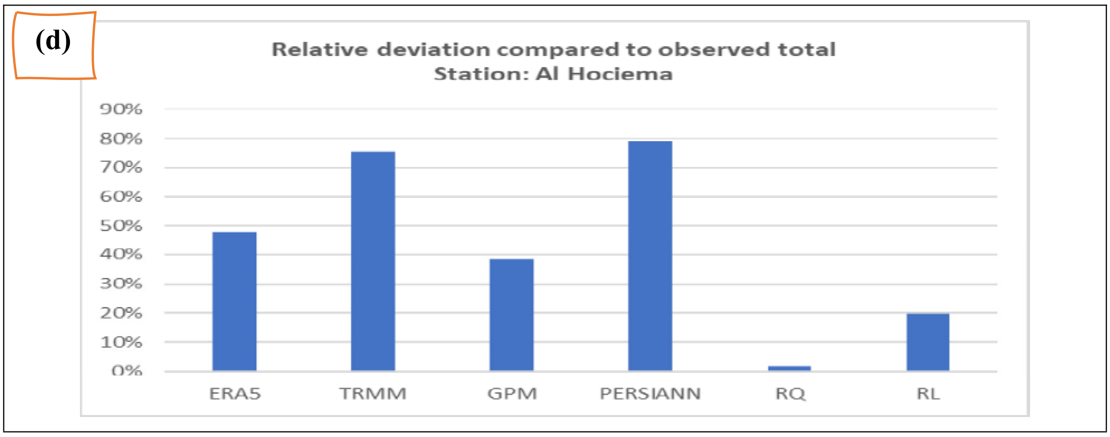


Figure 5: (a) Location of the Al Hoceima Station, (b) Observed Rainfall, Rainfall Calculated by Different Satellites, Rainfall Calculated by the RQ Model and RL Model, (c) Model, (d) Difference Compared to the Observed Cumulative Rainfall

Validation of RQ Model in Region 3

Region 3_1

The model we developed in region 3-1 allows the calculation of precipitation using the following equation:

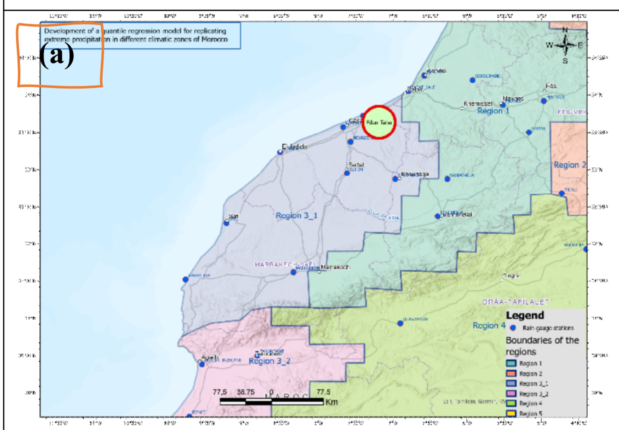
$$P_{RQ} = 0.88 * P_{(ERA5)} + 0.17 * P_{(GPM)} + 0.12 * P_{(TRMM)} + 0.21 * P_{(PERSIANN)}$$

- **Fdan Taba Station:** The observed event in November 2010 at Fedan Taba station, which is a station not used in constructing our RQ model, as shown in figure 6 (a), lasted for 6 days from 26/11/2010 to 01/12/2010, resulting in a cumulative rainfall of 219.6 mm.

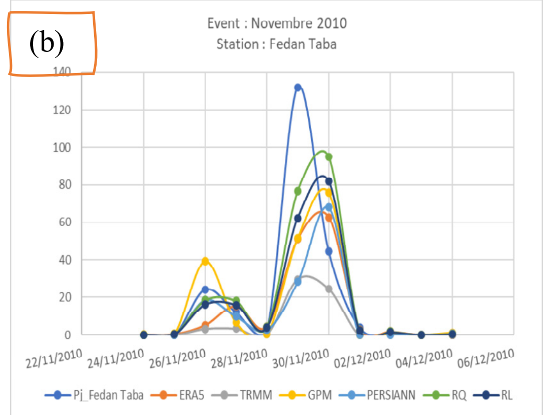
The use of our developed RQ model accurately reproduced this event with a precision of 1%, estimating a cumulative rainfall of 218.04 mm for the same period. In contrast, all other satellite products underestimated this event, as depicted in figure 6 (c).

Station : Fdan Taba	Lat = 33.61 ; Long = -7.84

(a)



(b)



(c)

Date	Pj_Fedan Taba	ERA5	TRMM	GPM	PERSIANN	RQ	RL
24/11/2010	0	0,04	0,00	0,52	0,00	0,17	0,14
25/11/2010	0	0,40	0,00	0,71	0,00	0,54	0,47
26/11/2010	24	5,17	3,27	39,31	18,50	18,34	16,11
27/11/2010	12,1	15,28	3,36	6,51	9,84	17,99	15,75
28/11/2010	2,5	4,40	0,50	0,21	2,68	4,49	4,05
29/11/2010	132,1	51,17	29,59	51,60	28,36	76,88	62,43
30/11/2010	44,5	62,82	24,32	75,96	68,13	94,87	81,94
01/12/2010	3,9	2,79	0,00	0,00	0,00	2,49	2,21
02/12/2010	0,5	1,77	0,00	0,77	0,00	1,78	1,56
03/12/2010	0	0,03	0,00	0,00	0,00	0,03	0,03
04/12/2010	0	0,00	0,13	1,25	0,62	0,45	0,39
Cumul_Event	**219,6**	**143,88**	**61,17**	**176,84**	**128,12**	**218,04**	**185,09**
Différence/cumul_Event		**34%**	**72%**	**19%**	**42%**	**1%**	**16%**

(d)

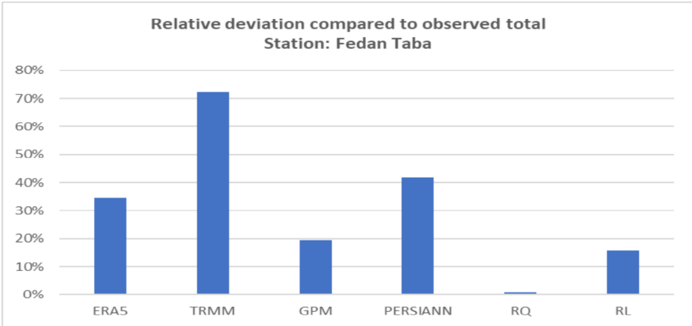


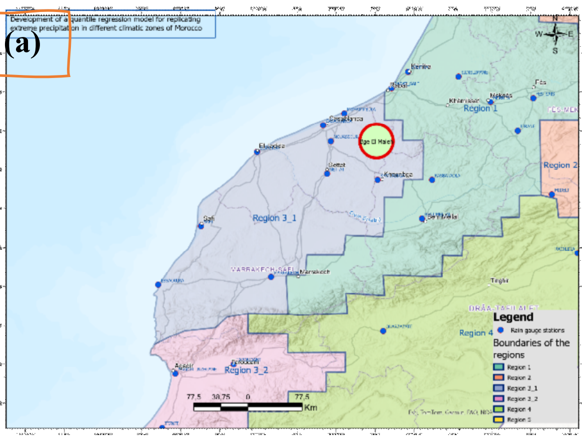
Figure 6: (a) Location of the Fdan Taba Station, (b) Observed Rainfall, Rainfall Calculated by Different Satellites, Rainfall Calculated by the RQ Model and RL Model, (c) Model, (d) Difference Compared to the Observed Cumulative Rainfall

- **El Maleh Dam:** We also utilized the most recent floods in January 2021, occurring in the Casablanca-Mohammedia regions, to apply our model to a real case that was not part of the data used for constructing the RQ model, which spans from January 1, 2000, to December 31, 2018. A cumulative rainfall of 123.9 mm (figure 7(c)) was recorded at the Mellah dam station from January 3, 2021, to January 11, 2021, located approximately 25 km upstream of the city of Mohammedia. For the same period, our RQ model allowed us to calculate a cumulative rainfall of 119.9 mm, resulting in a difference of 3%, which is very promising.

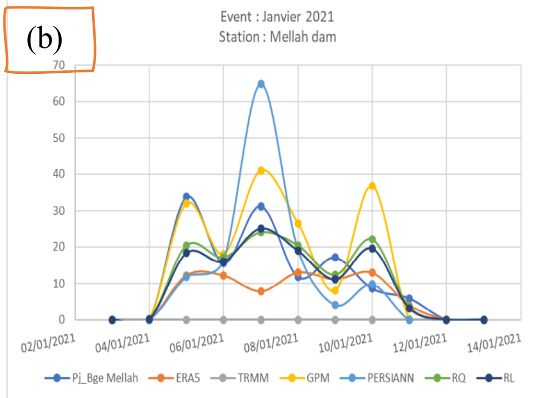
Station : El Maleh dam

Lat = 33.49 ; Long = -7.33

(a)



(b)



(c)

Date	Pj_Mellah dam	ERA5	TRMM	GPM	PERSIANN	RQ	RL
03/01/2021	0	0,01	0,00	0,00	0,00	0,01	0,01
04/01/2021	0	0,00	0,00	0,21	0,00	0,06	0,05
05/01/2021	33,7	12,22	0,00	31,97	11,81	20,37	18,33
06/01/2021	15,5	12,17	0,00	17,99	16,38	17,14	16,03
07/01/2021	31,3	7,79	0,00	40,92	64,88	24,06	25,12
08/01/2021	11,8	12,97	0,00	26,55	19,35	20,38	18,96
09/01/2021	17,1	11,15	0,00	7,98	4,03	12,40	11,14
10/01/2021	8,6	12,91	0,00	36,73	9,75	22,01	19,56
11/01/2021	5,9	3,90	0,00	0,07	0,00	3,49	3,10
12/01/2021	0	0,00	0,00	0,00	0,00	0,00	0,00
13/01/2021	0	0,00	0,00	0,00	0,00	0,00	0,00
Cumul_Event	**123,9**	**73,12**	**0,00**	**162,43**	**126,19**	**119,93**	**112,30**
Différence/cumul_Event		**41%**	**100%**	**-31%**	**-2%**	**3%**	**9%**

(d)

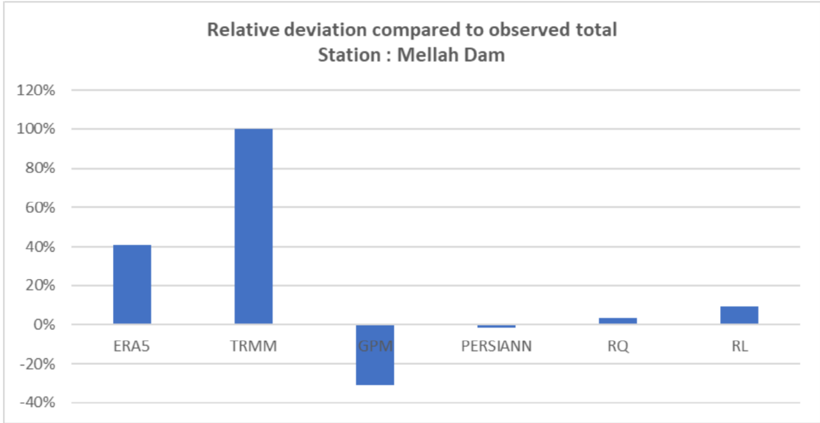


Figure 7: (a) Location of El Malleh Dam Station, (b) Observed Rainfall, Rainfall Calculated by Different Satellites, Rainfall by RQ Model and RL Model, (c) Model, (d) Difference Compared to the Observed Cumulative Rainfall

Region 3_2

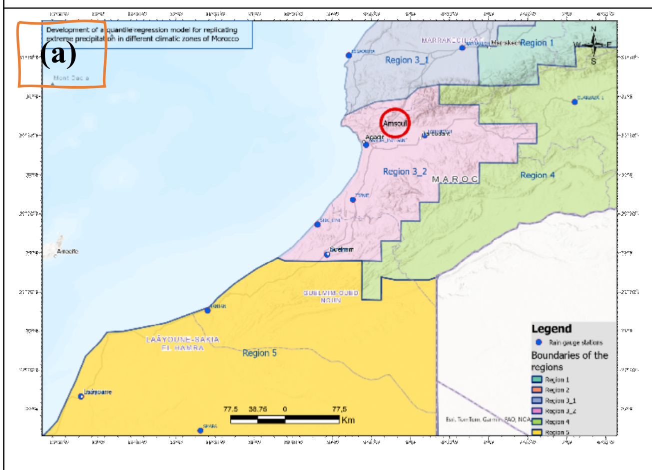
The model we developed in region 3-2 allows the calculation of precipitation using the following equation:

$$P_{RQ} = 0.69 * P_{(ERA5)} + 0.02 * P_{(GPM)} + 0.71 * P_{(TRMM)} + 0.39 * P_{(PERSIANN)}$$

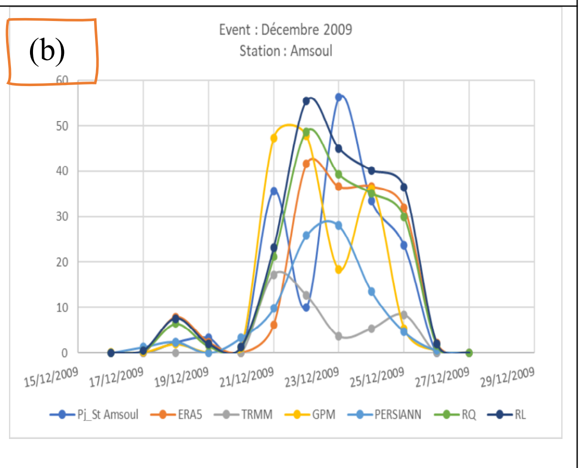
Amsoul station: The Amsoul station as illustrated in figure 8 (a), managed by DRPE, was not used in the construction of the Quantile Regression (RQ) model. We selected the December 2009 event, with a cumulative rainfall of 165 mm. Except for ERA5 and GPM, which show respective deviations of 0% and 3%, other satellite products significantly underestimate the cumulative rainfall of the event. The RQ model shows a deviation of -13% compared to ground data, while the Linear Regression (RL) model exhibits a deviation of -30%.

Station Amsoul | Lat = 30.69 ; Long = -9.22

(a)



(b)



(c)

Date	Pj_St Amsoul	ERA5	TRMM	GPM	PERSIANN	RQ	RL
16/12/2009	0	0,03	0,00	0,13	0,00	0,02	0,02
17/12/2009	0	0,00	0,00	0,15	1,30	0,51	0,51
18/12/2009	2,4	7,89	0,00	1,98	2,44	6,47	7,61
19/12/2009	3,4	2,32	0,00	0,12	0,00	1,62	1,97
20/12/2009	0	0,01	0,00	0,98	3,32	1,32	1,30
21/12/2009	35,7	6,22	17,30	47,39	9,81	21,17	23,20
22/12/2009	10,1	41,57	12,69	47,90	25,88	48,74	55,46
23/12/2009	56,3	36,59	3,71	18,38	28,04	39,28	45,01
24/12/2009	33,4	36,66	5,40	36,19	13,54	35,16	40,15
25/12/2009	23,7	31,95	8,40	5,42	4,64	30,01	36,46
26/12/2009	0	2,21	0,00	0,60	0,56	1,76	2,08
27/12/2009	0	0,00	0,00	0,07	0,00	0,00	0,00
Cumul_Ev	165,00	165,45	47,50	159,33	89,52	186,06	213,77
Diff par rapport Obs		0%	71%	3%	46%	-13%	-30%

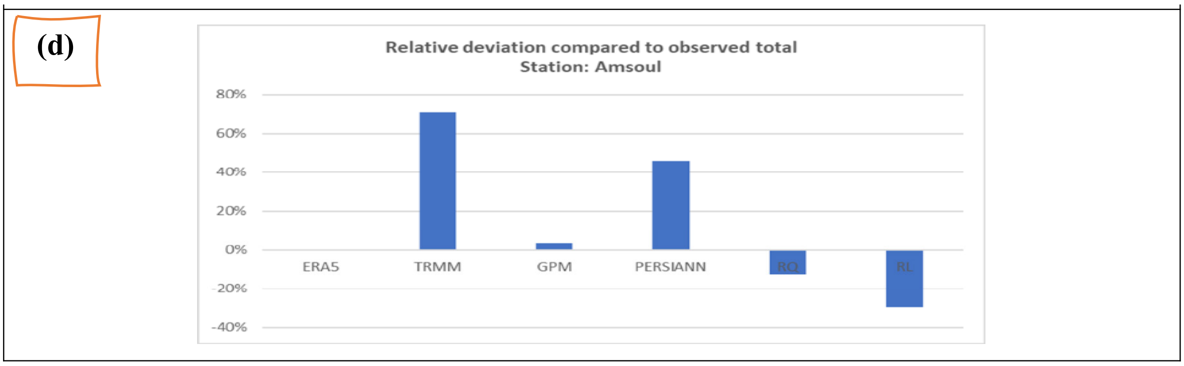


Figure 8: (a) Location of the Amsoul Station, (b) Observed Rainfall, Rainfall Calculated by Different Satellites, Rainfall by RQ Model and RL Model, (c) Model, (d) Difference Compared to the Observed Cumulative Rainfall

Validation of RQ Model in Region 4

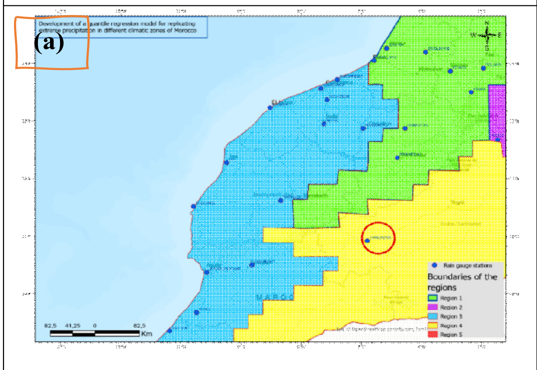
The model we developed in region 4 allows the calculation of precipitation using the following equation:

$$P_{RQ} = -\ 0.06 * P_{(ERA5)} + 0.90 * P_{(GPM)} + 0.74 * P_{(TRMM)} + 0.48 * P_{(PERSIANN)}$$

Ouarzazte Station

To validate the quantile regression (RQ) model in region 4, we selected the March 2002 event recorded at the Ouarzazate station, as illustrated in figure 9 (a). This event, lasting two days, accumulated 71.90 mm of precipitation. By analyzing the satellite product totals, we observe that all satellite products detected the event with varying discrepancies compared to ground-based measurements, ranging from 80% for TRMM to 24% for PERSIANN. The RQ model manages to minimize the discrepancy to -7% compared to ground data, while the linear regression (RL) model reduces the discrepancy to -15%.

Station : Ouarzazte | Lat = 30.93 ; Long = -6.9



(c)

Date	Obs	ERA5	TRMM	GPM	PERSIANN	RQ	RL
30/03/2002	0,00	0,00	0,00	0,11	3,49	1,76	2,06
31/03/2002	48,20	1,40	9,96	24,37	33,88	43,18	43,70
01/04/2002	23,70	38,05	4,32	29,41	13,33	29,77	34,20
02/04/2002	0,00	0,10	0,26	0,25	3,95	2,31	2,66
Cumul_Ev	71,90	39,55	14,55	54,15	54,65	77,02	82,62
Diff par rapport Obs		45%	80%	25%	24%	-7%	-15%

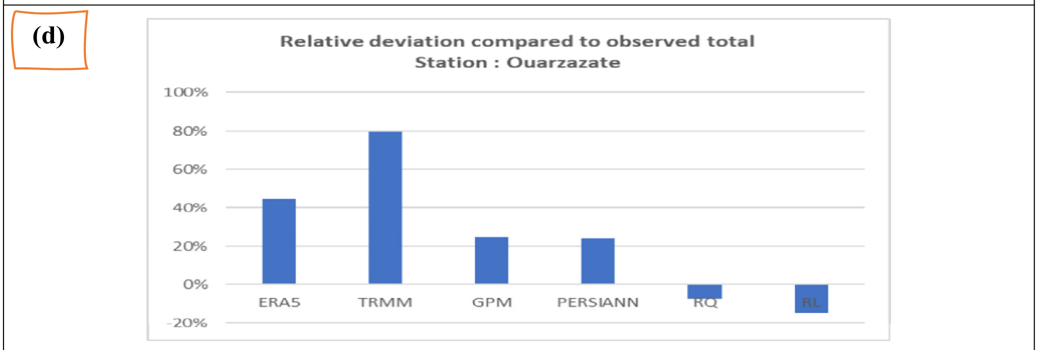


Figure 9: (a) Location of the Ouarzazate Station, (b) Observed Rainfall, Rainfall Calculated by Different Satellite Products, Rainfall from the RQ Model and the RL Model, (c) Model, (d) Discrepancy Compared to Observed Total

Validation of RQ Model in Region 5

The model we developed in region 3-1 allows the calculation of precipitation using the following equation:

$$P_{RQ} = 1.23 * P_{(ERA5)} - 1.20 * P_{(GPM)} + 0.84 * P_{(TRMM)} + 2.13 * P_{(PERSIANN)}$$

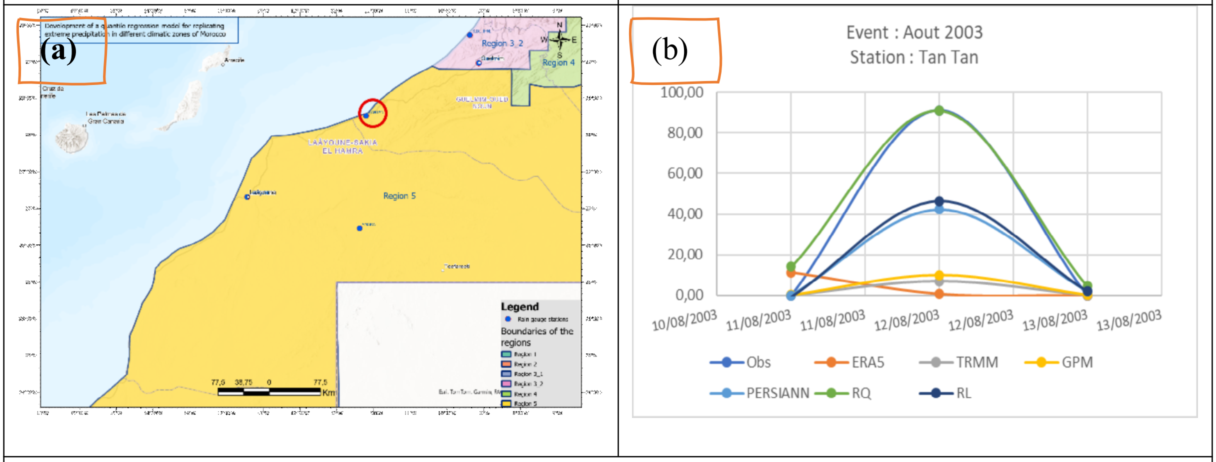
Region 5 is characterized by an arid climate with rare precipitation. Therefore, we used rainfall events exceeding 10 mm to build our RQ model, as opposed to the 20 mm threshold used for other regions. This decision aims to retain more data, as our preliminary data analysis showed that extreme events in this region are rare and do not exceed 50 mm. The validation of the RQ model yielded satisfactory results. In this region, only stations managed by the DGM are available.

Tan Tan Station

For the August 2003 event recorded at the Tan Tan station, as shown in figure 10 (a), we observe that ERA5 did not detect this event, as illustrated in Figure 9(a) and (b). TRMM and GPM detected the event but with very low totals (TRMM = 6.84 mm and GPM = 10.19 mm) compared to the ground-measured total of 91.10 mm, as shown in Figure 9(c). In contrast, PERSIANN estimated a total of 44.35 mm. Using our quantile regression (RQ) model, the total reached 109.47 mm, with a 20% overestimation compared to the ground data. This total is higher than all satellite products and the linear regression, which provided a total of 47.78 mm, resulting in a 48% underestimation compared to the ground data.

Station : Tan Tan

Lat = 28.26 ; Long = -11.6



(c)

Date	Obs	ERA5	TRMM	GPM	PERSIANN	RQ	RL
11/08/2003	0,00	11,29	0,00	0,06	0,00	13,94	-0,80
12/08/2003	91,10	0,62	6,84	9,95	42,33	91,08	46,51
13/08/2003	0,00	0,00	0,00	0,17	2,02	4,44	2,02
Cumul_Ev	**91,10**	**11,91**	**6,84**	**10,19**	**44,35**	**109,47**	**47,74**
Diff par rapport Obs	**87%**	**92%**	**89%**	**51%**	**-20%**	**48%**	

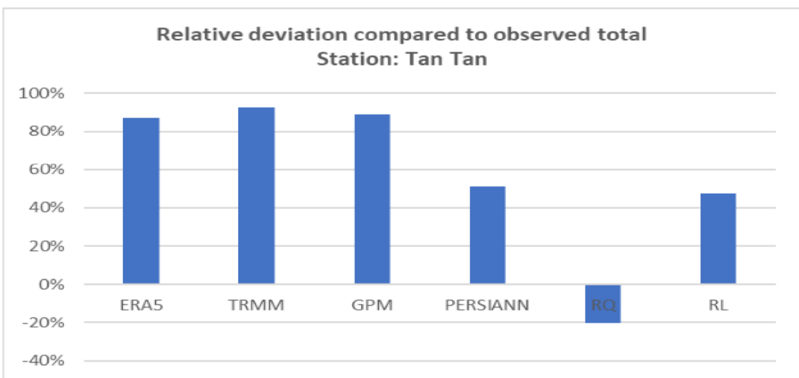


Figure 10: (a) Location of the Tan Tan Station, (b) Observed Rainfall, Rainfall Calculated by Different Satellite Products, Rainfall from the RQ Model and the RL Model, (c) Model, (d) Discrepancy Compared to Observed Total

The validation of the RQ model developed for all climatic regions yielded highly satisfactory results. The developed model will enable researchers interested in hydrological modeling, natural disaster modeling, climate trends, etc., to generate reliable precipitation series, especially for extreme precipitation in various areas of the territory. Our main contribution, beyond the construction of an RQ model, is the addition of a sixth region, corresponding to the subdivision of region 3 into two sub-regions, region 3_1 and region 3_2, as illustrated in figure 11.

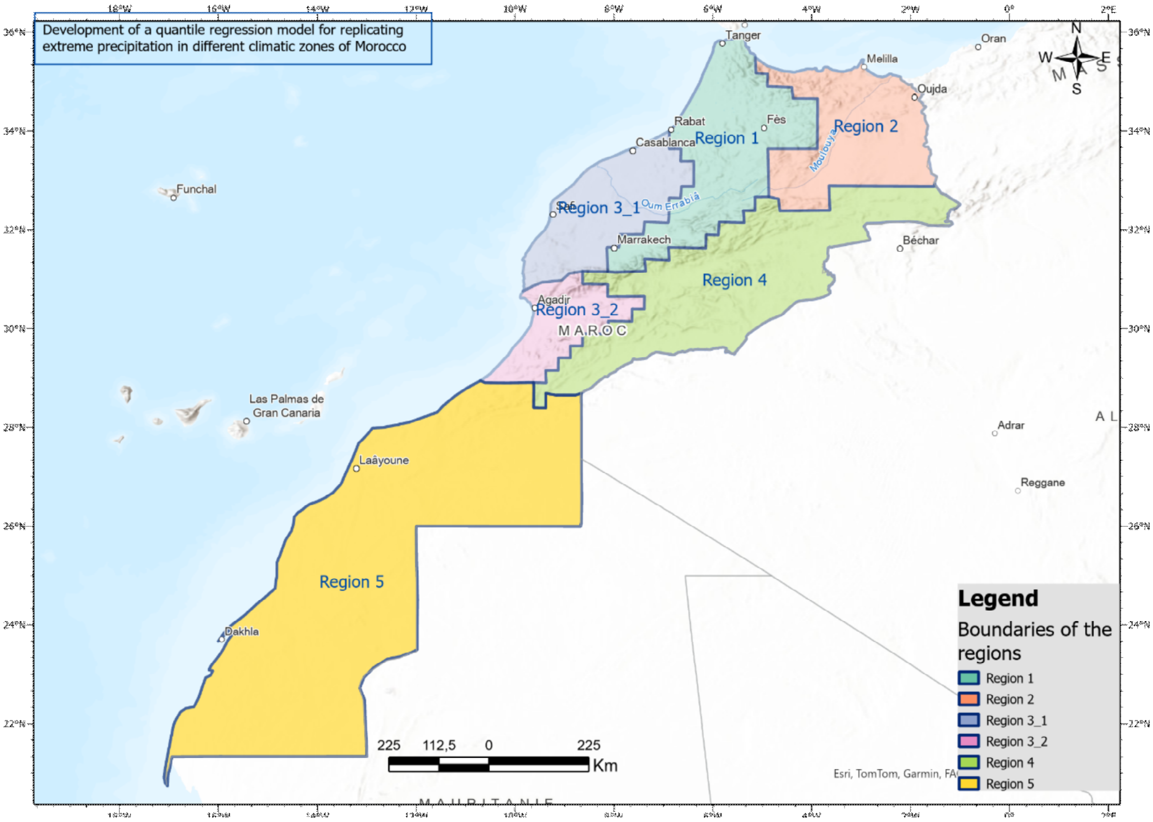


Figure 11: Final Division of Morocco into 6 Homogeneous Climatic Regions

To the best of our knowledge, no similar studies have been developed in other regions of the world, which prevented us from comparing our results with those of other works. One of the main advantages of our model is its ability to be transferred to other regions with a climate similar to that of Morocco. For regions with different climates, we recommend subdividing the territory into homogeneous climatic zones and recalculating the coefficients of the quantile regression (QR) model using the same methodology.

Conclusion

Through this work, we successfully developed a quantile regression (QR) model for each region of Morocco to reproduce extreme precipitation in any location. Additionally, we made a significant improvement to the research of Mr. Chaqdid by subdividing region 3 into two sub-regions: 3-1 and 3-2. This subdivision is justified by the presence of the High Atlas mountain range, which helped reduce the gap between the precipitation totals calculated by the QR model and the observed totals.

The developed model provides highly satisfactory results compared to those of a linear regression model and the use of a single satellite product. The discrepancies with the totals measured by ground stations are very encouraging, ranging from 1% to 13%, except in region 05, where the discrepancy reaches -20%. However, this is still better compared to the linear regression model's error (48%) and the discrepancies of satellite products (ERA5 = 87%, TRMM = 92%, GPM = 49%, and PERSIANN = 51%). These discrepancies are due to false alarms from certain satellite products, as a result of the very specific rainfall regime in this region.

Based on the model validation results for each region, we can affirm that the developed model represents a significant advancement in the study of extreme climatic events in arid and semi-arid regions. This model is based on two major pillars: the integration of multiple satellite products and the use of a model that accounts for extreme values.
The model can be utilized by insurance and reinsurance companies that primarily rely on satellite precipitation for flood risk modeling, aimed at risk transfer.

Our work opens promising perspectives for improving the modeling of extreme hydrological events in regions lacking dense measurement networks. In the future, integrating QR models with medium- and long-term climate forecasts could strengthen natural disaster risk management related to flooding. Furthermore, applying this approach to other types of extreme climatic events, such as droughts, could contribute to a more comprehensive management of water resources in the context of climate change. The inclusion of other environmental factors, such as temperature, humidity, evaporation, topography, and others, in the QR model could also further refine the accuracy of predictions [15].

Declarations

The authors declare that they have **no financial or non-financial conflicts of interest** in relation to this publication. The presented work did not receive any funding from public, commercial, or non-profit organizations. The authors used their own resources to conduct this study.

Authors' Contributions

- Oumechtaq Ismail: Article writing
- Oulidi Abderrahim: Discussion
- Bahaj Tarik and Oulidi Abderrahim: Review

References

1. Guo, H., Chen, S., Bao, A., Hu, J., Gebregiorgis, A. S., Xue, X., & Zhang, X. (2015). Inter-comparison of high-resolution satellite precipitation products over Central Asia. Remote Sensing, 7(6), 7181-7211.
2. Sultana, R., & Nasrollahi, N. (2018). Evaluation of remote sensing precipitation estimates over Saudi Arabia. Journal of Arid Environments, 151, 90-103.
3. Katiraie-Boroujerdy, P. S., Nasrollahi, N., Hsu, K. L., & Sorooshian, S. (2013). Evaluation of satellite-based precipitation estimation over Iran. Journal of arid environments, 97, 205-219.
4. Ouatiki, H., Boudhar, A., Tramblay, Y., Jarlan, L., Benabdelouhab, T., Hanich, L., ... & Chehbouni, A. (2017). Evaluation of TRMM 3B42 V7 rainfall product over the Oum Er Rbia watershed in Morocco. Climate, 5(1), 1.
5. Chaqdid, A., Tuel, A., El Fatimy, A., & El Moçayd, N. (2023). Extreme rainfall events in Morocco : Spatial dependence and climate drivers. Weather and climate extremes, 40, 100556.
6. Gomis-Cebolla, J., Rattayova, V., Salazar-Galán, S., & Francés, F. (2023). Evaluation of ERA5 and ERA5-Land reanalysis precipitation datasets over Spain (1951–2020). Atmospheric Research, 284, 106606.
7. Berges, J. C. (2019). GPM, le programme Global Precipitation Measurement (NASA/JAXA) –fondements théoriques et accès à la base de données. Climatologie, 16, 23-51.
8. Huffman, G. J. 2007. The TRMM Multisatellite Precipitation Analysis (TMPA) : Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology, 8(1), 38–55.
9. Hong, Y., Hsu, K. L., Sorooshian, S., & Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. Journal of Applied Meteorology, 43(12), 1834-1853.
10. Rasera, G., Anabor, V., Scremin Puhales, F., & Dal Piva, E. (2018). Developing an MCS index using the climatology of South America. Meteorological Applications, 25(3), 394-405.
11. Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific data, 2(1), 1-21.
12. Koenker, R., & Bassett, G. (1978). Regression Quantiles. In Econometrica (Vol. 46, Issue 1).
13. Roulston, M. S., & Smith, L. A. (2002). Evaluating probabilistic forecasts using information theory. Monthly Weather Review, 130(6), 1653-1660.
14. AghaKouchak, A., Behrangi, A., Sorooshian, S., Hsu, K., & Amitai, E. (2011). Evaluation of satellite-retrieved extreme precipitation rates across the central United States. Journal of Geophysical Research : Atmospheres, 116(D2).
15. Sultana, R., & Nasrollahi, N. (2018). Evaluation of remote sensing precipitation estimates over Saudi Arabia. Journal of Arid Environments, 151, 90-103.