EEET ECOLOGICAL ENGINEERING & ENVIRONMENTAL TECHNOLOGY

Ecological Engineering & Environmental Technology 2024, 25(4), 241–251
https://doi.org/10.12912/27197050/183791
ISSN 2719-7050, License CC-BY 4.0

Received: 2024.02.04
Accepted: 2024.02.20
Published: 2024.02.28

Evaluation of the Impact of Gap Filling Technology in Precipitation Series on the Estimation of Climate Trends, the Case of the Souss Massa Watershed

Oumechtaq Ismail[1], Laghzali Abbdelmajid[2*], Tarik Bahaj[3], Oulidi Abderrahim[1], Amghar Lamya[1], Allaoui Abdelhamid[1], Mouadil Manal[1], Mustapha Boualoul[4], Bachaoui El Mostafa[2], Elkhaldi Khalid[5]

[1] Solidarity Fund Against Catastrophic Events, Department of Studies and Risk Management
[2] Sultan Moulay Slimane University, Faculty of Science and Technology, Beni Mellal, Morocco
[3] Department of Geology, University FS Rabat, Rabat, Morocco
[4] Department of Geology, University FS Meknes, Meknes, Morocco
[5] Marine Geosciences and Soil Sciences Laboratory, Faculty of Sciences, Chouaib Doukkali University Jabran Khalil, Jabran Avenue B.P 299-24000 El Jadida Grand-Casablanca Morocco
* Corresponding author's e-mail: abdelmajid.laghzali@gmail.com

ABSTRACT

Accurate climatic data, especially precipitation measurements, play a critical role in various studies concerning the water cycle, particularly in modeling flood and drought risks. Unfortunately, these datasets often suffer from temporary gaps that are randomly dispersed over time. This study aims to assess the effectiveness of three imputation methods: KNN, MICE, and missForest, in impute missing values in climate series. The evaluation is conducted in two distinct rainfall regimes: the Moulouya basin and the Sous Massa basin. The performance analysis considers the percentage of missing data across the entire dataset. The imputed datasets are used to estimate annual precipitation, which are then subjected to statistical tests to identify potential trends and detect changepoints. The analysis focuses on the precipitation series within the Souss Massa watershed, encompassing 27 rainfall stations. Results indicate that data imputation has a highly positive impact on the study of rainfall series trends and change point detection. The study found that studying trends without data imputation could lead to questionable conclusions. The most significant breakpoints detected in the analyzed rainfall series were in the years 1988, 1991, 1997, 2007, and 2010. The decrease in precipitation at stations showing a downward trend varies between -60 mm and -137 mm using the MICE method, and between -40 mm and 186 mm using the missForest method.

Keywords: climate trends, change point, precipitation, data imputation, KNN, MICE, missForest, R software.

INTRODUCTION

Understanding the evolution of climatic parameters, such as temperatures and precipitation, requires the availability of sufficiently long and complete datasets, which is often not the case in most places around the world. We often encounter either short chronological series or gaps within these series. This significantly impacts trend analysis and the modeling of these phenomena. The presence of gaps in climatological datasets, including data on precipitation, temperatures, and evaporation, is a common occurrence due to various factors such as instrument malfunctions, site unavailability, data transmission issues, archiving problems, and more.

Missing data in time series have a significant impact and lead to errors during the analysis and interpretation of hydrological modeling results [Sapriza-Azuri et al., 2019] highlighted the challenges and potential errors associated

with this data gap. [Melki et al., 2020] examined the consequences of missing precipitation data on hydrological modeling, highlighting possible distortions in forecasts and hydrological analyses. Completed climatological data is crucial for obtaining accurate results [Evin et al., 2021]. Furthermore, the study by [Zhao et al., 2018] emphasized the importance of completing missing precipitation data for forecasting streamflow in ungauged basins, highlighting the impact on the reliability of these forecasts. In this work, we emphasize the importance of filling gaps in precipitation series before analyzing climate trends and detecting change points in the Souss Massa watershed.

Rapid advancements in the fields of computer science and scientific research have led to numerous imputation techniques, some of which require significant computational capabilities while others do not. This situation presents advantages but also raises two crucial questions: the choice of the best imputation technique to use and the impact of these techniques on the study of trends and the detection of change points in climate series.

In this study, we have focused on two main aspects: evaluating the performance of three imputation techniques (KNN, MIC, and missForest) under two different rainfall regimes (Moulouya basin and Sous Massa basin). We selected these three techniques due to their flexibility and wide range of applications [Van Buuren et al., 2011]. Subsequently, we will examine the impact of each imputation technique on the study of climate trends using the Mann Kendall test and the detection of change points in the Souss Massa basin. Like any gap-filling operation, calculating the percentage of missing values is a crucial step. We have computed this percentage per station for the Moulouya basin.

An analysis of the results obtained in the table below shows that only 16 of the 59 stations had no missing values, representing 27% of the total. It is quite logical to understand that to obtain good results when modeling a phenomenon that calls rainfall data, it is essential to have a clear knowledge of the variability of precipitation in different localities, which requires a significant number of stations. Consequently, it is necessary to fill in the gaps in the other stations.

MATERIALS AND METHODS

The study area

Rainfall patterns can significantly vary from one region to another, making it crucial to test different imputation techniques across various rainfall regimes to ensure their applicability in different geographical situations. With this aim, we

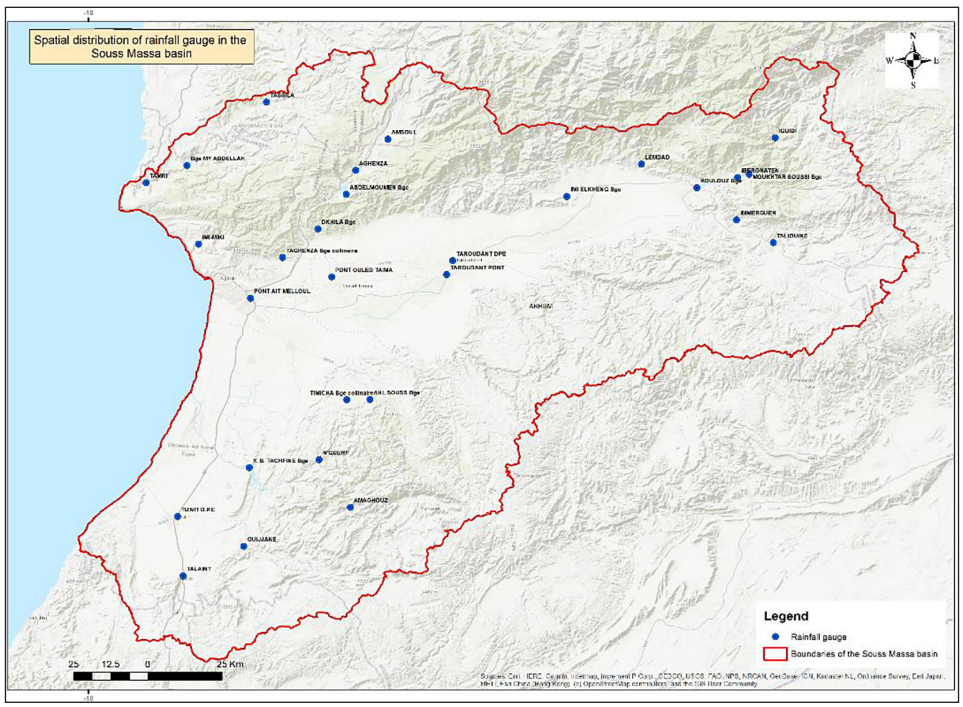

Fig. 1. Spatial distribution of rainfall stations in the Souss Massa watershed

selected two different watersheds: the Moulouya watershed Figure 1, characterized by a Mediterranean climate where annual precipitation can range from less than 100 mm to just over 600 mm [Driouech 2010], and the Souss Massa watershed Figure 2, characterized by variable rainfall from year to year, with an average annual precipitation of 327 mm.

Methodology

The methodology adopted is based on five stages:

- Step 1 – retrieval of daily precipitation data from the two watersheds, namely the Moulouya and Souss Massa watersheds.
- Step 2 – data formatting – during this stage, the collected data (in matrix format) was transformed into a date-value format to facilitate their utilization in the R software.
- Step 3 – visual verification aimed at eliminating irrelevant data such as text (rainfall traces, instrument malfunctions, observer leave, etc.) and symbols within the datasets.
- Step 4 – gap filling – in this stage, we initiated the filling of gaps detected in the series using three imputation techniques, subsequently evaluating the performance of each technique.
- Step 5 – based on the completed daily rainfall data, we generated annual rainfall series to assess the impact of each imputation technique on trend analysis and change point detection. These steps are summarized in Figure 3:

Data

The primary input data consists of daily precipitation measurements collected by rain gauge stations placed at various locations within the two watersheds.

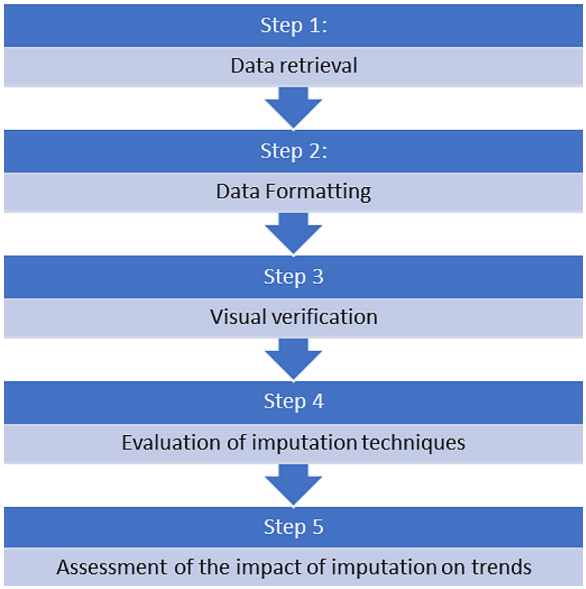

Fig. 3. Working methodology

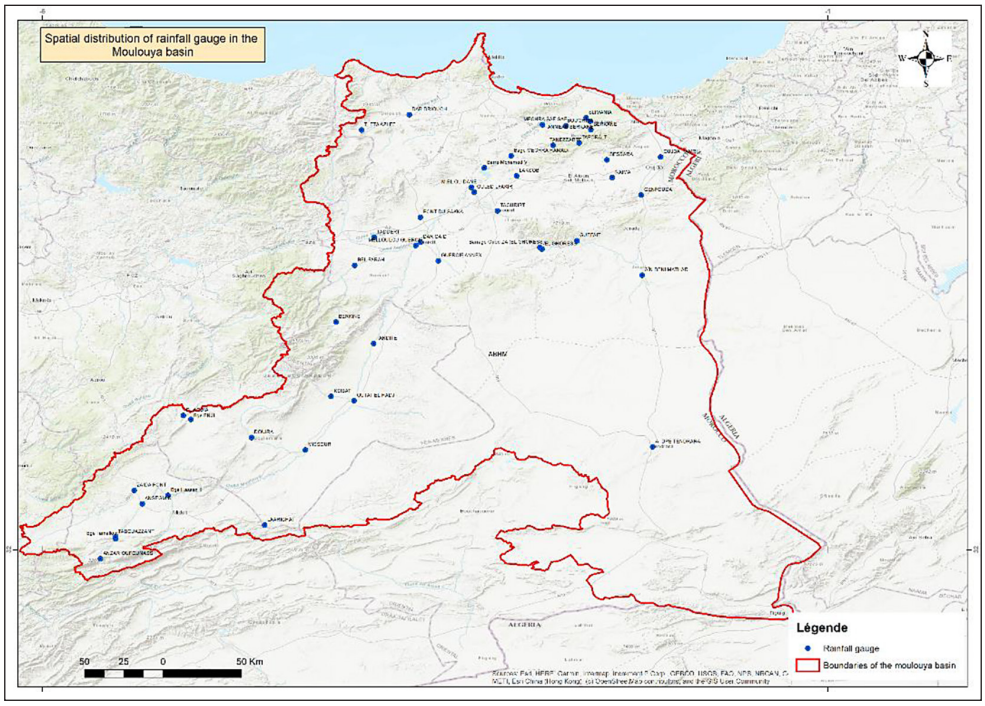

Fig. 2. Spatial distribution of rainfall stations in the Moulouya watershed

Materials

R software was used to perform all the required calculations, given its richness in terms of documentation and packages, as well as its performance in performing complex and repetitive operations.

Methods

Performance criteria

To evaluate the performance of the imputation methods, we used the following three statistical indicators: MAE, RMSE and CV RMSE. The model with the low values in these indicators would be the best.

- MAE: mean absolute error:

$$MAE = \frac{\sum_{i=1}^{h} \left| X_i^{obs} - X_i^{imp} \right|}{h} \quad (1)$$

- RMSE: square root mean square error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{h} \left| X_i^{obs} - X_i^{imp} \right|^2}{h}} \quad (2)$$

- CV RMSE: coefficient of variation of the square root of the root mean square errors:

$$CVRMSE = \frac{RMSE}{\overline{X_{obs}}} \sqrt{\frac{\sum_{i=1}^{h} \left| X_i^{obs} - X_i^{imp} \right|^2}{h}} \quad (3)$$

where: X_{obs} – the average of the values of the variable X observed on all the data studied.

Imputation techniques

The treatment of missing data has been widely studied in the statistical literature [Imbert et al., 2018, Niass et al., 2015, Rousseau et al., 2012]. Several methods for calculating missing data have been developed and can be distinguished into two domains: the domain of time series and the domain of periodic data analysis, univariate and multivariate, the reconstruction of missing data has been widely studied in the field of time series [Aissia 2014, Marlinda et al., 2010, Nejjari et al., 2020].

- KNN method – KNN is a useful compilation method for matching a point with its k nearest neighbors in multidimensional space. Originally introduced in 2001 by O. Troyanskaya for the study of gene expression [Troyanskaya et al., 2001]
  - Step 1 – calculation of the distances between the i and the n-1 records.
  - Step 2 – the average of the k nearest neighbors.
- MICE method – multiple imputation by chained equations (MICE), is based on a Monte-Carlo Markov Chain algorithm. In this imputation technique, many regression models are run such that the variable with missing data is modeled in terms of other variables in the data set [Bousri et. al., 2021].

The steps for applying the method are as follows:

- Step 1 – imputation by the mean
- Step 2 – missing values of a single variable
- Step 3 – regress by the other variables
- Step 4 – predict the missing values of this variable
- Step 5 – repeat for the other variables
- Step 6 – repeat m times
- Step 7 – merge the results m times

Statistical trend and breakout tests

- Trend detection test
  - Mann Kendall trend test – is used to determine with a nonparametric test whether a trend is identifiable in a time series that possibly includes a seasonal component. This nonparametric trend test is the result of an improvement of the test first studied by Mann (1945) then taken up by Kendall (1975) and finally optimized by Hirsch (1982, 1984) in order to take into account a component seasonal. The Pettitt (1979) and Mann-Kendall (1947) [Mann et al., 1947, Pettitt 1979] tests are so-called non-parametric statistical tests because they make no assumptions about the underlying distribution of the data. In particular, they apply to data not having a Gaussian distribution. They are therefore adapted to hydrometeorological data for which the distributions are often asymmetrical, these two tests have been used by several researchers in the study of climatic trends [Acharki et al., 2019].
  - Change point detection test – homogeneity tests bring together a large number of tests for which the null hypothesis is that a time series is homogeneous between two given times. As part of this work, we opted for the use of Pettitt's test since it is used by

several researchers in studies similar to our case [Acharki et al., 2019, Paturel et al., 1998, Paturel et al., 2004].

– Pettit's test – belongs to the category of non-parametric tests that do not require any assumptions about the distribution of the data [Pettitt 1979]. The test is an adaptation of the rank-based Mann-Whitney test to identify the time at which a change occurs. Several authors have used this test in the study of change points in the climatic series (precipitation and temperature series).

RESULT AND DISCUSSION

The results of this study are presented in a structured manner. We commenced by presenting the outcomes regarding the evaluation of the performance of missing data imputation techniques. Following that, we analyzed the trends at each station before and after filling the gaps using various techniques. Finally, we discussed the results pertaining to change point detection, corresponding to the dates of modification in rainfall patterns.

Evaluation of imputation techniques' performance

The first step we undertook was calculating the percentage of gaps within each dataset across all rainfall stations in both watersheds. The percentage of missing data across all stations in the Souss Massa watershed is 41.7%, while the available data represent 58.3% of the total. Concerning the Moulouya watershed, the overall percentage of missing data is 32.4%, with 62.7% of data available. The percentages of missing data per station are depicted in Figures 4 and 5. Upon

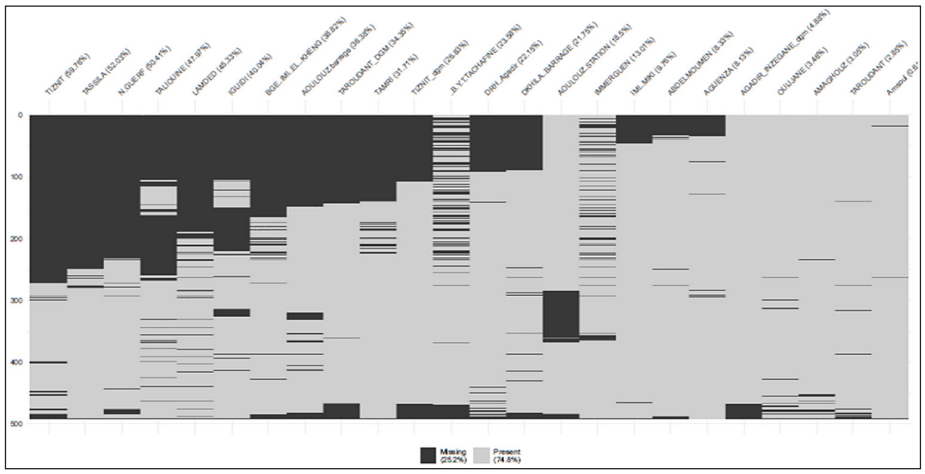

Fig. 4. Missing data in the basin of Souss-Massa

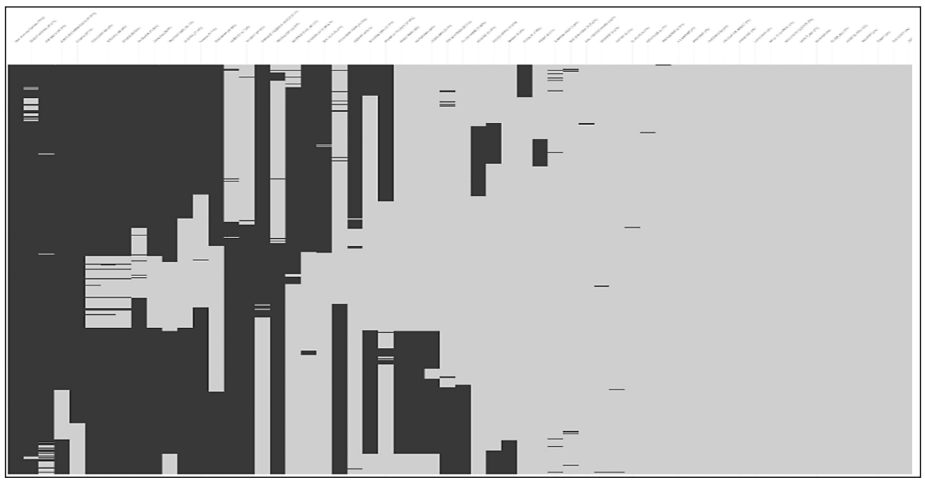

Fig. 5. Missing data in the basin of Moulouya

analyzing the results, it's evident that only 16 out of 59 stations have no missing values, accounting for 27% of the total. It's essential to comprehend that achieving accurate modeling for phenomena reliant on precipitation data necessitates a clear understanding of precipitation variability across different locations, mandating a significant number of stations. Hence, filling in the gaps in other stations becomes necessary.

Two approaches were employed. The first involved measuring three parameters MAE, RMSE, and CVRMSE at Souss Massa watershed stations. These results are presented in Table 1. The second approach measured the NRMSE parameter for varying percentages of missing values (10%, 20%, 30%, 40%, and 50%) in the Moulouya watershed stations. The results are provided in Table 2. This was done to assess the impact of the quantity of missing values on the performance of each technique.

The results demonstrate that regardless of the percentage of missing values, the missForest technique remains the most effective as it minimizes all indicators in both watersheds. This outcome aligns with other studies conducted in different regions worldwide, such as the study by Bousri et al. [Bousri, I., & al., 2021].

Trend study

After imputing the missing data, resulting in the creation of four (04) databases: a database without imputation, a database imputed by the KNN technique, a database imputed by the MICE technique, and a database imputed by the missForest technique, we examined the trends at each rainfall station within these four databases. The trend study results are summarized in Table 4:

- the symbol 0 – indicates no significant trend;
- the symbol + – indicates an upward trend;
- the symbol - – indicates a downward trend.

The analysis of the impact of filling gaps using three imputation techniques on precipitation series trends in the Souss Massa watershed revealed the following:

- trend analysis on raw data showed that 14 stations had no significant trend, 13 stations exhibited an upward trend, and no station showed a downward trend.
- imputation using the KNN method with K=5 maintained the same distribution of stations as the raw data, except for one station displaying a downward trend.
- on the other hand, imputation using the MICE method identified 5 stations with a downward trend, 22 stations with no significant trend, and no stations with an upward trend.
- imputation using the missForest method detected a downward trend for 3 additional stations compared to the MICE method, totaling 8 stations with a downward trend. Additionally, one station showed an upward trend.

Table 5 summarizes the results of the trend analysis obtained.

Change points study

The Pettitt test results indicate that data imputation enhances the detection of change points in rainfall series. When calculations were performed on raw data, we identified ten (10) groups

Table 1. Performance evaluation results of the techniques for selected stations in Souss Massa

Station	MAE			RMSE			CVRMSE		
	K_NN	MICE	missForest	K_NN	MICE	missForest	K_NN	MICE	missForest
AGADIR	7.3	5.1	5.4	10.8	8.5	8.1	0.7	0.5	0.5
OUIJJANE	8.2	10.2	5.3	14.8	17.6	8	0.8	0.9	0.4
AMAGHOUZ	20.4	24.2	13.3	25.3	32.3	18.4	0.8	1	0.6

Table 2. NRMSE calculation results for each percentage of missing data in the Moulouya basin

NRMSE_Bassin de la Moulouya					
Test	10%	20%	30%	40%	50%
MissForest	15.59	26.32	32.14	38.19	44.26
MICE	22.28	31.22	43.12	49.68	62.51
KNN	20.1	38.71	47.46	58.83	67.76

Table 3. Results of trend analysis using the Mann-Kendall test

Imputation	Sans imputation			Imputation KNN_K=5			Imputation par MICE			Imputation par missForest		
Station	tau	P-value	Tendance	tau	P-value	Tendance	tau	P-value	Tendance	tau	P-value	Tendance
S1	-0.192	0.073	0	-0.163	0.127	0	-0.081	0.450	0	-0.061	0.570	0
S2	-0.065	0.544	0	-0.065	0.544	0	-0.05	0.650	0	-0.043	0.690	0
S3	-0.072	0.503	0	-0.079	0.464	0	-0.167	0.120	0	-0.114	0.290	0
S4	-0.085	0.426	0	-0.051	0.638	0	-0.161	0.130	0	-0.205	0.050	0
S5	-0.017	0.884	0	-0.006	0.967	0	0.006	0.970	0	0.056	0.600	0
S6	-0.099	0.357	0	-0.092	0.391	0	-0.043	0.690	0	-0.019	0.870	0
S7	0.052	0.630	0	0.028	0.802	0	-0.012	0.920	0	-0.017	0.880	0
S8	0.144	0.180	0	0.008	0.950	0	-0.083	0.440	0	-0.152	0.150	0
S9	0.159	0.137	0	0.157	0.140	0	-0.118	0.270	0	-0.174	0.100	0
S10	0.34	0.002	+	0.322	0.003	+	-0.05	0.650	0	-0.032	0.770	0
S11	0.183	0.088	0	0.156	0.143	0	-0.118	0.270	0	-0.145	0.170	0
S12	0.3	0.006	+	0.222	0.037	+	-0.094	0.380	0	-0.134	0.210	0
S13	0.389	0.000	+	-0.238	0.025	-	-0.317	0.003	-	-0.411	0.000	-
S14	0.462	0.000	+	0.174	0.103	0	-0.226	0.030	-	-0.225	0.030	-
S15	0.496	0.000	+	0.497	0.000	+	-0.147	0.170	0	-0.236	0.030	-
S16	0.007	0.961	0	0.189	0.075	0	-0.176	0.100	0	-0.096	0.370	0
S17	0.439	0.000	+	0.451	0.000	+	-0.147	0.170	0	-0.260	0.010	-
S18	0.036	0.767	0	0.269	0.011	+	-0.129	0.230	0	0.070	0.520	0
S19	0.489	0.000	+	0.43	0.000	+	-0.099	0.360	0	-0.172	0.110	0
S20	0.484	0.000	+	0.34	0.001	+	-0.198	0.060	0	-0.209	0.050	-
S21	-0.008	0.963	0	-0.058	0.594	0	-0.254	0.020	-	0.342	0.000	+
S22	0.591	0.000	+	0.522	0.000	+	-0.07	0.520	0	-0.145	0.170	0
S23	0.627	0.000	+	0.542	0.000	+	-0.165	0.120	0	-0.229	0.030	-
S24	0.584	0.000	+	0.426	0.000	+	-0.257	0.020	-	-0.127	0.230	0
S25	0.606	0.000	+	0.324	0.002	+	-0.162	0.130	0	-0.183	0.090	0
S26	0.543	0.000	+	0.455	0.000	+	-0.054	0.620	0	-0.043	0.690	0
S27	0.175	0.184	0	0.355	0.001	+	-0.265	0.010	-	-0.207	0.050	-

Table 4. Distribution of the number of stations by type of trend

Test de Mann Kendall	Etat brute	Imp_KNN (K=5)	Imp_MICE	Imp_MissForest
Pas de tendance significative	14	13	22	18
Tendance vers la baisse	0	1	5	8
Tendance vers la hausse	13	13	0	1

of stations, as shown in Table 6; these groups vary in terms of homogeneity. For instance, in group 3 (G3), stations had change points at K = 14, others at K = 15, and further ones at K = 16. Note – a group represents a set of rainfall stations that share the same change point and exhibit the same trend either upward or downward after the change point.

However, after imputing missing data using KNN and MICE, the number of groups decreased from ten (10) to only four (04) highly homogenous groups regarding change points. Imputation using the missForest technique added another group compared to KNN and MICE, with a change point at k = 28, as seen in Table 6. Additionally, slight modifications were observed in station assignments to different groups and in the detected change point values.

In terms of the number of stations per group, group G3 contains the highest number of stations (15 stations) with a change point at K = 21, corresponding to the year 1997. Following that, group G2 consists of 07 stations with K = 14 (and station S1 with K=15), corresponding to the year 1991. Group G1 contains only 03 stations with

Table 5. Results of change point detection tests using different techniques

Etat brute				KNN				MICE				MissForest			
Groupe	Station	K	p_value	Groupe	Station	K	p_value	Groupe	Station	K	p_value	Groupe	Station	K	p_value
G1	S1	6	0.053	G1	S14	12	0.26	G1	S14	12	0.26	G1	S3	14	0.64
	S2	7	0.038		S21	12	0.17		S21	12	0.17		S6	14	1.00
G2	S3	10	0.105		S10	13	1.00		S10	13	1.00		S16	14	1.00
	S4	10	0.001	G2	S3	14	0.37	G2	S3	14	0.37		S22	14	0.62
	S5	10	0.001		S5	14	1.00		S5	14	1.00		S23	14	0.10
G3	S6	14	1.000		S6	14	1.00		S6	14	1.00		S1	15	1.00
	S7	14	0.767		S16	14	0.35		S16	14	0.35	G2	S4	21	0.30
	S8	14	0.077		S22	14	0.93		S22	14	0.93		S8	21	0.35
	S9	15	0.000		S24	14	0.06		S24	14	0.06		S9	21	0.33
	S10	15	0.048		S1	15	1.00		S1	15	1.00		S11	21	0.43
	S11	16	0.000	G3	S13	20	0.01	G3	S13	20	0.01		S12	21	0.62
	S12	16	0.000		S4	21	0.36		S4	21	0.36		S13	21	0.00
G4	S13	18	0.000		S8	21	0.77		S8	21	0.77		S14	21	0.40
	S14	18	0.433		S9	21	0.62		S9	21	0.62		S15	21	0.13
G5	S15	21	0.082		S11	21	0.69		S11	21	0.69		S17	21	0.08
	S16	21	0.873		S12	21	0.96		S12	21	0.96		S19	21	0.35
	S17	21	0.002		S15	21	0.29		S15	21	0.29		S20	21	0.16
G6	S18	22	0.000		S17	21	0.47		S17	21	0.47		S26	21	1.00
	S19	23	0.000		S18	21	0.54		S18	21	0.54	G3	S21	28	0.00
	S20	24	0.000		S19	21	0.69		S19	21	0.69		S25	28	0.06
	S21	25	0.000		S20	21	0.16		S20	21	0.16	G4	S5	31	0.82
G7	S22	28	0.000		S23	21	0.20		S23	21	0.20		S7	31	1.00
G8	S23	31	1.000		S25	21	0.22		S25	21	0.22		S10	31	1.00
G9	S24	34	0.819		S26	21	1.00		S26	21	1.00		S18	31	0.21
	S25	34	0.957		S27	21	0.01		S27	21	0.01	G5	S2	34	1.00
	S26	34	0.002	G4	S7	31	1.00	G4	S7	31	1.00		S24	34	0.64
G10	S27	38	1.000		S2	34	1.00		S2	34	1.00		S27	34	0.04

K = 12 (and K=13 for S10), corresponding to the year 1988. Finally, G4 contains 02 stations with K = 31 and 34, corresponding respectively to the years 2007 and 2010. The Pettitt test can serve as a critical criterion for determining homogeneous zones concerning rainfall patterns, especially by grouping stations that share the same change point. Therefore, it can be affirmed that stations within each group exhibit similar rainfall patterns. The annual rainfall trend of stations within group

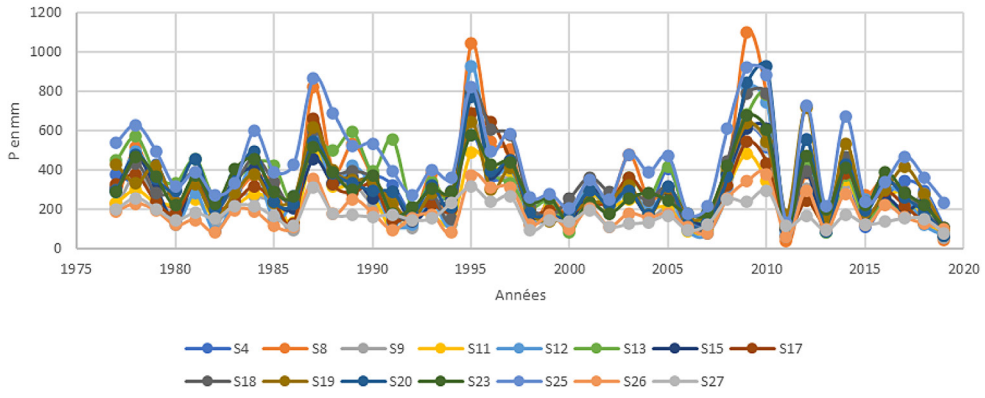

Fig. 6. Evolution of the annual rainfall of the G3 group stations (imputation by KNN)

G3, imputed by MICE, demonstrates that all stations exhibit a similar rainfall pattern throughout the observation period. The same observation applies to stations within group G2, imputed by missForest. Referencing Figures 6 to 9 provides an example of this. The Pettitt test results confirm the Mann-Kendall test outcomes by verifying the detected change date from the Pettitt test and the rainfall regime trend of the station. For instance, the graphical representation of annual

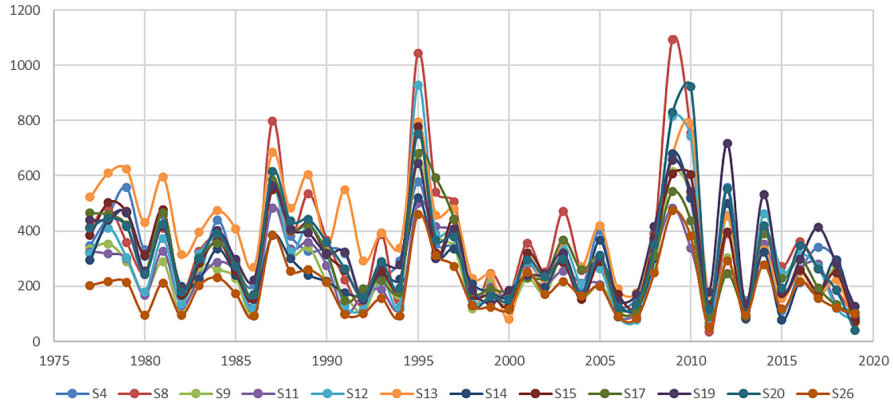

Fig. 7. Evolution of the annual rainfall of the G2 group stations (imputation by missForest)

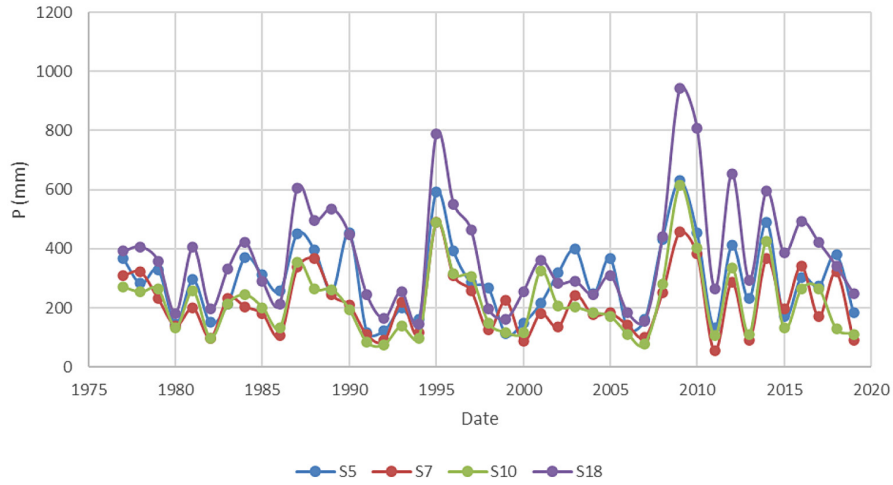

Fig. 8. Evolution of the annual rainfall of the G4 group stations (imputation by missForest)

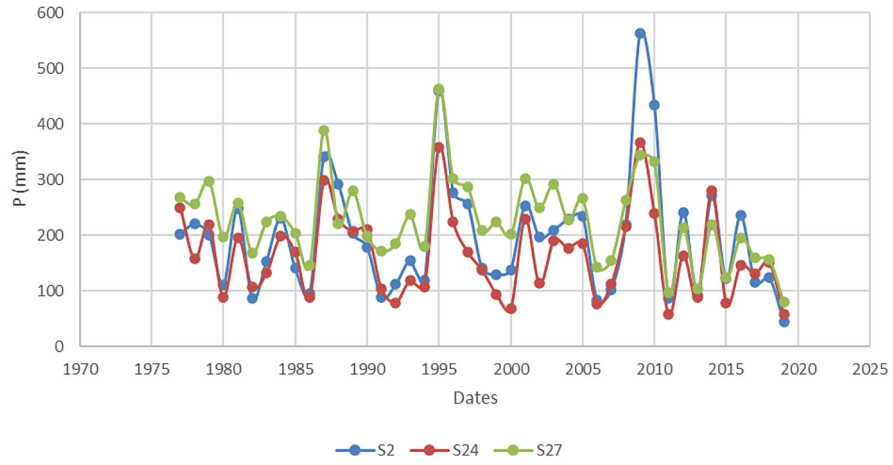

Fig. 9. Evolution of the annual rainfall of the stations of the G5 group (imputation by missForest)

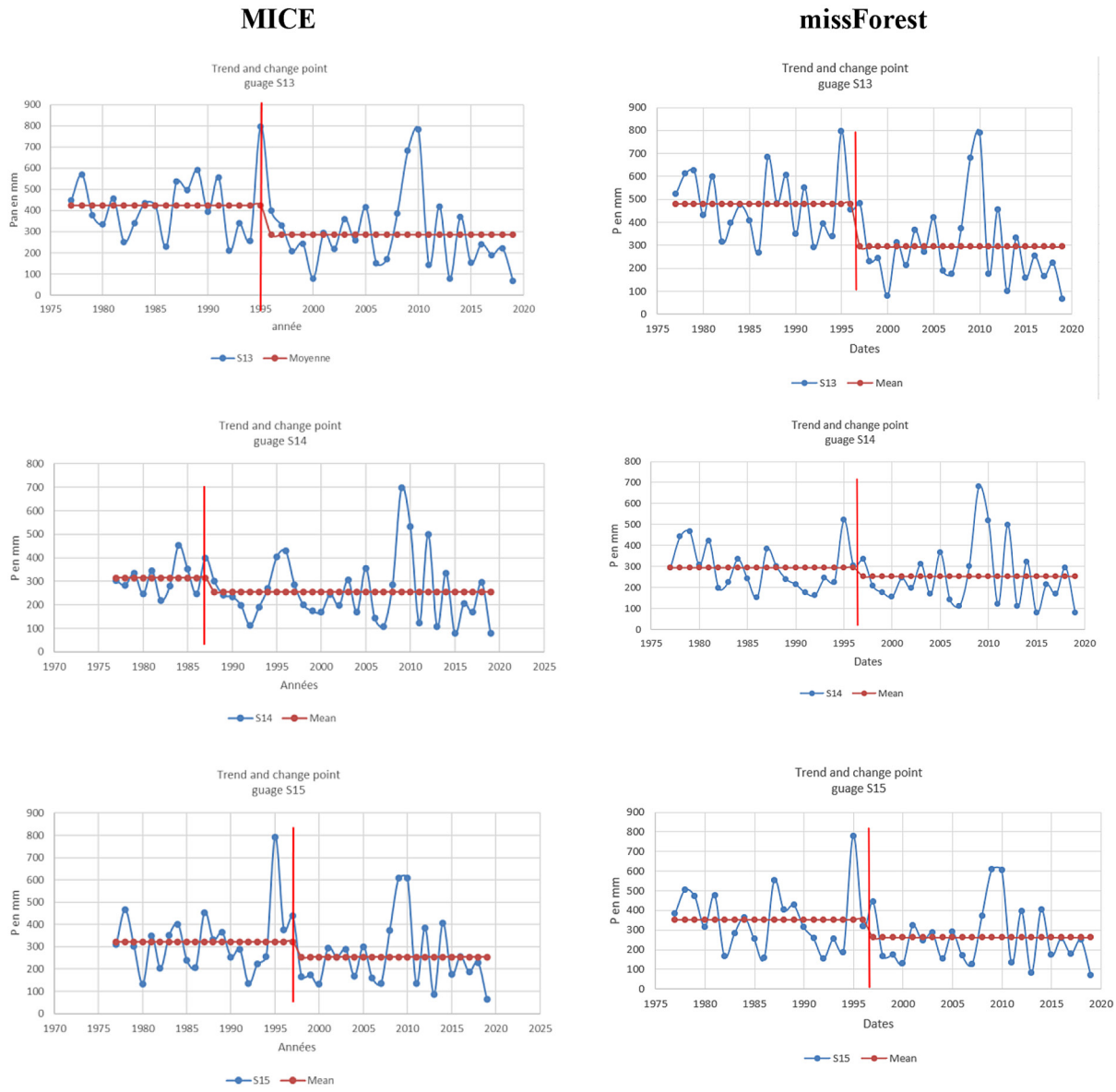

Fig. 10. Comparison of the impact of data imputation using MICE and misForest on change point detection and the amount of rainfall decrease

rainfall for station S13, seen in Figure 10a, clearly demonstrates a downward trend from the detected change point (year 1995). The decrease concerning the average is -137 mm when using the MICE technique and -186 mm when using the missForest technique.

The change points in rainfall series, particularly in stations displaying significant decreases, are illustrated in Figure 10.

CONCLUSIONS

In general, the missForest method proves to be the most effective, followed by the MICE method, while the K-MN method exhibits the poorest performance. These results hold true for both watersheds. The percentage of missing data does not influence the reliability of the applied techniques. For instance, the missForest imputation method consistently remains the most efficient regardless of the proportion of missing data.

Examining trends and change points without applying an imputation technique can lead to misleading conclusions about historical trends and change points. Regarding change point detection, the KNN and MICE methods yield similar results, identifying four groups with the same change point. On the other hand, the missForest method allows for the classification of stations

into five groups, showing slight variations in the change point for certain stations.

Significant improvement is observed in station grouping based on change point compared to the raw data, i.e., without any imputation. Applying the MICE method reveals that five stations in the Souss Massa basin exhibit a decreasing trend, no increasing trend, and 22 stations show no clear trend. Considering the results obtained from the missForest method, eight stations in the Souss Massa basin display a decreasing trend, one station shows an increasing trend, and 18 stations exhibit no clear trend. The most notable change point dates detected in the Souss Massa basin are 1988, 1991, 1997, 2007, and 2010. The decrease in precipitation for stations exhibiting a downward trend (S13, S14, and S15) ranges from -60 mm to -137 mm according to the MICE method, and from -40 mm to 186 mm according to the missForest method.

Acknowlegements

We would like to express our great gratitude to Professor Mohammed El idrissi -Polydisciplinary Faculty- Beni Mellal. for his valuable and constructive suggestions during the planning and development of this research work. His willingness to give his time so generously was much appreciated.

REFERENCES

1. Acharki, S., Amharref, M., El Halimi, R., Bernoussi, A.S. 2019. Assessment by statistical approach of climate change impact on water resources: Application to the Gharb perimeter (Morocco). Water Science Review, 32(3), 291–315. https://doi.org/10.7202/1067310ar
2. Aissia, M.A.B. 2014. Étude des variables hydrologiques dans un cadre multivarié et dans un contexte de changement. Ph.D. Thesis, Québec University, Québec.
3. Bousri, I., Salah, S.A., Arab, B.M. 2021. Validation d'une méthode d'imputation de données manquantes pour la reconstitution des séries de température. JAMA, 5, 28–32.
4. Driouech, F. 2010. Distribution of winter precipitation over Morocco in the context of climate change: downscaling and uncertainties. Ph.D. Thesis, Toulouse University, Toulouse.
5. Evin, H., Suetsugu, F., Kagabu, M. 2021. The Importance of Filling Missing Data in Hydrological Modeling: A Review of Methods and Impacts
6. Imbert, A., Vialaneix, N. 2018. Describing, accounting for, imputing and evaluating missing values in statistical studies: a review of existing approaches. Journal de la société française de statistique, 159(2), 1–55.
7. Zhao L., Hu X., et al. 2018. Importance of Filling Missing Rainfall Data in Streamflow Forecasting in Ungauged Catchments
8. Mann, H.B., Whitney, D.R. 1947. On a test of whether one of two random variables is stochastically larger than the other. The annals of mathematical statistics, 50–60.
9. Marlinda, Uyun, A.S., Miyazaki, T., Ueda, Y., Akisawa, A. 2010. Performance analysis of a double-effect adsorption refrigeration cycle with a silica gel/water working pair. Energies, 3(11), 1704–1720.
10. Melki S.S., Kariuki S.M. 2020. Impact of Missing Rainfall Data on Hydrological Modeling.
11. Nejjari, I., Abdelhai, S., Lebzar, B. 2020. Measuring human capital: dimensions and alternative methods. La Revue de Publicité et de Communication Marketing, 1(2).
12. Niass, O., Diongue, A.K., Touré, A. 2015. Analysis of missing data in sereo-epidemiologic studies. African Journal of Applied Statistics, 2(1), 29–37.
13. Paturel, J.E., Servat, E., Delattre, M.O., Lubes-Niel, H. 1998. Analysis of rainfall long series in non-Sahelian West and Central Africa within a context of climate variability. Hydrol. Sci. J., 43(6), 937–946.
14. Paturel, J.-E., Ibrahim, B. 2004. Sahelian Paradox View project Monthly rainfall gridded data set for Africa View project.
15. Pettitt, A.N. 1979. A non-parametric approach to the change-point problem. Journal of the Royal Statistical Society: Series C (Applied Statistics), 28(2), 126–135.
16. Rousseau, M., Simon, M., Bertrand, R., Hachey, K. 2012. Reporting missing data: a study of selected articles published from 2003–2007. Quality & Quantity, 46, 1393–1406.
17. Sapriza M.P., Azuri, J., Buytaert, B., Timbe, et al. 2019. implications of missing rainfall data on hydrological modelling: a case study in the Tropical Andes.
18. Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Altman, R.B. 2001. Missing value estimation methods for DNA microarrays. Bioinformatics, 17(6), 520–525.
19. Van Buuren, S., Groothuis-Oudshoorn, K. 2011. Journal of Statistical Software mice: Multivariate Imputation by Chained Equations in R, 45, 1–67.