

**SIMULATION DES PRECIPITATIONS MENSUELLES PAR UN
MODELE STATISTIQUE :
APPLICATION A LA STATION D'ALGER
SIMULATION OF MONTHLY PRECIPITATIONS BY A
STATISTICAL MODEL.
APPLICATION TO THE STATION OF ALGIERS**

Doudja SOUAG-GAMANE Université des Sciences et de la Technologie
Houari Boumediene, Faculté de Génie Civil, BP. 32 Bab Ezzouar
16111, Alger Algérie souagd@yahoo.fr

Noureddine DECHEMI Ecole Nationale Polytechnique d'Alger, 10
Avenue Hacene Badi, BP 16182 El Harrach, Alger, Algérie
ndechemi@yahoo.fr

RESUME : La génération des séries synthétiques trouve son application dans l'optimisation de la gestion des barrages réservoirs. En général la simulation est destinée au design, la planification, les études opérationnelles des schémas de gestion de la ressource en eau ainsi que les études des changements climatiques. La génération des précipitations exige des modèles dont la combinaison et la configuration dépendent des processus et des échelles temporelle et spatiale impliquées. Très peu de travaux ont été réalisés depuis 1985, dans le domaine de la génération stochastique des précipitations, plus particulièrement pour des climats semi-arides ayant des séries pluviométriques à caractère purement aléatoire et non stationnaires. Ce travail présente essentiellement une approche pour la génération de longues séries synthétiques de précipitations mensuelles sous un climat semi-aride (dans lequel s'inscrit la région du nord algérien), qui préservent les caractéristiques statistiques des séries historiques.

Mots clés : génération stochastique, Analyse en composantes principales, semi-aride, temporel, variabilité climatique.

ABSTRACT : The generation of synthetic series finds its application in the optimization of water resources management. Generally the simulation is intended for the design, the planning, the operational studies of water

resources management as well as the studies of climate change. The generation of the precipitation requires models with structures and configurations depending on processes and on the temporal and spatial scales involved. Very few works were realized since 1985, in the field of the stochastic generation of the precipitation, more particularly for semi-arid climates having rainfall series with purely random character and not stationary. This work presents essentially an approach for the generation of long synthetic series of monthly precipitation under a semi-arid climate (the case of the region of the North Algeria), which reproduces correctly the statistical characteristics of the historical series

Keywords : Stochastically generation, Principal component analysis, semi-arid, temporal, climatic variability

INTRODUCTION

Les enregistrements historiques ne donnent qu'une réalisation unique des événements climatiques, différents scénarii de données générées peuvent alors être utilisés pour évaluer l'impact de la variabilité climatique sur les ressources en eau et les systèmes agricoles. Les modèles de simulation produisent des séries synthétiques de précipitation à différentes échelles spatiales et temporelles.

Trois grandes catégories de modèles de simulation ont été développées et utilisées pour la simulation des variables hydrométéorologiques qui sont les modèles généraux de circulation, les modèles spatio-temporels stochastiques et les modèles stochastiques journalier. Le choix et l'utilisation des modèles devraient être justifiés par la précision désirée des résultats et de la capacité d'estimer entièrement les paramètres exigés.

Les précipitations mensuelles décadaires ou journalières peuvent être générées soit directement par un modèle au pas de temps correspondant soit par désagrégation des valeurs annuelles simulées (plus facilement modélisables).

Très peu de travaux ont été réalisés depuis 1985, dans le domaine de la génération stochastique des précipitations (Srikanthan et McMahan, 2001).

Thomas et Fiering (1962) ont été les premiers à avoir proposé un modèle de Markov d'ordre 1 pour générer des séries de débits. Srikanthan et McMahon (1985) ont ensuite recommandé ce modèle pour générer des données de précipitations. Il a été largement utilisé dans ce domaine sous différents climats (Lana et Burgueno, 1998; Wilks, 1999; Chineke et al., 1999; Castellvi et Stockle, 2001). Cependant, le principal inconvénient de ce modèle est qu'il ne peut pas modéliser correctement les longues périodes sèches ou humides (Racsco et al., 1991; Guttorp, 1995; Semenov et Porter, 1995); les paramètres de ce type de modèles ne prenaient en considération que les variations saisonnières et ignoraient totalement les variations interannuelles. De longues périodes sèches et humides ont été observées dans le passé. La prise en compte, dans la modélisation des précipitations, de ce phénomène est de ce fait devenue très importante. Thyer et Kuczecra (2000) ont alors développé un modèle de Markov à état caché (Hidden State Markov, HSM) qui suppose explicitement que le climat est composé de deux états, un état sec (années à faibles pluies) et un état humide.

Le modèle proposé dans le présent travail est un modèle statistique pour la génération des précipitations mensuelles qui tente de s'affranchir des différentes contraintes des modèles existants, telles que la chronologie, la longueur des séries et la normalité de la variable.

MATERIELS ET METHODES

Données utilisées

Pour la simulation de la pluviométrie au pas de temps mensuel, nous avons pris en considération une série historique de précipitations recueillies à la station météorologique d'Alger constituée de 50 années d'observations. Les caractéristiques statistiques mensuelles historiques (moyennes écart-types et coefficients de variation données dans le tableau 1 montrent bien l'irrégularité du régime pluviométrique de la région.

Tableau 1. Caractéristiques statistiques mensuelles des précipitations journalières.

Mois	Jan	Fév	Mar	Avr	Mai	Jui	Juil	Aoû	Sep	Oct	Nov	Déc
Moyennes	2.8	2.7	2.0	2.2	1.1	0.5	0.1	0.1	0.9	1.8	2.4	3.5
Ect. Typ	6.8	6.6	5.6	6.7	4.5	2.6	1.0	1.5	5.5	8.0	8.7	9.2
CV	2.4	2.4	2.7	3.0	4.0	5.1	10.3	19.2	6.1	4.3	3.6	2.6

Les caractéristiques des données mensuelles sont données dans le tableau 2. Nous pouvons constater, d'après le coefficient d'asymétrie C_s et le coefficient de variation C_v mensuels interannuels calculés, que les valeurs mensuelles de précipitations présentent une distribution fortement dissymétrique, avec une grande variabilité.

Tableau 2. Caractéristiques statistiques des précipitations mensuelles utilisées

Variable X	X_{moy}	S_x	C_v	C_s	X_{min}	X_{max}
Précipitation	200.00	36.27	1.649	1.508	140.2	0.0

Modèle de simulation

Le modèle proposé dans le présent travail est un modèle statistique pour la génération des précipitations journalières basé sur l'Analyse en Composantes Principales, une technique reconnue par sa puissance aussi bien dans le domaine de l'analyse descriptive que de l'analyse opérationnelle.

Le modèle de simulation est donné par l'expression (1)

$$\tilde{Y}_j = \sum_{l=1}^Q Cov(Y_j, C'_l) \cdot C'_l + \varepsilon_j \quad (1)$$

Avec ;

Y_j la variable simulée au temps j;

$Cov(Y_j, C'_l)$ covariance entre la variable étudiée et la l^{ième} composante principale (CP) normée;

Q nombre de composantes principales prises en considération;

Y_j variable étudiée;

ε_j variable résiduelle, donnée par la différence entre les variables observées et reconstituées à l'aide des Q composantes principales prises en considération. La variable ε_j est de moyenne nulle et d'écart-type σ_{ε_j} .

La génération des précipitations mensuelles par le modèle élaboré revient à la simulation des trois composantes du modèle à savoir: les covariances entre variables, les composantes principales et les résidus.

RESULTATS ET DISCUSSIONS

Le modèle élaboré a été utilisé pour la simulation d'un phénomène complexe et très difficile à cerner, particulièrement pour une région à climat semi-aride connu par ses saisons de pluies d'une extrême irrégularité et susceptibles d'interruptions qui peuvent durer plusieurs semaines, voir plusieurs mois ou années (séries chronologiques intermittentes). Tout cela rend le problème de simulation du phénomène très délicat, plus précisément à des pas de temps inférieurs à l'année (mois, décade, jour, ...).

Le principe de l'analyse en composantes principales étant de concentrer le maximum d'information dans un nombre réduit de variables, le choix du nombre de composantes principales à prendre en considération dépend du pourcentage de variance totale expliquée. L'ACP est effectuée sur une matrice constituée de 12 variables (valeurs de pluies mensuelles) sur 50 individus (nombre d'années d'observations). Les résultats montrent que plus de 70% de la variance totale est expliquée par les quatre premières composantes (33.3% des composantes principales sont nécessaires pour reconstituer 72% de l'inertie totale).

Tableau 3. Apports des composantes principales

	<i>C1</i>	<i>C2</i>	<i>C3</i>	<i>C4</i>	<i>C5</i>	<i>C6</i>
Valeur propre	2,72	2,49	1,84	1,66	1,13	0,83
% variance	22,69	20,74	15,32	13,83	9,38	6,96
% cumulé	22,69	43,43	58,74	72,57	81,95	88,90

La génération des précipitations mensuelles par le modèle revient à la simulation des trois groupes de paramètres: les covariances entre variables, les composantes principales et les résidus.

Les séries historiques des covariances et des composantes principales sont d'abord ajustées à différentes lois statistiques par la méthode du maximum de vraisemblance, les meilleurs ajustements ont été donnés par la loi Gumbel et la loi de Gauss. Ces deux paramètres ont été donc simulés dans ces deux lois ainsi que dans leurs fonctions de répartition en utilisant une interpolation par la méthode du spline cubique.

Les résidus suivent une loi normale de moyenne nulle et d'écart-type σ_ε , ils ont alors été simulés dans cette loi en plus des fonctions de répartition et des chaînes de Markov.

Plusieurs variantes, résultant de la combinaison de ces différentes méthodes de simulation des paramètres du modèle, sont testées et validées en utilisant les tests statistiques, qui consistent à comparer les moments de différents ordres, les fonctions de distributions de probabilités ainsi que des caractéristiques statistiques des périodes sèches et humides des séries historiques et simulées.

La figure 1 montre bien la bonne reproduction des moyennes et écart-types mensuels. La structure de dépendance séquentielle représentée par la fonction d'autocorrélation est aussi bien reproduite (mémoire du phénomène) par le modèle comme le montre la figure 2.

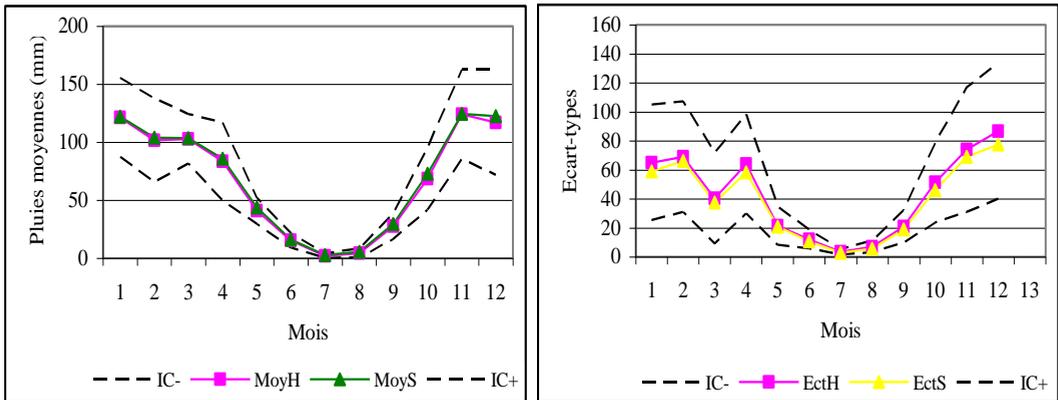


Fig. 1. Moyennes et écart-types des séries historique et simulée.

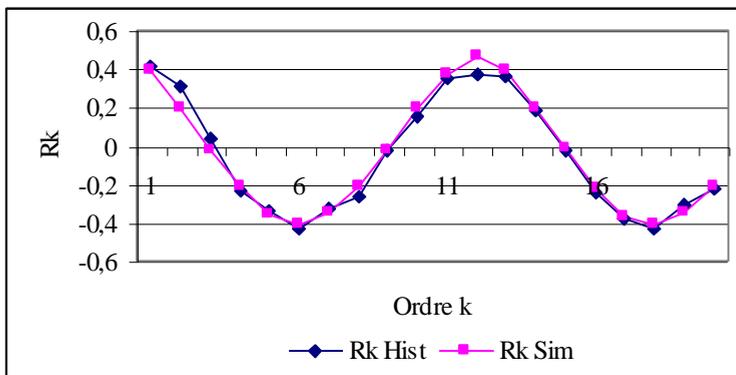


Fig. 2. Autocorrélogrammes historique et simulé.

De plus les caractéristiques statistiques des périodes extrêmes (sèches et humides) données par le tableau 4 montrent que la structure temporelle est reproduite de façon satisfaisante.

Tableau 4. Caractéristiques statistiques des périodes extrêmes historiques et simulées.

	Période sèche		Période Humide	
	Vol Moy (mm)	Vol max (mm)	Vol Moy (mm)	Vol max (mm)
Historiques	62.97	134.05	298.31	928.01
Simulées	68.82	186.77	348.42	1097.75

	Période sèche			Période Humide		
	Long.Moy (mois)	Long.Max (mois)	Ect-type (mois)	Long.Moy (mois)	Long.Max (mois)	Ect-type (mois)
Historiques	2.6	5.0	1.6	4.2	9.0	2.6
Simulées	3.1	8.0	1.7	4.8	11	2.8

Les modèles stochastiques généralement utilisés pour la simulation de telles séries exigent la stationnarité, d'où l'obligation de leur faire subir des transformations pour les rendre stationnaires.

L'avantage de la technique utilisée dans ce travail est que la stationnarité n'est plus une contrainte à prendre en considération.

Par ailleurs, pour l'utilisation des modèles Markoviens, ou des modèles autorégressifs (AR ou ARMA), par exemple, la chronologie des séries historiques s'impose ce qui conduit à réduire la taille de l'échantillon utilisé et par conséquent une perte d'information. Le modèle élaboré s'adapte bien à des séries tronquées puisque la chronologie ne pose aucun problème. En effet les résultats obtenus pour des séries tronquées d'une manière aléatoire, en ne gardant que 40 années d'observations sont aussi satisfaisants que ceux obtenus pour une série de 40 années successives (tous les tests utilisés ont été vérifiés).

CONCLUSION

Le modèle de simulation élaboré est basé sur l'analyse en composantes principales. Il présente de grands avantages par rapport aux modèles existants, il offre une variété importante de combinaisons de méthodes de génération et permet de s'affranchir des contraintes imposées par les autres modèles. Le type de variable, la stationnarité, la troncature des observations, ou la longueur de la série historique ne constituent en aucun cas un handicap à l'application du modèle élaboré.

REFERENCES BIBLIOGRAPHIQUES

- Lana X, Burgueno A (2000a): *Some statistical characteristics of monthly and annual pluviometric irregularity for the Spanish Mediterranean coast*. Theoretical and Applied Climatology. 65: 79-97
- Olsson, J., (1998): *Evaluation of a scaling cascade model for temporal rainfall disaggregation*. Hydrology and earth system sciences, 2(1), 19-30.
- Souag-Gamane D., Dechemi N., Bermad A. (2007): *Simulation de la pluviométrie journalière en zone semi-aride par l'analyse en composantes principales*. Sécheresse Volume 18, Numéro 2 : 97-105
- Srikanthan, R., McMahon, T.A., (2001): *stochastic generation of annual, monthly and daily climatic data; A review*. Hydrology and earth system sciences, 5(4), 653-670.
- Thomas, H., & Fiering M. B. (1962). *Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation*. In design of water resources systems (A. Mass et al., edition), pp. 459-493, Cambridge, Massachusetts, Harvard University Press.
- Thyer, M., Kuczera., G., (2000): *modelling long term persistence in hydro-climatic time series using a hidden state Markov model*. Water resources research, 36, 3301-3310.