

OPTIMISATION DE L'ÉCHANTILLONNAGE ET TRAITEMENT DES DONNÉES D'UNE RECHERCHE SUR LA SÉDIMENTOLOGIE ACTUELLE (ÉROSION-SÉDIMENTATION) SOUS CLIMAT INTERTROPICAL HUMIDE

J. Ph. MANGIN

Université de Nice

RÉSUMÉ

Environ 100 000 variables analytiques ont été retirées d'une recherche de six ans en Côte d'Ivoire portant sur les eaux naturelles, de la pluie à l'océan, et sur leur action d'érosion chimique et mécanique et de transport. L'optimisation de l'échantillonnage, compte tenu des contraintes de coût et d'éloignement des laboratoires d'analyses, a conduit à des *prélèvements aussi nombreux que possible au cours d'un même événement* (averse, crue, etc.) pour des événements se situant au début, au milieu et à la fin d'une séquence naturelle (saison des pluies par exemple).

Le traitement mathématique a porté sur deux aspects : *les corrélations entre elles* des variables physiques, chimiques, minéralogiques, etc. et la *reconstitution d'un phénomène complet* (ex. transport solide fluvial annuel) à partir des points de mesures obtenus au cours de ce phénomène. D'excellents résultats ont été obtenus au moyen d'une analyse factorielle par ordinateur qui fait apparaître les corrélations les moins visibles du tableau des données. La reconstitution du phénomène complet en partant de quelques points de mesures se fait le plus aisément en utilisant la méthode des variogrammes.

ABSTRACT

A six years research in Ivory Coast provided about 100 000 analytical variables taken from rain and run-off waters, rivers and soils in order to describe actual conditions of erosion, transport and sedimentation.

The *optimization of the sampling net* conducted to collect the greatest number of measurements or of samples during an « event » (flow, rain, a.s.o.) and to repeat this method for the events of the beginning, the middle and the end of a natural sequence (wet season or flow season a.s.o.).

The mathematical process had two aims : correlation between two (or more) physical, chemical, mineralogical variables and rebuilding of a phenomenon only known by several samples.

Factorial analysis with a computer gave us the most convenient results for the first point ; the use of variograms seems to be the best for the second purpose.

The model obtained by both the optimization and the mathematical process is fair similar to the natural one.

РЕЗЮМЕ

Оптимизация отбора проб и обработка данных исследования по современной седиментологии (эрозия-осадконакопление) в условиях тропического влажного климата.

Было извлечено примерно 100 000 аналитических переменных после шестилетнего исследования, на территории Берега Слоновой Кости, природных вод, от дождевых до океанских, и их действия в отношении химической и механической эрозии и переноса. Оптимизация отбора проб, принимая во внимание трудности связанные с затратами и отдалением лабораторий производящих анализы, привела к сбору возможно наибольшего количества проб в течение одного и того-же события (ливень, паводок и т. д.) для событий имеющих место в начале, середине и в конце одного естественного отрезка времени (например дождевой сезон).

Математическая обработка велась в двух направлениях: парные корреляции между физическими, химическими, минералогическими и иными переменными, и восстановление одного целого явления (напр. годовой, твердый речной сток), исходя из пунктов измерений полученных для данного явления. Были получены хорошие результаты путем факториального анализа при помощи электронной вычислительной машины, выявляющей наименее заметные корреляции в таблице данных. Восстановление явления в его целости, исходя из нескольких пунктов измерений, облегчается при использовании метода вариограмм.

Nous sommes à la fin d'une opération multidisciplinaire en Côte d'Ivoire sur un bassin versant de 100 000 km² (le fleuve Bandama axé N/S) drainant, dans sa partie nord, une savane sur granite et latérite et une forêt sur schistes dans sa partie sud. Durant six années, des prélèvements et des dosages ont été conduits sur l'eau naturelle à tous les stades de son circuit, de la pluie à l'eau de mer ; les sols et leur substratum, la couverture végétale ont été examinés et analysés à intervalles fixés par le déroulement des saisons. Le régime hydrologique du fleuve Bandama et de ses affluents a fait l'objet de mesures quotidiennes. Les données rassemblées sont en cours de traitement mathématique.

Bref, géologues des sédiments, hydrologues, pédologues, botanistes, océanographes, chimistes, micropaléontologistes et mathématiciens se sont associés sous l'égide de l'ORSTOM* pour :

- étudier aussi exactement que possible le comportement actuel d'un bassin soumis à l'altération et à l'érosion en climat intertropical contrasté ;
- évaluer l'exportation solide et soluble de ce bassin ;
- mettre en évidence la sédimentation mécanique et chimique au long du vecteur fluvial, dans son estuaire et à son débouché en mer.

Une telle opération devrait apporter des données brutes sur des phénomènes d'altération, d'érosion, de transport et de sédimentation.

En outre, ce modèle actuel, choisi en milieu naturel peu affecté par l'occupation humaine, pourrait sans doute faciliter la compréhension de certains milieux de sédimentation fossiles (MANGIN 1963).

**

Les données brutes et leur interprétation seront trouvées par ailleurs dans les travaux des quatre chercheurs géologues qui ont mené l'essentiel de cette campagne (J. LECOLLE, F. LENOIR, P. MATHIEU, et C. MONNET) et dans les mémoires des spécialistes qui les ont aidés. L'ensemble permettra sans doute les comparaisons souhaitées avec quelques milieux fossiles.

Le but, plus modeste, de cette note est de vérifier le bien-fondé de la méthode d'échantillonnage et d'évaluer l'intérêt d'un traitement mathématique des résultats recueillis.

(*) Office de la Recherche Scientifique et Technique Outre-Mer.

I. OPTIMISATION DE L'ÉCHANTILLONNAGE

Pour le sédimentologue, le bassin du Bandama comporte deux parties assez distinctes :

La savane recouvre, au Nord, des *granites* et un peu de *schistes* tandis qu'au Sud la *forêt* s'est installée sur des *schistes* avec quelques zones *granitiques*.

Une étude complète eût nécessité *quatre stations* : savane/granite, savane/schistes, forêt/schistes et forêt/granite. Des raisons d'économie nous ont contraint à nous limiter à *deux stations* principales d'observation : *savane/granite (Korhogo)* et *forêt/schistes (Tiassalé)*.

Sur chacune de ces stations ont été prélevés les échantillons d'eau de pluie, d'eau de ruissellement et d'eau de marigots affluents du Bandama ; sur ce fleuve ont été prélevés les échantillons d'eau en divers points du courant. Les sols soumis aux précipitations et au ruissellement ont fait l'objet, eux aussi, d'un échantillonnage dans les stations.

Dès le début, un programme précis a dû être dressé, nécessaire à l'établissement d'un budget en personnel et en moyens (MANGIN 1962) : c'était le *premier type de modèle*, de style intuitif, sans préjugement facile sur le comportement en nature. Ce *premier modèle* comportait un échantillonnage basé sur une hypothèse de cycle naturel.

Les prélèvements suivaient ce schéma :

- Les sols *avant et après la saison des pluies, aux mêmes points.*
- Les eaux de pluies *en savane, en clairière et sous couvert forestier en trois moments de l'averse (déterminée par lecture de pluviographe) début, maximum et fin pour les principales précipitations de la saison ses pluies (lessivage des poussières atmosphériques) ;*
- Les eaux de ruissellement *comme les eaux de pluie plus un prélèvement quelques heures après la fin d'averse pour le ressuiement des sols ;*
- Les eaux de marigots *comme les eaux de ruissellement ;*
- Les eaux de fleuve Bandama *en 3 points d'une section (RD, RG, axe à -0,50 m de la surface) selon l'hydrogramme des crues et des étiages mais systématiquement en début, sommet et fin d'événement ;*
- Les eaux de lagune et d'embouchure de fleuve, les eaux marines au débouché, *en quatre épisodes majeurs liant les variations haute mer/basse mer aux fluctuations fluviales crues/étiages.*

Plusieurs profondeurs, échelonnées en espace-log, étaient prévues afin de contrôler la stratification des eaux et de leur contenu.

Toutes les prises devaient s'accompagner de mesures physiques in situ de pH, redox, température, résistivité et oxygène dissous. Bref, *ce modèle était assez systématique* et permettait de fixer un calendrier approximatif de missions partant de la base proche d'Abidjan et de chiffrer un budget de prélèvements.

*
**

Après une année d'expérience il est apparu qu'un tel modèle comportait trop peu de points mesurés pour chaque « événement » (averse ou crue ou ressuiement etc.) pour suivre réellement le déroulement du phénomène.

Il devenait nécessaire de mesurer les événements « en continu » car des processus fugitifs échappaient à la maille élaborée précédemment, qui avait seulement permis de les entrevoir ; mais le budget ni l'effectif du personnel ne permettaient de procéder à cette mesure continue des paramètres pour tous les événements de chaque station. Il fallait donc choisir et, optimisation avant la lettre, ce choix a déterminé (MANGIN 1965) de prendre les premiers événements de chaque série (exemple : les premières averses de la saison

des pluies, les premières crues des divers éléments du réseau hydrographique etc.) puis ceux du milieu et de la fin de cette série. En outre, dans la mesure du possible, chaque fois qu'un chercheur était sur place il devait suivre *en continu*, avec les appareils de mesures in situ, quelle que soit l'heure, le déroulement complet d'une averse, d'une crue, d'une décrue, etc. Une contrainte majeure était le nombre réduit d'appareils de mesure in situ : deux sont restés à poste fixe sur les stations Korhogo et Tiassalé et un était affecté aux mesures « volantes » essentiellement sur les diverses stations-radeaux du Bandama*.

Ce dispositif n'a du être modifié que fort peu ; dès la quatrième année, les principales modalités régissant les phénomènes déjà mesurés étaient seulement contrôlées sporadiquement tandis que l'effort portait sur ceux des événements échappant à la maille du deuxième modèle.

Finalement ce mode d'échantillonnage s'est révélé optimisé par rapport à l'objet : *obtenir le plus de données possibles sur les phénomènes actuels d'érosion/sédimentation*, et aux contraintes :

- *Trois appareils de mesures in situ des paramètres physiques.*
- *Quatre géologues-sédimentologues à temps plein.*
- *Deux hydrologues à temps partiel.*
- *100 000 km² de terrain.*
- *Nombre limité d'analyses** devant être faites à Paris avec des résultats différés de trois à six mois par rapport aux dates de prélèvements.*

Je pense que pour toute opération de ce genre, une fois vérifiée la valeur (positive ou négative) du modèle intuitif, le modèle mettant en jeu une série non consécutive d'examen en temps réel d'un événement naturel est le seul qui ne laisse échapper aucun épisode et, de ce fait, ne peut prêter à confusion. La difficulté réside dans le choix de ces événements ; *l'expérience de cette opération montre qu'une maille basée sur les événements extrêmes de la série est la plus significative pour l'explication du phénomène global.* (Exemple : les averses de début, milieu et fin de saison des pluies et les crues de début, de milieu et de fin de l'époque des crues avec quelques mesures aux divers étiages pour rendre compte du débit solide et chimique annuel d'un fleuve).

2. TRAITEMENT DES DONNÉES

Deux aspects à ce problème :

- A) Les corrélations simples ou multiples entre les variables.
- B) La reconstitution d'un phénomène complet à partir de données ponctuelles (crue, érosion due à une averse, transport solide annuel, etc.).

A) Quelques 6 000 échantillons ont été prélevés accompagnés chacun de la notation d'un minimum de 4 paramètres physiques ; tous ont fait l'objet d'un minimum de 24 dosages chimiques auxquels se sont ajoutés des granulométries, des déterminations minéralogiques et diverses mesures pédologiques pour les sols et les sédiments.

Au total, c'est un ensemble d'environ 100 000 données *brutes* composant la « matrice des données » ou « matrice des variables ».

* Rien n'aurait pu être mené à bien sans l'aide dévouée et compétente apportée, de jour et de nuit, par le personnel de l'Hydrologie ORSTOM en Côte d'Ivoire.

** Un très gros effort a été fait par le laboratoire Central ORSTOM à Bondy, qui, sous la direction de Mr PINTA, a assuré de bout en bout une très lourde charge de dosages.

a) Il est bien évident qu'en première approximation la plupart de ces données apportent au bon sens, sans grand calcul, l'hypothèse d'un comportement au sein du processus général : c'est ainsi que peuvent être supposées les principales liaisons entre données qui feront l'objet d'un calcul soit « manuel » soit par machine.

Comme exemples, les relations entre charge solide annuelle et débit liquide, ou entre charge chimique et résistivité de l'eau ou, encore, entre charge chimique exportée et débit liquide etc. (Mangin, Lecolle, Mathieu, Monnet, Pinta et Sircoulon, 1966).

b) Un deuxième stade, plus élaboré, revient à mettre en évidence les corrélations possibles entre deux signaux (débit liquide/débit solide ; silice/chlorure ; alcalins/alcalino-terreux ; valeur de la charge solide/situation temporelle dans la série). Là, un calcul simple de corrélation peut être conduit « à la main » et met l'accent sur des relations plus ou moins attendues. Ces corrélations deux à deux sont intéressantes mais ne rendent compte évidemment que d'aspects ponctuels.

c) Plusieurs signaux peuvent être examinés pour vérifier une corrélation multiple : ce traitement doit être alors effectué par des méthodes statistiques classiques appuyées par un calcul à la machine. Des résultats plus complets apparaissent (DEFOSSEZ, MANGIN, VAN DEN DRIESSCHE, 1967) qui doivent cependant être vérifiés, au stade final, par une analyse factorielle portant sur l'ensemble des variables.

d) Ce stade final (dans l'état actuel de nos connaissances) consiste à établir une analyse factorielle poussée, à l'aide de l'entière matrice des données, qu'il est meilleur, toutefois, de fragmenter en divers épisodes homogènes afin de définir plus aisément les « facteurs » et leur poids réel (factor scores). (Exemple : matrice des données pour une crue forestière puis pour une crue de savane, puis pour l'ensemble. Ou encore : matrice des données pour les eaux de lagune de divers groupes de salinité, etc.).

Les résultats actuellement entre nos mains sont très intéressants et font ressortir en premier lieu une éclatante confirmation des approches obtenues par les calculs « manuels » de corrélations simples entre deux signaux (b). Mais la finesse de l'analyse VARIMAX (KULBICKI et SOURISSE 1968) permet de détecter des corrélations beaucoup plus subtiles, généralement obliérées par les évidences.

En bref, il est certain que les projets similaires à venir gagneront beaucoup au report systématique des données mesurées ou analysées sur support adéquat à l'ordinateur ; on pourra alors obtenir immédiatement (en temps-machine) la matrice des variables dans tous les ordres désirables et, en utilisant les programmes appropriés, le traitement convenable de ces données ; l'analyse factorielle étant, pour le moment, le plus directement rentable de ces traitements.

B) La reconstitution complète d'un événement dont on ne connaît que quelques points de mesure est l'autre aspect du traitement des données auquel notre équipe a dû faire face : charge solide annuelle ; répétition possible, d'une année à l'autre ou d'une crue à l'autre, du décalage entre l'hydrogramme et le diagramme de charge en suspension etc. sont parmi les problèmes posés.

En utilisant les techniques classiques du calcul des probabilités en matière d'extrapolation, il a été possible de reconstituer un modèle donnant :

— les variations des charges solides unitaires suivant un schéma identique pour des hydrogrammes dont les crues sont bien individualisées (MONNET, 1971).

— les variations dans la composition chimique à une station donnée qui sont le fruit, non d'une évolution progressive, mais du passage des eaux de provenances diverses (LENOIR, 1971).

— l'influence du milieu forestier sur les charges solubles (MATHIEU, 1971).

— l'influence du cycle saisonnier sur la sédimentation d'un bassin fluvio-marin (LECOLLE, 1971).

Mais on sait maintenant que le procédé du variogramme (MATHERON, 1969) aurait permis une plus grande précision. Pour construire le variogramme d'une séquence naturelle, il est obligatoire de prévoir des points de mesure très régulièrement espacés (le « pas ») ; c'est ainsi qu'il aurait fallu, pour suivre les probabilités de reproduction d'un phénomène de crue (ou d'averse ou d'érosion), mesurer chaque jour

à la même heure par exemple. Ce qui n'a pas été fait faute d'avoir envisagé à temps cette technique inconnue à l'époque du lancement du programme et, de surcroît, très onéreuse à pratiquer. Le « pas » de mesure n'est pas toujours, en outre, compatible avec l'étude de l'événement dont il peut laisser passer l'essentiel, s'il est fugitif (MANGIN et MIZIKOS, 1971).

Il faut donc conclure, pour obtenir du phénomène naturel établi un modèle aussi complet que possible, à la nécessité de combiner l'échantillonnage décrit plus haut comme adapté à l'événement et la régularité du pas indispensable à la construction de variogrammes. Cette possibilité devra être ajoutée à la liste des contraintes pour les optimisations relatives aux opérations de collecte similaires à la nôtre.

**

Les résultats concrets seront fournis par les chercheurs responsables de chaque tranche de l'opération ; je me contente ici, comme initiateur et directeur du projet, de dresser le bilan de la technique opératoire.

Le but de cette note sera atteint si elle évite à nos successeurs les innombrables tâtonnements qui ont été notre lot.

OUVRAGES CITÉS DANS CETTE NOTE

- DEFOSSEZ M., MANGIN J.Ph., PINTA M., VAN DEN DRIESSCHE R., (1967). — Répartition de quelques éléments trace dans les eaux de surface en zone intertropicale (Côte d'Ivoire). Bull. serv., Carte Géol. Als. Lorr., 20, 4, pp. 257-276.
- KULBICKI G., SOURISSE C. (1968). — Aspects actuels du traitement sur ordinateur des données géochimiques au CRP. Bull. Centre Rech. Pau, SNPA, 2, I, pp. 191-212.
- LECOLLE J., (1971). — Note préliminaire sur la sédimentation et les charges solubles dans un système paralique en milieu intertropical humide (Côte d'Ivoire), C.R. Acad. Sci. Paris, t. 273, pp. 37-40.
- LENOIR F., (1971). — A propos de la répartition saisonnière des éléments transportés en solution dans un fleuve de la zone intertropicale (Le Bandama, Côte d'Ivoire). C.R. Acad. Sci., Paris, t. 273, pp. 34-36.
- MANGIN J.Ph., (1962). — « Programme Mangin ». Côte d'Ivoire 1963-1966. Plan d'étude pour 1963. ORSTOM, Paris.
- MANGIN J.Ph., (1963). — Etude des phénomènes actuels d'érosion, de transport et de sédimentation conduisant à des dépôts alternants. C.R. Somm., Soc. Géol. France, pp. 153.-154.
- MANGIN J.Ph., (1965). — Etude du couple érosion-sédimentation en Côte d'Ivoire (« Opération Mangin »). Le point scientifique fin 1964. ORSTOM, Paris.
- MANGIN J.Ph., LECOLLE J., MATHIEU Ph., MONNET C., PINTA M., SIRCOULON J. (1966). — Géochimie des eaux naturelles. Le transport en solution par un fleuve de Côte d'Ivoire. CR. Acad. Sci., Paris, t. 262, pp. 2204-2206.
- MANGIN J.Ph., MIZIKOS J.P., (1971). — A critical study of methods for mathematical data processing in actual sedimentology. VIII^e Congrès de sédimentologie. Heidelberg, p. 64.
- MATHERON G., (1969). — Cours de Géostatistique. Ecole Nationale Supérieure des Mines. Paris, 82 p.
- MATHIEU Ph.) 1971.— Apports chimiques par les eaux de pluie et action des eaux de ruissellement en milieu forestier tropical (Côte d'Ivoire). C.R. Acad. Sci., Paris, t. 273, pp. 45-48.
- MONNET C., (1971). — Données sur le transport en suspension par un fleuve de zone intertropicale (Bandama, Côte d'Ivoire). C.R. Acad. Sci., Paris, t. 273, pp. 41-44.