

## Remerciements

Je tiens particulièrement à remercier M Célestin KOKONENDJI pour ses conseils et sa disponibilité, Mme Frédérique SEYLER pour son accueil et ses conseils dans l'élaboration de ce rapport, et Mlle Emily LEVIEL pour son aide.

J'adresse aussi mes remerciements à M Jean Loup Guyot pour m'avoir proposé ce stage ainsi qu'à M. Gérard Cochonneau pour m'avoir fourni les différents jeux de données hydrologiques qui ont servi dans ce travail..

**Formatado:** Sem sublinhado,  
Cor da fonte: Automática

**Formatado:** Sem sublinhado,  
Cor da fonte: Automática

# SOMMAIRE

<b>Remerciements</b> .....	<b>1</b>
<b>SOMMAIRE</b> .....	<b>2</b>
<b>Partie I : INTRODUCTION</b> .....	<b>5</b>
<b>Le LMTG</b> .....	<b>7</b>
<b>Les enjeux</b> .....	<b>7</b>
1. Enjeux de connaissance .....	7
2. Enjeux de société .....	8
3. Enjeux de développement .....	8
<b>Organisation</b> .....	<b>9</b>
<b>Organigramme de l'UMR 5563</b> .....	<b>10</b>
<b>L'IRD</b> .....	<b>12</b>
1. Carte d'identité .....	12
2. Objectifs et missions .....	12
3. Organisation .....	12
4. Moyens.....	12
<b>Le projet HYBAM (Hydrologie et Géochimie du Bassin Amazonien)</b> .....	<b>13</b>
<b>Partie II : MATERIELS ET METHODES</b> .....	<b>14</b>
1. <b>Objet de l'étude au sein du LMTG</b> .....	<b>14</b>
2. <b>Le bassin amazonien</b> .....	<b>15</b>
3. <b>Les stations</b> .....	<b>16</b>
4. <b>Les sous-bassins</b> .....	<b>17</b>
5. <b>Définitions des variables et extraction des données</b> .....	<b>18</b>
6. <b>Logiciels utilisés</b> .....	<b>23</b>
7. <b>Méthodes statistiques</b> .....	<b>23</b>

<b>Partie III : ANALYSE DES RESULTATS.....</b>	<b>24</b>	
<b>I. QUALITE D'ESTIMATION DES LARGEURS A PARTIR D'IMAGE JERS.....</b>	<b>24</b>	
1. La largeur mesurée sur le terrain.....	25	
2. La largeur mesurée par le logiciel Arc View.....	26	
3. Comparaison des deux largeurs .....	27	
4. Tests .....	28	
4.1 Test paramétrique .....	28	
4.2 Test non paramétrique.....	29	
5. Conclusion.....	29	
<b>II. RELATIONS STATISTIQUES ENTRE LE DEBIT ANNUEL MOYEN ET LES PARAMETRES GEOMORPHOLOGIQUES DU COURS D'EAU ET DU BASSIN .....</b>	<b>30</b>	
<b>1. Etude des variables .....</b>	<b>31</b>	
1.1 Statistique descriptive .....	31	
1.1.1 La sinuosité .....	31	
1.1.2 La largeur.....	31	
1.1.3 L'aire des sous bassins .....	32	
1.1.4 La pente .....	32	
1.2 Etude des corrélations .....	33	
<b>2. Analyse en Composante Principale .....</b>	<b>34</b>	
2.1 Statistique élémentaire .....	34	
2.2 Choix des axes .....	35	
2.3 Etude des corrélations.....	37	
2.4 Le nuage d'individus .....	39	
2.5 Conclusion.....	39	
<b>3. Régression linéaire simple .....</b>	<b>40</b>	
3.1 Le débit moyen en fonction de la largeur.....	40	
3.2 Le débit en fonction de la largeur.....	43	
3.3 La largeur en fonction du périmètre.....	46	
3.4 Le débit moyen en fonction du périmètre. ....	48	
3.5 Le débit moyen en fonction de l'aire de la section. ....	50	
<b>4. Régression linéaire multiple.....</b>	<b>51</b>	
4.1 Le débit moyen en fonction de la largeur, l'aire des sous bassins et de la sinuosité.....	51	
4.2 La largeur en fonction de l'aire de la section et de la profondeur moyenne.....	53	
4.3 Le débit moyen en fonction du rayon hydraulique et de l'aire de la section.....	56	
4.4 Le débit moyen en fonction du rayon hydraulique, de la pente et de la profondeur moyenne.....	59	Excluido: 58
4.5 Le débit moyen en fonction de la profondeur maximale, de la largeur et de l'aire de la section.....	62	Excluido: 61
4.6 Le débit moyen en fonction de la sinuosité , de la profondeur moyenne et du rayon hydraulique.....	65	Excluido: 64
4.7 Le débit moyen en fonction de la profondeur maximale, la largeur et le rayon hydraulique.....	68	Excluido: 67
<b>5. Conclusion.....</b>	<b>71</b>	Excluido: 70
<b>CONCLUSION.....</b>	<b>72</b>	Excluido: 71

<b>Lexique</b> .....	<b>73</b>	Excluido: 72
<b>Annexe</b> .....	<b>77</b>	Excluido: 76
1. Tableau d'ajustement des profondeurs .....	<b>78</b>	Excluido: 77
2. Décile des deux largeurs .....	<b>81</b>	Excluido: 80
3. Test de nullité du coefficient de Pearson .....	<b>81</b>	Excluido: 80
4. Test d'adéquation d'appartenance à une famille de loi .....	<b>82</b>	Excluido: 81
5. Test d'égalité des moyennes .....	<b>82</b>	Excluido: 81
6. Test d'égalité des médianes .....	<b>83</b>	Excluido: 81
7. Listing des résultats .....	<b>84</b>	Excluido: 82
8. Nuage des individus .....	<b>85</b>	Excluido: 83
9. Régression multiple .....	<b>89</b>	Excluido: 84
		Excluido: 87
<b>Bibliographie</b> .....	<b>91</b>	Excluido: 89

## **Partie I : INTRODUCTION**

Pour conclure ma formation en IUT, j'ai effectué un stage d'une durée de 10 semaines (du 31 mars au 9 juin) au Laboratoire des Mécanismes et Transferts en Géologie, Unité mixte (Université Paul Sabatier, CNRS, IRD) de recherche 5563, Ce stage se situe dans le cadre du projet HYBAM (Hydrologie et géochimie du Bassin amazonien).

Le bassin amazonien s'étend sur sept pays (la Bolivie, le Brésil, la Colombie, l'Equateur, la Guyane, le Pérou, et le Venezuela). Il est drainé par l'Amazone ainsi que par ses affluents. Il est découpé en plusieurs sous bassins.

Le but de cette étude est d'analyser le débit annuel moyen des principaux affluents de l'Amazone et de l'Amazone lui-même en fonction des caractéristiques géomorphologiques des bassins et des fleuves.

Dans ce but, le laboratoire a mis à ma disposition le logiciel Arc View qui m'a servi à obtenir certaines des r données nécessaires concernant le bassin amazonien (en particulier les données géomorphologiques, tandis que les autres données (en particulier les données hydrologiques) ont été extraites de la base de données Hybam. J'ai réalisé l'étude statistique à l'aide des logiciels Statgraphic's, SPAD, et Excel.

**LE LMTG**



# Le LMTG

Le LMTG, Laboratoire des Mécanismes et Transferts Géologiques, est situé à Toulouse, au 38 rue des 36 ponts. Le laboratoire dans son ensemble essaie de répondre à trois enjeux majeurs que sont les enjeux de **connaissance**, de **société** et de **développement** (partenariat fort avec des pays étrangers, rapprochement avec l'IRD par l'arrivée de deux Unités de cet Institut (dont Hybam) dans le cadre de cette UMR).

## Les enjeux

### 1. Enjeux de connaissance

Les chercheurs du LMTG ont pour objectif de comprendre le fonctionnement de la Terre en tant que système chimique, physique et biologique complexe. L'activité du laboratoire est essentiellement consacrée au rôle de l'eau dans ce système, plus exactement aux interactions entre l'eau et la terre solide, depuis le manteau jusqu'aux surfaces biologiques, mais aussi à la dynamique de l'ensemble. L'originalité du LMTG est de maîtriser l'ensemble des concepts scientifiques et une bonne partie des méthodes analytiques et expérimentales nécessaires pour contribuer de façon significative à ce difficile problème scientifique.

On trouve, au LMTG, 4 approches différentes du problème des interactions eau/terre solide et de la géodynamique :

**Une approche expérimentale**, qui essaie de comprendre, à l'échelle microscopique, les mécanismes physico-chimiques.

**Une approche par les bilans**, qui vise à quantifier les transferts de matière et d'énergie. Cette approche utilise divers outils : analyse des eaux, géomorphologie, géodynamique subactuelle. Elle produit des bilans de matière, des flux d'éléments. Cette approche est essentiellement présente au sein des thèmes Géodynamique et Eau-Sol-Erosion (thème auquel est rattaché le programme Hybam).

L'intégration de ces deux approches se fait au travers de la **modélisation**.

La troisième approche est une **approche historique**, qui tente, au travers de la reconstitution d'étapes de l'histoire de la Terre, de transférer les modèles issus de la Terre actuelle à l'évolution de la planète.

## 2. Enjeux de société

Par leurs compétences les chercheurs du LMTG participent à la réponse du monde scientifique à de grandes questions posées par l'ensemble de la société. Les quatre grands domaines où ils interviennent sont : les matériaux, l'environnement (pollution, déchets, effet de serre) les risques et la formation. Les compétences du laboratoire dans le domaine des fluides et des minéraux sont telles que les sollicitations sont spécialement nombreuses dans le domaine de l'environnement. La politique du laboratoire dans ce domaine est de privilégier les études à caractère fondamental, dont les retombées scientifiques sont clairement établies.

## 3. Enjeux de développement

Le laboratoire a depuis longtemps une tradition de partenariat fort avec des pays étrangers, avec une importante coopération dans les pays d'Afrique, d'Asie et d'Amérique latine.

La recherche pour le développement doit respecter, au Sud comme au Nord, les mêmes critères de qualité pour répondre aux enjeux de connaissance. En plus elle est destinée à faire émerger, dans les communautés du Sud, des équipes qui pourront acquérir une autonomie et remplir les besoins de leur pays en matière de recherche et de formation.



## Organisation

Le laboratoire forme une Unité Mixte de Recherche, l'UMR 5563, dirigée par Bernard Dupré (adjoints : JM. Montel et G. Hérail). Cette unité regroupe des chercheurs de différents organismes scientifiques (UPS, CNRS, IRD), et appartient à l'OMP (Observatoire Midi-Pyrénées). Ces thèmes sont les suivants : Physiochimie des minéraux, des interfaces et des solutions, Géochimie des interactions crustales, Géodynamique des chaînes et des surfaces continentales, et enfin Eaux, sol, érosion : milieu naturel et anthropisé.

Les effectifs totaux sont les suivants :

- 77 chercheurs et enseignants chercheurs
  - 40 doctorants
  - 7 post-doctorants
- ainsi que des techniciens, ITA, ATER, ...

Il existe d'autre part un certain nombre de services communs aux divers organismes (Instrumentation géochimique, salles blanche, atelier de Mécanique-Electrotechnique, Informatique, Instrumentation et Mesures physiques, Atelier roche).

Le laboratoire est organisé autour d'actions de recherche transverses :

- Andes-Bassin amazonien
- Modélisation des flux d'altération à l'échelle des bassins versants
- Planétologie
- Déchets
- Modélisation de la séquestration du CO<sub>2</sub> dans les formations géologiques
- Paléoclimat

La présente étude porte sur le thème Andes-Bassin amazonien dans lequel le laboratoire propose d'étudier les modalités de l'érosion dans les Andes en fonction de l'évolution tectonique et climatique, de quantifier les transferts de matière qui lui sont dus actuellement et dans un passé récent, de déterminer l'itinéraire suivi par les éléments issus de l'érosion de cette chaîne.

# Organigramme de l'UMR 5563

**UMR 5563**

Chercheurs CNRS : 19  
 Enseignants chercheurs : 26  
 Chercheurs : OMP : 3 - IRD : 27  
 ITA: 16 - ITARF :15 - ITA IRD : 10

**Directeur :** Bernard **DUPRÉ**  
**Directeurs adjoints:** Jean-Marc **MONTEL-** Gérard **HÉRAIL**  
 Assistante de Direction: \*C. BAZIN (TCE)

## Services généraux

<b>Secrétariat IRD :</b> *** B. BARBIN (TCN) <b>Aide Technique :</b> **V. FOURNIER (AST) 50 % **J. MIR (TCS)	<b>Gestion Financière (CNRS, Université, IRD) :</b> ***B. BARBIN (TCN) ** M. FILHASTRE (AGT) H. PEREA (CDD) * S. RAKOTONDRAMASY (TCN), Resp. ** F. RONCHINI (AST) ** M. TERRAL (AGT) 50 %
--	---

## Equipes

Physico-chimie des minéraux, des interfaces et des solutions	Eaux, sols, érosion: milieu naturel et anthropisé	Géochimie des interactions crustales	Géodynamique des chaînes et des surfaces continentales
Responsable: Jean- Marc Montel	Responsable: Anne Probst	Responsable: Gilles Berger	Responsable: Gérard Hérial
<b>Géochimie: transferts et mécanismes:</b> Responsable Jacques Schott 3 enseignants chercheurs 4 chercheurs CNRS 1 chercheur associé 6 doctorants 2 post-docs  <b>Minéralogie:</b> Responsable: Jannick Ingrin 3 enseignants chercheurs 5 chercheurs CNRS 1 chercheur associé 1 ingénieur valorisation 1 ATER 5 doctorants 2 post-docs  1 technicien rattaché à la thématique	2 enseignants chercheurs 4 chercheurs CNRS 1 physicien OMP 5 chercheurs IRD au laboratoire 6 chercheurs IRD à l'étranger 1 ITA rattaché à l'équipe 5 ITA IRD à l'étranger 1 ingénieur contractuel 2 ATER 6 doctorants 1 post-doc	1 enseignant chercheur 3 chercheurs 2 physiciens OMP 1 chercheur en délégation 1 ingénieur contractuel 5 doctorants	16 enseignants chercheurs 3 chercheurs CNRS 5 chercheurs IRD au laboratoire 9 chercheurs IRD à l'étranger 8 ITA rattachés à l'équipe 2 ITA IRD à l'étranger 2 professeurs invités 14 doctorants 2 post-docs

## Services communs

Chimie(AA, COD, HPLC, etc. ...)	Atelier de Mécanique-Electrotechnique	Instrumentation Mesures géochimiques	Atelier DAO, Cartographie
*J.CAPPAROS- ESCALIER (AI), Resp. *C. CAUSSERAND (T)  <b>Salle blanche:</b> **C. BOUCAYRAND (T) *S. GARDOLL (AI) *M. VALLADON (IR), Resp. 50 %	**J.C. HARRICHOURY (IE), Resp.	ICP-MS: *F. CANDAUDAP (AI) *R. FREYDIER (IR), Resp *M. VALLADON (IR) 50 % TIMS: *P. BRUNET (AI), Resp Microsonde: **Ph. de PARSEVAL (IE) Resp	*C.CAVARE HESTER (TCE) 20% *P. EICHENE (IE)

Atelier Roche	Instrumentation Mesures physiques	Informatique
***C. LAGANE (T) 50% **F. de PARSEVAL (ADT) **J.F. MENA (ADT)	Analyses thermiques-DRX: **M. THIBAUT (T) RX: *T. AIGOUY (TCE)	*H. FERRE (IE)

Participation aux activités d'enseignement
**E. DEBROAS (IE) 50% **V. FOURNIER (AST) 50% **N. GUERRERO (IE) 50% **M. TERRAL (AGT) 50%

IR = Ingénieur de Recherche - IE = Ingénieur d'Etudes - AI = Assistant Ingénieur - TCE/TCS/TCN = Technicien - ADT = Adjoint Technique

\* Personnel CNRS      \*\* Personnel Université      \*\*\*Personnel IRD

IRD : Institut de Recherche pour le Développement  
 CNRS : Centre National de Recherche Scientifique  
 Ce stage se déroule avec l'équipe « Eau, sol, érosion ».

# L'IRD

## 1. Carte d'identité

Etablissement à caractère scientifique et technologique, l'**Institut de Recherche pour le Développement** est placé sous la double tutelle des ministères de la Recherche et des Affaires Etrangères. Il a été créé en 1998, succédant à l'ORSTOM (Office de la recherche scientifique et technique d'Outre-mer).

Cet institut conduit des programmes scientifiques centrés sur les relations entre l'homme et son environnement, dont l'objectif est de contribuer au développement des pays du Sud. Il mène des programmes de recherches en Afrique, en Asie, dans l'Océan Indien, en Amérique latine et dans le Pacifique, les chercheurs intervenant dans une cinquantaine de pays.



### Objectifs et missions

- **Milieux et environnement** : La croûte terrestre, évolutions et risques naturels (*éruptions, séismes, inondations...*) ; Environnements continentaux, côtiers et marins ; Les climats, variabilité et impacts ; L'eau, ressources et gestion durable.
- **Ressources vivantes** : Biodiversité agricole et microbienne (*amélioration génétique, défense des plantes cultivées...*) ; Écologie aquatique et halieutique ; Écosystèmes et ressources terrestres.
- **Sociétés et santé** : Grandes endémies (*paludisme, sida, dengue...*) ; L'Homme dans son environnement ; Interactions sociétés et santé ; Dynamiques urbaines ; Politiques de développement et mondialisation.
- **Expertise et valorisation** : Réalisation d'expertises collégiales, aide à la création d'entreprise, propriété industrielle, conclusion de contrats de consultance, démarche qualité en recherche
- **Soutien et formation** : D'une part contribuer au renforcement des capacités de recherche des pays du Sud au moyen de bourses individuelles, de soutien à des équipes de recherche, et à des filières ou à des ateliers de formation ; d'autre part, favoriser leur meilleure insertion au niveau international.

## 3. Organisation

L'Institut dispose de 34 centres et représentations dans le monde, dont :

- 23 implantations dans des pays de la zone intertropicale
- 5 implantations en métropole
- 5 implantations dans les DOM-TOM

## 4. Moyens

Pour réaliser ses objectifs, l'IRD a plusieurs **partenaires**, qui sont :

- La communauté scientifique mondiale (Établissements de recherche et d'enseignement supérieur)
- Les instances européennes et organisations internationales
- Des industriels français et étrangers

Le **budget** moyen est de 177 millions d'euros ; les **effectifs** sont de 2187 agents, répartis comme il suit :

- 831 chercheurs dont 40% affectés à l'étranger
- 801 ingénieurs, techniciens, administratifs
- 553 agents de statuts divers (notamment personnel local)

Ils sont séparés en 97 unités de recherche, et 34 centres et représentations dans le monde.

## **Le projet HYBAM (Hydrologie et Géochimie du Bassin Amazonien)**

Le programme HYBAM (**Hydrologie et Géochimie du Bassin Amazonien**), a été lancé par l'IRD il y a une dizaine d'années, avec l'objectif d'étudier le comportement hydrologique du plus grand bassin versant existant (le bassin amazonien s'étend sur près de 6 millions de km<sup>2</sup>, et son débit moyen à l'exutoire est de 209000m<sup>3</sup>/s). Au fur et à mesure de l'avancement des recherches, de la mise en place de nouveaux partenariats, et ainsi de la naissance de nouvelles attentes, le projet a évolué pour ajouter au volet **hydrologie**, un volet **hydrodynamique** visant à expliquer le transport sédimentaire entre les Andes et l'Océan (modélisation); Un volet **géochimie** a également été intégré, afin d'étudier les phénomènes de transfert des éléments chimiques sur le bassin. Enfin, une équipe de **géologues** s'applique également à comprendre la dynamique actuelle du bassin. Cette équipe pluridisciplinaire sous la direction de JL Guyot est intégrée à l'UMR LMTG qui constitue l'unité de recherche 154 de l'IRD..

Depuis le début de l'année 2003, le projet HYBAM a le statut d'**Observatoire de Recherche en Environnement**(ORE), reconnu par le Ministère de la Recherche. L'intitulé actuel du projet est le suivant : « Contrôles géodynamique, hydrologique et bio géochimique de l'érosion/altération et des transferts de matière dans le bassin de l'Amazone; Impact des variations hydroclimatiques (ENSO) et des activités anthropiques (déforestations) ».

Les différents **pays** désormais associés dans ce programme sont le Brésil, la Bolivie, l'Equateur, la France, et le Pérou.

Les principaux **partenaires** actuels sont l'Agence de l'eau (ANA) et l'Université de Brasilia, institutions brésiliennes, ainsi que les Services Nationaux de Météorologie et d'Hydrologie de Bolivie et d'Equateur (respectivement SENAMHI et INAMHI).

## **Partie II : MATERIELS ET METHODES**

### **1. Objet de l'étude au sein du LMTG**

Cette étude m'a été proposée par le LMTG dans le but de mettre en évidence les relations entre le débit des principaux tributaires de l'Amazonie et certains paramètres mesurables par télédétection. En effet, il est dans certains cas très difficile d'estimer les débits à partir des hauteurs d'eau mesurées sur le terrain, les mesures effectuées n'étant pas forcément fiable et les relations entre débit et hauteur d'eau n'étant pas toujours univoques. Les mesures de terrain sont dans tous les cas très onéreuses.

Avec les mesures des satellites radar altimétriques Topex, Jason et Envisat on arrive désormais à connaître la hauteur des fleuves à l'endroit où passe le satellite. Il serait donc intéressant pour le laboratoire de trouver une manière de , d'estimer le débit à partir des paramètres géomorphologiques des bassins et des hauteurs d'eau afin de limiter ou de compléter les mesures sur le terrain.

Dans ce but, il m'a été confié le sujet suivant :

Déterminer les relations statistiques entre le débit annuel moyen ( $Q_m$ ) et certaines mesures géomorphologiques des principaux affluents de l'Amazonie.

Indices morphologiques du cours d'eau (IMce)

- largeur du cours d'eau  
→ à partir d'image JERS  
→ des jaugeages ADCP
- profondeur et section mouillée (à partir de jaugeages ADCP)
- pente du cours d'eau
- sinuosité (à partir de JERS)

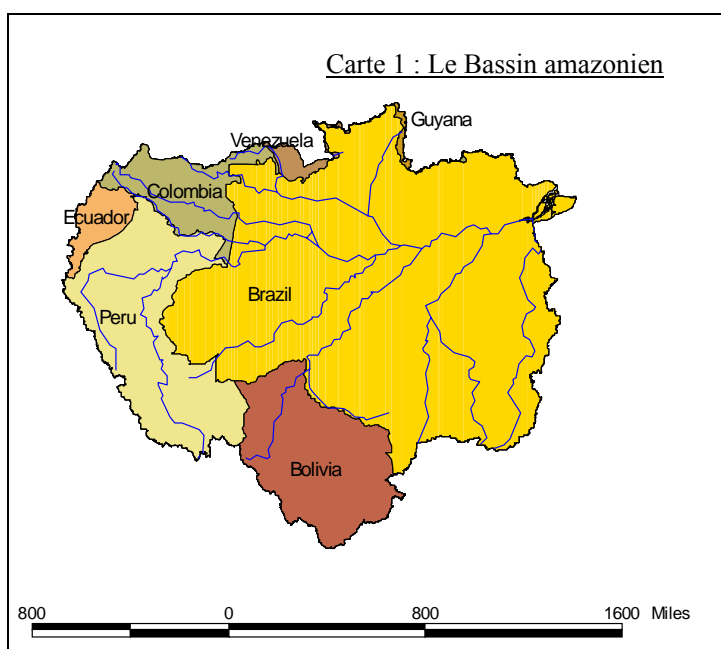
Indices morphologiques du bassin (IMb)

- Aire du bassin drainé

$$Q_m = f (I_{mce}, I_{mb})$$

## 2. Le bassin amazonien

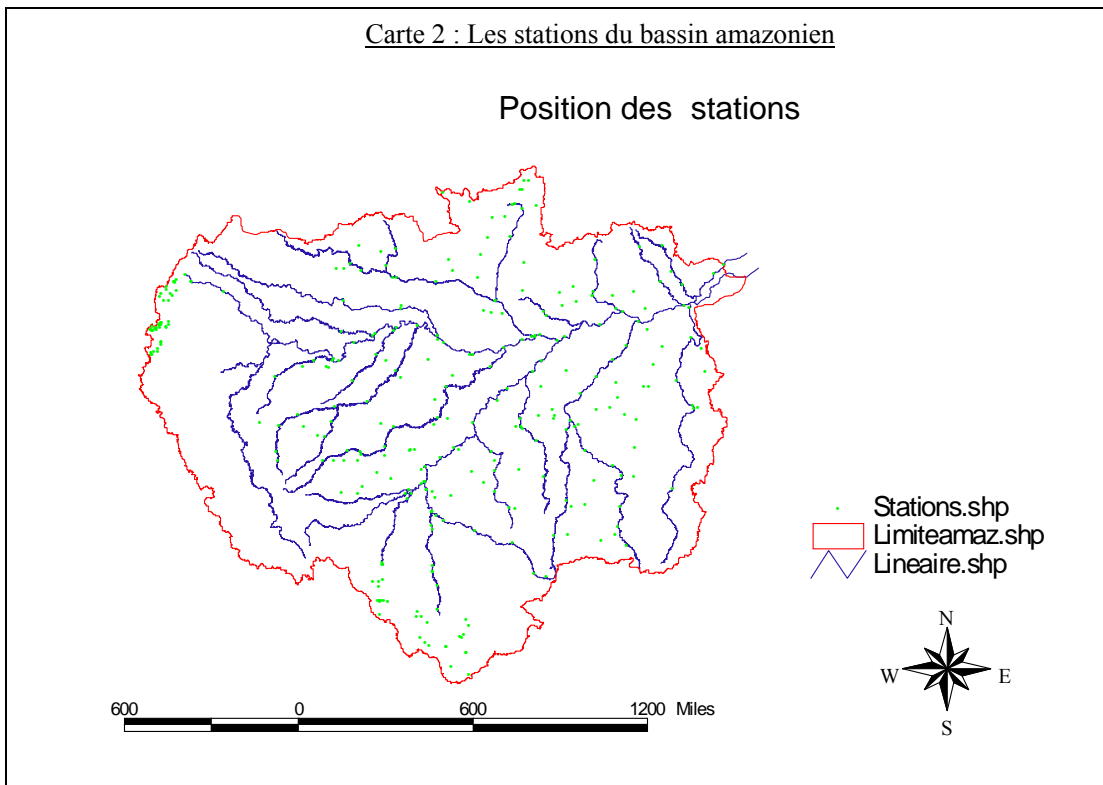
« Le bassin de l'Amazone s'étend sur sept pays (la Bolivie, le Brésil, la Colombie, l'Equateur, la Guyane, le Pérou, et le Venezuela). Le Brésil représente 68 % de la superficie de ce bassin, ensuite vient le Pérou (14%) et la Bolivie (10%), les quatre pays restant ne totalisant que 8%. » Extrait de la thèse « Hydrogéochimie des Fleuves de l'Amazonie bolivienne » de Jean Loup Guyot.



Situé entre le 5<sup>ème</sup> parallèle nord et le 20<sup>ème</sup> parallèle sud, le bassin versant de l'Amazone occupe une surface de quelques 6 millions de km<sup>2</sup>, soit 5% des terres émergées. Ce bassin versant est occupé par le plus long et le plus puissant fleuve de la planète, avec un réseau navigable de 25 000 km et un débit moyen annuel de 200 à 210 000 m<sup>3</sup>/sec. À son embouchure, ses apports représentent environ 15% des eaux douces déversées dans les océans par les fleuves de la planète. Outre ses caractéristiques hydrologiques particulières, il faut signaler la présence de la plus grande forêt tropicale avec une très grande richesse faunistique et floristique. Actuellement, ce bassin constitue un milieu menacé par les activités humaines.

### 3. Les stations

Il y a 269 stations sur le bassin amazonien. Une station est définie par un code station et par le nom de la station. Ci dessous on peut voir la position des stations sur le bassin.



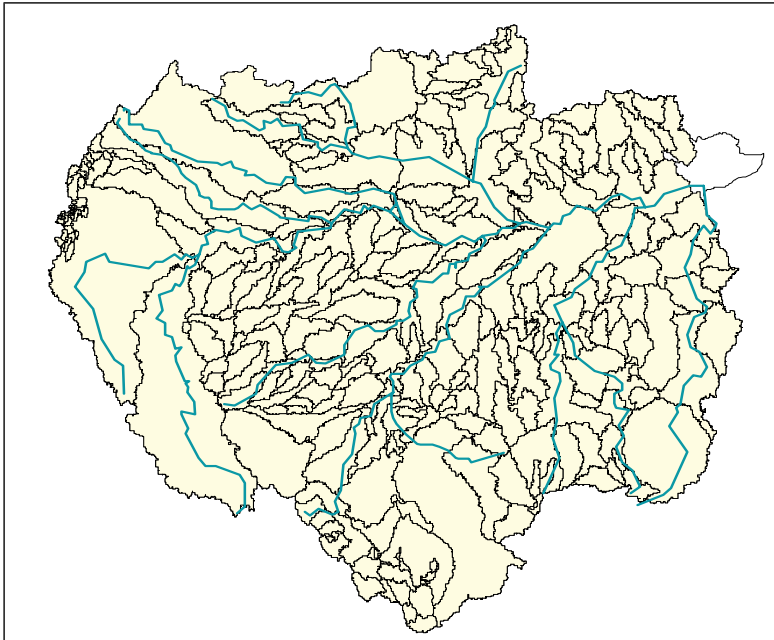
Toutes les stations sont situées le long d'un fleuve, ci dessus il n'apparaît que les fleuves principaux.



## 4. Les sous-bassins

Le bassin amazonien a été découpé en plusieurs sous bassins versants, à partir des stations de mesures. La délimitation des bassins s'est faite à partir du MNT (Modèle Numérique de Terrain) et de plusieurs réseaux hydrographiques. Elle prend en compte la direction d'écoulement pour chaque cellule du MNT en examinant l'altitude de ses huit voisins.

Carte 3 : Les sous bassins du Bassin amazonien



Ainsi, 392 sous bassins ont été délimités aux exutoires représentés par les stations de mesures hydrologiques.

## 5. Définitions des variables et extraction des données

Les paramètres géomorphologiques du cours d'eau sont :

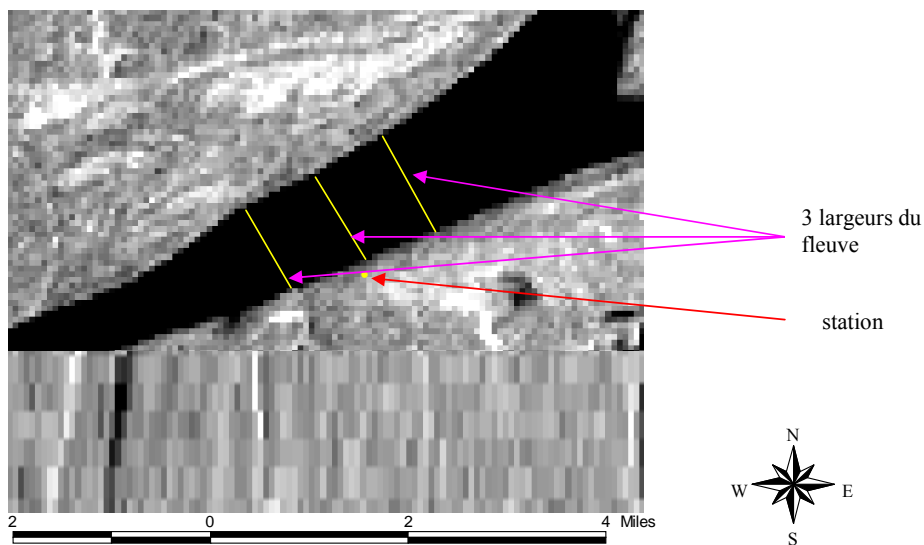
**La largeur du fleuve :** elle est mesurée en mètre et correspond à la distance en ligne droite entre les deux rives.

Elle a été estimée de deux façons différentes.

Une première méthode consiste à utiliser les images issues du satellite JERS et le logiciel Arc View. Plusieurs mesures ont été faites sur une même station. J'ai donc calculé la moyenne des largeurs pour chaque station.

Image 1 : image JERS

### La largeur du fleuve



Sur ces images, un pixel représente 100 mètres. De ce fait, la totalité des fleuves ayant une largeur inférieure à cette valeur n'ont pas pu être mesurés. 25 largeurs ont été mesurées par cette méthode.

Une autre manière de connaître la largeur du cours d'eau est le jaugeage ADCP (Acoustic Doppler Current Profiler). 62 largeurs ont été mesurées par cette méthode et les valeurs ont été extraites de la base de données Hybam.

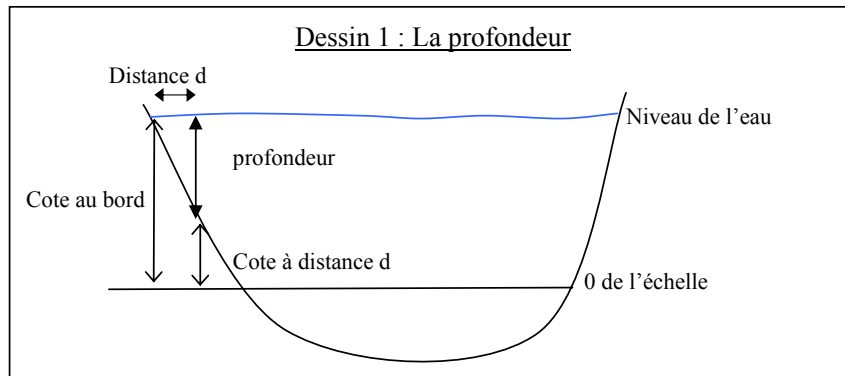
**La cte :** elle correspond à la hauteur d'eau (en mètre) entre le 0 de l'échelle et le niveau de l'eau. Le zéro de l'échelle est une valeur de référence pour la mesure de la cote. La cote est

mesurée à différentes distances de la rive, on a donc plusieurs cotes pour chaque station. Ces mesures ont été faites à partir de jaugeage ADCP.

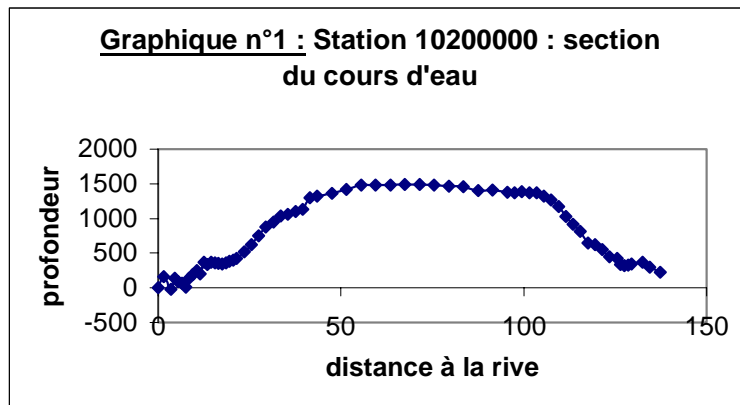
**La profondeur :** A partir de la cote, j'ai calculé la profondeur.

Profondeur<sub>d</sub> = cote au bord de la rive – cote à distance<sub>d</sub>

Certaines profondeurs sont négatives, ceci est lié au fait qu'il y a des vagues.



En traçant la profondeur en fonction de la distance au bord de la rive pour chaque station. On obtient la section du cours d'eau.



Pour l'analyse statistique, plusieurs profondeurs ont été calculées :

- **la profondeur maximale**
- **la profondeur médiane**
- **profondeur = aire/largeur**  
L'aire est la surface comprise entre le niveau de l'eau et la profondeur. La largeur est la largeur mesurée à partir de jaugeage ADCP.
- **La profondeur maximale à partir d'un ajustement de la section par un polynôme du second degré.**  
Il est nécessaire de souligner que pour certaines stations, la section n'est pas régulière. En effet, on peut voir des pics de profondeur. La profondeur maximale n'est pas représentative de la section. On a donc ajouté aux sections une courbe d'ajustement

qui lisserait ces irrégularités. L'ajustement par un polynôme du second degré de la section donne des coefficients de détermination significatif (voir annexe). Ainsi, l'ajustement est bon. Par le calcul de la dérivée de ces polynômes, on peut déterminer le maximum qui correspondra à la profondeur maximale ajustée.

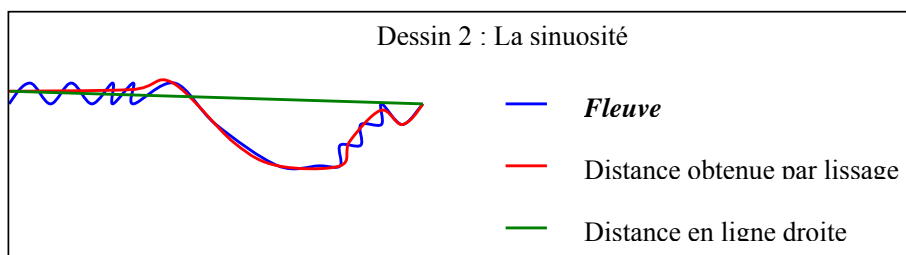
- **la profondeur moyenne pondérée par la distance à la rive**

**Le rayon hydraulique :** C'est le quotient entre la surface mouillée et le périmètre mouillé.

La surface mouillée est l'aire comprise entre le niveau de l'eau et la profondeur. C'est la section occupée par l'écoulement.

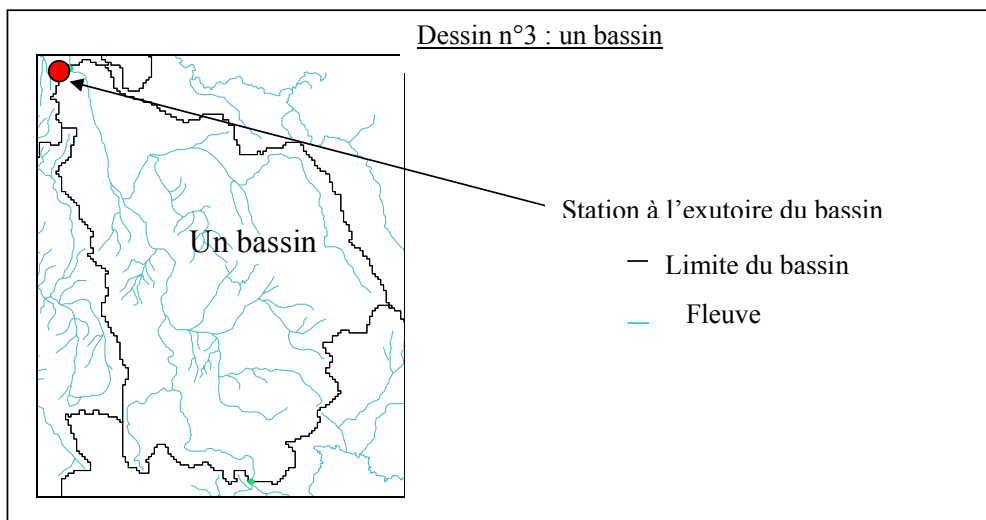
Périmètre mouillé = périmètre de la surface mouillée – largeur

**La sinuosité :** c'est le quotient de la distance obtenu par lissage du fleuve et de la distance en ligne droite.



Elle a été mesurée à partir du logiciel Arc View.

La sinuosité est mesurée par bassin. On considère que la sinuosité du bassin correspond à la station située à l'exutoire du bassin.



**Le débit :** C'est le volume d'eau écoulé en une année. Il se mesure en mètre cubes par seconde. A chaque station correspondent plusieurs mesures annuelles de débits entre 1987 et 2001. Le débit moyen annuel a donc été calculé. Ainsi, plus le nombre d'années est grand plus la valeur du débit a de signification. Ces données ont été extraites de la base de données hydrologiques du Brésil par Gérard Cochonneau sous Hydraccess.

**La pente :** à partir du MNT( voir lexique) et du logiciel Arc view, on relève les altitudes des stations. La pente entre deux stations est la différence des altitudes sur la longueur du cours d'eau entre les stations.

$$\text{Pente} = \frac{|\text{altitude1} - \text{altitude2}|}{\text{longueur}}$$

La pente fera référence à la station située à l'exutoire du bassin.

Les paramètres géomorphologiques du bassin sont :

**L'aire des sous-bassins :** C'est l'aire en km<sup>2</sup> des sous bassins.

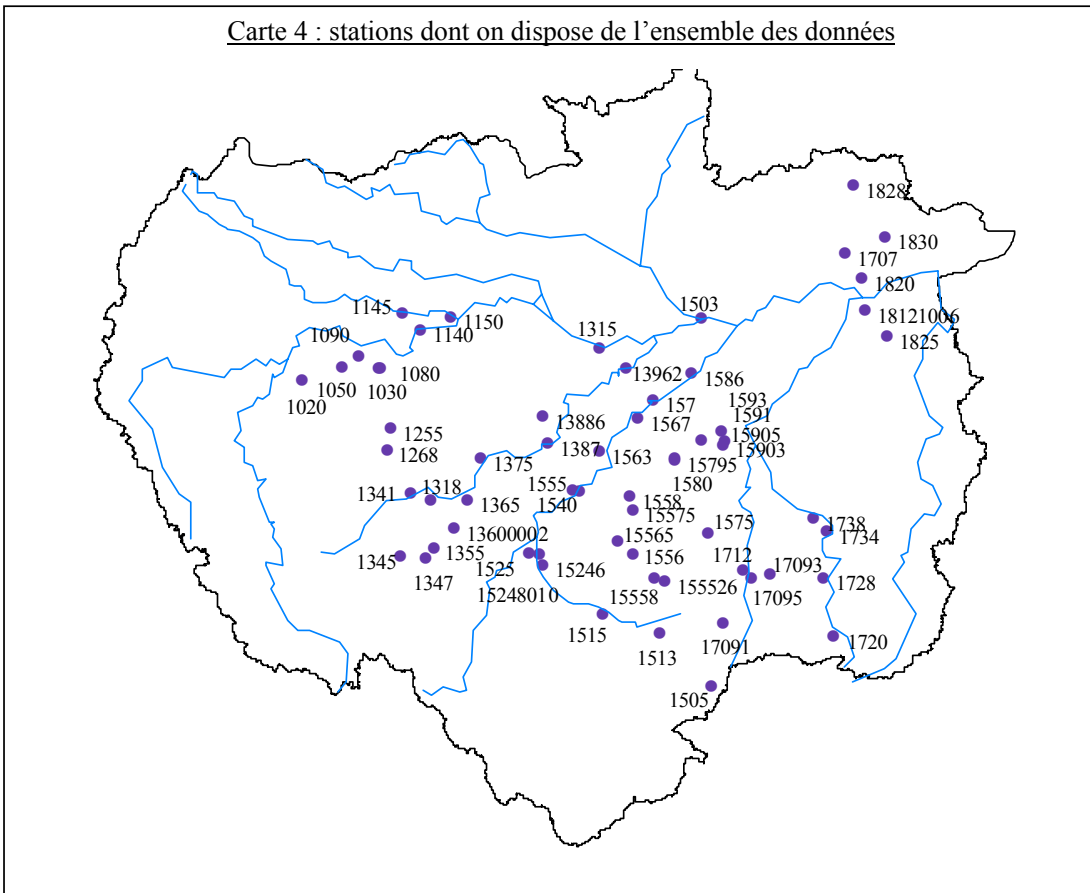
De la même manière que pour la sinuosité, l'aire d'un bassin correspondra à la station située à l'exutoire du bassin.

Après avoir manipulé l'ensemble de ces données, deux tableaux ont été mis en place.

Le premier regroupe les variables qui ne fluctuent pas en fonction du temps. C'est à dire l'aire des sous bassins, la largeur du fleuve, la sinuosité, la pente et le débit moyen. Les unités statistiques sont les 60 codes station.

Sur la carte ci-dessous, on peut voir les stations où l'on dispose de l'ensemble des données. Pour plus de lisibilité, les codes des stations ont été réduits à quatre chiffres. Par exemple pour la station 1145 le code est en fait 11450000.

Carte 4 : stations dont on dispose de l'ensemble des données



Un deuxième tableau est constitué des différentes profondeurs, la date de prélèvement, le rayon hydraulique, la surface mouillée(aire), le périmètre mouillé, l'aire des sous bassins, la largeur du fleuve, la sinuosité, la pente et le débit moyen. Les unités statistiques sont les 79 numéros d'identification de mesure de la profondeur car pour une même station il y a plusieurs mesures de section à des dates différentes.

## 6. Logiciels utilisés

Le logiciel Arc View :



Arc View GIS (Geographic Information System) permet de manipuler des vecteurs (tels le cours des rivières ou les limites des sous-bassins, des grilles régulières (comme le Modèle numérique de terrain) et des points (comme les stations) dans un environnement géoréférencé

Avec ce logiciel, j'ai pu extraire les largeurs des cours d'eau, puis j'ai mesuré la sinuosité, enfin j'ai calculé l'aire des sous bassins versants et les pentes.

Le logiciel Excel : l'étude graphique a été produite sous ce logiciel

Des macros ont été programmées sous Visual Basic pour la manipulation des données.

Le logiciel Statgraphics : grâce à ce logiciel, j'ai fait des tests statistiques (égalité des Moyennes, égalité des médianes)

Le logiciel SPAD : utilisé pour l'analyse en composante principale

Le rapport a été rédigé sous Word. L'exposé a été fait sous Power Point.

## 7. Méthodes statistiques

Les méthodes statistiques mises en application dans cette étude sont :

- statistiques descriptives
- analyse en composante principale
- régression multiple
- régression linéaire simple
- test d'hypothèse

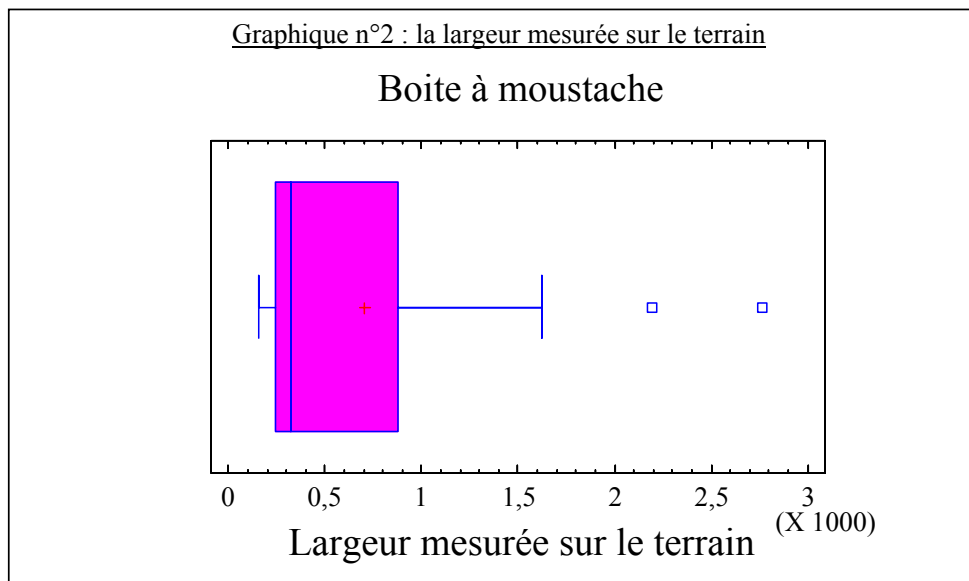
## **Partie III : ANALYSE DES RESULTATS**

### **I. QUALITE D'ESTIMATION DES LARGEURS A PARTIR D'IMAGE JERS**



Sur les 62 mesures prises sur le terrain, seulement 25 largeurs ont été relevées par le logiciel. On veut savoir la qualité d'estimation des largeurs sur des images de télédétection.

## 1. La largeur mesurée sur le terrain



```
Summary Statistics for largeur_terrain
Count = 25
Average = 700,296
Variance = 463466,0
Standard deviation = 680,784
Minimum = 153,8
Maximum = 2763,0
```

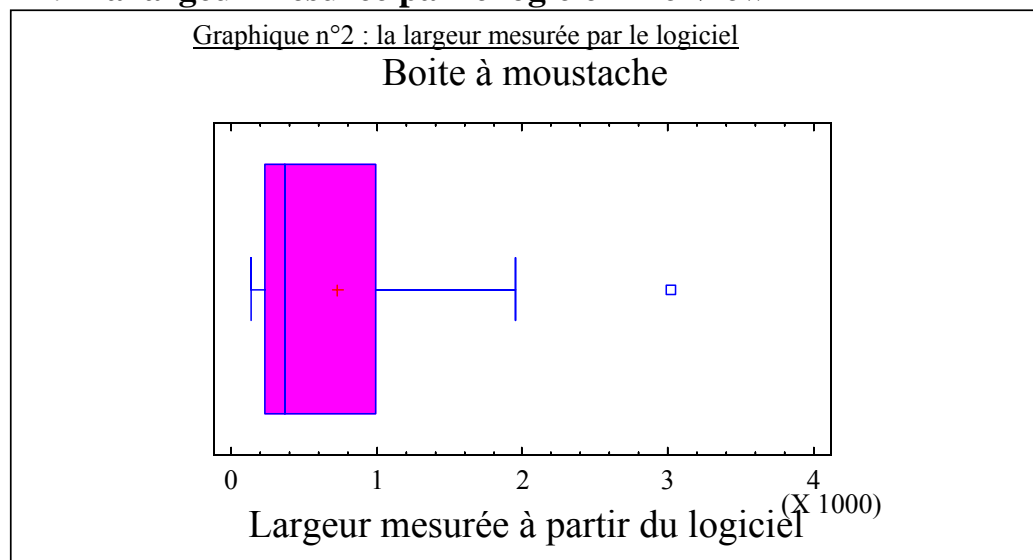
Ci dessus, on peut voir la répartition des 25 largeurs mesurées sur le terrain correspondant aux stations où l'on a estimé ces mêmes largeurs avec le logiciel (voir annexe).

La moitié des observations ont une largeur comprise entre 239.6m et 877.76m.

On peut noter deux valeurs extrêmes, c'est à dire que deux stations ont des largeurs supérieures à 1835 m.

La largeur médiane vaut 320.4m.

## 2. La largeur mesurée par le logiciel Arc View



### Summary Statistics for largeur\_jers

```
Count = 25  
Average = 728,284  
Variance = 515664,0  
Standard deviation = 718,097  
Minimum = 131,612  
Maximum = 3015,39
```

La moitié des observations ont une largeur comprise entre 232.238m et 994.756m.

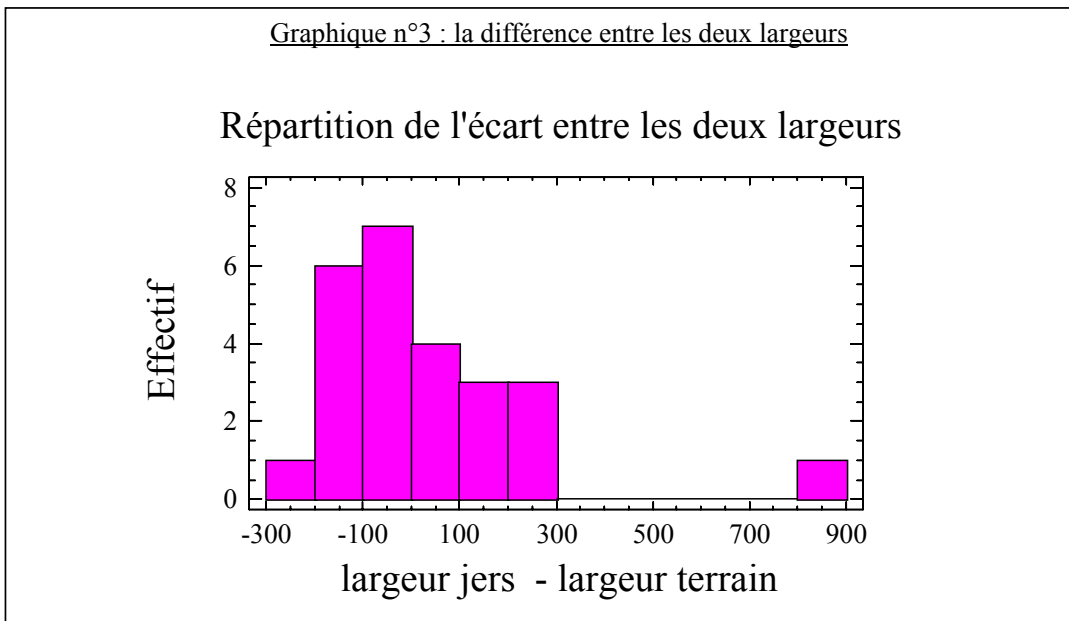
On peut noter une valeur extrême, c'est à dire qu'une station a une largeur supérieure à 2138.533 m.

La largeur médiane vaut 367.874m.

En comparant ces deux boîtes à moustache, on remarque que le premier quartile est identique, la valeur médiane est un peu plus élevée pour la largeur mesurée par le logiciel. Par contre, il existe un écart pour le troisième quartile qui s'élève à environ 120m (voir annexe).

### 3. Comparaison des deux largeurs

En calculant la différence entre les largeurs terrain et les largeurs du logiciel, on peut voir comment se répartissent les écarts de mesure.



Summary Statistics for largeur\_jers\_largeur\_terrain

```
Count = 25
Average = 27,9884
Variance = 48373,5
Standard deviation = 219,94
Minimum = -234,634
Maximum = 840,956
```

La répartition des écarts n'est pas symétrique. Dans 56% des cas, le logiciel sous estime les largeurs. On remarque une différence supérieure à 800 m.

80 % des écarts sont inférieurs à 2 pixels c'est à dire plus ou moins 200 m.

## 4. Tests

### 4.1 Test paramétrique

On veut savoir si les largeurs mesurées par le logiciel sont significativement différentes de celles relevées le long du fleuve.

Le test de comparaison des moyennes va nous montrer s'il y a égalité de distribution entre les deux largeurs. Il s'agit d'un appariement dans la mesure à chaque point de prélèvement correspondre une largeur estimée par Arc View et une mesurée sur le terrain.

Avant tout, testons la nullité du coefficient de Pearson.

Les hypothèses à tester sont :

$H_0 : \rho = 0$  » contre  $H_1 : \rho \neq 0$  »

L'estimation  $\rho^*$  de  $\rho$  est donnée par  $\rho^* = 0.9519$  et la probabilité critique est nulle (voir annexe).

**Ainsi, on rejette l'hypothèse nulle au seuil  $\alpha=0,05$ . Les deux largeurs sont donc dépendantes.**

De plus  $\rho$  est proche de 1, l'appariement est donc efficace.

Pour tenir compte de cette dépendance, on considère  $d_i = \text{largeur JERS}_i - \text{largeur terrain}_i$  pour  $i=1..25$

Le nombre d'observation (=25) est inférieur à 30 : il faut tester la normalité de distribution de la différence.

Les hypothèses à tester sont :

$H_0 : \text{« les observations suivent une loi normale } (\mu_d, \sigma_d) \text{ »}$  contre  $H_1 : \text{« } H_0 \text{ est fausse »}$

**La probabilité critique vaut 0.604383 (voir annexe), on accepte donc  $H_0$  au seuil  $\alpha=0,10$ .**

**La différence de largeur suit une loi normale de paramètres :**

$$\mu_d = 27.9884$$

$$\sigma_d = 219.94$$

Toutes les conditions sont valides pour effectuer le test d'égalité des moyennes.

Les hypothèses à tester sont :

$H_0 : \text{« la moyenne de la différence est nulle ; } \mu_d = 0 \text{ »}$  contre  $H_1 : \text{« } \mu_d \neq 0 \text{ »}$

Le test est bilatéral. La probabilité critique vaut  $0,530619 > 0.05$  (voir annexe).

**On accepte l'hypothèse nulle au seuil  $\alpha=0,05$  : la moyenne de la différence vaut 0.**

## 4.2 Test non paramétrique

### Test de signe

Les hypothèses à tester sont :

Ho : «  $md=0$  » contre H1 : «  $md \neq 0$  » avec m : médiane de l'échantillon

La probabilité critique vaut  $0.689153 > 0.05$  voir annexe

**Ainsi au seuil  $\alpha=0.05$  on accepte Ho : la médiane de la différence est nulle.**

### Test de rangs de Wilcoxon

Les hypothèses à tester sont :

Ho : «  $md=0$  » contre H1 : «  $md \neq 0$  »

La probabilité critique vaut  $0.999994 > 0.05$  voir annexe

**Ainsi au seuil  $\alpha=0.05$  on accepte Ho : la médiane de la différence est nulle.**

## 5. Conclusion

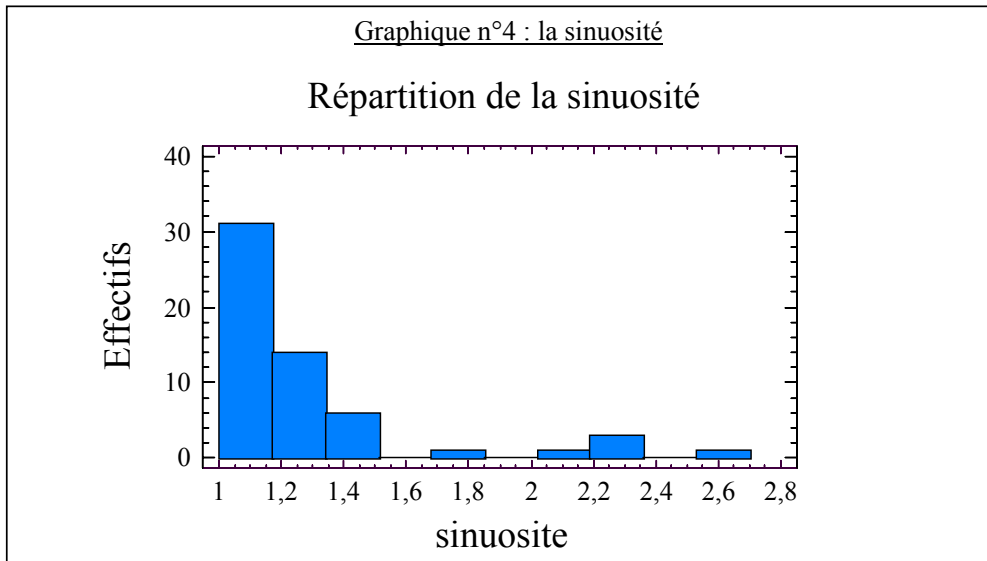
Tous les tests conduisent à accepter Ho : les largeurs mesurées par le logiciel ne sont pas significativement différentes de celles relevées le long du fleuve au seuil de 5%.

**II. RELATIONS STATISTIQUES ENTRE LE DEBIT  
ANNUEL MOYEN ET LES PARAMETRES  
GEOMORPHOLOGIQUES DU COURS D'EAU ET  
DU BASSIN**

# 1. Etude des variables

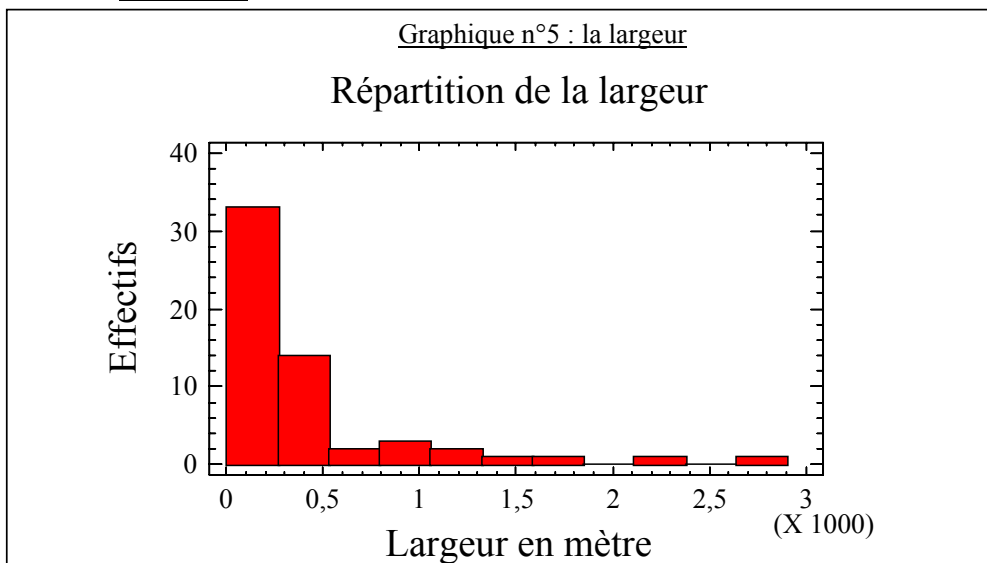
## 1.1 Statistique descriptive

### 1.1.1 La sinuosité



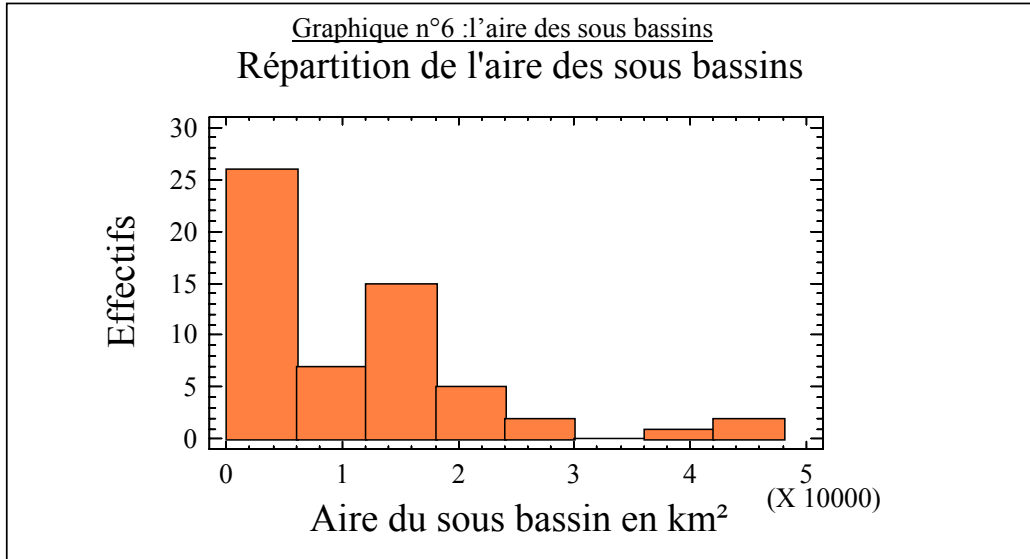
On peut voir que 31 fleuves ont une sinuosité inférieure à 1.2. C'est la classe modale. 4 fleuves prennent des valeurs de sinuosité supérieure à 2. C'est à dire que la longueur réelle du fleuve est deux fois plus importante que la longueur en ligne droite. Un fleuve est très sinueux, sa sinuosité à une valeur supérieure à 2.5. La sinuosité moyenne vaut 1.28.

### 1.1.2 La largeur



La majorité des largeurs sont inférieures à 250 m. Un peu moins de 80% des largeurs sont inférieures à 500 m. Une seule largeur est supérieure à 2500 m.

### 1.1.3 L'aire des sous bassins

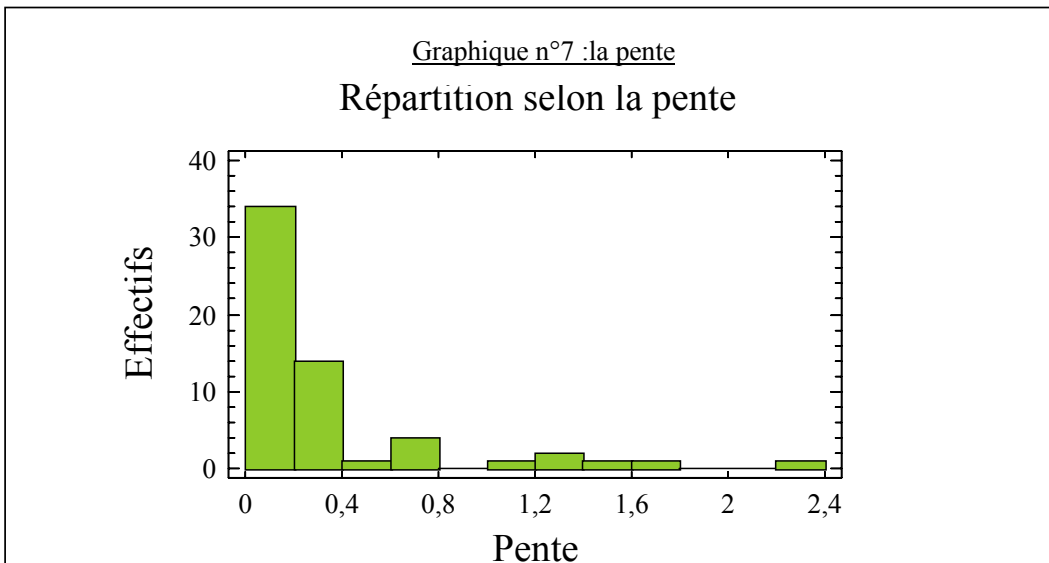


Comme on peut le noter, l'aire des sous bassins est très étendue.

Elle va de 532.61 km<sup>2</sup> à 44355,3 km<sup>2</sup>.

44% des superficies sont inférieures à 6000 km<sup>2</sup>. 3 bassins ont une aire supérieure à 36000 km<sup>2</sup>.

### 1.1.4 La pente



Plus de 80 % des pentes sont inférieures à 0,4 dont 56 % inférieures à 0,2.

Une station possède une pente de 2,21.



## 1.2 Etude des corrélations

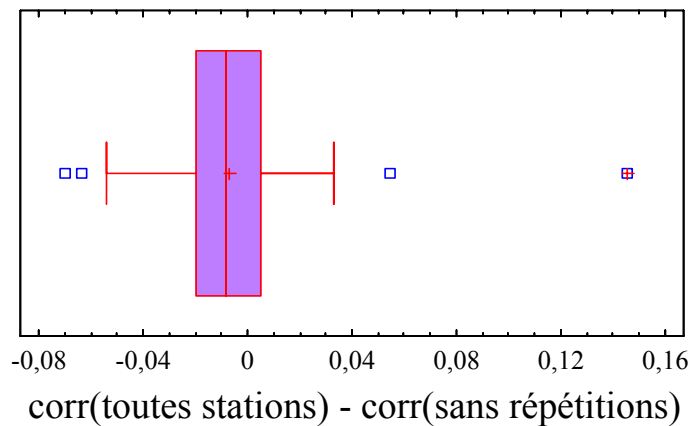
Avant de faire une analyse en composante principale, il faut souligner le fait que une même station a plusieurs valeurs qui concerne la section (profondeurs, aire, périmètre). Ces différentes valeurs auront peut être une influence sur les corrélations entre les variables. En calculant les corrélations entre les variables en ne prenant qu'une seule valeur par station et en les comparant avec les corrélations pour l'ensemble des stations, on peut voir si la répétition de donnée est influente.

Tableau n°1	sinuosité	prof estime max	profondeur max	profondeur médiane	profondeur moyenne	profondeur= aire/largeur	rayon hydraulique	Périmètre (m)	aire (m <sup>2</sup> )	largeur	aire bassin	débit moyen
sinuosité												
P estime max	-0,024											
P max	-0,029	-0,002										
P médiane	0,033	-0,008	-0,010									
P moyenne	0,008	-0,009	-0,007	0,002								
P=aire/largeur	0,005	-0,012	-0,012	0,008	-0,008							
rayon hydraulique	0,006	-0,009	-0,007	0,002	0,000	-0,008						
périmètre(m)	-0,021	-0,063	-0,054	-0,007	-0,017	-0,005	-0,017					
aire (m <sup>2</sup> )	-0,020	-0,032	-0,026	0,000	-0,007	0,007	-0,007	0,003				
largeur	-0,035	-0,028	-0,017	0,008	0,018	0,017	0,018	-0,002	0,016			
aire bassin	0,146	0,027	0,007	0,027	0,012	-0,028	0,011	-0,070	-0,049	-0,044		
débit moyen	-0,034	-0,016	-0,008	-0,007	0,003	0,016	0,003	-0,006	0,002	0,006	-0,029	
pente	-0,010	-0,019	-0,014	-0,030	-0,023	-0,012	-0,023	-0,008	-0,017	-0,022	-0,047	-0,017

On peut voir que l'écart entre les corrélations est très faible.

Graphique n°8 :

Boite à moustache de la différence des corrélations



Au vu de ces différences, l'analyse en composante principale se fera sur l'ensemble des données.

## 2. Analyse en Composante Principale

Une analyse en composante principale est faite sur 79 stations et l'ensemble des variables.

### 2.1 Statistique élémentaire

LIBELLE	MOYENNE	ECART-TYPE
sinuosité	1.29	0.40
profondeur estimée max	1436.77	761.02
profondeur maximale	1467.55	840.20
profondeur médiane	939.49	451.61
profondeur moyenne	858.29	459.19
profondeur=aire/largeur	277.32	201.53
rayon hydraulique	8.53	4.57
périmètre(m)	11740.14	17462.25
aire (m <sup>2</sup> )	150127.05	324406.13
largeur	364.24	445.49
aire bassin	12750.51	12115.61
débit moyen	4714.87	11878.48
pente	0.34	0.44

On constate que les variances des données sont très hétérogènes. C'est pourquoi une ACP normée est produite : il ne faut pas qu'une variable ait un poids plus important qu'une autre. Les stations 1315000, 1570000 et 15030000 sont placées en individus supplémentaires car elles présentent des valeurs très éloignées de la moyenne, ce sont des stations atypiques.

Tableau n°3

stations	15030000	13150000	15700000
sinuosité	1,30	1,39	1,14
profondeur maximale estimée	6242,19	6895,43	4573,97
profondeur maximale	6744,00	7547,00	4257
profondeur médiane	3552,00	4207,00	2142,5
profondeur moyenne	3406,08	4028,19	2805,47
profondeur=aire/largeur	1591,94	1122,02	2078,08
rayon hydraulique	34,03	40,00	27,98
périmètre(m)	129258,77	30151,58	72796,50
aire (m <sup>2</sup> )	4398543,04	1206172,45	2036515,75
largeur	2763,00	1075,00	980
aire bassin	5817,67	3268,99	12005,69
débit moyen	122495,24	80749,00	24565,2
pente	0,0560	0,0273	0,0598

## 2.2 Choix des axes

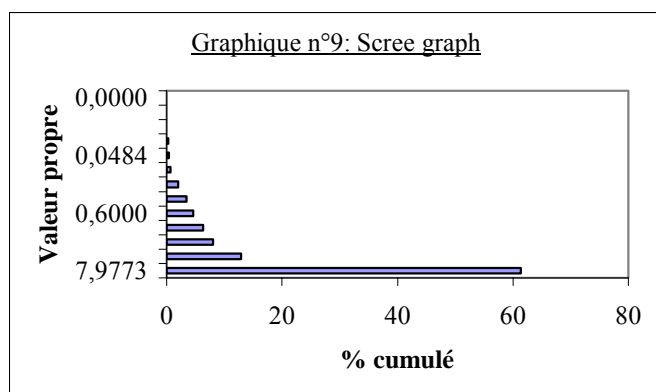


Tableau n°4

valeur propre	pourcentage	Pourcentage cumulé
7,9773	61,36	61,36
1,6782	12,91	74,27
1,0424	8,02	82,29
0,8156	6,27	88,57
0,6000	4,62	93,18
0,4445	3,42	96,60
0,2589	1,99	98,59
0,0821	0,63	99,22
0,0484	0,37	99,60
0,0327	0,25	99,85
0,0155	0,12	99,97
0,0043	0,03	100,00
0,0000	0,00	100,00

L'espace factorielle de dimension 3 restitue près de 82.29% de l'inertie initiale. Dans une ACP normée, on ne s'intéresse qu'aux valeurs propres supérieures à 1. L'analyse ne se fera que sur les trois premiers axes.

Tableau n°5	coordonnées				contributions			
	Axe 1	Axe 2	Axe 3	Axe 4	Axe 1	Axe 2	Axe 3	Axe 4
<b>sinuosité</b>	0,10	-0,69	0,40	0,27	0,13	28,36	15,68	8,76
<b>profondeur estimée max</b>	0,89	-0,22	-0,22	0,06	9,90	2,87	4,62	0,52
<b>profondeur max</b>	0,92	-0,13	-0,21	0,05	10,52	0,96	4,25	0,34
<b>profondeur médiane</b>	0,88	-0,19	-0,09	0,08	9,75	2,23	0,83	0,81
<b>profondeur moyenne</b>	0,94	-0,24	-0,16	0,10	11,02	3,49	2,60	1,15
<b>profondeur=aire/largeur</b>	0,84	-0,20	-0,05	0,12	8,95	2,44	0,23	1,78
<b>rayon hydraulique</b>	0,94	-0,23	-0,16	0,09	11,09	3,21	2,59	1,09
<b>périmètre(m)</b>	0,84	0,37	0,31	-0,07	8,83	8,19	9,51	0,64
<b>aire (m²)</b>	0,86	0,34	0,31	-0,01	9,21	6,81	9,31	0,02
<b>largeur</b>	0,87	0,37	0,21	-0,09	9,40	8,12	4,36	0,90
<b>aire bassin</b>	-0,08	-0,59	0,60	-0,28	0,07	20,78	34,96	9,28
<b>débit moyen</b>	0,83	0,38	0,25	-0,02	8,72	8,61	6,01	0,07
<b>pente</b>	-0,44	0,26	0,23	0,78	2,41	3,91	5,04	74,66

Toutes les profondeurs ont une forte coordonnée positive sur le premier axe. Le périmètre, l'aire de la section ainsi que le débit moyen ont également de fortes coordonnées positives. Ces variables contribuent environ toutes pour 10% à l'existence de l'axe 1. La somme de leurs contributions restitue près de 97.4% de l'information qu'apporte le premier axe. L'axe 1 traitera des paramètres « hydrologiques » du cours d'eau. Les stations situées à droite de l'axe 1 auront des paramètres « hydrologiques » très élevés ainsi qu'un fort débit.

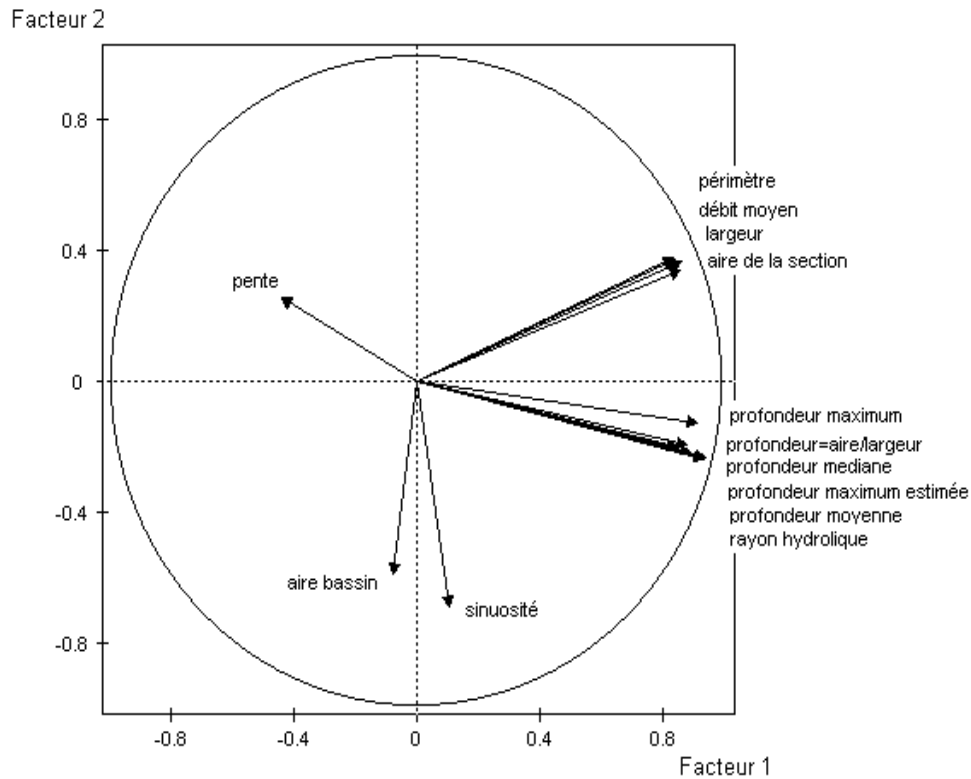
Seule la sinuosité et l'aire des bassins drainés ont une forte coordonnée négative. A elles deux, elles restituent 66.6% l'information. Le débit, la largeur et le périmètre ont des coordonnées positives. Le cumul des contributions de l'ensemble de ces variables vaut 73.3%. L'axe 2 traitera essentiellement des paramètres géomorphologique des bassins. Les stations ayant de fortes coordonnées négatives sur l'axe 2, auront une forte sinuosité et seront situées à l'exutoire de bassins de grande superficie.

Le troisième axe ne nous apporte aucune information supplémentaire.

La pente contribue à 74.66% de l'existence du quatrième axe. Cette variable possède une forte coordonnée positive. Ainsi à droite de cet axe, nous verrons apparaître les stations situées sur des cours d'eau ayant de forte pente.

### 2.3 Etude des corrélations

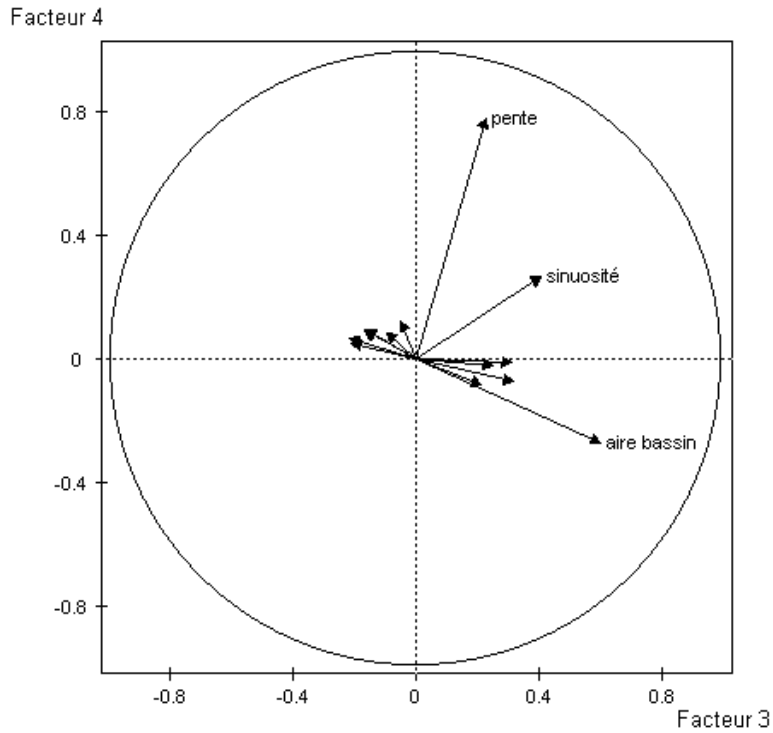
Graphique n°10 : cercle des corrélations



On peut voir d'après la longueur des vecteurs des variables que la sinuosité, la pente ainsi que l'aire du bassin sont les moins bien représentées sur le premier plan factoriel. Les autres vecteurs ont des longueurs proches de 1. On peut voir une opposition entre les profondeurs, le rayon hydraulique et le débit moyen, la largeur, le périmètre et l'aire de la section sur l'axe 2.

On remarque une forte corrélation positive entre le débit moyen et la largeur. En effet l'angle formé entre ces deux vecteurs est presque nul. Le débit moyen est également corrélé avec l'aire de la section ainsi qu'avec le périmètre. Les profondeurs sont fortement corrélées positivement entre elles. Cependant, on peut noter que la profondeur maximale se distingue des autres. Enfin on peut voir également une corrélation entre l'aire du bassin et la sinuosité.

Graphique n°11 :cercle des corrélations



Le cercle des corrélations ci dessus nous permet de visualiser les coordonnées des variables sur les axes 3 et 4. La variable pente est la mieux représentée. Ensuite viennent la sinuosité et l'aire des bassins. Les autres vecteurs ont des longueurs très faibles. Ils ne caractérisent pas ce plan.

#### **2.4 Le nuage d'individus**

En regardant le nuage des stations sur le premier plan factoriel (voir annexe):

Les stations 11400000, 11500000, 15860000, 15630000, 13962000, 1567000 ont des largeurs, profondeurs, section, et débits très élevés. Elles sont situées sur de grands fleuves profonds qui sont : Rio Surumu, Rio Solimoes, Rio Papagaio, Rio Trombetas, Rio Paru De Este, Rio Guapore

On peut voir que les stations 18280000 et 15670000 ont des fortes sinuosités mesurées sur de grands bassins : Oriximina et Boa Esperanca.

Ces stations sont placées sur les fleuves Rio Trombetas et Rio Fresco. La station 11500000 a un fleuve peu sinueux situé dans un bassin de petite superficie (Vila Surumu).

Les autres stations sont positionnées au centre des deux axes, elles ont des valeurs moyennes pour l'ensemble des variables.

Les individus supplémentaires ont de fortes coordonnées positives sur l'axe 1. Ce sont des stations situées sur de grands fleuves qui sont Rio Jauaperi, Rio Abacaxis et Rio Cotingo.

Sur le nuage des stations avec les axes 3 et 4 (voir annexe), on peut visualiser les stations positionnées sur des fleuves à fortes pentes.

Les stations 15795000, 13470000 et 15670000 ont de fortes coordonnées positives sur la quatrième axe, elle ont donc des valeurs de pente très élevées qui correspondent aux fleuves : Rio Trombetas, Rio Demeni, Rio Comemoração.

Les trois stations supplémentaires ont aussi de fortes pentes. Elles sont situées en haut de l'axe 4.

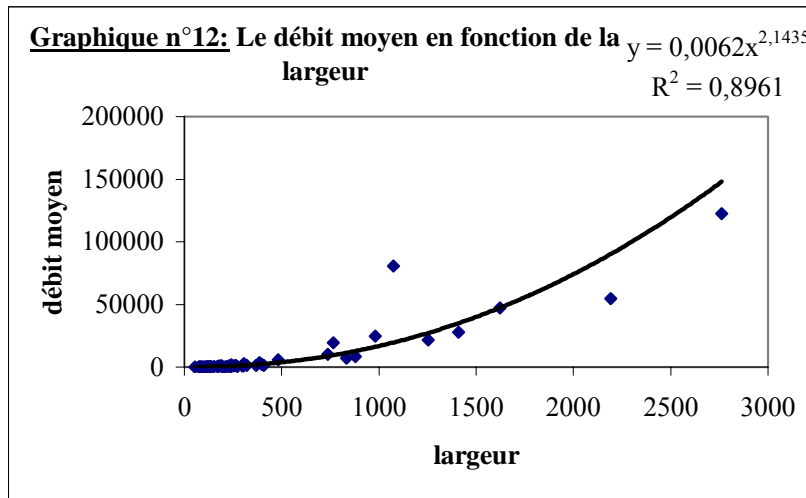
#### **2.5 Conclusion**

Cette ACP a permis de dégager des associations entre les paramètres « hydrologiques » du cours d'eau et les paramètres « géomorphologiques » du bassin. Ainsi, on peut distinguer deux grands groupes de variables. Le premier regroupe les variables concernant la section, le débit et la largeur. Il existe de fortes corrélations positives entre ces variables.

Un second groupe est composé de l'ensemble des profondeurs. Ce groupe n'est pas étonnant dans la mesure où c'est à partir des mêmes données qu'on a extraites les profondeurs. Enfin, on peut voir que la pente se distingue des autres variables. Ce n'est qu'au quatrième axe que on l'a analysé.

### 3. Régression linéaire simple

#### 3.1 Le débit moyen en fonction de la largeur.



Le modèle approprié est  $Q_m = a \text{ larg}^b$  avec  
Q<sub>m</sub> : débit moyen  
Larg : largeur

En prenant le logarithme,

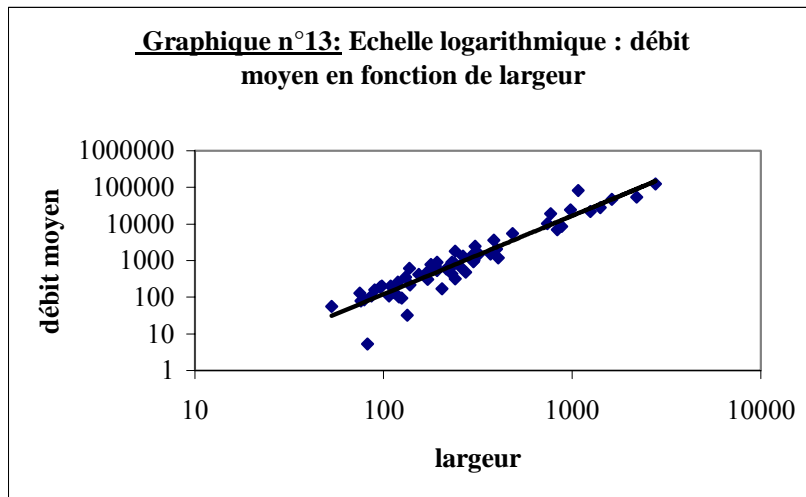
$$\ln(Q_m) = \ln(a \text{ larg}^b)$$

$\ln(Q_m) = \ln(a) + b \ln(\text{larg})$  on obtient une forme linéaire.

$$Y = A + BX + e_i \quad \text{avec } i = 1..60$$

$e_i$  est une réalisation de variable aléatoire de variance  $\Phi^2$  d'espérance nulle et deux à deux non corrélées.





En faisant la régression linéaire de  $\ln(Q_m)$  sur  $\ln(\text{largeur})$ , on obtient les résultats suivant :

Tableau n°6	Degré de liberté	Somme des carrés	Moyenne des carrés	F	Valeur critique de F
Régression	1	222,2452356	222,2452356	500,14615	3,35016E-30
Résidus	58	25,77291393	0,444360585		
Total	59	248,0181496			

Tableau n°7	Coefficients	Erreur-type	Statistique t	Probabilité
A*	-5,079987632	0,537522707	-9,450740526	2,413E-13
B*	2,14354688	0,095848324	22,36394754	3,3502E-30

Le test global va nous permettre de savoir si l'ajustement est bon. Les hypothèses à tester sont :  $H_0$  « il n'y a pas de liaison linéaire entre X et Y » contre  $H_1$  : «  $H_0$  est fausse »

La probabilité critique vaut  $3,35016E-30 < 0,05$ .

Ainsi, on rejette  $H_0$  au seuil de 5%.

L'estimation de A est  $A^* = -5,08$

L'estimation de B vaut  $B^* = 2,1435$

Donc,  $\ln(Q_m) = -5,08 + 2,1435 \ln(\text{larg})$

Et  $Q_m = e(-5,08 + 2,1435 \ln(\text{larg})) = e(-5,08) e(2,1435 \ln(\text{larg})) = 0,0062 e(\ln(\text{larg})^{2,1435})$

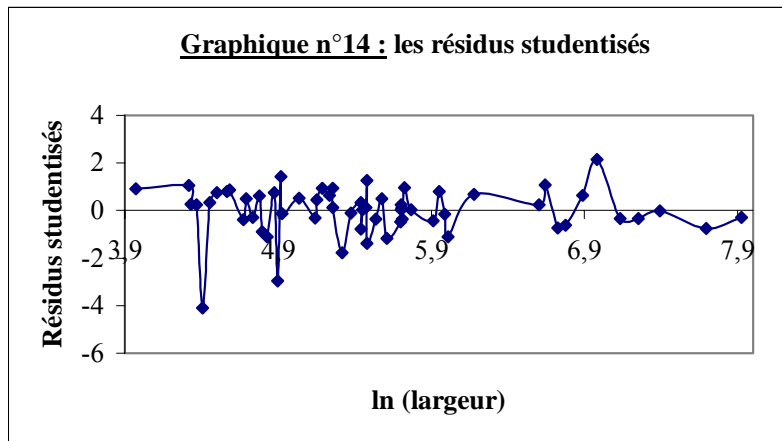
$$Q_m = 0,0062 \text{ larg}^{2,1435}$$

On retrouve les estimations de a et de b.

$$a^* = e(-4,6947) = 0,0062$$

$$b^* = 2,1435$$

Etude des résidus:



Les résidus ne présentent pas de forme particulière. 3 résidus ne sont pas compris entre  $-2$  et  $2$ . Ces résidus correspondent aux stations 18250000, 17070000 et 13150000.

En refaisant cette analyse sans ces trois stations on obtient un meilleur  $r^2=0.9395$ .

Le nouveau nuage de points est :

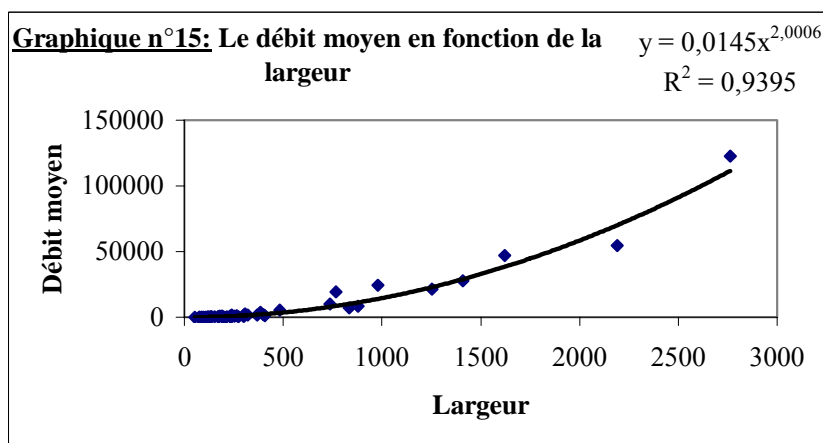
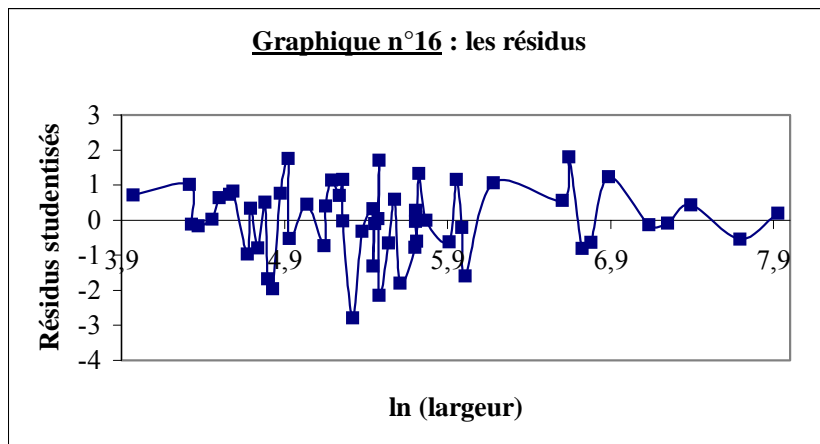


Tableau n°8	Degré de liberté	Somme des carrés	Moyenne des carrés	F	Valeur critique de F
Régression	1	178,5116414	178,5116414	853,956559	3,49049E-35
Résidus	55	11,49723622	0,209040659		
Total	56	190,0088776			

Tableau n°9	Coefficients	Erreur-type	Statistique t	Probabilité
A*	-4,231101202	0,384177587	-11,0133994	1,5702E-15
B*	2,000638892	0,068462195	29,22253512	3,4905E-35

Le test global conclut que l'ajustement est bon.

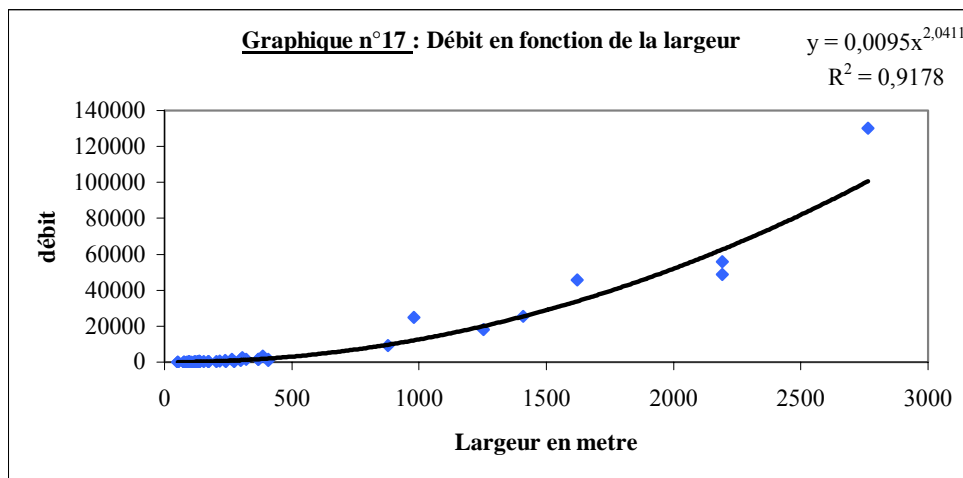


Seulement 2 résidus ne sont pas compris entre  $-2$  et  $2$ . L'ajustement est meilleur. Le débit moyen peut donc s'écrire en fonction de la largeur.

$$Q_m = 0.0145 \text{ larg}^{2.0006}$$

### 3.2 Le débit en fonction de la largeur.

La même étude a été réalisée mais avec le débit annuel.



Le modèle approprié est  $Q = a \text{ larg}^b$  avec

Q : débit

Larg : largeur

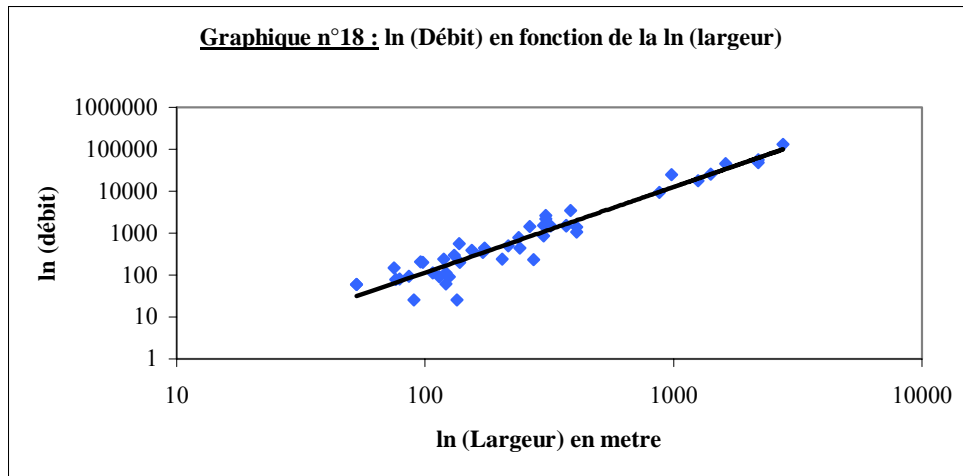
En prenant le logarithme,

$$\ln(Q) = \ln(a \text{ larg}^b)$$

$\ln(Q) = \ln(a) + b \ln(\text{larg})$  on obtient une forme linéaire.

$$Y_i = A + Bx_i + e_i \quad \text{avec } i = 1..46$$

$e_i$  est une réalisation de variable aléatoire de variance  $\Phi^2$  d'espérance nulle et deux à deux non corrélées.



En faisant la régression linéaire de  $\ln(Q)$  sur  $\ln(\text{largeur})$ , on obtient les résultats suivant :

Tableau	Degré de liberté	Somme des carrés	Moyenne des carrés	F	Valeur critique de F
n°10					
Régression	1	197,3760872	197,3760872	491,140417	1,67959E-25
Résidus	44	17,6824133	0,40187303		
Total	45	215,0585005			

Tableau	Coefficients	Erreur-type	Statistique t	Probabilité
n°11				
A*	-4,655432135	0,514352642	-9,051051274	1,318E-11
B*	2,04113166	0,092101814	22,16168804	1,6796E-25

Le test global va nous permettre de savoir si l'ajustement est bon. Les hypothèses à tester sont :  $H_0$  « il n'y a pas de liaison linéaire entre X et Y » contre  $H_1$  : «  $H_0$  est fautive »

La probabilité critique vaut  $1,67959 \cdot 10^{-25} < 0.05$ .

Ainsi, on rejette  $H_0$  au seuil de 5%.

L'estimation de A est  $A^* = -4.6554$

L'estimation de B vaut  $B^* = 2.0411$

Donc,  $\ln(Q) = -4,6554 + 2.0411 \ln(\text{larg})$

Et  $Q = e(-4,6554 + 2.0411 \ln(\text{larg})) = e(-4,6554) e(2.0411 \ln(\text{larg}))$

$$= 0.0095 e(\ln(\text{larg}^{2.0411}))$$

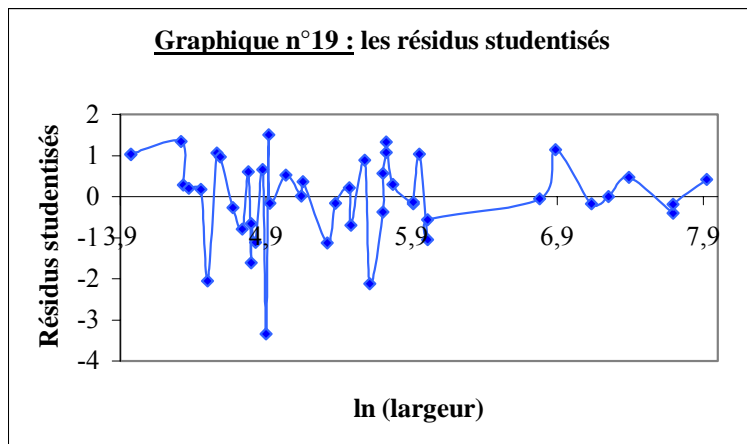
$$Q = 0.0095 \text{ larg}^{2.0411}$$

On retrouve les estimations de a et de b.

$$a^* = e(-4,6554) = 0.0095$$

$$b^* = 2.0411$$

Etude des résidus:



Les résidus ne présentent pas de forme particulière. Seulement 3 résidus ne sont pas compris entre  $-2$  et  $2$ . Ces trois points correspondent aux stations 17450000, 18250000 et 1830000. On refait la même analyse en supprimant ces trois points. Le  $r^2$  a augmenté, il est passé de 0.9178 à 0.9516.

Le nuage de points est :

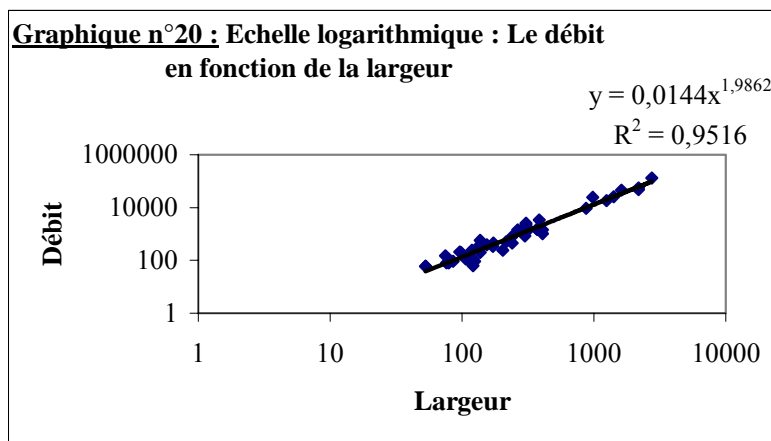
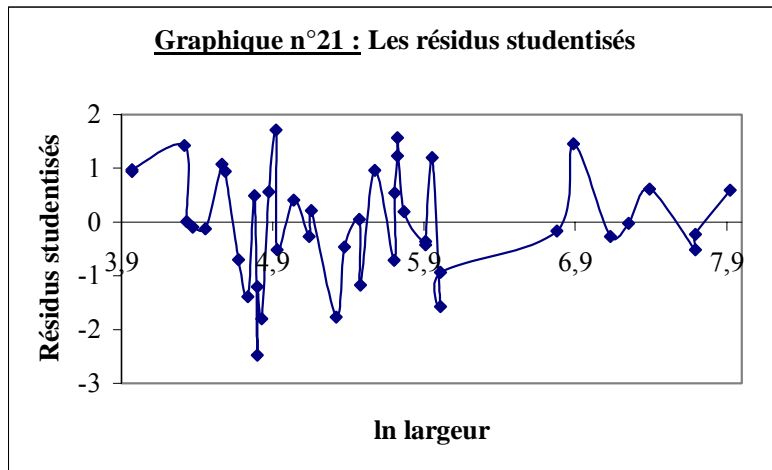


Tableau n°12	Degré de liberté	Somme des carrés	Moyenne des carrés	F	Valeur critique de F
Régression	1	181,370469	181,370469	806,581666	1,37247E-28
Résidus	41	9,219387874	0,224863119		
Total	42	190,5898569			

Tableau n°13	Coefficients	Erreur-type	Statistique t	Probabilité
A*	-4,242330818	0,393154708	-10,79048713	1,5012E-13
B*	1,986195587	0,069935525	28,40038144	1,3725E-28

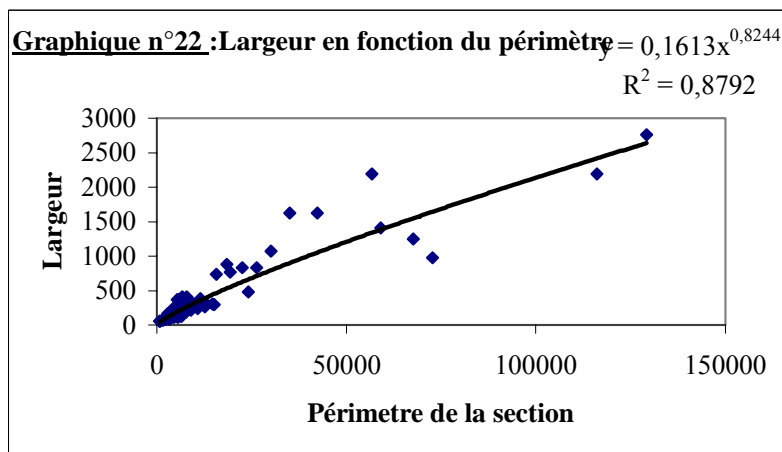


L'ajustement est meilleur, il ne reste plus qu'un seul résidu inférieur à  $-2$ .  
Le modèle est approprié. Le débit peut donc s'écrire en fonction de la largeur.

$$Q = 0.0144 \text{ larg}^{1.9862}$$

On retrouve la même forme d'équation qu'avec le débit moyen, seule la valeur des paramètres change.

### 3.3 La largeur en fonction du périmètre.



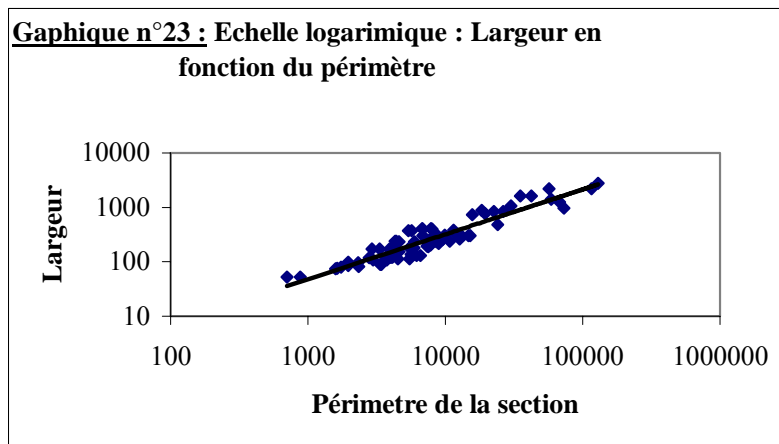
Le modèle approprié est  $\text{Larg} = a \text{ Peri}^b$  avec  
Larg : Largeur  
Peri : Périmètre

En prenant le logarithme,  
 $\text{Ln}(\text{Larg}) = \text{Ln}(a \text{ Peri}^b)$

$\ln(\text{Larg}) = \ln(a) + b \ln(\text{Peri})$  on obtient une forme linéaire.

$$Y = A + BX + e_i \quad \text{avec } i = 1..79$$

$e_i$  est une réalisation de variable aléatoire de variance  $\Phi^2$  d'espérance nulle et deux à deux non corrélées.



En faisant la régression linéaire de  $\ln(\text{Larg})$  sur  $\ln(\text{Peri})$ , on obtient les résultats suivant :

Tableau n°14	Degré de liberté	Somme des carrés	Moyenne des carrés	F	Valeur critique de F
Régression	1	57,10664791	57,10664791	560,41957	4,40559E-37
Résidus	77	7,8462854	0,10189981		
Total	78	64,95293331			

Tableau n°15	Coefficients	Erreur-type	Statistique t	Probabilité
A*	-1,824516747	0,312777296	-5,833277449	1,2114E-07
B*	0,824422707	0,034825174	23,6731825	4,4056E-37

Le test global va nous permettre de savoir si l'ajustement est bon. Les hypothèses à tester sont :  $H_0$  « il n'y a pas de liaison linéaire entre X et Y » contre  $H_1$  : «  $H_0$  est fausse »

La probabilité critique vaut  $4,40559E-37 < 0.05$ .

Ainsi, on rejette  $H_0$  au seuil de 5%.

L'estimation de A est  $A^* = -1.82$

L'estimation de B vaut  $B^* = 0.824$

Donc,  $\ln(\text{Larg}) = -1.82 + 0.8244 \ln(\text{Peri})$

Et  $\text{Larg} = e(-1.82 + 0.8244 \ln(\text{Peri})) = e(-1.82) e(0.8244 \ln(\text{Peri})) = 0.1613 e(\ln(\text{Peri})^{0.8244})$

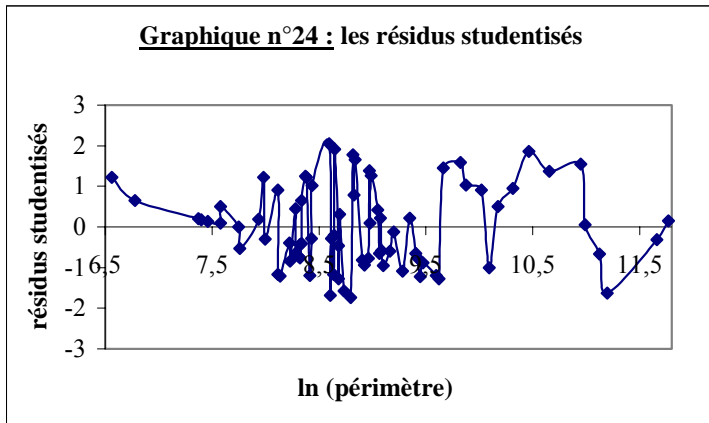
$$\text{Larg} = 0.1613 \text{ Peri}^{0.8244}$$

On retrouve les estimations de a et de b.

$$a^* = e(-1.8245) = 0.1613$$

$$b^* = 0.8244$$

Etude des résidus:

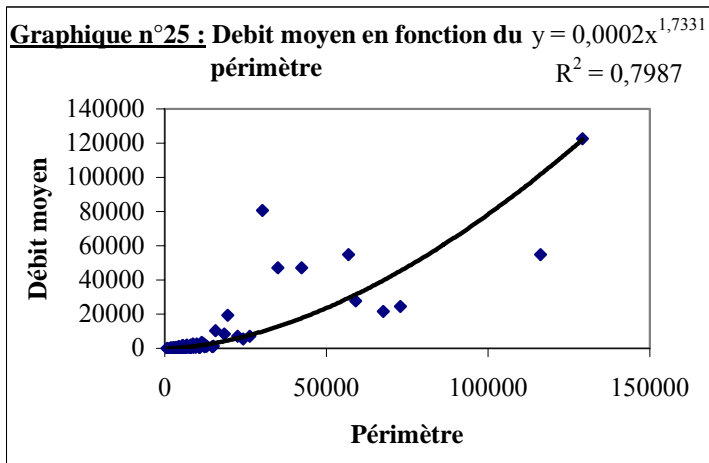


Les résidus ne présentent pas de forme particulière. Un seul résidu n'est pas compris entre  $-2$  et  $2$ . On peut donc conclure que l'ajustement est bon.

La largeur peut donc s'écrire en fonction du périmètre.

$$\text{Larg} = 0.1613 \text{Peri}^{0.8244}$$

### 3.4 Le débit moyen en fonction du périmètre.



Le modèle approprié est  $Q_m = a \text{Peri}^b$  avec  
 $Q_m$  : débit moyen  
 $\text{Peri}$  : Périmètre

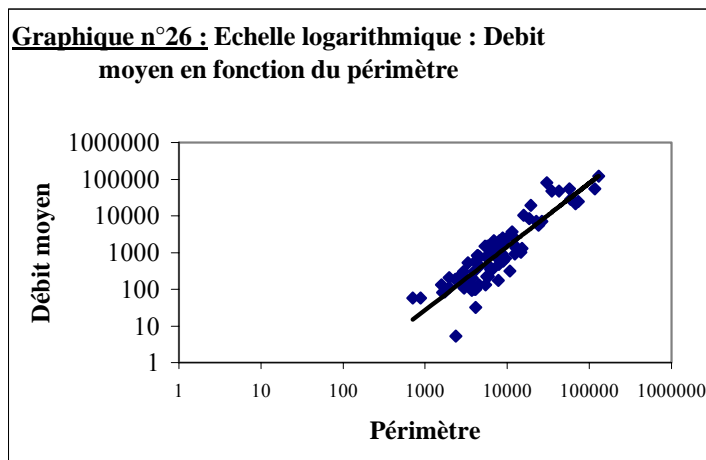
En prenant le logarithme,  
 $\text{Ln}(Q_m) = \text{Ln}(a \text{Peri}^b)$



$\ln(Q_m) = \ln(a) + b \ln(\text{Peri})$  on obtient une forme linéaire.

$$Y = A + BX + e_i \quad \text{avec } i = 1..60$$

$e_i$  est une réalisation de variable aléatoire de variance  $\Phi^2$  d'espérance nulle et deux à deux non corrélées.



En faisant la régression linéaire de  $\ln(Q_m)$  sur  $\ln(\text{largeur})$ , on obtient les résultats suivant :

Tableau n°16	Degré de liberté	Somme des carrés	Moyenne des carrés	F	Valeur critique de F
Régression	1	252,3618465	252,3618465	305,5410466	1,59039E-28
Résidus	77	63,5982052	0,825950717		
Total	78	315,9600517			

Tableau n°17	Coefficients	Erreur-type	Statistique t	Probabilité
A*	-8,683915573	0,890482933	-9,751916913	4,30606E-15
B*	1,733079395	0,099147936	17,47973245	1,59039E-28

Le test global va nous permettre de savoir si l'ajustement est bon. Les hypothèses à tester sont :  $H_0$  « il n'y a pas de liaison linéaire entre X et Y » contre  $H_1$  : «  $H_0$  est fausse »

La probabilité critique vaut  $1,59039E-28 < 0.05$ .

Ainsi, on rejette  $H_0$  au seuil de 5%.

L'estimation de A est  $A^* = -8.68$

L'estimation de B vaut  $B^* = 1.7331$

Donc,  $\ln(Q_m) = -8.68 + 1.73 \ln(\text{Peri})$

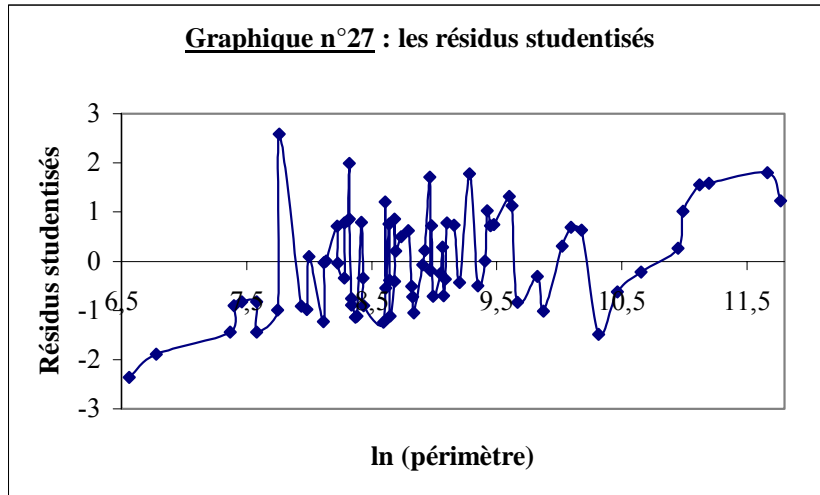
$$Q_m = 0.0002 \text{ Peri}^{1.7331}$$

On retrouve les estimations de a et de b.

$$a^* = e(-8.68) = 0.0002$$

$$b^* = 1.7331$$

Etude des résidus:



Les résidus ne présentent pas de forme particulière. 2 résidus ne sont pas compris entre -2 et 2. On peut donc conclure que l'ajustement est bon.

Le débit moyen peut donc s'écrire en fonction du périmètre.

$$Q_m = 0.0002 \text{ Peri}^{1.7331}$$

### 3.5 Le débit moyen en fonction de l'aire de la section.

La régression linéaire simple de  $\ln(Q_m)$  en fonction de  $\ln(\text{aire de la section})$  donne les résultats suivant :

Tableau n°18 :

Dependent variable: lndebitmoyen

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	-5,74749	0,793707	-7,24133	0,0000
ln airesection	1,13976	0,0715244	15,9353	0,0000

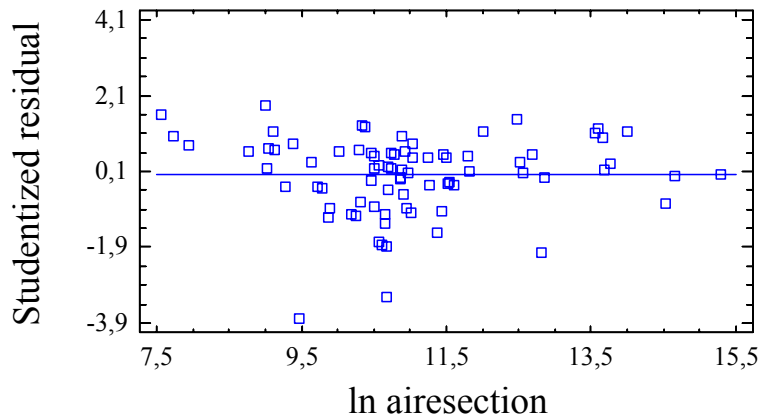
Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	242,444	1	242,444	253,93	0,0000
Residual	73,5162	77	0,954756		
Total (Corr.)	315,96	78			

Donc le débit moyen peut s'écrire en fonction de l'aire de la section.

$$Q_m = e(-5.74749)\text{aire\_section}^{1.13976} = 0.00319 \text{ aire\_section}^{1.13976}$$

**Graphique n°28 : Residual Plot**



Le graphique des résidus studentisés ne présente aucune forme particulière. Seulement 2 résidus n'appartiennent pas à l'intervalle  $[-2 ; 2]$ . La probabilité critique est nulle  $< 0.05$ . On accepte donc le modèle au seuil de 5%.

$$Q_m = 0.00319 \text{ aire\_section}^{1.13976}$$

#### **4. Régression linéaire multiple**

##### **4.1 Le débit moyen en fonction de la largeur, l'aire des sous bassins et de la sinuosité.**

On veut expliquer le débit moyen en fonction de la largeur, l'aire des sous bassins et de la sinuosité.

Le modèle choisi est :

$$\ln(\text{débit})_i = B_1 \ln(\text{sinuosité})_i + B_2 \ln(\text{aire bassin})_i + B_3 \ln(\text{largeur})_i + e_i$$

$i$ , allant de 1 à 60.

Les  $e_i$  sont des réalisations de variable aléatoire d'espérance mathématique nulle de variance  $\Phi^2$  et deux à deux non corrélées.

Testons la qualité d'ajustement de ce modèle.

Les hypothèses à tester sont :

$H_0$  : «  $B_1=B_2=B_3=0$  » contre  $H_1$ : « Il existe un  $j$  tel que  $B_j$  différent de 0 »

La probabilité critique vaut  $0.000 < 0.05$ . Ce qui conduit à rejeter  $H_0$  au seuil de 5% (voir annexe).

**Ainsi le modèle est approprié, l'ajustement est bon.**

L'apport de la variable  $\ln(\text{sinuosité})$  une fois les autres variables introduites dans le modèle est-il significatif ?

Les hypothèses à tester sont :

$H_0$  : «  $B_1=0$  » contre  $H_1$ : «  $B_1$  différent de 0 »

La probabilité critique vaut  $0.3490 > 0.05$ . Ainsi, on accepte  $H_0$  voir annexe.

**L'apport de la variable ln sinuosité n'est pas significatif.**

L'apport de la variable ln (aire bassin) une fois les autres variables introduites dans le modèle est-il significatif ?

Les hypothèses à tester sont :

Ho : «  $B_2 = 0$  » contre H<sub>1</sub>: «  $B_2$  différent de 0 »

La probabilité critique vaut  $0.00 < 0.05$ . Ainsi, on rejette Ho au seuil de 5% voir annexe.

**L'apport de la variable ln aire bassin est significatif.**

**La largeur et l'aire du bassin suffisent pour expliquer le débit moyen.**

Le nouveau modèle est :

$$\ln(\text{débit})_i = B_1 \ln(\text{aire bassin})_i + B_2 \ln(\text{largeur})_i + e_i$$

,i allant de 1 à 60.

Les  $e_i$  sont des réalisations de variable aléatoire d'espérance mathématique nulle de variance  $\Phi^2$  et deux à deux non corrélées.

Testons la qualité d'ajustement de ce modèle.

Les hypothèses à tester sont :

Ho : «  $B_1 = B_2 = 0$  » contre H<sub>1</sub>: « Il existe un j tel que  $B_j$  différent de 0 »

La probabilité critique vaut  $0.000 < 0.05$ . Ce qui conduit à rejeter Ho au seuil de 5% voir annexe.

**Ainsi le modèle est approprié, l'ajustement est bon.**

L'apport de la variable ln (aire bassin) une fois les autres variables introduites dans le modèle est-il significatif ?

Les hypothèses à tester sont :

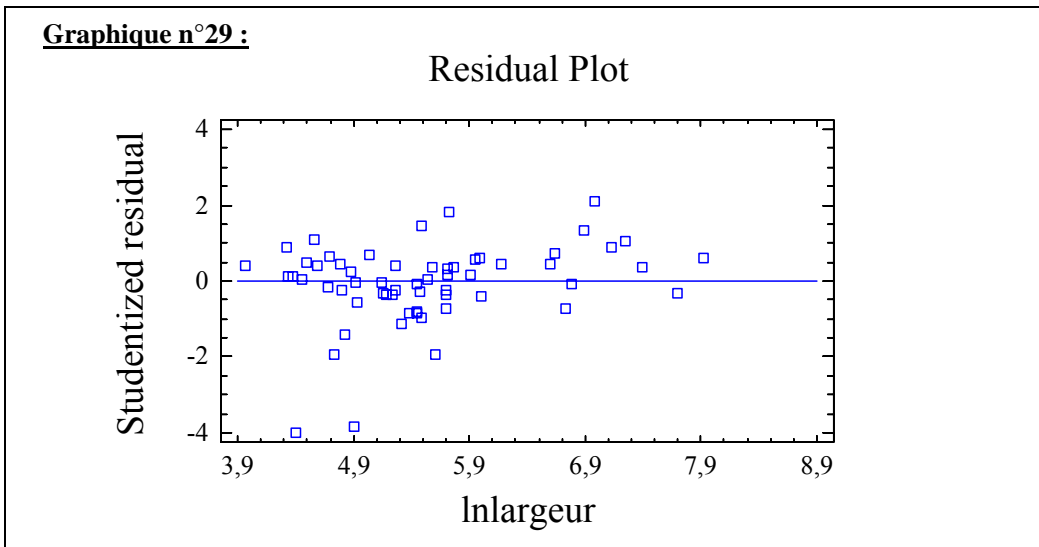
Ho : «  $B_1 = 0$  » contre H<sub>1</sub>: «  $B_1$  différent de 0 »

La probabilité critique vaut  $0.00 < 0.05$ . Ainsi, on rejette Ho au seuil de 5% voir annexe.

**L'apport de la variable ln aire bassin est significatif.**

Le  $r^2$  vaut 0.98.

**Graphique n°29 :**



Seulement 3 résidus studentisés ne sont pas compris entre  $-2$  et  $2$ .  
Les résidus ne présentent aucune forme particulière. Ainsi le modèle est accepté avec  $r^2=0.98$  au seuil de 5 %.

On a donc :

$$\ln(\text{débit moyen}) = 1.83624 \ln(\text{largeur}) - 0.377122 \ln(\text{aire bassin})$$

$$\text{Débit moyen} = \text{largeur}^{1.83634} / \text{aire bassin}^{0.377122}$$

#### 4.2 La largeur en fonction de l'aire de la section et de la profondeur moyenne.

Tableau n°19 :

## Multiple Regression Analysis

Dependent variable: lnlargeur

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	2,06043	0,480417	4,28883	0,0001
ln aire section	0,829808	0,05445	15,2398	0,0000
lnprofondeurmoyen	-0,845701	0,141481	-5,97748	0,0000

## Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	57,2155	2	28,6078	281,00	0,0000
Residual	7,7374	76	0,101808		
Total (Corr.)	64,9529	78			

R-squared = 88,0877 percent

R-squared (adjusted for d.f.) = 87,7742 percent

Standard Error of Est. = 0,319074

Mean absolute error = 0,266122

Durbin-Watson statistic = 0,986944 (P=0,0000)

Lag 1 residual autocorrelation = 0,505567

Comme on peut le voir ci dessus la largeur peut s'exprimer en fonction de l'aire de la section et de la profondeur moyenne.

Cependant ce modèle n'est pas valide. En effet la statistique du test de Durbin Watson montre que les résidus sont auto corrélés avec un coefficient de 0.505567 . Ainsi il faut re-estimer les paramètres en prenant en compte cette auto corrélation.

Tableau n°20 :

## Multiple Regression Analysis

Dependent variable: lnlargeur

Cochrane-Orcutt transformation applied: autocorrelation = 0,54808

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	1,70981	0,409893	4,17136	0,0001
ln aire section	0,750102	0,0442417	16,9546	0,0000
lnprofondeurmoyen	-0,661211	0,113674	-5,81672	0,0000

## Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	47,8542	2	23,9271	325,11	0,0000
Residual	5,51969	75	0,0735959		
Total (Corr.)	53,3739	77			

R-squared = 89,6584 percent

R-squared (adjusted for d.f.) = 89,3827 percent

Standard Error of Est. = 0,271286

Mean absolute error = 0,213996

Durbin-Watson statistic = 2,0697

Lag 1 residual autocorrelation = -0,0405953

La statistique du test de Durbin Waston montre que les résidus ne sont plus auto corrélés.

Le modèle choisi est :

$$\ln(\text{largeur})_i = B_0 + B_1 \ln(\text{profondeur moyenne})_i + B_2 \ln(\text{aire section})_i + e_i$$

$i$ , allant de 1 à 79.

Les  $e_i$  sont des réalisations de variable aléatoire d'espérance mathématique nulle de variance  $\Phi^2$  et deux à deux non corrélées.

Testons la qualité d'ajustement de ce modèle.

Les hypothèses à tester sont :

$H_0$  : «  $B_0=B_1=B_2=0$  » contre  $H_1$  : « Il existe un  $j$  tel que  $B_j$  différent de 0 »

La probabilité critique vaut  $0.000 < 0.05$ . Ce qui conduit à rejeter  $H_0$  au seuil de 5%.

**Ainsi le modèle est approprié, l'ajustement est bon.**

L'apport de la variable  $\ln$  (aire section) une fois les autres variables introduites dans le modèle est-il significatif ?

Les hypothèses à tester sont :

$H_0$  : «  $B_2=0$  » contre  $H_1$  : «  $B_2$  différent de 0 »

La probabilité critique vaut  $0.000 < 0.05$ . Ainsi, on rejette  $H_0$  au seuil de 5 %.

**L'apport de la variable  $\ln$  aire section est significatif.**

L'apport de la variable  $\ln$  (profondeur moyenne) une fois les autres variables introduites dans le modèle est-il significatif ?

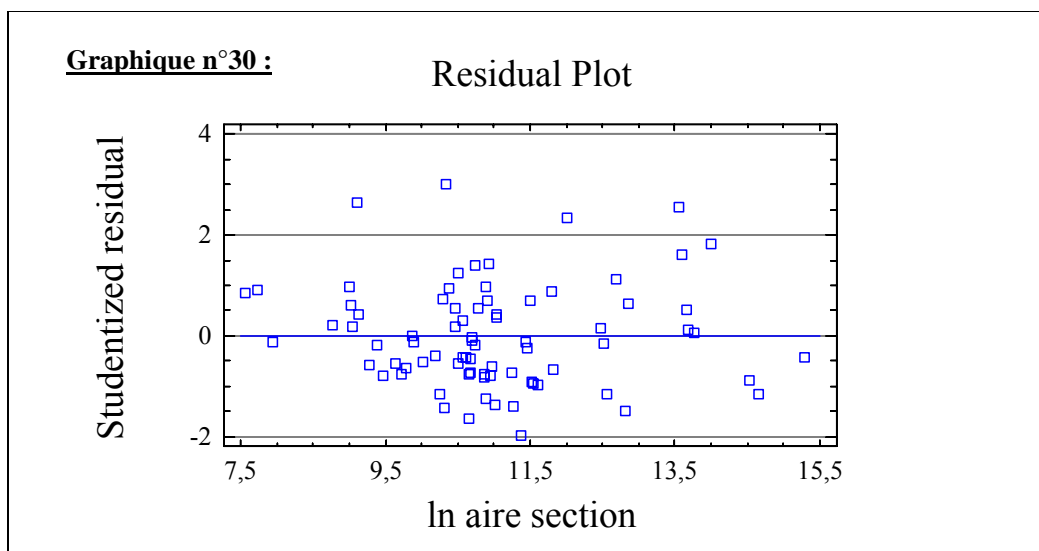
Les hypothèses à tester sont :

$H_0$  : «  $B_1=0$  » contre  $H_1$  : «  $B_1$  différent de 0 »

La probabilité critique vaut  $0.00 < 0.05$ . Ainsi, on rejette  $H_0$  au seuil de 5%.

**L'apport de la variable  $\ln$  profondeur moyenne est significatif.**

**Les deux variables sont significatives.**



Le graphique des résidus studentisés montre quatre valeurs supérieure à 2. Ce quatre valeurs correspondent aux stations 11400000, 15250000, 15750000, 17093000. On refait la même étude sans ces points.

Tableau n°21 :

Multiple Regression Analysis

Dependent variable: lnlargeur

Cochrane-Orcutt transformation applied: autocorrelation = 0,541508

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	1,57709	0,392517	4,01789	0,0001
ln aire section	0,748197	0,0429797	17,4082	0,0000
lnprofondeurmoyen	-0,641412	0,109268	-5,87007	0,0000

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	47,676	2	23,838	347,99	0,0000
Residual	4,93219	72	0,0685027		
Total (Corr.)	52,6081	74			

R-squared = 90,6247 percent

R-squared (adjusted for d.f.) = 90,3642 percent

Standard Error of Est. = 0,26173

Mean absolute error = 0,211252

Durbin-Watson statistic = 2,2139

Lag 1 residual autocorrelation = -0,113597

Le test global et les tests d'ajouts de variable sont concluants. Le graphique des résidus donne deux résidus supérieur à 2. Le  $r^2$  est meilleur, il vaut 0.90. On accepte l'ajustement.

La largeur peut s'exprimer en fonction de l'aire de la section et de la profondeur moyenne.

$\ln(\text{largeur}) = 1.57709 + 0.748197 \ln(\text{aire de la section}) - 0.641412 \ln(\text{profondeur moyenne})$

$\text{Largeur} = 4.8408 \text{ aire}^{0.748197} / \text{profondeur moyenne}^{0.641412}$
--

**4.3 Le débit moyen en fonction du rayon hydraulique et de l'aire de la section.**

Tableau n°22 :



-----  
Dependent variable: lndebitmoyen  
-----

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	-8,31083	1,0379	-8,00737	0,0000
lnrayonhydraulique	-1,45102	0,411588	-3,52541	0,0007
ln aire section	1,64605	0,158364	10,3941	0,0000

-----  
Analysis of Variance  
-----

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	252,777	2	126,388	152,03	0,0000
Residual	63,1835	76	0,831362		
Total (Corr.)	315,96	78			

R-squared = 80,0027 percent  
R-squared (adjusted for d.f.) = 79,4764 percent  
Standard Error of Est. = 0,911791  
Mean absolute error = 0,723315  
Durbin-Watson statistic = 1,62957 (P=0,0417)  
Lag 1 residual autocorrelation = 0,179066

La régression multiple de ln du débit en fonction de ln (rayon hydraulique) et de ln (aire de la section) donne des résidus autocorrélés. En effet, la probabilité critique du test de Durbin-Watson vaut 0.0417 et est inférieure à 0.05. De ce fait, on rejette l'hypothèse de non corrélations des résidus. La régression multiple avec les mêmes variables en prenant en compte l'auto corrélations donne les résultats suivants.

Tableau n°23 :

Dependent variable: lndebitmoyen

Cochrane-Orcutt transformation applied: autocorrelation = 0,200744  
-----

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	-7,70536	1,06286	-7,24968	0,0000
lnrayonhydraulique	-1,26322	0,412596	-3,06163	0,0031
ln aire section	1,55469	0,159427	9,75171	0,0000

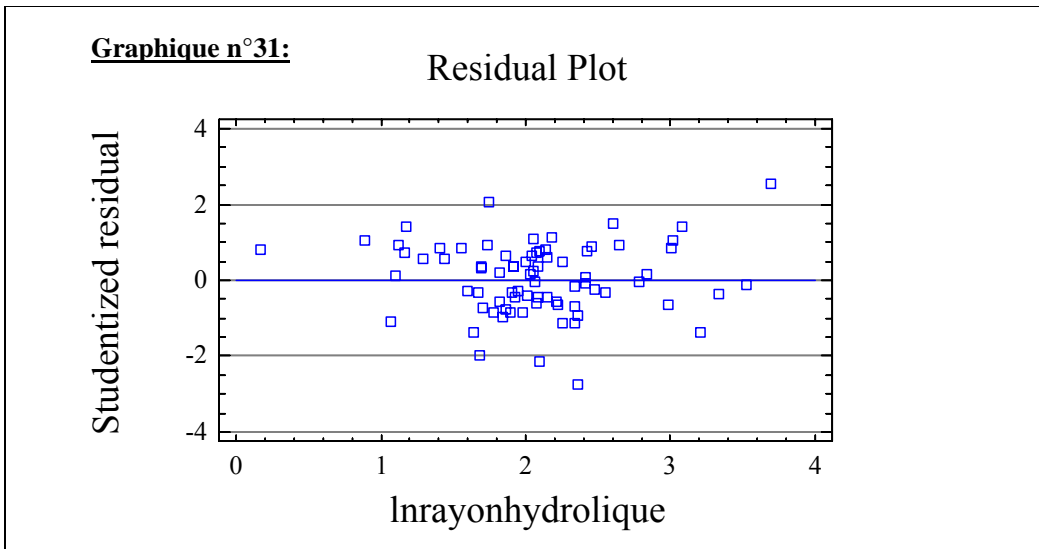
-----  
Analysis of Variance  
-----

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	213,826	2	106,913	132,42	0,0000
Residual	60,5548	75	0,807397		
Total (Corr.)	274,381	77			

R-squared = 77,9304 percent  
R-squared (adjusted for d.f.) = 77,3419 percent  
Standard Error of Est. = 0,898553  
Mean absolute error = 0,689642  
Durbin-Watson statistic = 2,06639  
Lag 1 residual autocorrelation = -0,0344717

La probabilité critique(=0.000<0.05) du test global conclut que l'ajustement est bon.

De plus, les probabilités critiques (respectivement  $0.031 < 0.05$  et  $0.000 < 0.05$ ) correspondant aux variables  $\ln$  (rayon hydraulique) et  $\ln$  (aire section) montre que l'ajout de ces deux variables est significatif.



Le graphique des résidus studentisés n'a pas de forme particulière. On peut voir 4 résidus non compris entre -2 et 2. Cependant deux de ces résidus sont proches de -2 et 2. Donc, on peut conclure que l'ajustement est de bonne qualité avec un  $r^2 = 0.779$ .

Le débit moyen peut donc s'estimer en fonction de ces deux variables.

$\ln$  (débit moyen) =  $-7.70536 - 1.26322 \ln$  (rayon hydraulique) +  $1.55469 \ln$  (aire de la section)

$$Q_m = 0.00045 \text{ aire section}^{1.55469} / \text{ rayon hydraulique}^{1.26322}$$

#### 4.4 Le débit moyen en fonction du rayon hydraulique, de la pente et de la profondeur moyenne.

Tableau n°24 :

Multiple Regression Analysis

Dependent variable: lndebitmoyen

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	774,272	65,7258	11,7803	0,0000
lnprofondeurmoyen	-167,483	14,2556	-11,7486	0,0000
lnrayonhydrolique	169,612	14,2808	11,8769	0,0000
lnpente	-0,2842	0,104682	-2,71489	0,0082

#### Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	267,631	3	89,2103	137,40	0,0000
Residual	48,0456	74	0,649264		
Total (Corr.)	315,677	77			

R-squared = 84,7801 percent  
R-squared (adjusted for d.f.) = 84,1631 percent  
Standard Error of Est. = 0,805769  
Mean absolute error = 0,541801  
Durbin-Watson statistic = 2,09279 (P=0,2932)  
Lag 1 residual autocorrelation = -0,0509415

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between lndebitmoyen and 3 independent variables. The equation of the fitted model is

$$\lndebitmoyen = 774,272 - 167,483 \cdot \lnprofondeurmoyenne + 169,612 \cdot \lnrayonhydrolique - 0,2842 \cdot \lnpente$$

La probabilité critique du test de Durbin-Watson vaut 0.2932 et est supérieure à 0.05. De ce fait, on accepte l'hypothèse de non corrélations des résidus.

Le modèle choisi est :

$$\ln(\text{débit moyen})_i = B_0 + B_1 \ln(\text{profondeur moyenne})_i + B_2 \ln(\text{rayon hydraulique})_i + B_3 \ln(\text{pente})_i + e_i$$

$i$ , allant de 1 à 79.

Les  $e_i$  sont des réalisations de variable aléatoire d'espérance mathématique nulle de variance  $\Phi^2$  et deux à deux non corrélées.

Testons la qualité d'ajustement de ce modèle.

Les hypothèses à tester sont :

$H_0$  : «  $B_0=B_1=B_2=B_3=0$  » contre  $H_1$ : « Il existe un  $j$  tel que  $B_j$  différent de 0 »

La probabilité critique vaut  $0.000 < 0.05$ . Ce qui conduit à rejeter  $H_0$  au seuil de 5%.

**Ainsi le modèle est approprié, l'ajustement est bon.**

L'apport de la variable ln (rayon hydraulique) une fois les autres variables introduites dans le modèle est-il significatif ?

Les hypothèses à tester sont :

Ho : «  $B_2 = 0$  » contre H<sub>1</sub>: «  $B_2$  différent de 0 »

La probabilité critique vaut  $0.000 < 0.05$ . Ainsi, on rejette Ho au seuil de 5 %.

**L'apport de la variable ln rayon hydraulique est significatif.**

L'apport de la variable ln (profondeur moyenne) une fois les autres variables introduites dans le modèle est-il significatif ?

Les hypothèses à tester sont :

Ho : «  $B_1 = 0$  » contre H<sub>1</sub>: «  $B_1$  différent de 0 »

La probabilité critique vaut  $0.00 < 0.05$ . Ainsi, on rejette Ho au seuil de 5%.

**L'apport de la variable ln profondeur moyenne est significatif.**

L'apport de la variable ln (pente) une fois les autres variables introduites dans le modèle est-il significatif ?

Les hypothèses à tester sont :

Ho : «  $B_3 = 0$  » contre H<sub>1</sub>: «  $B_3$  différent de 0 »

La probabilité critique vaut  $0.0082 < 0.05$ . Ainsi, on rejette Ho au seuil de 5%.

**L'apport de la variable ln pente est significatif.**

**Les trois variables sont significatives.**

```
Step 0:
-----
0 variables in the model. 77 d.f. for error.
R-squared = 0,00% Adjusted R-squared = 0,00% MSE = 4,0997

Step 1:
-----
Adding variable lnrayonhydraulique with F-to-enter = 81,0164
1 variables in the model. 76 d.f. for error.
R-squared = 51,60% Adjusted R-squared = 50,96% MSE = 2,01047

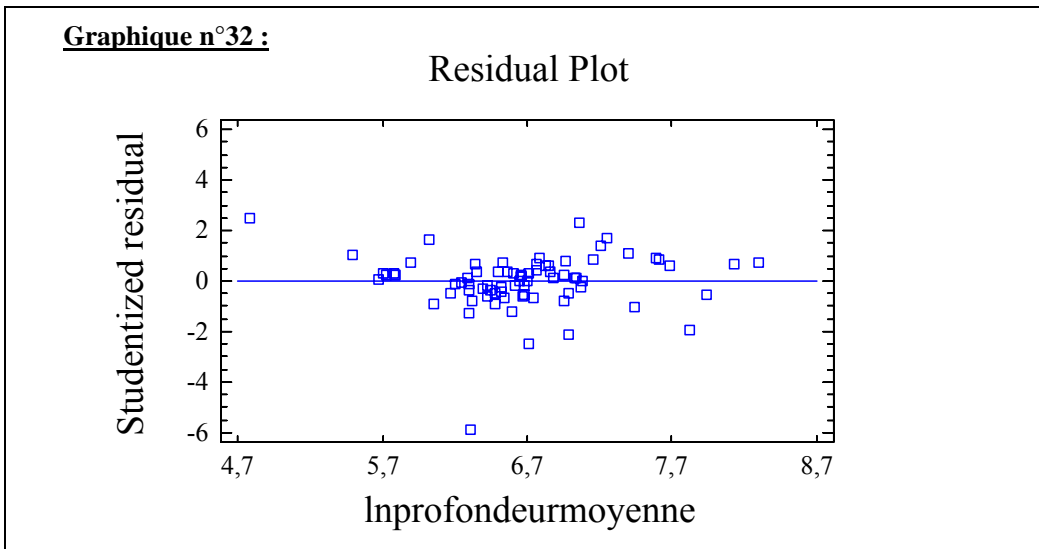
Step 2:
-----
Adding variable lnprofondeurmoyenne with F-to-enter = 141,912
2 variables in the model. 75 d.f. for error.
R-squared = 83,26% Adjusted R-squared = 82,82% MSE = 0,704414

Step 3:
-----
Adding variable lnpenne with F-to-enter = 7,37063
3 variables in the model. 74 d.f. for error.
R-squared = 84,78% Adjusted R-squared = 84,16% MSE = 0,649264
```

Lorsque l'on ajoute la variable profondeur moyenne au modèle débit moyen en fonction du rayon hydraulique, le  $r^2$  passe de 0.5160 à 0.8326 et le carré moyen des erreurs diminue (il passe de 2.01047 à 0.704414). La profondeur moyenne améliore donc le modèle.

Lorsque l'on ajoute la pente, le  $r^2$  augmente encore ( $r^2=0.8478$ ) et le carré moyen des erreurs diminue, il vaut 0.649264.

**Graphique n°32 :**



Le graphique des résidus montre 5 valeurs non comprise entre -2 et 2. Ce graphique conduit à rejeter le modèle. Cependant, le  $r^2$  vaut 0.847801 et le test global est positif. Donc le débit moyen peut quand même s'exprimer en fonction du rayon hydraulique, de la pente et de la profondeur moyenne.

$\ln(\text{débit moyen}) = 774.272 - 167.483 \ln(\text{profondeur moyenne}) + 169.612 \ln(\text{rayon hydraulique}) - 0.2842 \ln(\text{pente})$

$$\text{Débit moyen} = e(774.272) \text{profondeur moyenne}^{167.483} \text{rayon hydraulique}^{169.612} / \text{pente}^{0.2842}$$

#### 4.5 Le débit moyen en fonction de la profondeur maximale, de la largeur et de l'aire de la section.

Tableau n°25 :

Dependent variable: lndebitmoyen

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	-9,02685	1,09303	-8,25857	0,0000
lnprofondeurmax	1,22211	0,303409	4,02794	0,0001
lnlargeur	2,22484	0,180456	12,329	0,0000
ln aire section	-0,483569	0,172926	-2,79639	0,0066

##### Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	291,686	3	97,2287	300,41	0,0000
Residual	24,2738	75	0,323651		
Total (Corr.)	315,96	78			

R-squared = 92,3174 percent

R-squared (adjusted for d.f.) = 92,0101 percent

Standard Error of Est. = 0,568903

Mean absolute error = 0,40764

Durbin-Watson statistic = 1,43178 (P=0,0034)

Lag 1 residual autocorrelation = 0,27289

La probabilité critique du test de Durbin-Watson vaut 0.0034 et est inférieure à 0.05. De ce fait, on rejette l'hypothèse de non corrélations des résidus. La régression multiple avec les mêmes variables en prenant en compte l'auto corrélations donne les résultats suivants.

Tableau n°26 :

Dependent variable: lndebitmoyen

Cochrane-Orcutt transformation applied: autocorrelation = 0,319678

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	-8,71316	1,03888	-8,38711	0,0000
lnprofondeurmax	1,18115	0,273603	4,31701	0,0000
lnlargeur	2,43648	0,194447	12,5303	0,0000
ln aire section	-0,593204	0,16478	-3,59998	0,0006

##### Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	240,187	3	80,0624	273,78	0,0000
Residual	21,6402	74	0,292436		
Total (Corr.)	261,828	77			

R-squared = 91,7349 percent

R-squared (adjusted for d.f.) = 91,3999 percent

Standard Error of Est. = 0,540773

Mean absolute error = 0,367749

Durbin-Watson statistic = 2,14053

Lag 1 residual autocorrelation = -0,0748207

Le modèle choisi est :

$$\ln(\text{débit moyen})_i = B_0 + B_1 \ln(\text{profondeur maximale})_i + B_2 \ln(\text{largeur})_i + B_3 \ln(\text{aire de la section})_i + e_i$$

,i allant de 1 à 79.

Les  $e_i$  sont des réalisations de variable aléatoire d'espérance mathématique nulle de variance  $\Phi^2$  et deux à deux non corrélées.

Testons la qualité d'ajustement de ce modèle.

Les hypothèses à tester sont :

$H_0$  : «  $B_0=B_1=B_2=B_3=0$  » contre  $H_1$ : « Il existe un j tel que  $B_j$  différent de 0 »

La probabilité critique vaut  $0.000 < 0.05$ . Ce qui conduit à rejeter  $H_0$  au seuil de 5%.

**Ainsi le modèle est approprié, l'ajustement est bon.**

L'apport des 3 variables une fois les autre ajoutées dans le modèle est significatif dans la mesure ou les probabilités critiques sont inférieures à 0.05.

**Les trois variables sont significatives.**

```
Step 0:
-----
0 variables in the model. 77 d.f. for error.
R-squared = 0,00% Adjusted R-squared = 0,00% MSE = 3,40036

Step 1:
-----
Adding variable lnlargeur with F-to-enter = 656,727
1 variables in the model. 76 d.f. for error.
R-squared = 89,63% Adjusted R-squared = 89,49% MSE = 0,357333

Step 2:
-----
Adding variable lnprofondeurmax with F-to-enter = 5,09379
2 variables in the model. 75 d.f. for error.
R-squared = 90,29% Adjusted R-squared = 90,03% MSE = 0,339069

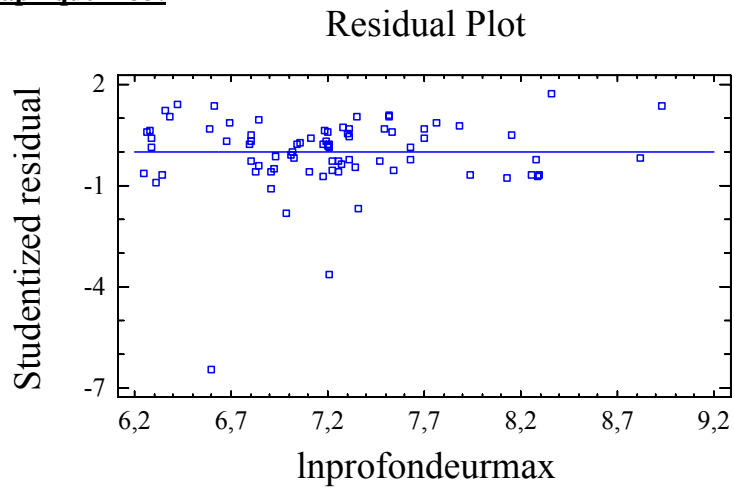
Step 3:
-----
Adding variable ln aire section with F-to-enter = 12,9598
3 variables in the model. 74 d.f. for error.
R-squared = 91,73% Adjusted R-squared = 91,40% MSE = 0,292436
```

Lorsque l'on ajoute la variable profondeur maximale au modèle débit moyen en fonction de la largeur, le  $r^2$  passe de 0.8963 à 0.9029 et le carré moyen des erreurs diminue (il passe de 0.357333 à 0.3390069). La profondeur maximale améliore donc le modèle.

Lorsque l'on ajoute l'aire de la section, le  $r^2$  augment encore ( $r^2=0.9173$ ) et le carré moyen des erreurs diminue, il vaut 0.29436.

**L'ajout de ces deux variables à la somme des carrés est significatif.**

**Graphique n°33:**



Le graphique des résidus studentisés donne seulement deux résidus non compris entre  $-2$  et  $2$ . Ces deux résidus correspondent aux stations 17070000 et 18250000. On accepte le modèle. Le débit moyen peut donc s'exprimer en fonction de la profondeur maximale, de l'aire de la section et de la largeur.

$\ln(\text{débit moyen}) = -8.71316 + 1.18115 \ln(\text{profondeur maximale}) + 2.43648 \ln(\text{largeur}) - 0.593204 \ln(\text{aire section})$   
Donc

$$\text{débit moyen} = 0.000164 \text{ profondeur maximale}^{1.18115} \text{ largeur}^{2.43648} / \text{aire section}^{0.593204}$$



## 4.6 Le débit moyen en fonction de la sinuosité , de la profondeur moyenne et du rayon hydraulique

Tableau n°27 :

Multiple Regression Analysis

Dependent variable: lndebitmoyen

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	861,71	71,2671	12,0913	0,0000
lnsinuosité	0,894447	0,427289	2,09331	0,0397
lnprofondeurmoyen	-186,518	15,46	-12,0645	0,0000
lnrayonhydraulique	188,951	15,4585	12,2231	0,0000

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	265,661	3	88,5536	131,68	0,0000
Residual	49,7632	74	0,672475		
Total (Corr.)	315,424	77			

R-squared = 84,2234 percent

R-squared (adjusted for d.f.) = 83,5838 percent

Standard Error of Est. = 0,820046

Mean absolute error = 0,549962

Durbin-Watson statistic = 1,86888 (P=0,2402)

Lag 1 residual autocorrelation = 0,0567612

Le modèle choisi est :

$\ln(\text{débit moyen})_i = B_0 + B_1 \ln(\text{profondeur moyenne})_i + B_2 \ln(\text{sinuosité})_i + B_3 \ln(\text{rayon hydraulique})_i + e_i$   
 $i$ , allant de 1 à 79.

Les  $e_i$  sont des réalisations de variable aléatoire d'espérance mathématique nulle de variance  $\Phi^2$  et deux à deux non corrélées.

Testons la qualité d'ajustement de ce modèle.

Les hypothèses à tester sont :

$H_0$  : «  $B_0=B_1=B_2=B_3=0$  » contre  $H_1$ : « Il existe un  $j$  tel que  $B_j$  différent de 0 »

La probabilité critique vaut  $0.000 < 0.05$ . Ce qui conduit à rejeter  $H_0$  au seuil de 5%.

**Ainsi le modèle est approprié, l'ajustement est bon.**

L'apport des 3 variables une fois les autre ajoutées dans le modèle est significatif dans la mesure ou les probabilités critiques sont inférieures à 0.05.

**Les trois variables sont significatives.**

Stepwise regression

-----  
Method: forward selection  
F-to-enter: 4,0  
F-to-remove: 4,0

Step 0:

-----

0 variables in the model. 77 d.f. for error.  
R-squared = 0,00% Adjusted R-squared = 0,00% MSE = 4,09641

Step 1:

-----

Adding variable lnrayonhydrolique with F-to-enter = 80,9752  
1 variables in the model. 76 d.f. for error.  
R-squared = 51,58% Adjusted R-squared = 50,95% MSE = 2,00939

Step 2:

-----

Adding variable lnprofondeurmoyenne with F-to-enter = 142,293  
2 variables in the model. 75 d.f. for error.  
R-squared = 83,29% Adjusted R-squared = 82,84% MSE = 0,702799

Step 3:

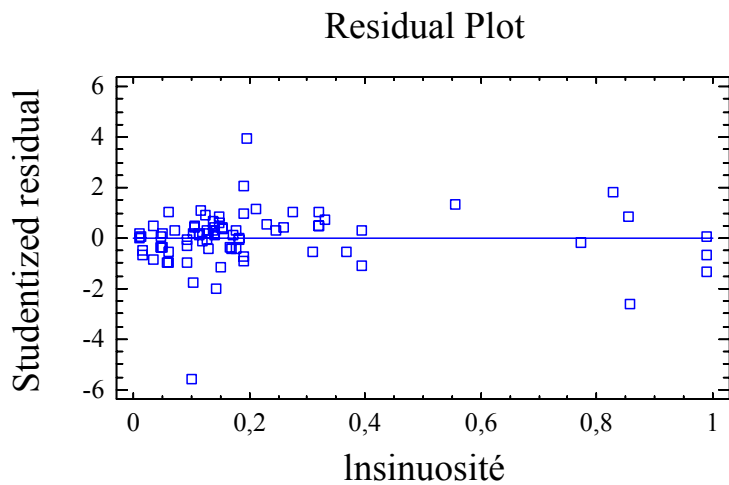
-----

Adding variable lnsinuosité with F-to-enter = 4,38194  
3 variables in the model. 74 d.f. for error.  
R-squared = 84,22% Adjusted R-squared = 83,58% MSE = 0,672475

Final model selected.

Lorsque l'on ajoute la variable profondeur moyenne au modèle débit moyen en fonction du rayon hydraulique, le  $r^2$  passe de 0.5158 à 0.8329 et le carré moyen des erreurs diminue (il passe de 2.00939 à 0.702799). La profondeur moyenne améliore donc le modèle. Lorsque l'on ajoute la sinuosité, le  $r^2$  augmente encore ( $r^2=0.8422$ ) et le carré moyen des erreurs diminue, il vaut 0.672475.

**L'ajout de ces deux variables à la somme des carrés est significatif.**



Le graphique des résidus studentisés donne 5 résidus non compris entre  $-2$  et  $2$ .  
Le nombre de résidus conduit au rejet du modèle.

Cependant le  $r^2$  vaut  $0.8422$  et le test global est positif, on peut quand même accepter le modèle.

Le débit moyen peut donc s'exprimer en fonction de la profondeur moyenne, de la sinuosité et du rayon hydraulique.

$$\ln(\text{débit moyen}) = 861.71 + 0.894447 \ln(\text{sinuosité}) - 186.518 \ln(\text{profondeur moyenne}) + 188.951 \ln(\text{rayon hydraulique})$$

Donc

$$\text{débit moyen} = e(861.71)\text{sinuosité}^{0.894447} \text{ rayon hydraulique}^{188.951} / \text{profondeur moyenne}^{186.518}$$

## 4.7 Le débit moyen en fonction de la profondeur maximale, la largeur et le rayon hydraulique.

Tableau n°28 :

Multiple Regression Analysis

Dependent variable: lndebitmoyen

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	-12,0139	1,82859	-6,57001	0,0000
lnprofondeurmax	1,50401	0,367962	4,08742	0,0001
lnlargeur	1,80891	0,112552	16,0718	0,0000
lnrayonhydraulique	-0,996144	0,331352	-3,00631	0,0036

### Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	292,038	3	97,346	305,20	0,0000
Residual	23,922	75	0,31896		
Total (Corr.)	315,96	78			

R-squared = 92,4288 percent

R-squared (adjusted for d.f.) = 92,1259 percent

Standard Error of Est. = 0,564765

Mean absolute error = 0,38584

Durbin-Watson statistic = 1,63341 (P=0,0378)

Lag 1 residual autocorrelation = 0,170857

La probabilité critique du test de Durbin-Watson vaut 0.00378 et est inférieure à 0.05. De ce fait, on rejette l'hypothèse de non corrélations des résidus. La régression multiple avec les mêmes variables en prenant en compte l'auto corrélations donne les résultats suivants.

Multiple Regression Analysis

Tableau n°29 :

Dependent variable: lndebitmoyen

Cochrane-Orcutt transformation applied: autocorrelation = 0,173067

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	-11,53	1,8478	-6,23986	0,0000
lnprofondeurmax	1,39193	0,371105	3,75078	0,0003
lnlargeur	1,84856	0,117461	15,7377	0,0000
lnrayonhydraulique	-0,951228	0,333136	-2,85538	0,0056

### Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	255,911	3	85,3037	278,43	0,0000
Residual	22,6715	74	0,306372		
Total (Corr.)	278,583	77			

R-squared = 91,8618 percent

R-squared (adjusted for d.f.) = 91,5319 percent

Standard Error of Est. = 0,553508

Mean absolute error = 0,360775

Durbin-Watson statistic = 2,04875

Lag 1 residual autocorrelation = -0,0282376

Le modèle choisi est :

$$\ln(\text{débit moyen})_i = B_0 + B_1 \ln(\text{profondeur max})_i + B_2 \ln(\text{largeur})_i + B_3 \ln(\text{rayon hydraulique})_i + e_i$$

,i allant de 1 à 79.

Les  $e_i$  sont des réalisations de variable aléatoire d'espérance mathématique nulle de variance  $\Phi^2$  et deux à deux non corrélées.

Testons la qualité d'ajustement de ce modèle.

Les hypothèses à tester sont :

$H_0$  : «  $B_0=B_1=B_2=B_3=0$  » contre  $H_1$ : « Il existe un j tel que  $B_j$  différent de 0 »

La probabilité critique vaut  $0.000 < 0.05$ . Ce qui conduit à rejeter  $H_0$  au seuil de 5%.

**Ainsi le modèle est approprié, l'ajustement est bon.**

L'apport des 3 variables une fois les autre ajoutées dans le modèle est significatif dans la mesure ou les probabilités critiques sont inférieures à 0.05.

```
0 variables in the model. 77 d.f. for error.
R-squared = 0,00% Adjusted R-squared = 0,00% MSE = 3,61796

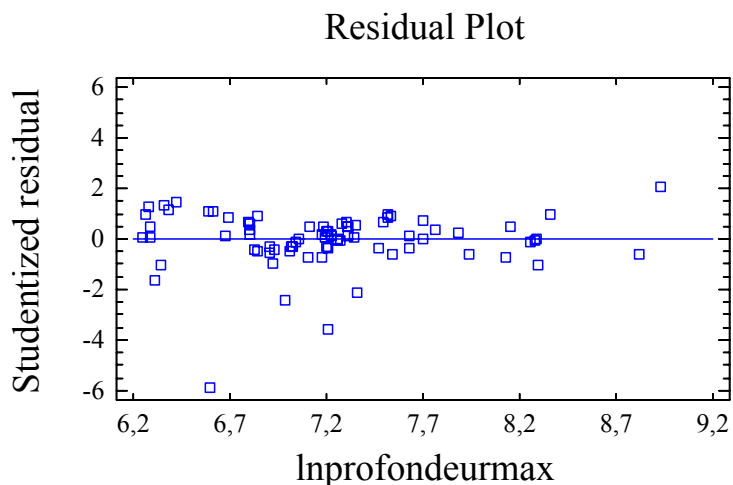
Step 1:
-----
Adding variable lnlargeur with F-to-enter = 701,641
1 variables in the model. 76 d.f. for error.
R-squared = 90,23% Adjusted R-squared = 90,10% MSE = 0,358241

Step 2:
-----
Adding variable lnprofondeurmax with F-to-enter = 6,12917
2 variables in the model. 75 d.f. for error.
R-squared = 90,97% Adjusted R-squared = 90,72% MSE = 0,335592

Step 3:
-----
Adding variable lnrayonhydraulique with F-to-enter = 8,15319
3 variables in the model. 74 d.f. for error.
R-squared = 91,86% Adjusted R-squared = 91,53% MSE = 0,306372
```

Lorsque l'on ajoute la variable profondeur max au modèle débit moyen en fonction du largeur, le  $r^2$  passe de 0.9023 à 0.9097 et le carré moyen des erreurs diminue (il passe de 0.358241 à 0.335592). La profondeur max améliore donc le modèle.

Lorsque l'on ajoute le rayon hydraulique, le  $r^2$  augmente encore ( $r^2=0.9186$ ) et le carré moyen des erreurs diminue, il vaut 0.306372.



Le graphique des résidus studentisés donne 5 résidus non compris entre  $-2$  et  $2$ .  
Le nombre de résidus conduit au rejet du modèle.

Cependant le  $r^2$  vaut  $0.9186$  et le test global est positif, on peut quand même accepter le modèle.

Le débit moyen peut donc s'exprimer en fonction de la profondeur max, de la largeur et du rayon hydraulique.

$$\ln(\text{débit moyen}) = -11.53 + 1.39193 \ln(\text{profondeur max}) + 1.84856 \ln(\text{largeur}) + 0.951228 \ln(\text{rayon hydraulique})$$

Donc

$$\text{débit moyen} = 0.0000098 \text{ profondeur max}^{1.39193} \text{ largeur}^{1.84856} \text{ rayon hydraulique}^{0.951228}$$

## 5. Conclusion

A elle seule la largeur explique près de 90% du débit moyen annuel. De plus, on peut voir que la largeur peut s'exprimer en fonction d'autres paramètres. On trouve :

$$\text{Largeur} = 4.8408 \text{ aire section}^{0.748197} / \text{profondeur moyenne}^{0.641412}$$

$$\text{Larg} = 0.1613 \text{ Peri}^{0.8244}$$

De ce fait, il n'a pas été évident d'intégrer d'autres variables pour modéliser le débit. Cependant, en y ajoutant la profondeur maximale et le rayon hydraulique, on arrive à un  $r^2$  de 0.9186. Le rayon hydraulique prenant en compte la surface mouillée et le périmètre mouillé, le débit moyen annuel est donc expliqué par l'ensemble des variables concernant la section du cours d'eau.

$$\text{débit moyen} = 0.0000098 \text{ profondeur max}^{1.39193} \text{ largeur}^{1.84856} \text{ rayon hydraulique}^{0.951228}$$

L'intégration des variables sinuosité, pente et aire du bassin dans un modèle ont permis d'obtenir les résultats suivant :

$$\text{débit moyen} = e(861.71) \text{ sinuosité}^{0.894447} \text{ rayon hydraulique}^{188.951} / \text{profondeur moyenne}^{186.518}$$

$$\text{Débit moyen} = e(774.272) \text{ profondeur moyenne}^{167.483} \text{ rayon hydraulique}^{169.612} / \text{pente}^{0.2842}$$

$$\text{Débit moyen} = \text{largeur}^{1.83634} / \text{aire bassin}^{0.377122}$$

Cependant, on peut noter que malgré des  $r^2$  significatif, le graphique des résidus contient des valeurs peu habituelles. Ce qui met en doute ces résultats. Enfin, il a été impossible d'intégrer ces trois variables dans un même modèle.

## **CONCLUSION**

Ce stage au LMTG m'a donné la possibilité d'acquérir des connaissances dans le domaine de l'hydrologie et d'appréhender de nouvelles méthodes de travail. J'ai pu aussi développer une spécialité de ma formation : les statistiques.

Une grande partie de ce stage fut consacré à l'extraction des données avec le logiciel Arc View et avec le logiciel Excel avec lequel j'ai pu mettre en application mes connaissances informatiques en créant des macros. Plusieurs résultats ont été dégagés de l'étude statistique:

- Avec les images JERS et le logiciel Arc View, on peut estimer les largeurs du fleuve.
- Le débit annuel moyen peut se simuler avec les paramètres géomorphologiques du fleuve et du bassin. La largeur est le paramètre qui explique le mieux le débit. En y ajoutant les caractéristiques de la section du cours d'eau, on arrive à une bonne estimation du débit annuel moyen.

Il serait maintenant intéressant de voir si ces relations statistiques s'appliquent aux débits mensuels et journaliers.

Ce stage a été très enrichissant car il m'a permis de découvrir l'importance de l'étude de l'eau et du sol.

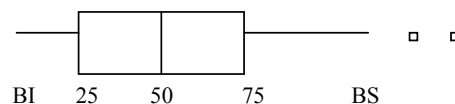


## Lexique

**ACP** : Technique d'analyse statistique, portant sur au moins deux variables numériques, permettant une représentation graphique des liens entre variables numériques et des positions des individus par rapport aux vecteurs de ces variables.

**Actions anthropiques** : Recense toutes les interventions directes ou indirectes des activités humaines: terrassements, remblais, constructions, digues, terril, carrières.

**Boite à moustache** :



BI : borne inférieure = la plus petite valeur plus grande que  $Q1 - 1.5(Q3 - Q1)$

BS : borne supérieure = la plus grande valeur, plus petite que  $Q3 + 1.5(Q3 - Q1)$

$Q1 = 1^{\text{er}}$  quartile

Les valeurs correspondent aux pourcentages d'individu dépassant une certaine valeur, les petits carrés représentent les valeurs extrêmes non comprises entre les bornes inférieure et supérieure.

**Classe modale** : Classe incluant le plus grand nombre d'individus

**Déciles** : représentent les neuf valeurs d'une variable numérique permettant une partition de la population en 10 groupes égaux.

**Coefficient de Pearson** : Indicateur permettant de juger la qualité de la liaison entre deux variables numériques. Une valeur positive (maximum = +1) indique une variation simultanée dans le même sens, une valeur négative (minimum = -1) signale une variation simultanée en sens inverse, une valeur nulle indique l'absence de liens entre les deux variables. Un coefficient de corrélation est caractérisé par son sens et par sa force.

**Ecart type** : indice de dispersion d'une variable numérique.

**Exutoire d'un bassin**: C'est le point de sortie de tous les écoulements pour ce bassin versant.

**IRD** : Institut de Recherche pour le Développement

**JERS** : Japanese Earth Resources Satellite

Jers-1 est un satellite d'observation de la terre pour couvrir le domaine global de terre pour l'enquête nationale de terre, l'agriculture, la sylviculture, et la pêche, la protection de l'environnement, la protection de désastre, et la surveillance côtière, etc. se concentrant sur l'observation autour du monde et l'exploitation de ressource.

**LMTG** : Laboratoire de Mécanisme et Transfert en Géologie

**Matrice des corrélations** : Tableau symétrique présentant les coefficients de corrélations entre les variables numériques analysées.

**Médiane** : La médiane correspond à la valeur d'une variable numérique permettant de diviser la population en deux effectifs égaux.

**Moyenne** : quotient de la somme des valeurs par leur nombre.

**Probabilité critique** : Cette valeur  $\alpha$ , fournie par la plupart des logiciels de tests, comparée au risque  $\alpha$  de première espèce qu'on accepte de prendre permet de conclure : si  $\alpha \leq \alpha$ , on rejette  $H_0$  ; si  $\alpha > \alpha$ , on ne rejette pas  $H_0$ .

**Quartile** : Les quartiles sont les trois valeurs d'une variable numérique divisant la population en 4 groupes égaux.

**R<sup>2</sup> ( ou coefficient de détermination )** : mesure la qualité de l'ajustement des estimations de l'équation de régression. Il est utilisé à la fois en régression simple et en régression multiple. Il permet d'avoir une idée globale de l'ajustement du modèle. Il s'interprète comme la part de la variance de la variable Y expliquée par la régression, varie entre 0 et 1 et s'exprime souvent en pourcentage. En régression simple, un R<sup>2</sup> proche de 1 est suffisant pour dire que l'ajustement est bon.

**Régression linéaire simple** : Équation, ou représentation graphique (la Droite des Moindres Carrés), permettant l'étude de la liaison entre deux variables numériques. L'équation de la droite est de type  $y = ax + b$  avec  $a =$  pente,  $b =$  la valeur de  $y$  quand  $x = 0$ . La pente de la droite ( $a$ ) indique le sens de la liaison entre les deux variables. Le coefficient de détermination  $r^2$  permet d'évaluer la force de la liaison. En cas de liaisons positives ou négatives entre les deux variables, la droite de régression devient un indice pronostique qui permet d'estimer les valeurs de Y en fonction des valeurs de X

**Régression multiple** : est définie comme un outil permettant d'étudier et de mesurer la relation existant entre une variable (Y), dite variable expliquée, et d'autres variables ( $X_i$ ), dites variables explicatives. Effectuer une régression multiple consiste à se baser sur les données d'un échantillon afin de déterminer une estimation de la relation mathématique entre la variable expliquée et les variables explicatives. Les résultats

d'une régression multiple sont donnés d'une part à travers l'équation de régression qui détermine la relation entre la variable expliquée et les variables explicatives, et d'autre part à travers divers coefficients et graphiques qu'il s'agira d'analyser afin :

d'établir le degré de fiabilité de l'estimation

d'examiner la significativité des résultats

d'examiner si la relation entre la variable expliquée et chaque variable explicative est réelle ou seulement apparente

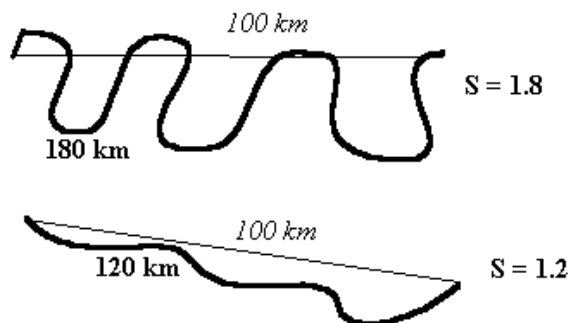
La régression multiple peut faire l'objet de différents usages. Elle peut être utilisée afin de décrire des relations entre les variables explicatives et d'analyser leur action sur la variable expliquée, afin de faire des estimations ou des prédictions, ou bien même afin de pouvoir utiliser l'équation de régression comme un outil de contrôle (si on souhaite par exemple savoir de quelle manière modifier la valeur d'une variable explicative afin d'obtenir une valeur fixée de Y).

**Les résidus :** ou " erreurs observées " sont définis comme étant les différences entre les valeurs observées et les valeurs estimées par un modèle de régression, ils ont la particularité de représenter la partie non expliquée par l'équation de régression. On les note communément comme suit :

$$e_i = Y_i - \hat{Y}_i$$

L'analyse des résidus a pour objectif de tester la validité d'un modèle de régression. Elle permet de déceler les défaillances d'un modèle, c'est pourquoi il est nécessaire de l'effectuer avant toute analyse de régression.

**Sinuosité :** Analyse de la sinuosité des cours d'eau



L'indice de sinuosité est le rapport entre la longueur du talweg et la distance à vol d'oiseau entre ses extrémités. Toutefois cette méthode est peu fiable (dépend de l'échelle à laquelle on effectue la mesure).

**Test non paramétrique :** Se dit de tout test d'hypothèse qui n'est pas un test paramétrique : Cela peut être un test d'adéquation, d'indépendance ou autre. De façon générale, le terme non paramétrique fait référence à des procédures de statistique inférentielle qui ne nécessitent aucune présupposition de forme particulière de la loi de probabilité des observations.

**Test de Kolmogorov :** Une distribution parfaitement normale étant pratiquement introuvable dans la réalité, le test de Kolmogorov permet d'estimer la probabilité pour laquelle une distribution est quasi-normale ou non. En tant qu'épreuves statistiques non paramétriques, ce test sert également à contrôler la normalité de distributions de variables ordinales.  
Hypothèse nulle : la distribution de la variable suit une Loi normale (courbe en forme de cloche).

**Test paramétrique :** C'est un test portant sur la valeur d'un (ou de plusieurs) paramètre(s) d'une population ; ou sur la comparaison des paramètres de deux populations.

**Test de Student :** c'est un test de significativité qui peut être employé lors de la comparaison de deux moyennes (ce test ne peut être utilisé qu'à deux conditions : les distributions des moyennes sont normales, c'est-à-dire décrivent une courbe de Gauss, et leurs variances sont de même taille) pour tester la significativité d'un coefficient de régression  
En ce qui concerne la comparaison de deux moyennes, on se pose la question suivante : la différence entre la moyenne observée et la moyenne de la population est-elle significative ou non ?  
Le test t de Student permet de répondre à ce genre de question. Le test t est calculé en effectuant le rapport de la différence des moyennes sur l'erreur standard et on obtient alors une valeur appelée "Valeur de t" ou t observée.

**Topex :** Le satellite Topex-Poséidon utilise des systèmes avancés d'altimétrie radar qui permettent d'effectuer des mesures précises et répétitives du niveau des océans, afin d'établir des cartes de topographie dynamique. Ce satellite est aussi utilisé pour le bassin de l'Amazone.

**UMR :** Unité mixte de recherche

## **Annexe**

## 1. Tableau d'ajustement des profondeurs

Registre identification	code station	Date	r <sup>2</sup>	sta fisher	fractile	qualité	équation	profondeur maximum estimée
7154	10200000	12/02/95	0,9468	569,50	3,13-3,15	l'ajustement est bon	$y = -0,3387x^2 + 48,877x - 255,17$	1508
7155	10300000	24/06/96	0,9103	360,26	3,11-3,13	l'ajustement est bon	$y = -0,1827x^2 + 36,604x - 274,63$	1559
7156	10500000	09/02/95	0,8566	158,30	3,15-3,18	l'ajustement est bon	$y = -0,0764x^2 + 23,409x + 289,33$	2082
7157	10500000	23/06/96	0,9056	283,00	3,15-3,18	l'ajustement est bon	$y = -0,0753x^2 + 22,866x + 250,22$	1986
7158	10500000	10/09/00	0,8886	307,10	3,11-3,13	l'ajustement est bon	$y = -0,0722x^2 + 21,816x + 306,79$	1955
7159	10800000	17/02/95	0,847	196,53	3,11-3,13	l'ajustement est bon	$y = -0,1669x^2 + 31,194x - 100,12$	1357
7160	10800000	20/06/96	0,8129	123,82	3,15-3,18	l'ajustement est bon	$y = -0,152x^2 + 28,807x - 42,06$	1323
7161	10900000	18/02/95	0,7955	149,76	3,11-3,13	l'ajustement est bon	$y = -0,1526x^2 + 25,931x + 183,3$	1285
7162	10900000	26/06/96	0,7719	103,21	3,13-3,15	l'ajustement est bon	$y = -0,136x^2 + 22,779x + 280,77$	1235
7164	11400000	31/01/95	0,7505	76,70	3,15-3,18	l'ajustement est bon	$y = -0,0037x^2 + 6,6143x + 230,93$	3187
7165	11400000	31/05/96	0,5297	32,10	3,15-3,18	l'ajustement est bon	$y = -0,0032x^2 + 5,5446x + 408,81$	2811
7166	11450000	28/01/95	0,6117	58,29	3,11-3,13	l'ajustement est bon	$y = -0,0043x^2 + 4,0505x + 458,77$	1413
7167	11450000	28/05/96	0,5736	38,34	3,15-3,18	l'ajustement est bon	$y = -0,0066x^2 + 6,1716x + 693,89$	2137
7168	11500000	24/01/95	0,612	37,86	3,18-3,23	l'ajustement est bon	$y = -0,0017x^2 + 3,5063x + 863,21$	2671
7169	11500000	26/05/96	0,3008	18,71	3,1-3,11	l'ajustement est bon	$y = -0,0009x^2 + 1,8045x + 1451,4$	2356
7227	12550000	07/09/00	0,9503	688,35	3,11-3,13	l'ajustement est bon	$y = -0,0595x^2 + 16,875x + 106,79$	1303
7189	12680000	10/09/00	0,8088	211,51	3,09	l'ajustement est bon	$y = -0,1496x^2 + 36,073x - 293,22$	1881
7235	13150000	23/07/96	0,9017	293,53	3,13-3,15	l'ajustement est bon	$y = -0,023x^2 + 24,149x + 556,58$	6895
7239	13180000	29/08/96	0,798	138,27	3,13	l'ajustement est bon	$y = -0,098x^2 + 21,84x + 5,9366$	1223
7244	13410000	30/08/96	0,7858	166,92	3,09-3,1	l'ajustement est bon	$y = -0,0938x^2 + 20,944x + 628,66$	1798
7248	13450000	18/09/00	0,9386	519,75	3,13-3,15	l'ajustement est bon	$y = -0,3771x^2 + 34,239x + 25,657$	803
7249	13450000	07/07/96	0,946	578,11	3,13-3,15	l'ajustement est bon	$y = -0,379x^2 + 35,485x + 66,439$	897
7253	13470000	16/09/00	0,9145	454,58	3,10-3,11	l'ajustement est bon	$y = -0,3844x^2 + 43,101x - 165,1$	1043
7254	13470000	05/07/96	0,9296	481,97	3,10-3,11	l'ajustement est bon	$y = -0,3828x^2 + 43,472x - 181,91$	1052
7258	13550000	15/09/00	0,9102	405,43	3,11	l'ajustement est bon	$y = -0,2884x^2 + 42,196x - 278,96$	1264

7259	13550000	04/07/96	0,9061	381,16	3,11-3,13	l'ajustement est bon	$y = -0,2977x^2 + 40,581x - 162,69$	1220
7270	13600002	12/09/00	0,9627	1251,77	3,09-3,1	l'ajustement est bon	$y = -0,4174x^2 + 55,823x - 204,99$	1661
7271	13600002	09/07/96	0,9605	1118,56	3,09-3,1	l'ajustement est bon	$y = -0,4062x^2 + 54,172x - 174,9$	1631
7274	13650000	10/06/96	0,9193	472,75	3,1-3,11	l'ajustement est bon	$y = -0,1677x^2 + 36,416x - 249,95$	1727
7283	13750000	31/05/96	0,9044	274,35	3,15-3,18	l'ajustement est bon	$y = -0,0594x^2 + 22,052x + 209,41$	2256
7289	13870000	11/09/00	0,8382	279,75	3	l'ajustement est bon	$y = -0,0357x^2 + 17,565x + 221,95$	2383
7294	13886000	13/07/96	0,5727	29,49	3,18-3,23	l'ajustement est bon	$y = -0,0347x^2 + 10,023x + 470,83$	1195
7298	13962000	20/07/96	0,8965	164,57	3,24	l'ajustement est bon	$y = -0,0247x^2 + 18,444x + 418,05$	3861
7409	15030000	29/08/96	0,8563	294,97	3,09-3,1	l'ajustement est bon	$y = -0,0028x^2 + 7,6282x + 1046,7$	6242
7422	15050000	25/07/95	0,9087	109,48	3,44	l'ajustement est bon	$y = -0,8122x^2 + 45,086x - 138,4$	487
7423	15050000	06/08/96	0,9247	171,92	3,34	l'ajustement est bon	$y = -0,8164x^2 + 46,09x - 154,89$	496
7438	15130000	21/10/93	0,7167	51,86	3,18-3,23	l'ajustement est bon	$y = -0,0828x^2 + 15,394x - 215,48$	500
7441	15150000	20/08/93	0,4941	38,58	3,11-3,13	l'ajustement est bon	$y = -0,0256x^2 + 8,9351x + 16,295$	796
7451	15246000	08/02/94	0,9667	464,48	3,29	l'ajustement est bon	$y = -0,5734x^2 + 40,038x - 173,33$	526
7455	15248010	13/09/93	0,8293	153,03	3,13-3,15	l'ajustement est bon	$y = -0,1814x^2 + 19,577x + 6,3991$	535
7459	15250000	12/09/93	0,8217	108,30	3,18-3,23	l'ajustement est bon	$y = -0,0069x^2 + 6,1426x + 203,04$	1570
7472	15400000	26/11/94	0,9408	405,24	3,15-3,18	l'ajustement est bon	$y = -0,0224x^2 + 16,708x + 505,91$	3622
7497	15550000	23/11/94	0,8827	316,06	3,10-3,11	l'ajustement est bon	$y = -0,1231x^2 + 29,605x - 500,01$	1280
7502	15552600	23/10/93	0,8773	150,15	3,18-3,23	l'ajustement est bon	$y = -0,3052x^2 + 26,451x + 45,429$	619
7519	15558000	17/10/93	0,9239	242,81	3,23	l'ajustement est bon	$y = -0,3077x^2 + 30,634x + 11,939$	774
7528	15560000	27/08/93	0,873	240,59	3,13	l'ajustement est bon	$y = -0,0382x^2 + 9,5131x + 234,19$	826
7532	15565000	26/08/93	0,8508	111,20	3,23-3,24	l'ajustement est bon	$y = -0,5803x^2 + 58,49x - 72,941$	1401
7536	15575000	25/08/93	0,8368	99,99	3,23-3,24	l'ajustement est bon	$y = -0,4572x^2 + 32,624x + 110,54$	693
7541	15580000	03/09/93	0,4473	34,40	3,1-3,11	l'ajustement est bon	$y = -0,0306x^2 + 8,8508x + 464,42$	1104
7549	15630000	25/05/96	0,8097	212,74	3,09	l'ajustement est bon	$y = -0,0051x^2 + 6,3096x + 409,67$	2361
7559	15670000	14/08/96	0,9467	817,04	3,09-3,1	l'ajustement est bon	$y = -0,1225x^2 + 43,481x - 491,14$	3367
7563	15700000	11/08/96	0,8942	638,11	3	l'ajustement est bon	$y = -0,0151x^2 + 15,961x + 356,2$	4574
7568	15750000	24/08/95	0,7882	50,24	3,35	l'ajustement est bon	$y = -0,0709x^2 + 13,425x - 59,337$	576
7569	15750000	04/07/96	0,7737	59,83	3,26-3,28	l'ajustement est bon	$y = -0,069x^2 + 13,286x - 55,312$	584
7572	15795000	03/06/96	0,6804	54,29	3,15-3,18	l'ajustement est bon	$y = -0,0563x^2 + 13,686x + 288,34$	1120
7576	15800000	01/06/96	0,8047	98,89	3,18-3,23	l'ajustement est bon	$y = -0,0481x^2 + 15,202x + 253,97$	1455
7587	15860000	07/08/96	0,9108	464,59	3,09-3,1	l'ajustement est bon	$y = -0,0064x^2 + 9,0482x + 411,51$	3610
7590	15903000	10/06/96	0,8163	122,20	3,15-3,18	l'ajustement est bon	$y = -0,6385x^2 + 54,913x + 60,375$	1241

7592	15905000	11/06/96	0,9106	331,03	3,13-3,15	l'ajustement est bon	$y = -0,4994x^2 + 55,227x - 45,721$	1481
7595	15910000	09/06/96	0,8745	202,08	3,15-3,18	l'ajustement est bon	$y = -0,2123x^2 + 31,96x + 184,9$	1388
7598	15930000	08/06/96	0,7687	76,44	3,18-3,23	l'ajustement est bon	$y = -0,2808x^2 + 26,079x + 69,356$	675
7605	17070000	24/06/87	0,8834	189,41	3,18	l'ajustement est bon	$y = -0,402x^2 + 36,802x - 176,4$	666
7613	17091000	29/08/95	0,7472	45,81	3,29-3,32	l'ajustement est bon	$y = -0,3885x^2 + 30,978x + 10,729$	628
7614	17091000	28/06/96	0,6916	37,00	3,28-3,29	l'ajustement est bon	$y = -0,3423x^2 + 34,093x - 245,63$	603
20756336	17091000	20/07/99	0,7069	42,21	3,26-3,28	l'ajustement est bon	$y = -0,2895x^2 + 30,741x - 193,66$	622
7623	17093000	19/08/95	0,7629	40,22	3,39	l'ajustement est bon	$y = -0,0224x^2 + 8,5412x + 134,21$	948
7624	17093000	05/07/96	0,8277	60,05	3,39	l'ajustement est bon	$y = -0,0207x^2 + 8,4646x + 82,642$	948
7628	17095000	26/08/95	0,8912	147,44	3,26	l'ajustement est bon	$y = -0,1054x^2 + 20,91x - 85,482$	952
7633	17120000	19/09/95	0,7646	60,09	3,24-3,26	l'ajustement est bon	$y = -0,0503x^2 + 12,308x + 200,78$	954
7640	17200000	01/09/96	0,917	220,96	3,23	l'ajustement est bon	$y = -0,2418x^2 + 31,38x - 60,095$	958
7652	17280000	22/09/95	0,8154	70,67	3,29	l'ajustement est bon	$y = -0,0869x^2 + 21,728x - 29,292$	1329
7653	17280000	26/06/96	0,8268	78,77	3,28-3,29	l'ajustement est bon	$y = -0,086x^2 + 21,644x - 29,565$	1332
7663	17340000	27/09/95	0,7897	60,08	3,29	l'ajustement est bon	$y = -0,0176x^2 + 7,537x + 186,87$	994
7664	17340000	04/07/96	0,8063	74,93	3,26	l'ajustement est bon	$y = -0,0187x^2 + 7,7433x + 176,13$	978
7678	17380000	28/09/95	0,7998	71,91	3,26	l'ajustement est bon	$y = -0,0296x^2 + 11,85x + 175,19$	1361
7679	17380000	08/07/96	0,7357	51,50	3,24-2,26	l'ajustement est bon	$y = -0,0258x^2 + 10,508x + 239,65$	1310
7685	18121006	29/10/86	0,5324	50,10	3,10-3,11	l'ajustement est bon	$y = -0,0755x^2 + 12,727x + 332,8$	869
7686	18200000	23/10/86	0,917	303,83	3,15-3,18	l'ajustement est bon	$y = -0,504x^2 + 52,264x - 211,21$	1144
7687	18250000	21/06/87	0,8283	144,72	3,15	l'ajustement est bon	$y = -0,2495x^2 + 33,28x + 314,67$	1424
7688	18280000	21/06/87	0,8283	144,72	3,15	l'ajustement est bon	$y = -0,2495x^2 + 33,28x + 314,67$	1424
20817334	18280000	16/10/97	0,8102	140,87	3,13-3,15	l'ajustement est bon	$y = -0,2479x^2 + 30,666x + 356,97$	1305
20817335	18280000	20/09/98	0,8112	137,49	3,13-3,15	l'ajustement est bon	$y = -0,2554x^2 + 31,413x + 425,32$	1391
7689	18300000	02/03/87	0,6441	49,77	3,15-3,18	l'ajustement est bon	$y = -0,0376x^2 + 10,806x + 222,21$	999



## 2. Décile des deux largeurs

Percentiles for largeur\_terrain

1,0% = 153,8  
 5,0% = 179,3  
 10,0% = 192,2  
 25,0% = 239,6  
 50,0% = 320,4  
 75,0% = 877,76  
 90,0% = 1621,7  
 95,0% = 2192,0  
 99,0% = 2763,0

Percentiles for largeur\_jers

1,0% = 131,612  
 5,0% = 137,536  
 10,0% = 137,927  
 25,0% = 232,238  
 50,0% = 367,874  
 75,0% = 994,756  
 90,0% = 1862,99  
 95,0% = 1957,37  
 99,0% = 3015,39

## 3. Test de nullité du coefficient de Pearson

Correlations

	largeur_jers	largeur_terrain
largeur_jers		0,9519 ( 25) 0,0000
largeur_terrain	0,9519 ( 25) 0,0000	

Correlation

(Sample Size)

P-Value

The StatAdvisor

-----  
 This table shows Pearson product moment correlations between each pair of variables. These correlation coefficients range between -1 and +1 and measure the strength of the linear relationship between the variables. Also shown in parentheses is the number of pairs of data values used to compute each coefficient. The third number in each location of the table is a P-value which tests the statistical significance of the estimated correlations. P-values below 0.05 indicate statistically significant non-zero correlations at the 95% confidence level. The following pairs of variables have P-values below 0.05:

largeur\_jers and largeur\_terrain

#### 4. Test d'adéquation d'appartenance à une famille de loi

Estimated Kolmogorov statistic DPLUS = 0,152713  
Estimated Kolmogorov statistic DMINUS = 0,121397  
Estimated overall statistic DN = 0,152713  
Approximate P-Value = 0,604383

EDF Statistic	Value	Modified Form	P-Value
Kolmogorov-Smirnov D	0,152713	0,788001	>=0.10*
Anderson-Darling A^2	1,10334	1,14041	0,0056*

\*Indicates that the P-Value has been compared to tables of critical values specially constructed for fitting the currently selected distribution. Other P-values are based on general tables and may be very conservative.

#### 5. Test d'égalité des moyennes

t-test  
-----

Null hypothesis: mean = 0,0  
Alternative: not equal

Computed t statistic = 0,636274  
P-Value = 0,530619

Do not reject the null hypothesis for alpha = 0,05.

## 6. Test d'égalité des médianes

sign test

-----

Null hypothesis: median = 0,0

Alternative: not equal

P-Value = 0,689153

Do not reject the null hypothesis for alpha = 0,05.

signed rank test

-----

Null hypothesis: median = 0,0

Alternative: not equal

P-Value = 0,999994

Do not reject the null hypothesis for alpha = 0,05.

The StatAdvisor

-----

This pane displays the results of three tests concerning the center of the population from which the sample of `largeur_jers-largeur_terrain` comes. The first test is a t-test of the null hypothesis that the mean `largeur_jers-largeur_terrain` equals 0,0 versus the alternative hypothesis that the mean `largeur_jers-largeur_terrain` is not equal to 0,0. Since the P-value for this test is greater than or equal to 0,05, we cannot reject the null hypothesis at the 95,0% confidence level. The second test is a sign test of the null hypothesis that the median `largeur_jers-largeur_terrain` equals 0,0 versus the alternative hypothesis that the median `largeur_jers-largeur_terrain` is not equal to 0,0. It is based on counting the number of values above and below the hypothesized median. Since the P-value for this test is greater than or equal to 0,05, we cannot reject the null hypothesis at the 95,0% confidence level. The third test is a signed rank test of the null hypothesis that the median `largeur_jers-largeur_terrain` equals 0,0 versus the alternative hypothesis that the median `largeur_jers-largeur_terrain` is not equal to 0,0. It is based on comparing the average ranks of values above and below the hypothesized median. Since the P-value for this test is greater than or equal to 0,05, we cannot reject the null hypothesis at the 95,0% confidence level. The sign and signed rank tests are less sensitive to the presence of outliers but are somewhat less powerful than the t-test if the data all come from a single normal distribution.

## 7. Listing des résultats

COORDONNEES, CONTRIBUTIONS DES INDIVIDUS  
INDIVIDUS ACTIFS (AXES 1 A 5)

IDENTIFIC	INDIVIDUS		COORDONNEES					CONTRIBUTIONS				
	P.REL	DISTO	1	2	3	4	5	1	2	3	4	5
10200000	1.32	7.53	-0.42	-1.26	-0.28	0.36	2.28	0.0	1.3	0.1	0.2	11.4
10300000	1.32	2.44	-0.40	0.39	-1.27	-0.45	0.28	0.0	0.1	2.0	0.3	0.2
10500000	1.32	9.93	0.99	-1.87	0.55	-0.93	-2.00	0.2	2.7	0.4	1.4	8.8
10500000	1.32	8.02	0.33	-1.52	0.84	-1.13	-1.76	0.0	1.8	0.9	2.1	6.8
10500000	1.32	9.39	1.01	-1.76	0.72	-0.94	-1.93	0.2	2.4	0.7	1.4	8.1
10800000	1.32	2.11	-0.69	0.56	-1.12	0.07	0.08	0.1	0.2	1.6	0.0	0.0
10800000	1.32	2.05	-0.52	0.46	-1.18	0.13	0.01	0.0	0.2	1.8	0.0	0.0
10900000	1.32	1.68	-0.68	0.40	-0.85	0.11	0.30	0.1	0.1	0.9	0.0	0.2
10900000	1.32	1.99	-0.91	0.46	-0.85	0.05	0.35	0.1	0.2	0.9	0.0	0.3
11400000	1.32	63.21	7.54	0.91	-0.55	0.33	-0.09	9.4	0.6	0.4	0.2	0.0
11400000	1.32	59.36	7.38	1.30	-0.11	0.20	0.05	9.0	1.3	0.0	0.1	0.0
11450000	1.32	3.91	1.32	0.48	-0.02	-0.23	1.12	0.3	0.2	0.0	0.1	2.7
11450000	1.32	16.22	3.55	-0.62	-0.94	0.45	0.42	2.1	0.3	1.1	0.3	0.4
11500000	1.32	60.99	6.74	2.70	1.26	-0.39	0.72	7.5	5.7	2.0	0.2	1.1
11500000	1.32	152.22	10.84	3.95	3.26	-0.04	0.64	19.4	12.2	13.4	0.0	0.9
12550000	1.32	1.07	-0.24	-0.46	-0.14	-0.43	-0.07	0.0	0.2	0.0	0.3	0.0
12680000	1.32	1.78	-0.19	0.52	-0.72	-0.06	0.06	0.0	0.2	0.6	0.0	0.0
13180000	1.32	2.36	-0.96	0.39	-0.61	-0.63	0.63	0.2	0.1	0.5	0.6	0.9
13410000	1.32	1.89	0.50	-0.09	-0.96	-0.14	-0.06	0.0	0.0	1.2	0.0	0.0
13450000	1.32	4.55	-2.03	0.38	0.06	-0.39	0.04	0.7	0.1	0.0	0.2	0.0
13450000	1.32	3.60	-1.81	0.27	-0.04	-0.33	-0.04	0.5	0.1	0.0	0.2	0.0
13470000	1.32	12.91	-2.07	1.47	-0.12	2.40	-0.71	0.7	1.7	0.0	9.3	1.1
13470000	1.32	12.18	-1.90	1.36	-0.22	2.46	-0.77	0.6	1.5	0.1	9.7	1.3
13550000	1.32	1.91	-0.85	0.28	-0.83	-0.29	0.11	0.1	0.1	0.9	0.1	0.0
13550000	1.32	2.38	-1.13	0.41	-0.74	-0.38	0.20	0.2	0.1	0.7	0.2	0.1
13600002	1.32	2.00	0.17	-0.60	-0.84	0.43	0.24	0.0	0.3	0.9	0.3	0.1
13600002	1.32	1.68	0.06	-0.56	-0.81	0.39	0.26	0.0	0.2	0.8	0.2	0.1
13650000	1.32	2.75	0.47	0.06	-1.35	0.48	-0.19	0.0	0.0	2.3	0.4	0.1
13750000	1.32	4.08	1.26	-0.22	-1.39	-0.16	0.05	0.3	0.0	2.4	0.0	0.0
13870000	1.32	8.41	2.29	0.12	-1.25	-0.08	-0.15	0.9	0.0	2.0	0.0	0.0
13886000	1.32	2.37	-0.49	0.46	-1.09	-0.65	0.42	0.0	0.2	1.5	0.7	0.4
13962000	1.32	45.70	5.65	-1.34	-2.69	0.64	-1.17	5.3	1.4	9.1	0.7	3.0
15050000	1.32	14.69	-3.27	1.07	0.86	1.34	-0.40	1.8	0.9	0.9	2.9	0.4
15050000	1.32	14.51	-3.25	1.06	0.86	1.34	-0.41	1.7	0.9	0.9	2.9	0.4
15130000	1.32	12.97	-2.79	0.39	0.46	-0.58	2.10	1.3	0.1	0.3	0.5	9.7
15150000	1.32	3.10	-1.15	0.92	-0.41	-0.39	0.39	0.2	0.7	0.2	0.2	0.3
15246000	1.32	21.96	-3.41	-0.38	2.19	1.61	1.51	1.9	0.1	6.1	4.2	5.0
15248010	1.32	17.15	-3.21	1.72	0.57	1.80	-0.48	1.7	2.3	0.4	5.2	0.5
15250000	1.32	2.61	0.78	0.56	-0.03	-0.52	0.65	0.1	0.3	0.0	0.4	0.9
15400000	1.32	27.47	3.47	-1.41	-0.32	0.93	2.16	2.0	1.6	0.1	1.4	10.2
15550000	1.32	1.07	-0.15	-0.20	-0.48	-0.59	-0.05	0.0	0.0	0.3	0.6	0.0
15552600	1.32	6.92	-2.40	0.45	0.09	-0.80	0.43	1.0	0.2	0.0	1.0	0.4
15558000	1.32	5.81	-2.07	0.27	-0.17	-0.84	0.77	0.7	0.1	0.0	1.1	1.3
15560000	1.32	3.61	-1.47	0.29	0.20	-1.07	0.10	0.4	0.1	0.1	1.8	0.0
15565000	1.32	2.95	0.00	-1.07	-0.73	-0.48	-0.43	0.0	0.9	0.7	0.4	0.4
15575000	1.32	11.10	-2.77	0.55	1.03	1.20	-0.76	1.3	0.2	1.3	2.3	1.3
15580000	1.32	1.15	-0.26	0.23	-0.41	-0.60	-0.12	0.0	0.0	0.2	0.6	0.0
15630000	1.32	34.66	5.28	0.58	1.85	-0.51	-0.33	4.6	0.3	4.3	0.4	0.2
15670000	1.32	78.26	6.66	-4.84	-1.40	2.09	-0.05	7.3	18.4	2.5	7.0	0.0
15750000	1.32	8.86	-2.71	0.63	0.56	-0.87	0.04	1.2	0.3	0.4	1.2	0.0
15750000	1.32	8.61	-2.67	0.64	0.58	-0.86	0.03	1.2	0.3	0.4	1.2	0.0
15795000	1.32	19.82	-1.67	1.27	0.39	3.59	-1.46	0.5	1.3	0.2	20.8	4.7
15800000	1.32	0.81	-0.52	0.16	-0.31	-0.16	-0.31	0.0	0.0	0.1	0.0	0.2
15860000	1.32	52.39	6.56	0.49	1.13	-0.53	-1.35	7.1	0.2	1.6	0.4	4.0
15905000	1.32	2.37	-0.59	-0.79	-0.16	-0.90	-0.69	0.1	0.5	0.0	1.3	1.0
15910000	1.32	1.55	-0.83	-0.49	-0.13	-0.69	-0.30	0.1	0.2	0.0	0.8	0.2
15930000	1.32	11.63	-2.68	-0.37	1.70	-0.56	-1.02	1.2	0.1	3.6	0.5	2.3
17070000	1.32	5.60	-2.11	0.03	0.44	-0.77	-0.41	0.7	0.0	0.2	1.0	0.4
17093000	1.32	3.20	-1.57	0.39	0.35	-0.37	-0.16	0.4	0.1	0.2	0.2	0.1
17093000	1.32	3.33	-1.61	0.42	0.38	-0.38	-0.14	0.4	0.1	0.2	0.2	0.0
17095000	1.32	3.48	-1.66	-0.18	0.40	-0.60	0.23	0.5	0.0	0.2	0.6	0.1
17120000	1.32	3.58	-1.41	0.60	-0.64	-0.39	0.71	0.3	0.3	0.5	0.3	1.1
17200000	1.32	4.34	-1.98	0.41	0.17	0.37	-0.19	0.6	0.1	0.0	0.2	0.1
17280000	1.32	1.50	-0.65	0.18	-0.84	0.11	0.29	0.1	0.0	0.9	0.0	0.2
17280000	1.32	1.59	-0.76	0.24	-0.80	0.08	0.32	0.1	0.0	0.8	0.0	0.2
17340000	1.32	2.50	-1.18	0.28	0.11	-0.84	-0.15	0.2	0.1	0.0	1.1	0.1
17340000	1.32	2.84	-1.29	0.37	0.21	-0.88	-0.11	0.3	0.1	0.1	1.3	0.0

17380000	1.32	2.09	-0.65	-0.27	0.29	-0.82	-0.68	0.1	0.1	0.1	1.1	1.0
17380000	1.32	1.90	-0.58	-0.26	0.31	-0.81	-0.69	0.1	0.1	0.1	1.0	1.0
18121006	1.32	4.46	-1.62	0.15	0.36	-1.10	-0.43	0.4	0.0	0.2	2.0	0.4
18200000	1.32	2.98	-1.42	-0.59	0.60	0.31	-0.32	0.3	0.3	0.5	0.2	0.2
18250000	1.32	2.85	0.11	-0.19	-1.41	0.13	0.02	0.0	0.0	2.5	0.0	0.0
18280000	1.32	20.31	0.16	-3.81	2.19	0.55	0.66	0.0	11.4	6.0	0.5	1.0
18280000	1.32	19.62	-0.25	-3.61	2.36	0.43	0.79	0.0	10.2	7.0	0.3	1.4
18280000	1.32	20.04	0.03	-3.75	2.23	0.52	0.70	0.0	11.0	6.3	0.4	1.1
18300000	1.32	3.09	-1.24	0.84	-0.39	0.68	0.21	0.3	0.6	0.2	0.7	0.1

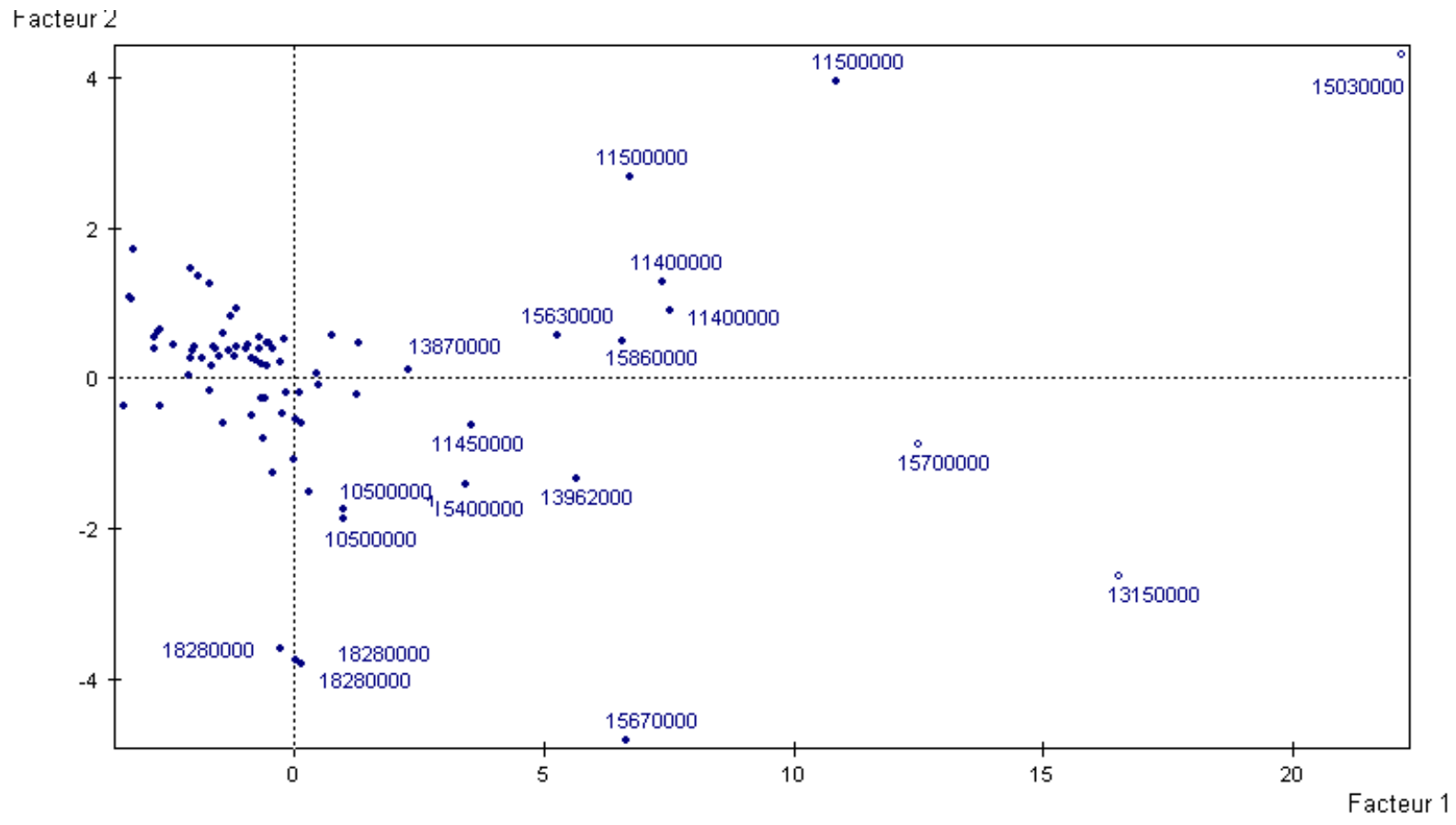
INDIVIDUS ILLUSTRATIFS (AXES 1 A 5)

13150000	1.32	325.27	16.52	-2.62	-3.42	2.85	-1.90
15030000	1.32	562.13	22.15	4.30	3.86	1.52	-0.42
15700000	1.32	202.36	12.51	-0.87	-0.39	1.64	-1.73

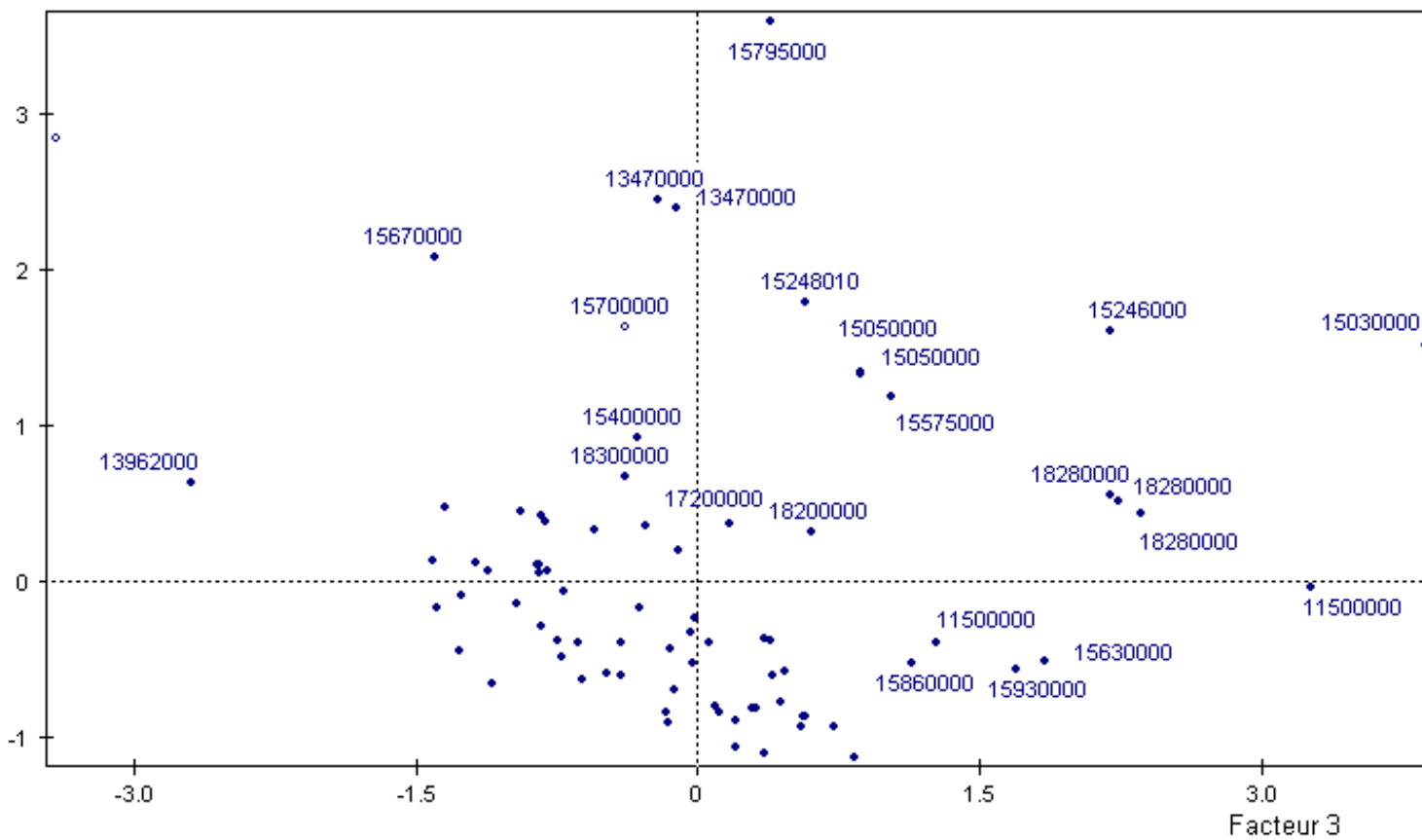
## 8. Nuage des individus

Voir page suivante.





Facteur 4





## 9. Régression multiple

Multiple Regression Analysis

Dependent variable: lndebitmoyen

Parameter	Estimate	Standard Error	T Statistic	P-Value
lnlargeur	1,84351	0,0827875	22,2679	0,0000
lnsinuosité	0,42422	0,449171	0,944452	0,3490
lnairebassin	-0,39411	0,0532778	-7,39726	0,0000

### Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	2921,96	3	973,986	1664,56	0,0000
Residual	32,7674	56	0,585131		
Total	2954,73	59			

R-squared = 98,891 percent

R-squared (adjusted for d.f.) = 98,8514 percent

Standard Error of Est. = 0,764939

Mean absolute error = 0,533939

Durbin-Watson statistic = 1,93325

Lag 1 residual autocorrelation = 0,00141494

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between lndebitmoyen and 3 independent variables. The equation of the fitted model is

$$\text{lndebitmoyen} = 1,84351 \cdot \text{lnlargeur} + 0,42422 \cdot \text{lnsinuosité} - 0,39411 \cdot \text{lnairebassin}$$

Since the P-value in the ANOVA table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains 98,891% of the variability in lndebitmoyen. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 98,8514%. (Note: since the model does not contain a constant, you should be careful in interpreting the R-Squared values. Do not compare these R-Squared values with those of models which do contain a constant.) The standard error of the estimate shows the standard deviation of the residuals to be 0,764939. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 0,533939 is the average value of the residuals. The Durbin-Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file.

Multiple Regression Analysis

-----  
 Dependent variable: lndebitmoyen  
 -----

Parameter	Estimate	Standard Error	T Statistic	P-Value
lnlargeur	1,83634	0,0834203	22,0131	0,0000
lnairebassin	-0,377122	0,0523685	-7,20131	0,0000

-----  
 Analysis of Variance  
 -----

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	2976,51	2	1488,25	2497,61	0,0000
Residual	34,5605	58	0,59587		
Total	3011,07	60			

-----  
 R-squared = 98,8522 percent  
 R-squared (adjusted for d.f.) = 98,8324 percent  
 Standard Error of Est. = 0,771926  
 Mean absolute error = 0,532893  
 Durbin-Watson statistic = 2,08148  
 Lag 1 residual autocorrelation = -0,0705735  
 -----

The StatAdvisor

-----  
 The output shows the results of fitting a multiple linear regression model to describe the relationship between lndebitmoyen and 2 independent variables. The equation of the fitted model is

$$\text{lndebitmoyen} = 1,83634 * \text{lnlargeur} - 0,377122 * \text{lnairebassin}$$

Since the P-value in the ANOVA table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains 98,8522% of the variability in lndebitmoyen. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 98,8324%. (Note: since the model does not contain a constant, you should be careful in interpreting the R-Squared values. Do not compare these R-Squared values with those of models which do contain a constant.) The standard error of the estimate shows the standard deviation of the residuals to be 0,771926. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 0,532893 is the average value of the residuals. The Durbin-Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0,0000, belonging to lnairebassin. Since the P-value is less than 0.01, the highest order term is statistically significant at the 99% confidence level. Consequently, you probably don't want to remove any variables from the model.

## **Bibliographie**

JL Guyot : HYDROGEOCHIMIE DES FLEUVES DE L'AMAZONIE BOLIVIENNE  
1993, p19 à 37

Arc View GIS :The geographic Information System for everyone 1996

B. LE Guennec : cours HYDRAULIQUE 1994 (Ecole nationale d'électrotechnique,  
d'électronique, d'informatique et d'hydraulique de Toulouse)

Tables Statistiques : 1995 p 20 à 27

F Seyler, F. Muller, G Cochonneau, JL Guyot : Résolution et précision des outils numériques  
( MNT et image satellite) nécessaires à la délimitation de bassins versants en région  
tropicale 2003