

Université Pierre et Marie Curie

**Master 2 Sciences de l'Univers, Environnement, Ecologie
Parcours Hydrologie-Hydrogéologie**

**Caractérisation des erreurs, et tentative d'amélioration des
performances d'un modèle de réseau hydrographique**

Axel Patindé Belemtougri

Directrice de recherche : Agnès Ducharne

Co-encadrante : Marie Silvestre



**UMR Sisyphe 7619
Université P. et M. Curie
Paris, France**

Septembre 2017



Abstract

The precise characterization of watercourses is a major challenge to better understand the environment and participate in the good management of water resources. This characterization is increasingly done using DEM (Digital Elevation Model). This intership is based on the study of (Schneider et al. 2017) which proposes using a 15 " DEM from the HydroSHEDS database and environmental parameters (Litholgy, Climate, Slope), a series of values of critical areas (minimum contributory surface to initiate a stream) in order to better characterize the rivers. These critical areas are derived from a calibration in France using Carthage database and in Australia using the Australian Hydrological Geospatial Fabric Database (AHGF), but the global hydrographic network model derived from these critical areas (LCS) shows some errors. By focusing on France, our analysis shows that the artificialisation and intermittency of the hydrographic networks accounts for 40% of the variance of the errors of the model and that a large part of the errors relates on the choice of reference layer for calibration of the LCS model in France. The recalibration of the model taking into account the layer "Tronçons Hydrographiques" which is better adapted, allowed us to have better performances.

Key words : Hydrography, water resources, lithology, climate, topography, DEM

Remerciements

Je ne saurais commencer ce rapport sans adresser mes sincères remerciements à mon encadrante Agnès DUCHARNE pour sa disponibilité, les conseils prodigués tout au long de ce stage, ses remarques pertinentes ainsi que la confiance qu'elle a placée en moi en m'attribuant ce stage.

Je tiens particulièrement à remercier ma co-encadrante Marie Silvestre qui n'a ménagé aucun effort pour répondre à toutes mes interrogations dans divers domaines durant ce stage notamment en SIG et en traitement des données.

Mes remerciements s'adressent également à Ana Schneider qui a su prendre de son temps pour me donner de plus amples détails sur son article à Ardalan Tootchi et à Ludovic Oudin pour m'avoir fourni des données climatiques.

Je remercie toutes les personnes du bureau qui m'ont apporté de l'aide pour l'analyse des données avec le logiciel R.

Je ne saurais oublier de remercier tout le personnel du laboratoire Metis pour leur accueil et enthousiasme.

Enfin, merci aux professeurs et aux étudiants du M2HH de l'université Pierre et Marie Curie.

Merci également à tous ceux qui de près ou de loin ont contribué à la réussite de ce stage.

Tables des matières

1	Introduction	7
2	Synthèse Bibliographique	8
2.1	Réseaux LC et LCS (Schneider et al. 2017)	8
2.1.1	Description	8
2.1.2	Principes et Performances de la calibration	8
2.2	Artificialisation et intermittence des réseaux hydrographiques	11
2.3	Bilan de la revue bibliographique et problématique	12
3	Matériels et Méthodes	13
3.1	Caractéristiques des données et outils	13
3.2	Calculs et interpolations à la résolution de 7.5'	17
3.3	Critères de qualité	18
3.4	Procédures de recalibration	19
4	Résultats et interprétations	22
4.1	Analyse des erreurs du modèle de réseaux hydrographiques LCS	22
4.1.1	Comparaisons de réseaux de référence de la BD Carthage	22
4.1.2	Caractérisation des erreurs du modèle LCS	23
4.1.3	Facteurs explicatifs potentiel des erreurs	24
4.1.4	Régression linéaire multivariée	26
4.2	Recalibrage du modèle LCS	28
4.2.1	Couche de lithologie	28
4.2.2	Relation entre les Aires critiques et les densités de drainage	30
4.2.3	Relations entre les densités de drainage et les paramètres environnementaux	36
4.2.4	Extraction des réseaux hydrographiques	41
5	Conclusion et Perspectives	45
6	Annexes	46
7	Références Bibliographiques	48
8	Résumé	50

Liste des figures

Figure 1 : Paramètres statistiques de comparaison entre les densités de drainage calculées et celle de référence dans plusieurs pays (Schneider et al. 2017).....	10
Figure 2 : Densité de drainage en France (km/km ²) : a) Réseau de référence ; b) LCS (Schneider et al. 2017).	10
Figure 3 : Zoom sur des réseaux artificiels présents dans le réseau hydrographique français dans le département des Landes (Schneider et al. 2017).....	10
Figure 4 : Comparaison des densités de drainage des couches cours d'eau et tronçons hydrographiques de la BD Carthage pour un maillage de 7,5'	22
Figure 5 : Différence de densité de drainage entre le réseau LCS et les réseaux de références de la France pour un maillage de 7,5'	23
Figure 6 : Comparaison entre la densité des tronçons hydrographiques (intermittents) et la différence (LCS-TH_total) au niveau de la France pour un maillage de 7.5'	24
Figure 7 : Comparaison entre la densité de drainage des tronçons hydrographiques (artificiels) et la différence (LCS-TH_total) au niveau de la France pour un maillage de 7.5'	25
Figure 8 : Taux de superficie irriguée et drainée par maille au niveau de la France pour un maillage de 7.5'	26
Figure 9 : Représentation des classes strictement majoritaires de lithologies et des classes majoritaires à 70% au sein des mailles de 7,5'	29
Figure 10 : Répartition de la densité de drainage des tronçons hydrographiques totaux en fonction des lithologies majoritaires à 70%. La largeur des classes est proportionnelle à la racine carrée du nombre de mailles associés.....	29
Figure 11 : Régression linéaire des valeurs d'aires critiques (A _{cr} totales) les plus proches des réseaux en fonction de leur densité de drainage (km/km ²).....	30
Figure 12 : Régression linéaire des valeurs d'aires critiques (A _{cr} <4 km ²) les plus proches des réseaux en fonction de leur densité de drainage (km/km ²).....	31
Figure 13: Régression logarithmique des valeurs d'aires critiques (A _{cr} totales) les plus proches des réseaux en fonction de leur densité de drainage (km/km ²)	32
Figure 14 : Régression logarithmique des valeurs d'aires critiques (A _{cr} <4km ²) les plus proches des réseaux en fonction de leur densité de drainage (km/km ²)	32
Figure 15 : Régression logarithmique des valeurs d'aires critiques (A _{cr} totales) les plus proches des réseaux en fonction de leur densité de drainage (km/km ²)	34
Figure 16 : Régression logarithmique des valeurs d'aires critiques (A _{cr} < 4 km ²) les plus proches des réseaux en fonction de leur densité de drainage (km/km ²)	35
Figure 17 : Régression linéaire entre les densités de drainage des réseaux (km/km ²) et certains paramètres environnementaux au niveau de la classe (Mt)	37
Figure 18 : Régression linéaire entre les densités de drainage des réseaux (km/km ²) et certains paramètres environnementaux au niveau de la classe (Sm)	38
Figure 19 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques total simulé 1 avec le réseau original.	40
Figure 20 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques naturel simulé 1 avec le réseau original.....	40
Figure 21 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques naturel intermittent simulé 1 avec le réseau original.....	40
Figure 22 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques artificiel simulé 1 avec le réseau original.....	41

Figure 23 : Représentation des réseaux Thnat_interm simulés 2 et Thnat_interm observés	41
Figure 24 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques naturel simulé 2 avec le réseau original.	42
Figure 25 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques total simulé 2 avec le réseau original.	42
Figure 26 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques naturel intermittent simulé 2 avec le réseau original.	42
Figure 27 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques artificiel simulé 2 avec le réseau original.	43

Liste des tableaux

Tableau 1 : Récapitulatif des données utilisées.....	13
Tableau 2 : Les Différents types de réseaux issus de la couche « tronçons hydrographiques » pris en compte.....	15
Tableau 3 : Récapitulatif des régressions effectuées.....	27
Tableau 4 : Nombre de mailles et pourcentages occupé par les classes de lithologies, en France	28
Tableau 5 : Liste des lithologies retenues.....	30
Tableau 6 : Récapitulatif des coefficients de détermination obtenus pour chaque équation et en fonction des différents réseaux (A_{cr} totales).....	36
Tableau 7 : Récapitulatif des coefficients de détermination obtenus pour chaque équation et en fonction des différents réseaux ($A_{cr} < 4\text{km}^2$).....	36
Tableau 8 : Récapitulatif des régressions linéaires multiples entre les réseaux Th_naturel et Th_total et les paramètres en fonction des différentes classes de lithologies	39
Tableau 9 : Analyses statistiques entre les réseaux simulés et les réseaux observés	43
Tableau 11 : Performances des réseaux LCS ,Th_total_sim 1, Th_total_sim 2.....	44

Table des annexes

Annexe 1 : Résultats de la calibration A_{cr} (km^2) définissant le modèle LCS en France et en Australie pour tous les paramètres environnementaux sélectionnés (Schneider et al., 2017). -----	46
Annexe 2 : Carte à l'échelle globale de la densité de drainage issue du modèle LCS (Schneider et al. 2017).-----	46
Annexe 3 : Récapitulatif des régressions linéaires multiples entre les réseaux Thnat_interm et Th_artificiel et les paramètres en fonction des différentes classes de lithologies -----	47

1 Introduction

Les volumes d'eau présents à tout instant dans toutes les rivières sont une faible portion de la ressource en eau existante sur terre 0,00012% (De Marsily 2009). L'importance des cours d'eau pour l'homme en termes d'utilisation pour l'alimentation en eau potable et de la biodiversité est grande. Cela explique pourquoi protéger les cours d'eau, la qualité de l'eau et la biodiversité sont vital pour l'homme. Le bon état écologique des rivières et lacs est l'objectif visé par les politiques sur l'eau. Au sein de l'Union Européenne, la directive cadre sur l'eau (DCE) fixe des objectifs et des méthodes pour atteindre ce bon état des eaux (Barraqué 2001). Mais que savons-nous exactement des cours d'eau qui nous entourent ? L'étude de (Bishop et al. 2008) montre que de nombreux petits cours d'eau qu'il dénomme « Aqua Incognita » n'apparaissent pas dans les bases de données cartographiques notamment dans plein de régions du globe ou seules les cours d'eau principaux sont représentés. Cartographier et caractériser de façon précise le réseau hydrographique c'est contribuer à une meilleure connaissance du milieu et à une bonne gestion de la ressource en eau.

Pour cela, l'utilisation des modèles numériques de terrain (MNT) à haute résolution permet de dériver un réseau hydrographique cohérent qui traduit la réalité du terrain. Depuis quelques années, des bases de données MNT à plus fines résolutions et globales issues de satellites ou de radars sont disponibles, notamment HydroSHEDS¹ (Lehner, Verdin, and Jarvis 2008) qui propose des MNT à des résolutions de 3'',15'',30'' et qui sont convenable pour des applications hydrologiques. Les outils de systèmes d'information géographiques (SIG) et les logiciels de programmation permettent par la suite d'extraire rapidement les réseaux d'écoulement. Cette étape passe généralement par la définition d'une aire critique (A_{cr}), qui représente la surface contributive minimum pour initier un cours d'eau, et qui est souvent définie assez arbitrairement et de manière unique sur de grands domaines, par manque d'information, ce qui fait que les cours d'eau générés n'ont souvent pas la même longueur que ceux sur le terrain (Passalacqua, Tarolli, and Foufoula-Georgiou 2010).

C'est dans cette optique que l'étude de (Schneider et al. 2017) propose un jeu de plusieurs valeurs d' A_{cr} en fonction de la lithologie, du climat et de la pente LCS (Lithology, Climate and Slope). Elle propose aussi une méthode d'identification des rivières intermittentes afin de mieux caractériser les cours d'eau. Bien que cette étude constitue une avancée majeure pour décrire les petits cours d'eau dans des applications globales, ou dans des régions où les données hydrographiques manquent, il subsiste de nombreuses erreurs, dont certaines sont probablement liées à la non prise en compte d'informations sur l'artificialisation présente dans les réseaux hydrographiques. La structure de réseaux artificiels est en effet très variable dans l'espace et difficile à caractériser (Cazorzi et al. 2013).

Dans cette étude, nous nous focaliserons sur la France continentale et nous tenterons en premier lieu de caractériser finement les erreurs du modèle de réseau hydrographique LCS de (Schneider et al. 2017). Ensuite, nous étudierons les corrélations entre ces erreurs et plusieurs facteurs tels que l'artificialisation et l'intermittence des réseaux. En définitive nous tenterons de corriger ces erreurs en recalibrant le réseau LCS.

¹ Données disponibles sur www.hydrosheds.org

2 Synthèse Bibliographique

Cette partie s'appuie sur les travaux menés par (Schneider et al. 2017) pour l'extraction du réseau LC et LCS et donne des détails sur l'artificialisation et l'intermittence des réseaux hydrographiques.

2.1 Réseaux LC et LCS (Schneider et al. 2017)

2.1.1 Description

La caractérisation précise des cours d'eau est essentielle pour de nombreuses applications notamment en hydrologie pour la simulation des écoulements (Pella et al. 2012). A l'aide d'un MNT, cette caractérisation passe par le choix d'une aire critique (A_{cr}) qui représente la surface contributive minimum pour initier un cours d'eau. Dans le milieu naturel, l'initiation des rivières dépend du climat, de la lithologie, de la pente du terrain, du couvert végétal et donc de paramètres environnementaux (Vogt, Colombo, and Bertolo 2003), faisant de la densité de drainage un paramètre qui varie continuellement dans l'espace. En utilisant un MNT à 15'' (environ 500 m à l'équateur) issue de la base de données HydroSHEDS, Schneider et al.(2017) propose une calibration de la variabilité spatiale de l'aire critique A_{cr} en fonction de paramètres environnementaux tels que la lithologie, le climat et la pente. L'étude propose deux modèles de calibration de l'aire critique : LC (Lithology and Climate) qui prend en compte la lithologie et le climat ensuite LCS (Lithology Climate and Slope) qui tient compte de la lithologie, du climat et de la pente.

2.1.2 Principes et Performances de la calibration

La calibration s'est focalisée sur la France et l'Australie et les réseaux pris comme référence sont respectivement ceux de la BD Carthage² et de Geofabric³, tous deux produits à l'échelle de 1 :50000. Le traitement des données est effectué à l'aide des outils de SIG. En utilisant un MNT de 15'' issue de la base de données HydroSHEDS, des réseaux hydrographiques sont extraits avec des aires critiques allant de 0.25 à 200 km². Pour les différents réseaux, la densité de drainage est calculée au sein des mailles de 7.5' (environ 12 km à l'équateur) en faisant la somme des longueurs des cours d'eau dans la maille divisée par sa surface.

Pour tenir compte de l'influence de la lithologie, les auteurs utilisent la base de données de (Hartmann and Moosdorf 2012) à l'échelle moyenne de 1 :3750000, avec 12 classes. Celle du climat provient de WFDEI⁴ (Weedon et al. 2014). Elle contient des données de précipitation moyenne interannuelle de (1980 à 2009) à la résolution de 0,5 degrés réparties en 5 classes. La pente extraite du MNT est regroupée en 5 classes. Une intersection spatiale de ces paramètres environnementaux donne 300 zones uniformes pour la calibration du modèle LCS. Dans chacune de ces zones, une moyenne de la densité de drainage est calculée pour chacun des réseaux extraits ainsi que pour le réseau de

2 Base de Données sur la CARTographie THématique des AGences de l'eau

3 Australian Hydrological Geospatial Fabric

4 Watch Forcing Data methodology applied to ERA-Interim

référence. La valeur d' A_{cr} choisie correspond à celle minimisant la différence de densité de drainage entre les réseaux extraits et le réseau de référence. La même procédure est utilisée pour la calibration du modèle LC mais en ne tenant compte que de la Lithologie et du Climat. Cette opération donne ainsi des valeurs calibrées d' A_{cr} sur le territoire Français et Australien. Des regroupements sont faits au sein des classes de lithologie, climat et de pente afin d'obtenir des valeurs uniques par classe pour les deux pays. Les valeurs d' A_{cr} ainsi calibrées pour le modèle LCS sont récapitulées en **Annexe 1**. Ces valeurs sont par la suite utilisées pour extraire les réseaux hydrographiques globaux et donc la densité de drainage en tenant compte des paramètres environnementaux **Annexe 2**.

L'étude montre que ces deux modèles (LC et LCS) présentent des performances satisfaisantes comparés au réseau hydrographique de référence en France, Australie, USA, et Brésil (**Figure 1**). En France par exemple, la moyenne de la densité de drainage du modèle LCS ($0,71 \text{ km/km}^2$) est proche de celle du réseau de référence ($0,75 \text{ km/km}^2$) avec un biais de $-0,04 \text{ km/km}^2$. Celle des autres réseaux dont le modèle LC et le réseau de la FAO, sont respectivement de $0,68$ et $0,08 \text{ km/km}^2$. Le réseau de la FAO (Food and Agriculture Organisation) est issu de la base de données AQUAMAPS⁵ (FAO 2014). Les densités de drainage du réseau de la FAO sont en général faibles car les aires critiques utilisées sont trop grandes. Les coefficients de corrélation par rapport au réseau de référence sont de l'ordre de $0,32$ pour le modèle LCS, $0,30$ pour le modèle LC et $0,01$ pour les réseaux de la FAO. Cela montre que parmi les modèles étudiés, le modèle LCS est celui qui obtient les meilleures performances. Mais de nombreuses erreurs subsistent à quelques endroits, ou le modèle LCS sous-estime fortement les densités drainage (**Figure 2**). C'est le cas par exemple au niveau du département des Landes en France où on remarque une forte présence de réseaux artificiels. La (**Figure 3**) illustre la présence de cette artificialisation dans le réseau hydrographique. Sur cette figure on remarque des patterns circulaires et très rectilignes qui ne sont pas caractéristiques des réseaux naturels et qui sont probablement des réseaux de drainage et d'irrigation.

L'étude de (Schneider et al. 2017) propose aussi une méthode de premier ordre pour l'identification des réseaux intermittents. Cette méthode considère les rivières qui sont localisées sous un certain seuil de précipitation comme intermittentes. Celles qui sont localisées au-delà de ce seuil sont classées comme pérennes. A cet effet (Raymond et al. 2013) montrent que la largeur des cours d'eau et les précipitations sont très corrélées. Le calibrage est fait en Australie à partir des réseaux de la FAO (AQUAMAPS) dans lequel 69% des rivières de l'Australie sont intermittentes. Le seuil de précipitation retenu est celui dont les isohyètes chevauchent le mieux les limites des réseaux intermittents de la FAO. Le seuil calibré est de 500 mm/an . Avec ce seuil, 42% des réseaux LCS en Australie sont intermittents. Ce taux est plus faible par rapport au taux de la FAO (69%) et s'explique par le fait que LCS a plus de réseaux que la FAO.

5 Base de données de la FAO sur l'eau et l'agriculture

Region	Parameter	FAO	LC	LCS	Ref.	Others
AU	Mean	0.08	0.42	0.45	0.43	-
	SD	0.07	0.32	0.32	0.43	-
	Bias	-0.36	-0.01	0.02	0.00	-
	Corr.	-0.01	0.57	0.59	1.00	-
	RMSE	0.56	0.36	0.34	0.00	-
FR	Mean	0.08	0.68	0.71	0.75	0.34
	SD	0.08	0.24	0.30	0.36	0.18
	Bias	-0.66	-0.07	-0.04	0.00	-0.41
	Corr.	0.01	0.30	0.32	1.00	0.19
	RMSE	0.81	0.40	0.39	0.00	0.56
US	Mean	0.07	0.72	0.78	0.75	1.30
	SD	0.07	0.28	0.30	0.37	0.84
	Bias	-0.68	-0.03	0.03	0.00	0.55
	Corr.	0.09	0.09	0.11	1.00	0.24
	RMSE	0.77	0.47	0.47	0.00	1.00
BR	Mean	0.08	0.93	0.92	0.52	-
	SD	0.07	0.16	0.16	0.29	-
	Bias	-0.44	0.41	0.40	0.00	-
	Corr.	0.12	0.14	0.14	1.00	-
	RMSE	0.49	0.49	0.48	0.00	-
Europe	Mean	0.08	0.59	0.63	-	0.31
	SD	0.08	0.29	0.32	-	0.15
Global	Mean	0.08	0.69	0.74	-	-
	SD	0.08	0.32	0.34	-	-
% Intermittent		34	27	29	-	-

^aRef. = δ from river network reference data (AU = Geofabric and FR = CARTHAGE used for calibration and US = NHD and BR = IBGE [2015] used for validation); Others = δ from Luo *et al.* [2016] over the U.S. and from CCM2 [Vogt *et al.*, 2007] over France and Europe; Mean = mean δ (km^{-1}); SD = standard deviation of δ (km^{-1}); Bias = bias between model and Ref. (km^{-1}); Corr. = correlation coefficient between model and Ref. (dimensionless); RMSE = root-mean-square error between model and Ref. (km^{-1}). % Intermittent corresponds to the ratio of total length of intermittent rivers to the total river length.

Figure 1 : Paramètres statistiques de comparaison entre les densités de drainage calculées et celle de référence dans plusieurs pays (Schneider et al. 2017).

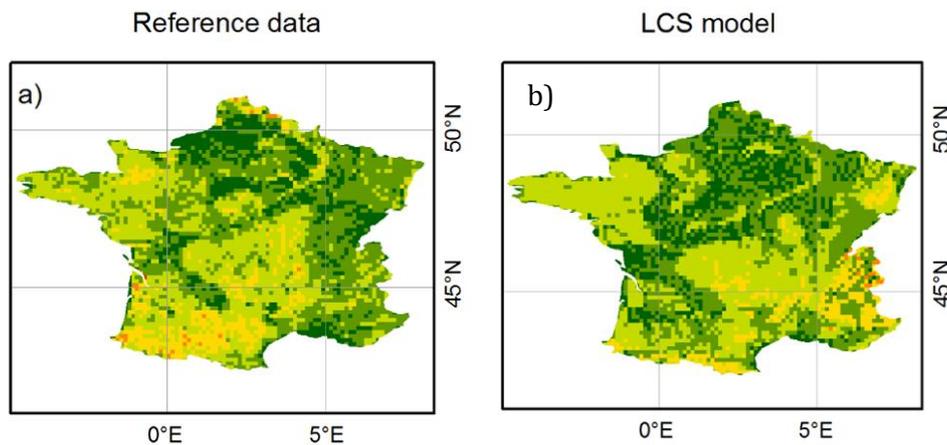


Figure 2 : Densité de drainage en France (km/km^2) : a) Réseau de référence ; b) LCS (Schneider et al. 2017).

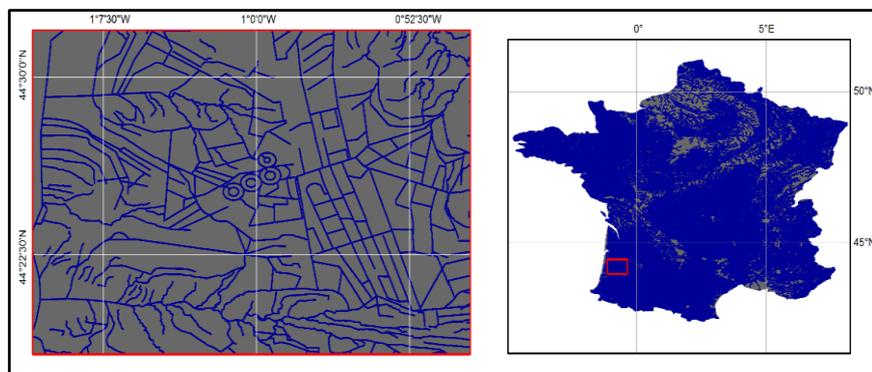


Figure 3 : Zoom sur des réseaux artificiels présents dans le réseau hydrographique français dans le département des Landes (Schneider et al. 2017).

2.2 Artificialisation et intermittence des réseaux hydrographiques

L'artificialisation des réseaux hydrographiques correspond principalement à la présence de tronçons artificiels en plus des tronçons naturels. A cet effet (Dunn and Mackay 1996) affirment qu'en plus de l'existence de réseaux naturels, au Royaume-Uni un pourcentage très élevé de la superficie des terres est drainé par des réseaux artificiels. Le drainage est une opération qui consiste à favoriser l'évacuation de l'eau gravitaire de la zone racinaire vers un cours d'eau à la suite d'une précipitation. Ce drainage se fait sous de nombreuses formes différentes, des drains de sous-sols ou des fossés ouverts. (Lagacherie et al. 2006) indique que les réseaux artificiels sont des composantes importantes des paysages agricoles, en particulier dans les régions françaises intensivement cultivées où elles sont traditionnellement utilisées pour drainer les excès d'eau. Les travaux réalisés pour le drainage des parcelles agricoles et l'irrigation concernent un linéaire hydrographique long. Pour (Herzon and Helenius 2008), le drainage dans l'agriculture vise à l'amélioration générale de l'utilisation des terres agricoles à travers l'apport d'eau pour l'irrigation, la lutte contre les inondations par l'élimination des excès d'eau des terres, où la lutte contre l'érosion. La lutte contre les inondations des terres agricoles à travers l'évacuation des excès d'eau dans des canaux de drainage et la création de nouveaux canaux pour l'irrigation sont parmi les principaux éléments qui contribuent à l'artificialisation des réseaux hydrographiques.

L'urbanisation a aussi un impact sur l'altération des réseaux hydrographiques. Elle est le deuxième facteur d'altération des réseaux hydrographiques après l'agriculture. La densité de drainage des réseaux naturel en milieu urbain décroît drastiquement car les réseaux sont comblés ou placés sous canaux. Par contre la densité de drainage des réseaux artificiels augmente dû à la présence de nombreux canaux d'évacuation des eaux (Paul and Meyer 2001), elle influe aussi sur les débits d'écoulement (Salavati et al. 2016).

Partout dans le monde, on trouve des rivières qui a un certain moment de l'année sont à sec et ces rivières intermittentes sont parfois plus nombreuses que les rivières pérennes. Cette intermittence est souvent due à une exploitation excessive des cours d'eau ou à la variabilité saisonnière de la pluviométrie. Nous utilisons dans cette étude le terme de rivières intermittentes pour désigner tous les cours d'eau temporaires, éphémères, saisonniers et épisodiques. Snelder et al. (2013) cité dans Datry, Larned, and Tockner (2014) indique que 25 à 40% de la longueur totale des rivières en France est intermittente. La prise en compte de l'artificialisation et de l'intermittence dans cette étude pourrait nous apporter des éléments de réponse sur la structure des réseaux étudiés.

2.3 Bilan de la revue bibliographique et problématique

(Schneider et al. 2017) proposent une méthode d'identification des réseaux intermittents et une série de valeur d'Aires critique pour l'extraction d'un réseau hydrographique global cohérent à l'échelle 1:50000, en fonction de paramètres environnementaux. Cette calibration s'est faite en se focalisant principalement sur les réseaux de référence de la France et de l'Australie qui ont été jugés de bonne qualité. Cette calibration apporte des résultats probants car elle permet d'obtenir des densités de drainage extrait qui sont proches des densités de drainage de référence. Dans la suite de l'étude nous considérerons uniquement le modèle LCS qui montre plus de performance que le modèle LC. Il est cependant loin d'être parfait et montre des densités de drainage très inférieures ou supérieures aux réseaux de référence dans de nombreuses régions du globe, dont la France. Ces erreurs sont peut-être dues à la présence de réseaux artificiels dans les réseaux de référence. Les études montrent que l'hydraulique agricole serait une des sources de cette artificialisation des cours d'eau (Cacas et al. 1986). Cette artificialisation se manifeste à travers la création de drains agricole pour l'évacuation des excès d'eau ou la mise sous canaux de cours d'eau pour l'irrigation des terres. L'intermittence est aussi une source d'erreur potentielle en rendant plus difficile l'identification des cours d'eau.

Dans la suite de ce mémoire, nous essaierons d'expliquer les principales erreurs du modèle LCS en France afin de proposer un modèle plus réaliste. La partie 3 présente les données utilisées, et les méthodes. La partie 4 présente les différents résultats issus de l'étude sur les erreurs du modèle LCS ainsi que les différents facteurs pouvant expliquer ces erreurs. Une recalibration est effectuée en vue d'avoir une meilleure représentation du réseau.

3 Matériels et Méthodes

3.1 Caractéristiques des données et outils

Les bases de données utilisées lors de cette étude sont récapitulées dans le **Tableau 1** et sont relatives au réseau hydrographique de référence en France métropolitaine, aux données climatiques, au découpage administratif du territoire, au drainage, à l'irrigation agricoles et à la lithologie. Les différentes données exploitées dans ces bases de données sont détaillées ci-dessous.

Tableau 1 : Récapitulatif des données utilisées

Base de données	Description	Période	Autres
BD Carthage (Hydrographie)	La BD CARTHAGE (Base de Données sur la CARTographie THématique des AGences de l'eau et du ministère chargé de l'environnement) est la base de données complète du réseau hydrographique français. Elle est une œuvre de collaboration entre l'IGN, le ministère chargé de l'environnement et les agences de l'eau. Avec le découpage du territoire en zones hydrographiques d'une part et la codification de ces zones et du réseau hydrographique d'autre part, la BD CARTHAGE constitue un référentiel hydrographique, couvrant l'ensemble du territoire métropolitain.	2014	Echelle 1 :50000 Version 3.0
Source	http://www.sandre.eaufrance.fr/atlas/srv/fre/catalog.search#/metadata/3409c9c3-9836-43be-bac3-b110c82b3a25		
GEOFLA (Limites Administratives)	Produite par l'IGN, le jeu de donnée GEOFLA décrit l'ensemble des unités administratives nationales (France métropolitaine et DROM) en régions, départements, communes, arrondissements, et cantons.	2016	Echelle 1 : 1 000 000 Version 2.2
Source	http://professionnels.ign.fr/geofla		
Agrete (Drainage et Irrigation)	Agrete est une plateforme en ligne présentant les données issues des enquêtes et recensements agricoles français réalisées par le Service de la Statistique et de la Prospective (SSP) du Ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt.	2010	Départements et cantons
Source	http://agreste.agriculture.gouv.fr/recensement-agricole-2010/les-cartes-en-ligne-727/		
MESAN (Précipitation et Température)	MESAN (MESoscale ANalysis system) est un système opérationnel d'analyse de données météorologiques. Il fournit des données maillées notamment utilisées en entrée des modèles. Ces données à l'échelle de l'Europe sont entre autres la précipitation, la température, le rayonnement solaire.	1979 – 2014	Résolution 22 km

Global Lithological Map (Lithologie)	GLiM (Global Lithological Map database) est une carte lithologique globale haute résolution qui a été assemblée à partir de cartes géologiques régionales existantes traduites en informations lithologiques à l'aide de la littérature régionale (Hartmann and Moosdorf 2012).	2012	Version 1.1 Echelle 1 : 3750000 Résolution 0.5°
--------------------------------------	---	------	---

❖ BD Carthage

La BD Carthage contient 11 grands thèmes (couches) dont seules les couches « Cours d'eau » et « Tronçons hydrographiques » sont prises en compte dans le présent rapport. La BD Carthage est un référentiel spatial de l'hydrographie en France et comme le décrit (Pella, Sauquet, and Chandesris 2006) ce référentiel géographique identifie sans ambiguïté chaque cours d'eau et chaque plan d'eau sur le territoire. Les réseaux des couches « Cours d'eau » et « Tronçons hydrographiques » sont donc des « réseaux de référence ». Ils se déclinent de la manière suivante selon le rapport descriptif de la BD Carthage (IGN 2006):

- **Couche « Cours d'eau » :** Un cours d'eau est une portion connexe du réseau hydrographique lié à un toponyme, possédant une source ou origine et un confluent ou embouchure. Un cours d'eau est composé de tronçons hydrographiques élémentaires.
- **Couche « Tronçons hydrographiques » :** Il fait référence aux tronçons hydrographiques élémentaires, aux cours d'eau, et aux relations de superposition entre tronçons.

En France continentale, le linéaire hydrographique que nous avons calculé pour les couches « Cours d'eau » et « Tronçons hydrographiques » est respectivement de 417918 km et 493567 km. Chaque cours d'eau peut être défini par un ou plusieurs tronçons élémentaires. Cependant, certains tronçons présents dans la couche « Tronçons hydrographiques » ne sont pas pris en compte dans la couche « Cours d'eau », qui peut donc être par endroit "plus riche" que la couche « Cours d'eau ». Le rapport descriptif de la BD Carthage (IGN 2006) souligne ce fait en disant que la géométrie des « Cours d'eau » n'est pas forcément complète. La couche « Cours d'eau » ne contient que les éléments de la couche « Tronçons hydrographiques » qui disposent d'un code hydrographique. Or cette codification se fait localement et se sont les agences de l'eau qui collectent et contrôlent les demandes d'évolution de la BD Carthage. Les principaux cours d'eau sont codés et donc regroupés dans la couche « cours d'eau » et le reste des tronçons est ajouté en fonction des besoins (Pella, Sauquet, and Chandesris 2006).

Il est important de noter que la calibration du réseau LCS (Schneider et al. 2017) en France s'est faite à l'aide de la couche « Cours d'eau ». Pour notre étude, nous nous servons au contraire de la couche « Tronçons hydrographique » qui contient plus de linéaire hydrographique et qui est plus réaliste selon l'analyse ci-dessus. De plus, les éléments de la couche « Tronçons hydrographiques » sont caractérisés par des attributs qui permettent d'identifier les réseaux artificiels et intermittents ce qui n'est pas le cas de la couche « Cours d'eau ».

Ainsi, les réseaux artificiels et naturels présent dans la couche « Tronçons hydrographiques » peuvent être identifiés par le champ « Nature » qui marque le caractère naturel ou artificiel du cours d'eau. Il contient les éléments suivants :

- **Aqueduc, conduite forcée** : tuyau ou chenal artificiel conçu pour le transport de l'eau
- **Canal, chenal** : voie d'eau artificielle
- **Cours d'eau naturel** : linéaire d'écoulement naturel
- **En attente de mise à jour** : linéaire en cours d'actualisation
- **Estuaire** : écoulement d'un cours d'eau dans la zone d'estran
- **Sans objet** : linéaire dont l'état est inconnu ou fictif

Pour l'étude, les réseaux pris en compte comme artificiels sont les éléments « Aqueduc, conduite forcée » et « Canal, chenal », car les autres éléments potentiellement artificiels sont inconnus ou en cours d'actualisation.

Le champ « Etat » présent dans la couche « Tronçons hydrographiques » permet d'identifier les différents types d'écoulement et est composé des éléments suivants :

- **En attente de mise à jour**
- **Inconnu** : L'existence d'un écoulement est certaine, mais le tracé n'est pas connu avec précision
- **Permanent** : Écoulement permanent
- **Intermittent** : Écoulement intermittent
- **Axe fictif** : Tronçons crée pour assurer la continuité des cours d'eau à la traversée des éléments surfaciques (Eau douce permanente ou Eau salée permanente) ou lorsque le tracé n'est pas connu avec précision
- **A sec** : Canal abandonné, à sec

Les différents types de réseau issus de la couche « Tronçons hydrographiques » pris en compte dans cette étude sont consignés dans le **Tableau 2** suivant, avec leur abréviation.

Tableau 2 : Les Différents types de réseaux issus de la couche « tronçons hydrographiques » pris en compte

Noms	Abréviations
Tronçons hydrographiques complets	Th_total
Tronçons hydrographiques naturels	Th_nat
Tronçons hydrographiques artificiels	Th_art
Tronçons hydrographiques permanents	Th_perm
Tronçons hydrographiques intermittents	Th_interm
Tronçons hydrographiques naturels et permanents	Thnat_perm
Tronçons hydrographiques naturels et intermittents	Thnat_interm
Tronçons hydrographiques artificiels et permanents	Thart_perm
Tronçons hydrographiques artificiels et intermittents	Thart_interm

❖ Irrigation et drainage

Les données de drainage et d'irrigation utilisées sont disponibles sur la plateforme Agreste et issues du recensement agricole 2010. Elles se répartissent comme suit :

- **Part de la superficie irriguée dans la superficie agricole utilisée (SAU) :** Les superficies irriguées comprennent les superficies irriguées au moins une fois au cours de la campagne agricole, quel que soit le mode d'irrigation.
- **Part de la superficie drainée dans la superficie agricole utilisée (SAU) :** il s'agit ici de la superficie drainée ou assainie par un réseau de drains enterrés. Ne sont pas prises en compte les superficies dont les drains n'évacuent plus, ni celles qui ne sont drainées que ponctuellement (captage de mouillères), à l'aide d'un seul drain et non d'un réseau

❖ Données climatiques

La précipitation et la température moyenne journalière sont issues des données MESAN disponibles sur l'Europe à la résolution de 22 km sur la période de 1979–2014 (Dahlgren et al. 2016) . La précipitation moyenne annuelle (mm/an) et la température moyenne annuelle (°C) sont calculées en faisant la moyenne pluriannuelle des précipitations et des températures. Pour chacun des mois de l'année des moyennes pluriannuelles de précipitations (mm/j) et de températures (°C) sont aussi calculées. La formule de calcul de l'ETP (ÉvapoTranspiration Potentielle) de (Oudin, Michel, and Anctil 2005) prenant en compte la température moyenne journalière et le rayonnement extraterrestre est utilisée pour calculer les valeurs d'ETP journalières. Les moyennes pluriannuelles d'ETP sont calculées pour chacun des 12 mois de l'année et pour l'année complète.

La pluie efficace moyenne annuelle ainsi que la pluie efficace moyenne mensuelle pour chacun des 12 mois de l'année sont obtenues en faisant la différence entre les valeurs de précipitations et d'ETP correspondantes.

Enfin, la Dry Season Length (DSL) quantifie le nombre de mois dans l'année où la pluie efficace est inférieure à 0.

❖ Outils

Le traitement de ces données a été effectué grâce au logiciel Arcgis. Arcgis est un logiciel de SIG (Système d'information Géographique) qui permet de collecter, organiser, gérer, et analyser des informations géographiques. Il permet de créer des cartes non seulement pour afficher des informations, mais aussi pour trouver et comprendre les tendances et les relations, réaliser des analyses et une modélisation pour résoudre des problèmes spécifiques, visualiser et suivre l'état, permettre la saisie et la compilation de données (Esri 2017).

Les analyses statistiques sont faites avec le logiciel R qui est un logiciel libre de statistique. Il est à la fois un langage informatique et un environnement de travail : les commandes sont exécutées grâce à des instructions codées dans un langage relativement simple. Il sert à manipuler des données, à tracer des graphiques et à faire des analyses statistiques sur ces données (Micheaux, Drouilhet, and Liquet 2011).

3.2 Calculs et interpolations à la résolution de 7.5'

Le territoire Français a été divisé en des mailles de 7,5' afin de mieux caractériser les hétérogénéités spatiales des valeurs de paramètres utilisés. Chaque maille de 7,5' contient une et une seule valeur. C'est sur la base de ce maillage que s'effectueront les différents calculs.

❖ Densité de drainage

La densité de drainage peut être définie selon (Horton 1945) comme la longueur totale du réseau hydrographique dans un bassin versant divisé par sa surface.

$$Dd = \frac{\sum L}{S}$$

avec Dd la densité de drainage (km/km²), L la longueur totale du réseau hydrographique (km) et S la surface du bassin versant (km²).

Couramment les densités de drainage se calculent en tenant compte de la surface des bassins versant mais l'idée ici n'est pas de comparer des bassins versant mais de quantifier de façon plus fine les variations spatiales de densité de drainage. A cet effet (Pella, Chandesris, and Wasson 2004) stipulent que le calcul de la densité de drainage par maille permet d'éditer une carte rendant l'information plus directement lisible. Ainsi, les réseaux sélectionnés au sein des différentes couches sont intersectés avec le maillage de 7,5' de la France. La densité de drainage est ensuite calculée automatiquement en faisant la somme des longueurs des réseaux au sein d'une maille divisée par la superficie de la maille.

❖ Taux d'irrigation et de drainage

Les parts de superficies irriguées et drainées dans la SAU (Superficie Agricole Utilisable) à l'échelle du canton sont multipliées par la superficie de la SAU ce qui nous donne les superficies irriguées et drainées à l'échelle du canton. Ces valeurs sont intersectées avec le maillage de 7,5' de la France. Il arrive dans ces cas que certaines mailles recoupent plusieurs cantons et qu'il y ait plusieurs valeurs de taux. Dans ces cas, la valeur de la maille est obtenue en faisant la moyenne pondérée des taux en fonction de leurs superficies.

❖ Paramètres

Les valeurs des paramètres issues des données MESAN à la résolution de 22 km tels que la précipitation moyenne annuelle, la pluie efficace moyenne annuelle, l'ETP moyenne annuelle, la précipitation, la pluie efficace, et l'ETP moyenne mensuelle pour chacun des 12 mois de l'année, et la DSL sont interpolés sur le maillage de 7.5'.

3.3 Critères de qualité

Pour évaluer les erreurs du modèle LCS et du nouveau modèle issu de ce stage, nous avons utilisé plusieurs indicateurs d'écart statistique dont notamment la RMSE et l'erreur absolue. La relation entre ces erreurs et certains facteurs explicatifs a été évaluée à l'aide des coefficients de corrélations.

❖ La RMSE (Root Mean Square Error)

Elle est la racine carrée de l'erreur quadratique moyenne et est fréquemment utilisée pour mesurer la différence entre les valeurs prédites par un modèle et les valeurs réelles observées. Ces différences individuelles sont également appelées résidus, et la RMSE sert à les agréger en une seule mesure de pouvoir prédictif. Les valeurs de RMSE sont comprises entre 0 et l'infini, avec 0 comme meilleur score de qualité. Plus cette valeur de RMSE est petite plus les valeurs prédites s'approchent des valeurs réelles. Dans notre cas la RMSE mesure la différence entre le réseau simulé et le réseau de référence et se formule comme suit :

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{sim,i} - X_{reference,i})^2}{n}}$$

avec $X_{sim,i}$: les densités de drainage du réseau simulé ; $X_{reference,i}$: les densités de drainage du réseau de référence et n : le nombre de mailles

❖ L'erreur absolue

Cette erreur, aussi appelée biais, correspond à la différence entre une valeur calculée et une valeur théorique et dans notre cas :

$$Erreur\ absolue\ (Biais) = (X_{sim,i} - X_{reference,i})$$

Des erreurs absolues positives traduisent une surestimation du réseau simulé par rapport aux réseaux de référence et inversement.

❖ Coefficient de corrélation de Pearson et coefficient de détermination

Le coefficient de corrélation de Pearson r mesure l'intensité et le sens de la liaison linéaire entre deux variables. Elle nous permettra d'étudier la corrélation entre les distributions spatiales de deux variables sur le maillage. La valeur du coefficient varie entre -1 et 1. Un r proche de 0 traduit a priori une absence de relation entre les variables tandis qu'un r proche de 1 en valeur absolue traduit une forte relation linéaire. Dans notre cas, un coefficient r proche de zéro en valeur absolue entre le réseau simulé et les réseaux de référence indique une erreur de distribution spatiale du réseau simulé.

Le test de significativité comme son nom l'indique nous permet de tester la significativité du coefficient de corrélation r . On considèrera dans la suite de l'étude un coefficient significatif si la p-value est inférieure à 0,05.

Le coefficient de détermination R^2 est une mesure de la qualité de la prédiction d'une régression linéaire. Il quantifie la part de variance d'une variable expliquée par une ou des variables explicatives. Dans le cas d'une régression linéaire simple (à une variable explicative), le coefficient de détermination est égal au carré du coefficient de corrélation. Dans notre étude, il donne une indication de la part de variance de l'erreur du modèle expliquée par les facteurs potentiels. Il nous donne aussi une mesure de qualité de relation entre les aires critiques calibrées et les densités de drainage.

3.4 Procédures de recalibration

La recalibration vise à améliorer le modèle LCS de sorte que les réseaux « simulés » soient proches des réseaux observés. Cette recalibration se décline de la manière suivante.

Etape1 : Définir une relation entre les densités de drainage observées (Dd_{obs}) par maille et les Aires critiques (A_{cr}). Cette relation est de type $A_{cr} = f(Dd_{obs})$.

Pour cela 9 types réseaux observés, issus de la couche « Tronçons hydrographiques », sont étudiés à savoir : Th_total, Th_nat, Th_art, Th_perm, Thnat_perm, Thart_perm, Th_interm, Thnat_interm, Thart_interm. Pour chacun d'entre eux :

1. Plusieurs réseaux hydrographiques ont été extraits en utilisant un MNT de 15'' issues de la base de données HydroSHEDS avec des aires critiques définies par un nombre de pixels n_{cr} allant de (1 à 1000), ce qui correspond à des Aires critiques A_{cr} de (0.215 km² à 215km² à l'équateur). Au total 29 réseaux sont extraits.
2. Pour chaque pixel de chacun des réseaux extraits précédemment, la densité de drainage est calculée en faisant la somme totale de la longueur de réseau dans une maille de 7.5' divisée par sa surface.
3. À l'aide de la couche de lithologie de (Hartmann and Moosdorf 2012) nous avons identifiées toutes les mailles dont la lithologie dominante couvre au moins 70 % de la surface. C'est sur ce maillage correspondant aux lithologies supérieures à 70% que la calibration se fera, pour que l'influence de la lithologie soit plus claire, en éliminant les mailles où plusieurs lithologies sont présentes en proportions équivalentes (ex 40%,30%,30%).
4. Pour chaque pixel retenu, la valeur d' A_{cr} sélectionnée est définie par le réseau extrait (parmi les 29 possibilités) qui approche le mieux la densité observée. Ainsi l' A_{cr} du pixel sélectionné est celle minimisant dans la maille la différence de densité de drainage entre les réseaux extraits et les réseaux sélectionnés.
5. Les A_{cr} ci-dessus sont exprimées en nombre de pixel (n_{cr}) et transformées en km² en utilisant la formule $A_{cr} = n_{cr} * A_p$ avec A_{cr} : Aires critiques (km²) ; n_{cr} : Nombre de pixel amont drainé ; A_p : Aire d'un pixel (km²). L'aire d'un pixel de 15'' varie en fonction de la latitude et se calcule selon la formule $A_p \cong 0.215 * \cos \lambda$ avec λ : la latitude, ou 0.215 km² est la surface d'un pixel de 15'' à l'équateur.

6. Des régressions linéaires simples, exponentielles ou logarithmiques sont par la suite réalisées afin d'établir les différentes relations entre les A_{cr} de chacun des 9 types de réseau et les densités de drainage correspondantes. Une seule équation est par la suite choisie pour résumer les différentes relations.

Etape 2 : Définir des relations entre les densités de drainage observées et les facteurs explicatifs comme la pluie, la pente, et ce sur chacune des couches de lithologie majoritaire à 70 %. L'objectif est d'établir une équation du type $Dd_{obs} = g(\text{facteurs})$. Cette étape est réalisée indépendamment pour chacun des 9 réseaux observés.

1. Pour chacune des 7 classes de lithologie, des régressions linéaires simples sont effectuées entre la densité de drainage et les paramètres environnementaux tels que la pluie efficace, l'Etp, la précipitation, la pente, la Dry season length, le taux superficie irrigué par maille, le taux de superficie drainée par maille. Le but est de sélectionner les paramètres ayant un pouvoir explicatif.
2. Une régression linéaire multiple est par la suite réalisée entre les Dd_{obs} et les facteurs ayant un pouvoir explicatif ce qui nous donne pour chaque type de réseau 7 équations pour chacun des 9 réseaux observés.
3. Ces équations constituent un modèle permettant d'approximer les densités de drainages observées. Une comparaison est faite entre les densités simulées et celles observées.

Etape 3 : Définir des relations entre les aires critiques et les paramètres environnementaux du type $A_{cr} = f(g(\text{facteurs}))$.

1. En tenant compte des deux premières étapes, les valeurs des aires critiques (km^2) sont calculées en transformant les valeurs des densités de drainages simulées à l'aide des équations $A_{cr} = f(Dd_{obs})$ et $Dd_{obs} = g(\text{facteurs})$. On obtient une valeur d' A_{cr} par maille de 7.5 ' (en prenant en compte la lithologie dominante).
2. Les aires critiques A_{cr} en km^2 sont transformées en nombre de pixels en vue de l'extraction d'un réseau sous Arcgis. Ceci est obtenu en divisant A_{cr} par la superficie des pixels du MNT (15 "), en prenant soin de les arrondir à l'entier le plus proche $n_{cr} = \frac{A_{cr}}{A_p}$ avec A_{cr} : Aires critiques (km^2) ; n_{cr} : Nombre de pixel amont drainé ; A_p : Aire d'un pixel de 15 " (km^2), définie en fonction de la latitude comme à l'étape 1.
3. On obtient donc une carte de n_{cr} à la résolution de 7.5' qui est utilisée pour extraire un réseau hydrographique à partir du MNT à 15 ".
4. Un script écrit par (Schneider et al. 2017) est utilisé par la suite pour s'assurer que les réseaux générés en amont existent en aval, en ignorant les interruptions du réseau qui peuvent se produire si n_{cr} est plus grand à l'aval qu'à l'amont.

5. Les rivières sont par la suite intersectées avec le maillage de 7.5' et des nouvelles densités de drainage sont calculées.
6. Une deuxième évaluation est faite en comparant les densités de drainages simulées et les densités de drainage observées à l'aide de critère de qualité comme la RMSE et le biais.

4 Résultats et interprétations

4.1 Analyse des erreurs du modèle de réseaux hydrographiques LCS

4.1.1 Comparaisons de réseaux de référence de la BD Carthage

La spatialisation de la densité de drainage des couches « Cours d'eau » et « Tronçons Hydrographiques » de la BD Carthage (**Figure 4a et 4b**) montre que dans la partie Nord, Nord-Ouest et Centre de la France, les densités de drainage sont presque similaires.

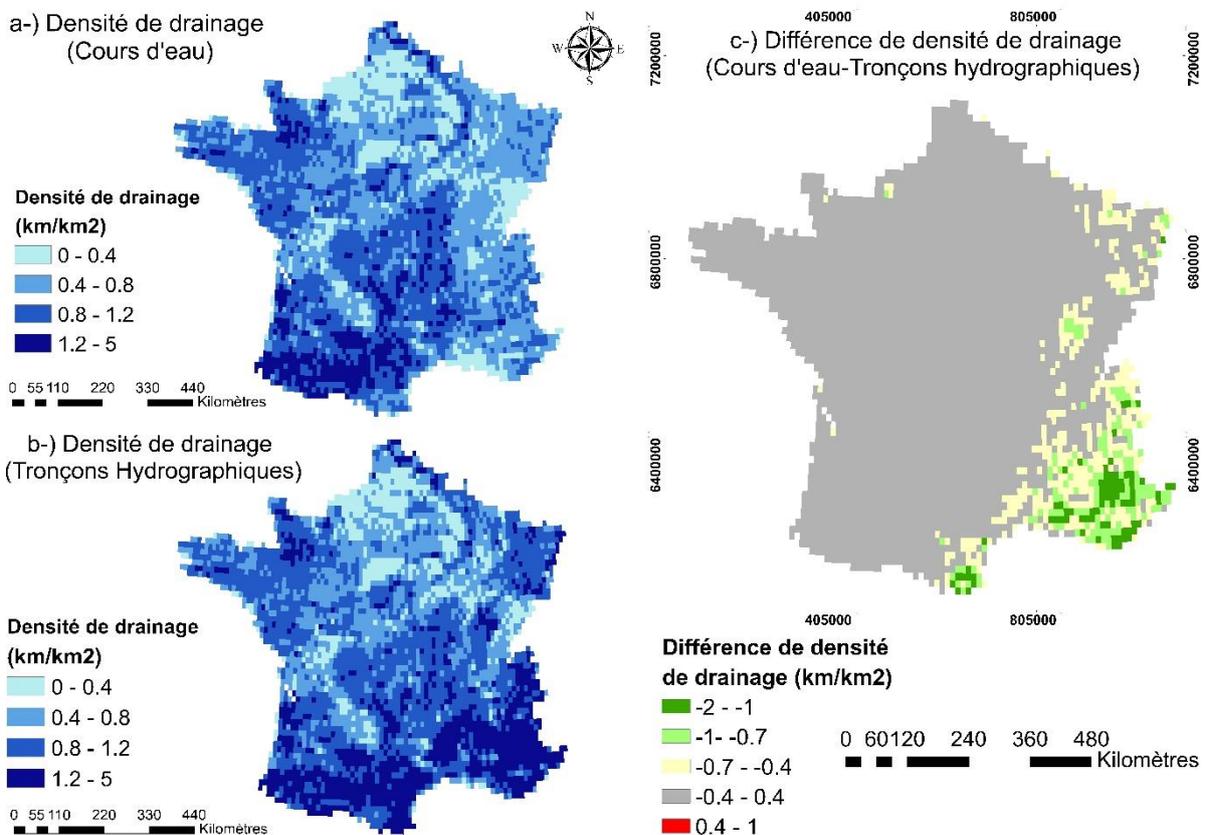


Figure 4 : Comparaison des densités de drainage des couches cours d'eau et tronçons hydrographiques de la BD Carthage pour un maillage de 7,5'

La carte des différences entre les densités de drainage des couches « Cours d'eau » et « Tronçons Hydrographiques » (**Figure 4c**) montre en revanche une différence fortement négative dans le Sud-Est et l'Est de la France et quelques différences positives sur le reste du territoire. Ces valeurs négatives traduisent le fait que la densité de drainage de la couche « Tronçons Hydrographiques » est supérieure à celle de la couche « Cours d'eau » à ces endroits. Cela confirme le fait que la couche « Tronçons hydrographiques » fournit une description plus complète du réseau hydrographique.

4.1.2 Caractérisation des erreurs du modèle LCS

La caractérisation de l'erreur du modèle LCS se fait en analysant les différences entre le modèle LCS et deux couches de référence de la BD Carthage la couche CDO : « Cours d'eau » utilisée par (Schneider et al. 2017) en **Figure 5a** et la couche Th_total : « Tronçon hydrographique » plus réaliste **Figure 5b**. La **Figure 5a** montre que la densité de drainage LCS est fortement supérieure à celle de la couche « cours d'eau » dans une partie du Sud-Est et de l'Est de la France (ou la densité CDO est sous-estimée comme montré ci-dessus). La densité LCS est inférieure à celle CDO dans une partie du Sud-Ouest notamment au niveau des Landes et au centre ou CDO est comparable à la couche « Tronçon hydrographique ». Les valeurs de densités de drainage de l'erreur absolue (LCS-CDO) évoluent entre -1,46 et 1,40 km/km². La RMSE entre le modèle LCS et les réseaux de la couche CDO est de 0,39 (Schneider et al. 2017) .

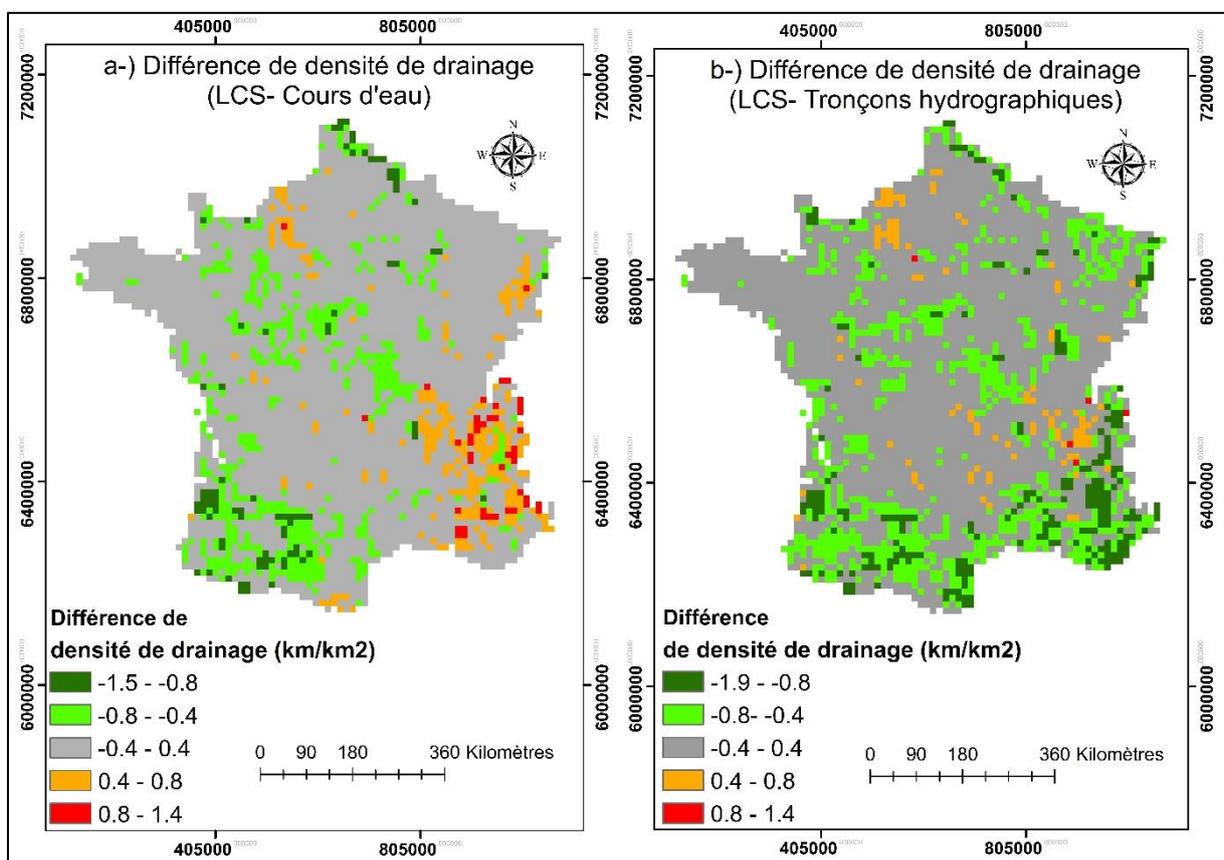


Figure 5 : Différence de densité de drainage entre le réseau LCS et les réseaux de références de la France pour un maillage de 7,5'

Si on calcule les erreurs du modèle LCS par rapport à la couche « Tronçon hydrographique » (qui est plus complète) on trouve une dominance des erreurs négatives (**Figure 5b**). Les erreurs de densité de drainage évoluent entre -1,85 et 1,31 km/km² pour une valeur de RMSE égale à 0,42. La répartition plus importante des valeurs négatives des erreurs se localise au Sud de la France, ce qui implique que le modèle LCS sous-estime plus la densité de drainage au Sud. Sur le reste du territoire, les densités de drainage se répartissent de façon hétérogène et en passant de la référence CDO à la référence complète TH_total, on augmente donc la sous-estimation de la densité de drainage par le modèle LCS ce qui augmente un peu la RMSE. Ceci aboutit à une sous-estimation beaucoup

systématique de la densité de drainage, qui s'accompagne d'une bien meilleure corrélation spatiale entre la densité LCS et la densité de référence ($r=0.32$ avec CDO et $r=0.48$ avec TH_total).

Au vu de l'analyse ci-dessus, il convient d'utiliser l'erreur absolue (LCS-TH_total) pour identifier les facteurs pouvant expliquer les erreurs du modèle LCS et permettant donc peut-être de les réduire dans la phase de recalibration.

4.1.3 Facteurs explicatifs potentiels des erreurs

❖ Tronçons hydrographiques intermittents

La **Figure 6** montre de forte densité de drainage des réseaux intermittents au Sud de la France qui coïncident à peu près avec les fortes différences (LCS-TH_total).

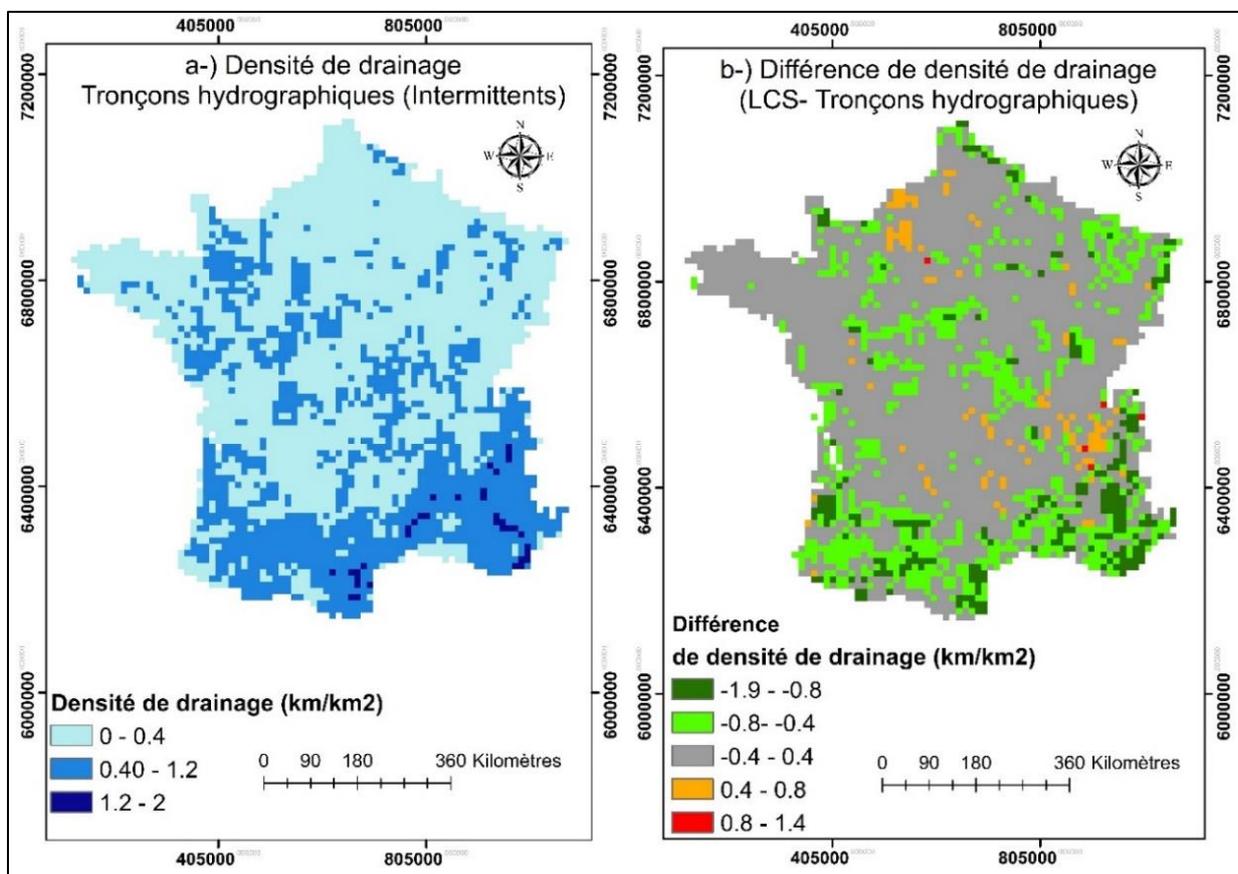


Figure 6 : Comparaison entre la densité les tronçons hydrographiques (intermittents) et la différence (LCS-TH_total) au niveau de la France pour un maillage de 7.5'

La corrélation entre l'intermittence et l'erreur absolue (LCS-TH_total) donne un coefficient négatif $r = -0,54$. Ce coefficient traduit le fait que plus l'intermittence augmente plus on a des erreurs. Le test de significativité donne une p-value qui est $< 2,2 * 10^{-16} < 0,05$. Au risque de 5% on peut dire que ce coefficient de corrélation est significatif. Le modèle LCS sous-estime la densité de drainage d'autant plus que l'intermittence est élevée.

❖ Tronçons hydrographiques artificiels

La **Figure 7** montre que les fortes densités de réseaux artificiels (TH_artificiel) coïncident avec des valeurs négatives de l'erreur absolue (LCS-TH_total). Le coefficient de corrélation correspondant est de $r = -0,40$ et significatif avec une p-value $< 2,2 * 10^{-16} < 0,05$. Ainsi plus cette artificialisation augmente et plus le modèle LCS tend à sous-estimer la densité de drainage, mais cette relation est moins marquée qu'avec l'intermittence.

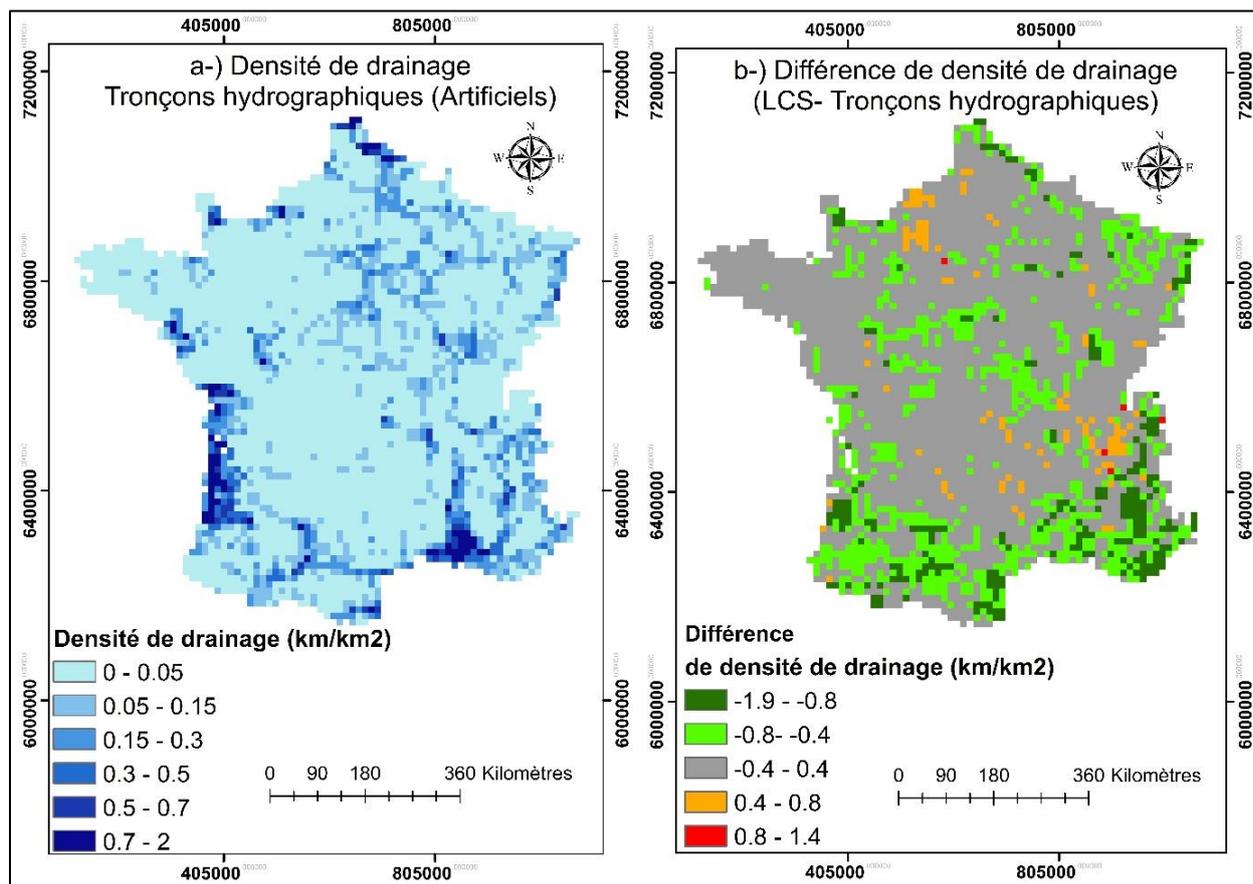


Figure 7 : Comparaison entre la densité de drainage des tronçons hydrographiques (artificiels) et la différence (LCS-TH_total) au niveau de la France pour un maillage de 7.5'

L'artificialisation des cours d'eau est due à plusieurs facteurs dont certaines sont d'origine agricole. Le drainage agricole des terres à l'aide de canaux et l'irrigation sont parmi ces éléments principaux pouvant influencer l'artificialisation des réseaux hydrographiques.

❖ Réseaux agricoles (Drainage et Irrigation)

La répartition au sein des mailles de 7,5' du taux de superficie irriguée par maille nous permet de repérer de fortes valeurs à l'Est, au Centre et au niveau du Sud-Ouest (notamment dans les Landes) **Figure 8a**. La répartition du taux de superficie drainée par maille montre des fortes valeurs à l'extrême Nord, au centre, au niveau du Sud-Ouest(Landes), et à certains endroits de l'Est **Figure 8b**. Ensemble, ces deux cartes semblent expliquer une bonne partie des tronçons artificiels.

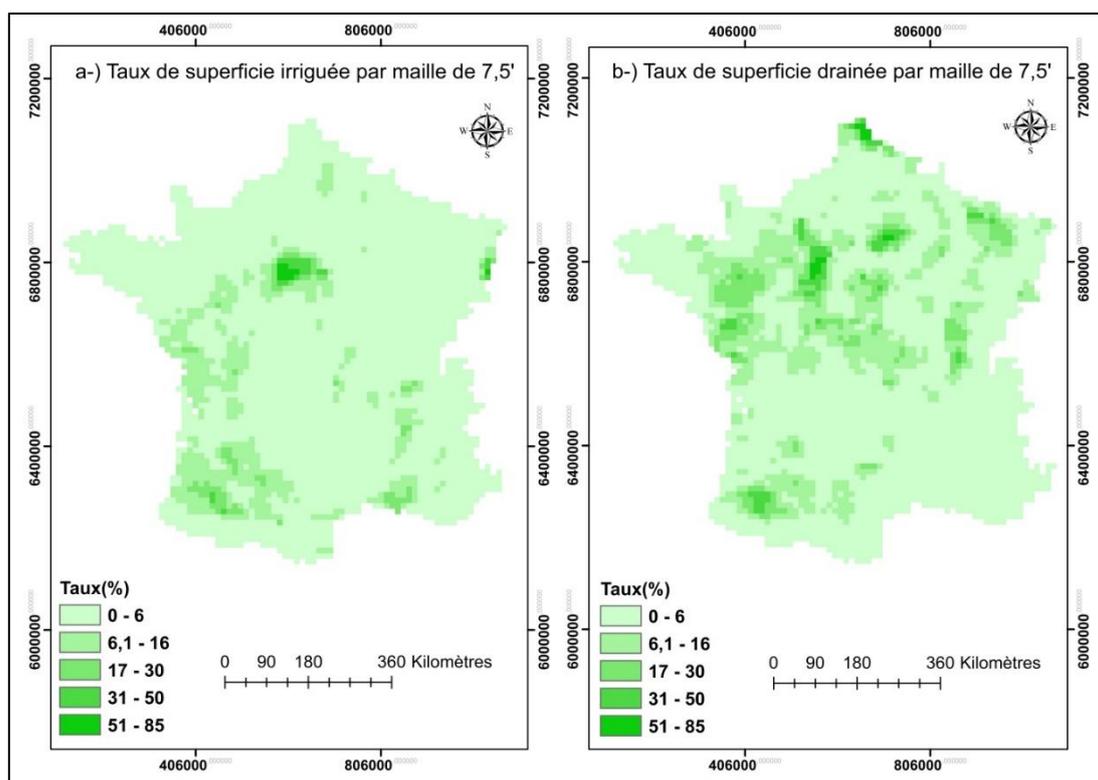


Figure 8 : Taux de superficie irriguée et drainée par maille au niveau de la France pour un maillage de 7,5'

L'analyse des relations entre les erreurs absolue (LCS-TH_total) et chacun des taux nous donne respectivement des coefficients de corrélation de $r = -0,17$ pour la part de superficie irriguée par maille et $r = -0,15$ pour la part de superficie drainée par maille. La valeur des p-value de chacun des taux est inférieure $2,2 * 10^{-16}$ et au risque de 5% on peut dire que ces coefficients de corrélation sont significatifs. Ces corrélations sont cohérentes avec celle obtenue entre LCS-TH_total et TH_artificiel.

4.1.4 Régression linéaire multivariée

A la suite de ces analyses, nous avons réalisé une régression linéaire multiple entre l'erreur du modèle et les variables (TH_intermittent) et (TH_artificiel) dans le but de déterminer la part de variance de l'erreur expliquée par ces réseaux. Cette régression nous donne l'équation suivante :

$$Y = -0,66 * X_1 - 0,72 * X_2 + 0,1$$

Y : Erreur du modèle (LCS-TH_total) en Km/km²
 X₁ : Densité de drainage (TH_intermittent) en Km/km
 X₂ : Densité de drainage (TH_artificiel) en Km/km²

La corrélation entre l'erreur (LCS-TH_total) et chacune des variables (TH_Intermittent) et (TH_Artificiel) est linéaire négatives, le modèle sous-estime la densité de drainage pour une augmentation des réseaux intermittents et artificiels. Le coefficient de détermination R² de la régression est de 0,4 et s'interprète comme la proportion de l'erreur (LCS-TH_total) linéairement expliqué par (TH_Intermittent) et (TH_artificiel). Ainsi 40% de la variance de l'erreur est expliquée par (TH_Intermittent) et (TH_Artificiel). Des régressions linéaires multiples avec d'autres facteurs explicatifs potentiels sont récapitulées dans le **Tableau 3** suivants.

Tableau 3 : Récapitulatif des régressions effectuées

Equations	r ²
Y= -0,72*X ₁ +0,06 Y= erreur du modèle (LCS- TH_total) en Km/km ² X ₁ : Densité de drainage (TH_intermittent) en Km/km ²	0,29
Y = -0,66*X ₁ -0,72*X ₂ +0,1 X ₁ : Densité de drainage (TH_intermittent) en Km/km ² X ₂ : Densité de drainage (TH_artificiel) en Km/km ²	0,40
Y= -0,66*X ₁ -0,73*X ₂ +0,0004*X ₃ +0,09 X ₁ : Densité de drainage (TH_intermittent) en Km/km ² X ₂ : Densité de drainage (TH_artificiel) en Km/km ² X ₃ : Taux de superficie irriguée par maille en (%)	0,40
Y= -0,66*X ₁ -0,71*X ₂ +0,0007*X ₃ -0,002*X ₄ +0,11 X ₁ : Densité de drainage (TH_intermittent) en Km/km ² X ₂ : Densité de drainage (TH_artificiel) en Km/km ² X ₃ : Taux de superficie irriguée par maille en % X ₄ : Taux de superficie drainée par maille en %	0,41

L'ajout d'autres facteurs comme le taux de superficie irriguée par maille et le taux de superficie drainée par maille augmente très légèrement le R². À la vue de nos analyses, nous pouvons dire que parmi les facteurs potentiels que nous avons étudiés, l'intermittence et l'artificialisation sont les facteurs les plus pertinents bien que la part de la variance expliquée ne soit que de 40%.

L'artificialisation et l'intermittence des réseaux hydrographiques n'expliquent pas à elles seules les erreurs du modèle LCS. La calibration du modèle LCS s'est appuyée sur les réseaux de la base de données AHGF en Australie et de la couche « Cours d'eau » de la BD Carthage en France. Nos études à travers la différence (CDO-TH_total) nous montrent que la couche « Tronçons Hydrographiques » est plus adaptée que la couche « Cours d'eau » pour le calibrage du modèle en France. Une part non négligeable des erreurs du modèle LCS pourrait venir du choix de la couche de référence. Pour plus d'efficacité, il serait alors nécessaire recalibrer le réseau à l'aide de la couche « Tronçons Hydrographiques ».

4.2 Recalibrage du modèle LCS

Cette partie de recalibration vise à proposer des A_{cr} spatialement variables en fonction de paramètres environnementaux comme la précipitation, la lithologie, la pente et l'ETP. Ce qui devrait permettre d'extraire de façon adéquate différents types de réseaux hydrographiques français à partir d'un MNT de 15''.

4.2.1 Couche de lithologie

La couche de Lithologie utilisée provient de la base de données de (Hartmann and Moosdorf 2012), répartie en 13 classes. Une intersection de cette couche avec le maillage de la France continentale à 7.5' nous a permis d'identifier les lithologies strictement majoritaires et les lithologies majoritaires à 70% en fonction du pourcentage de surface occupée dans la maille. Les lithologies strictement majoritaires sont les lithologies qui sont majoritaires au sein de chaque maille en pourcentage de surface. Celles majoritaires à 70 % sont des lithologies dont le pourcentage de surface occupée est supérieur ou égal à 70 %. Le **Tableau 3** ci-dessous résume le nombre de mailles et le pourcentage occupé par chaque type de lithologie. Nous utiliserons pour le calibrage la couche de lithologie supérieure à 70% afin de nous assurer que la densité de drainage observée reflète au mieux la lithologie.

Tableau 4 : Nombre de mailles et pourcentages occupé par les classes de lithologies, en France

Classe de lithologie	Abréviations	Lithologies majoritaires		Lithologies majoritaires à 70 %	
		Nombres de mailles occupées	Pourcentage (%)	Nombres de mailles occupées	Pourcentage (%)
Metamorphic Rocks	Mt	344	8,45	143	6,54
Acid Plutonic Rocks	Pa	347	8,52	148	6,76
Basic Plutonic Rocks	Pb	7	0,17	0	0,00
Intermetiate Plutonic Rocks	Pi	3	0,07	2	0,09
Carbonate Sedimentary Rocks	Sc	1613	39,60	1064	48,63
Mixed Sedimentary Rocks	Sm	589	14,46	234	10,69
Siliciclastic Sedimentary Rocks	Ss	540	13,26	288	13,16
Unconsolidated Sediments	Su	560	13,75	288	13,16
Acid Volcanic Rocks	Va	26	0,64	5	0,23
Basic Volcanic Rocks	Vb	44	1,08	16	0,73
Intermediate Volcanic Rocks	Vi	0	0,00	0	0,00
Water Bodies	Wb	0	0,00	0	0,00
Ice And Glaciers	Ig	0	0,00	0	0,00
Total		4073		2188	

Le **Tableau 4** montre que, les classes de lithologies Vi, Wb, et Ig ne sont majoritaires à 70% dans aucune des mailles, et que Pb, Pi ne sont majoritaires à 70% que dans un très faible nombre de mailles, ce qui nous amène à les supprimer et à nous en tenir au reste des lithologies. La **Figure 9** représente ci-dessous ces différents types de répartition.

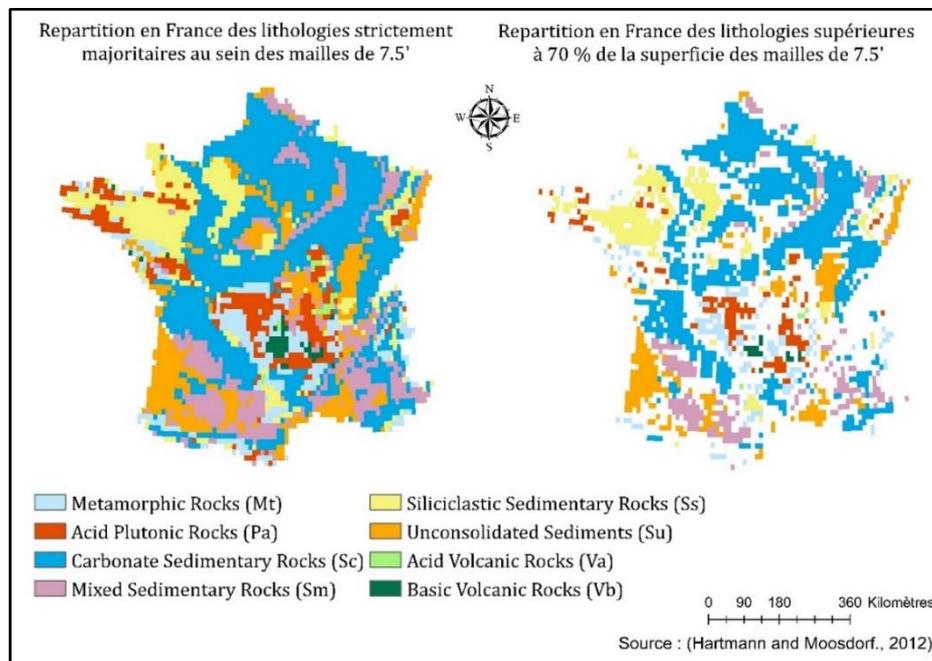


Figure 9 : Représentation des classes strictement majoritaires de lithologies et des classes majoritaires à 70% au sein des mailles de 7,5'

Les couches Va (Acid Volcanic Rocks) et Vb (Basic Volcanic Rocks) occupent aussi un très faible nombre de mailles (**Tableau 4**). Ce sont des roches volcaniques et leur répartition sur le réseau Th_total montre qu'elles sont associées à des moyennes de densité de drainage proches **Figure 10**. Nous les avons regroupés en une seule classe (Va+Vb). La suite des analyses se fera donc sur 7 couches de lithologies, récapitulées dans le (**Tableau 5**).

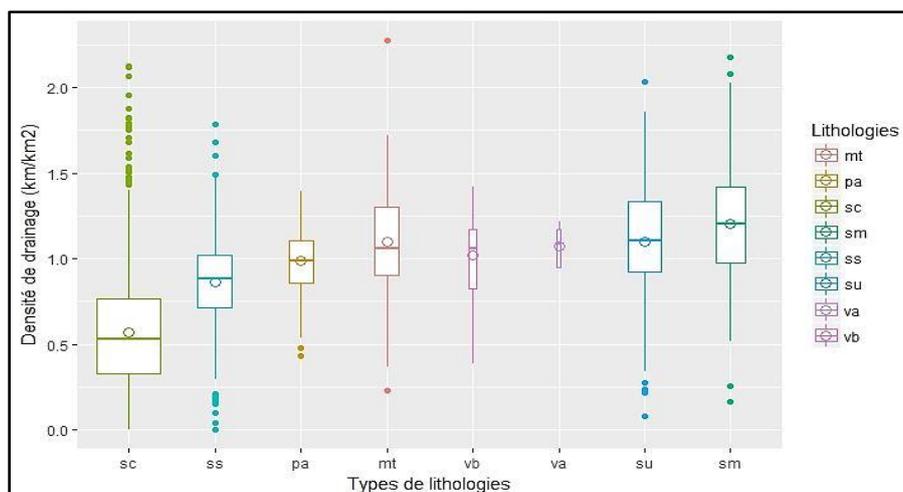


Figure 10 : Répartition de la densité de drainage des tronçons hydrographiques totaux en fonction des lithologies majoritaires à 70%. La largeur des classes est proportionnelle à la racine carrée du nombre de mailles associés.

Tableau 5 : Liste des lithologies retenues

Classe de lithologie	Abréviation
Metamorphic Rocks	Mt
Acid Plutonic Rocks	Pa
Carbonate Sedimentary Rocks	Sc
Mixed Sedimentary Rocks	Sm
Siliciclastic Sedimentary Rocks	Ss
Unconsolidated Sediments	Su
Acid Volcanic Rocks + Basic Volcanic Rocks	Va+Vb

4.2.2 Relation entre les Aires critiques et les densités de drainage

Comme détaillé à l'étape 1 de la section 3.4, au sein des couches de lithologies retenues et pour chacun des 9 types de réseaux, des régressions linéaires sont effectuées afin d'établir des relations entre les A_{cr} calibrées (A_{cr} totales) et les densités de drainage (Figure 11).

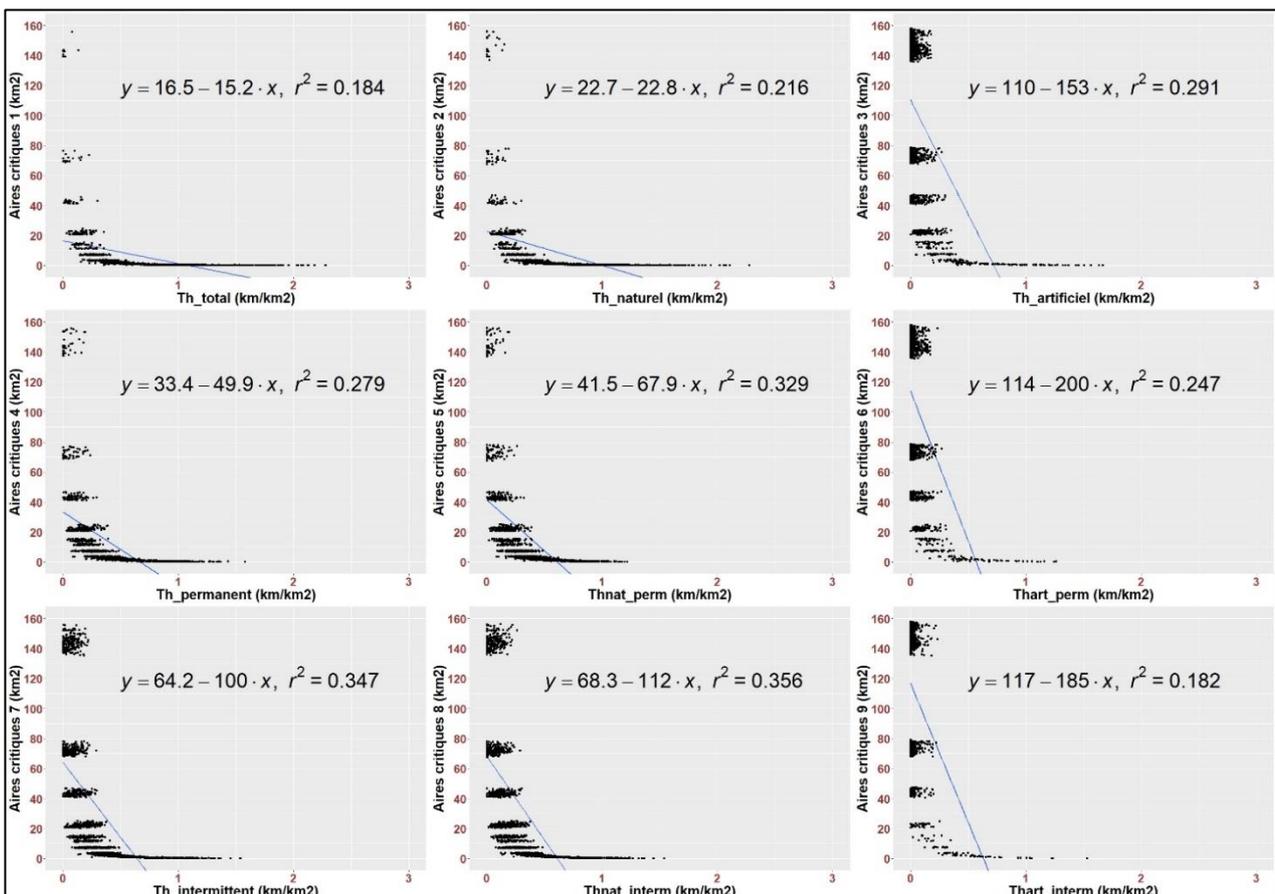


Figure 11 : Régression linéaire des valeurs d'aires critiques (A_{cr} totales) les plus proches des réseaux en fonction de leur densité de drainage (km/km²)

Dans l'ensemble des régressions sur la **Figure 11**, les coefficients de détermination sont faibles. Le coefficient de détermination r^2 le plus élevé est obtenu avec le réseau Thnat_interm (0.36) et le plus faible avec Thart_interm (0.18).

On remarque sur cette **Figure 11** que les relations sont non linéaires et qu'une grande majorité de points se trouvent sous un certain seuil d'aires critiques comme dans le cas du réseau Th_total. Afin de mieux comprendre les relations à ce niveau nous avons entrepris de faire une analyse jusqu'au seuil de 4 km² (**Figure 12**).

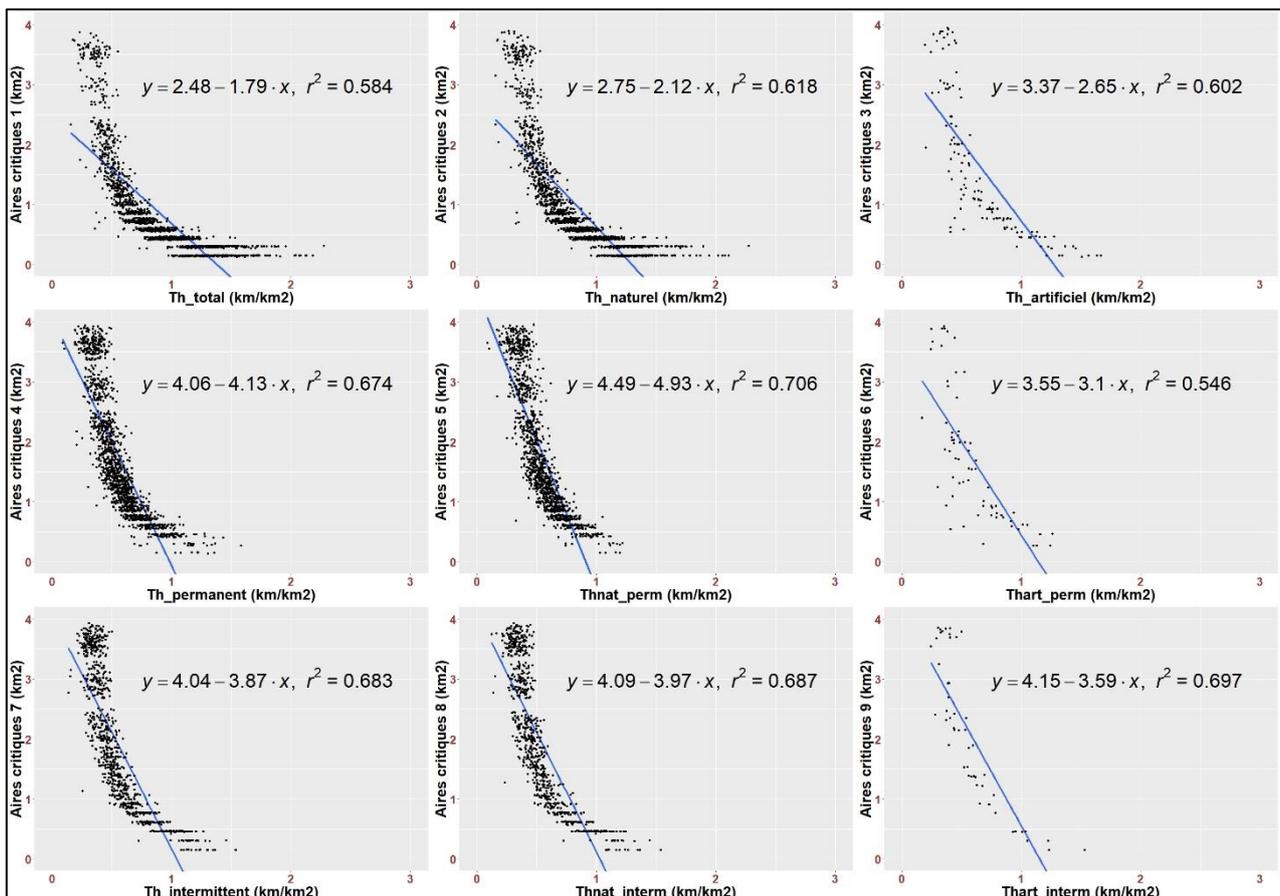


Figure 12 : Régression linéaire des valeurs d'aires critiques ($A_{cr} < 4 \text{ km}^2$) les plus proches des réseaux en fonction de leur densité de drainage (km/km^2)

Sur la **Figure 12**, les coefficients de détermination augmentent en général, les relations ne suivent pas un ajustement linéaire. Ainsi nous avons testé plusieurs ajustements dont l'exponentiel, logarithme népérien, inverse, puissance. L'ajustement bi-logarithmique est celui qui donne les meilleures performances (**Figure 13** et **Figure 14**).

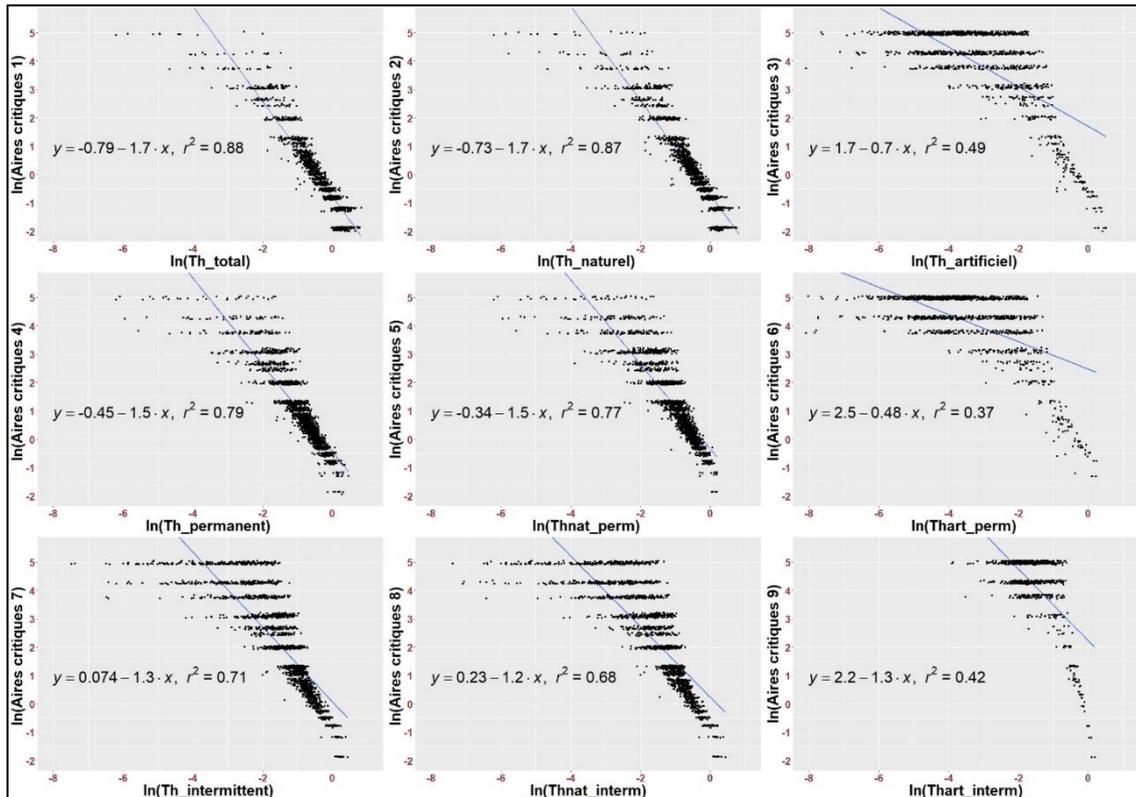


Figure 13: Régression logarithmique des valeurs d'aires critiques (A_{cr} totales) les plus proches des réseaux en fonction de leur densité de drainage (km^2/km^2)

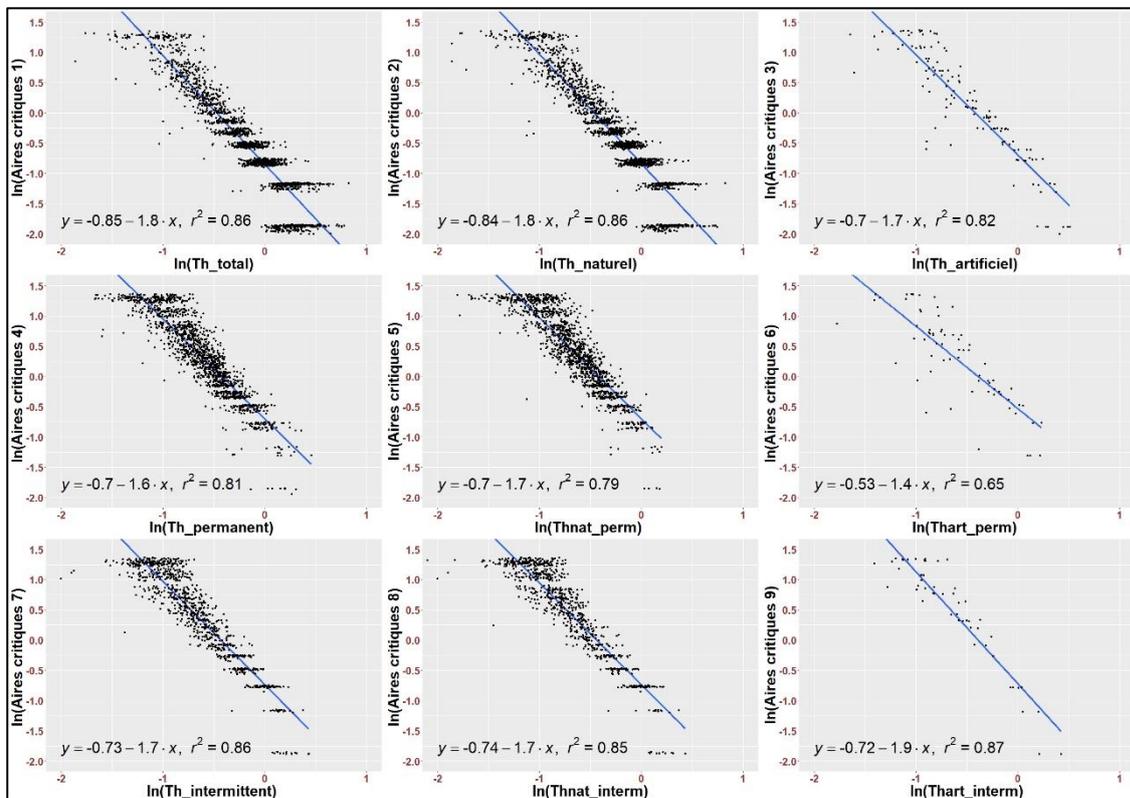


Figure 14 : Régression logarithmique des valeurs d'aires critiques ($A_{cr} < 4 km^2$) les plus proches des réseaux en fonction de leur densité de drainage (km^2/km^2)

Au niveau de la **Figure 13**, on voit que l'ajustement est bien meilleur qu'en linéaire, et ce pour tous les réseaux. Les performances (r^2) des réseaux artificiels sont moins bonnes et sont de l'ordre de 0.4 tandis que pour les autres réseaux il tourne autour de 0.7 et 0.8. Cela est peut-être dû aux Aires critiques généralement élevées des réseaux artificiels et donc des densités de drainage faibles de ces réseaux ce qui rend les rend difficile à caractériser. Au niveau de la **Figure 14**, la régression logarithmique nous donne de très bonnes performances avec des coefficients de détermination r^2 variant entre 0.6 et 0.8.

Dans l'optique d'avoir une seule équation prenant en compte toutes ces relations, des ajustements sont faits avec différentes équations (**Figure 15** et **Figure 16**).

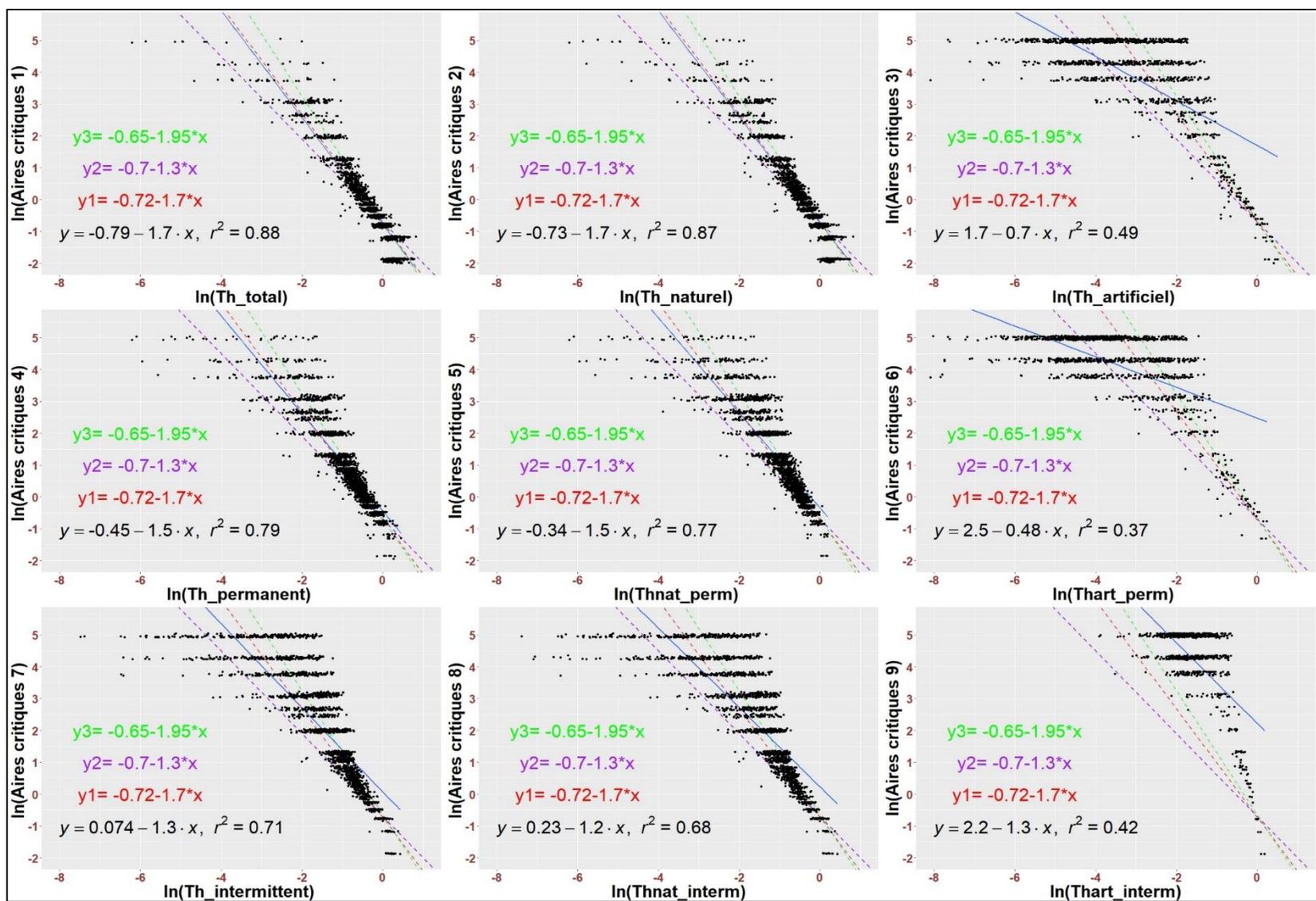


Figure 15 : Régression logarithmique des valeurs d'aires critiques (A_{cr} totales) les plus proches des réseaux en fonction de leur densité de drainage (km/km^2)

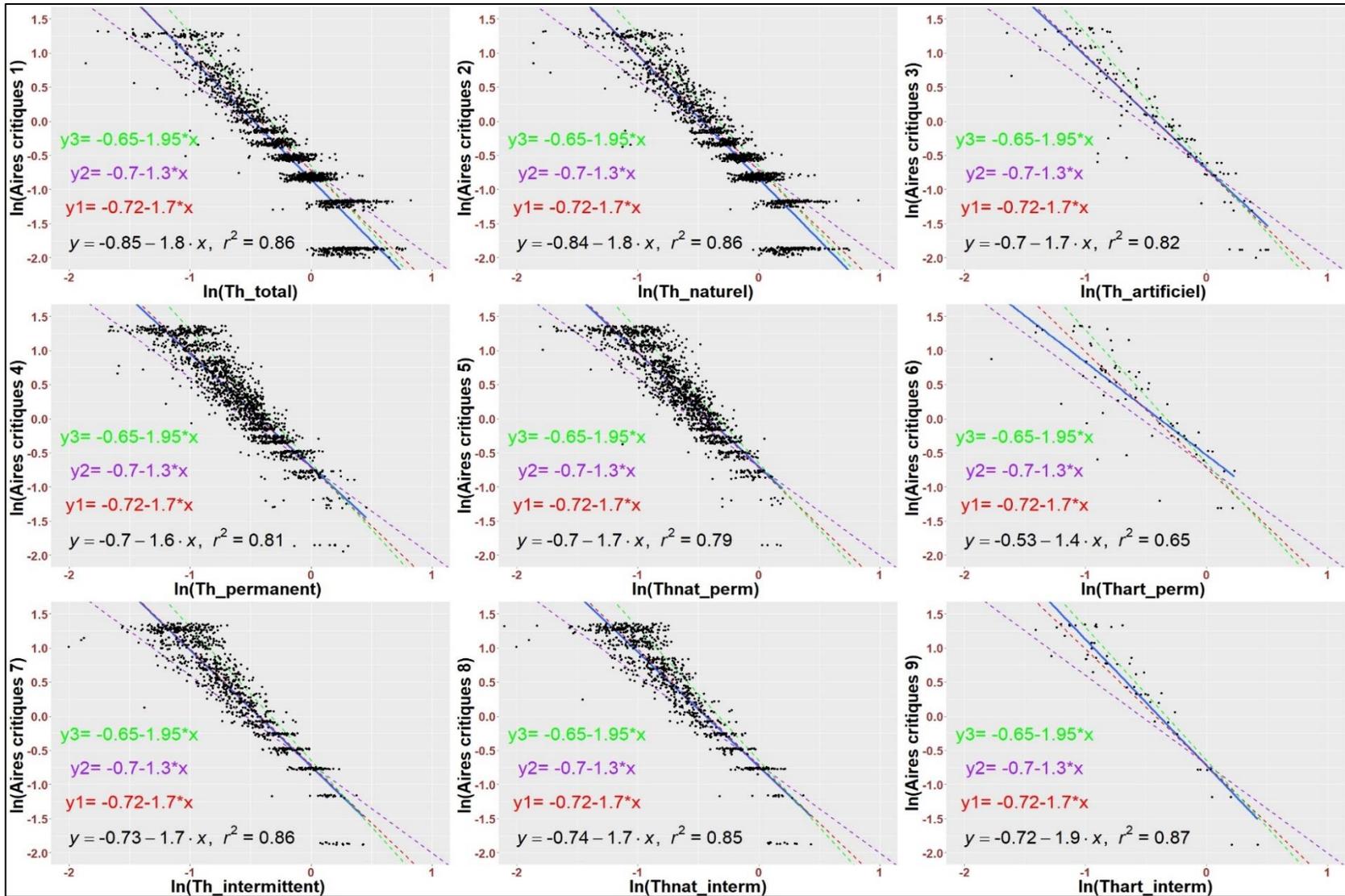


Figure 16 : Régression logarithmique des valeurs d'aires critiques ($A_{cr} < 4 \text{ km}^2$) les plus proches des réseaux en fonction de leur densité de drainage (km^2/km^2)

Les différents r^2 obtenu pour chaque équation par rapport aux différents réseaux sont cosignés dans le **Tableau 7** et **Tableau 7** suivant.

Tableau 6 : Récapitulatif des coefficients de détermination obtenus pour chaque équation et en fonction des différents réseaux (A_{cr} totales)

	Th_ total	Th_ nat	Th_ art	Th_ perm	Thnat_ perm	Thart_ perm	Th_ interm	Thnat_ interm	Thart_ interm
$Y_{\text{regression}}$	0.88	0.87	0.49	0.79	0.77	0.42	0.71	0.68	0.42
$Y=-0.72-1.7*X$	0.87	0.86	0.06	0.77	0.76	0.03	0.65	0.60	0.01
$Y_2=-0.7-1.3*X$	0.84	0.83	0.16	0.68	0.63	0.08	0.52	0.48	0.00
$Y_3=-0.65-0.95*X$	0.83	0.83	0.01	0.73	0.71	0.00	0.55	0.47	0.01

Tableau 7 : Récapitulatif des coefficients de détermination obtenus pour chaque équation et en fonction des différents réseaux ($A_{cr} < 4\text{km}^2$)

	Th_ total	Th_ nat	Th_ art	Th_ perm	Thnat_ perm	Thart_ perm	Th_ interm	Thnat_ interm	Thart_ interm
$Y_{\text{regression}}$	0.86	0.86	0.82	0.81	0.80	0.65	0.86	0.85	0.87
$Y_1 = -0.72 - 1.7 * X$	0.84	0.85	0.81	0.80	0.79	0.61	0.85	0.84	0.85
$Y_2 = -0.7 - 1.3 * X$	0.79	0.80	0.74	0.67	0.61	0.57	0.71	0.71	0.63
$Y_3 = -0.65 - 0.95 * X$	0.78	0.78	0.74	0.66	0.63	0.41	0.73	0.69	0.84

Sur le tableau 6, les équations testées présentent de bonnes performances sur les réseaux à l'exception des artificiels (Th_art, Thart_perm, Thart_interm) alors que sur le tableau 7 les performances sont bonnes pour tous les réseaux.

L'équation qui donne les meilleures performances pour l'ensemble des réseaux est :
 $Y_1 = -0.72 - 1.7 * X$

La régression étant logarithmique on aura $\ln(Y_1) = -0.72 - 1.7 * \ln(X)$

$Y_1 = e^{-0.72} * X^{-1.7}$ avec $Y_1 = A_{cr}$ et $X = \text{Densité de drainage (Dd) en km/km}^2$

$$A_{cr} = e^{-0.72} * Dd^{-1.7}$$

Ainsi on a une première équation donnant les Aires critiques en fonction de la densité de drainage. Cette équation et les autres relations qui seront faites entre les densités de drainage et les paramètres environnementaux nous permettront d'obtenir des équations d'aires critiques en fonction de paramètres environnementaux.

4.2.3 Relations entre les densités de drainage et les paramètres environnementaux

Comme mentionné à l'étape 2 de la section 3.4, pour chacune des 7 classes de lithologie, des régressions linéaires multivariées sont établies entre les densités de drainage des différents réseaux et les paramètres environnementaux, pour obtenir 7 modèles de densité de drainage. Les **Figure 17** et **Figure 18** présentent ainsi les relations linéaires entre les densités de drainage et les paramètres environnementaux pour les classe de lithologies Mt (Metamorphic rocks) et Sm (Mixed sedimentary rocks). Les paramètres pris en compte sont : l'Etp, la pluie efficace, la pente, et la Dry season length.

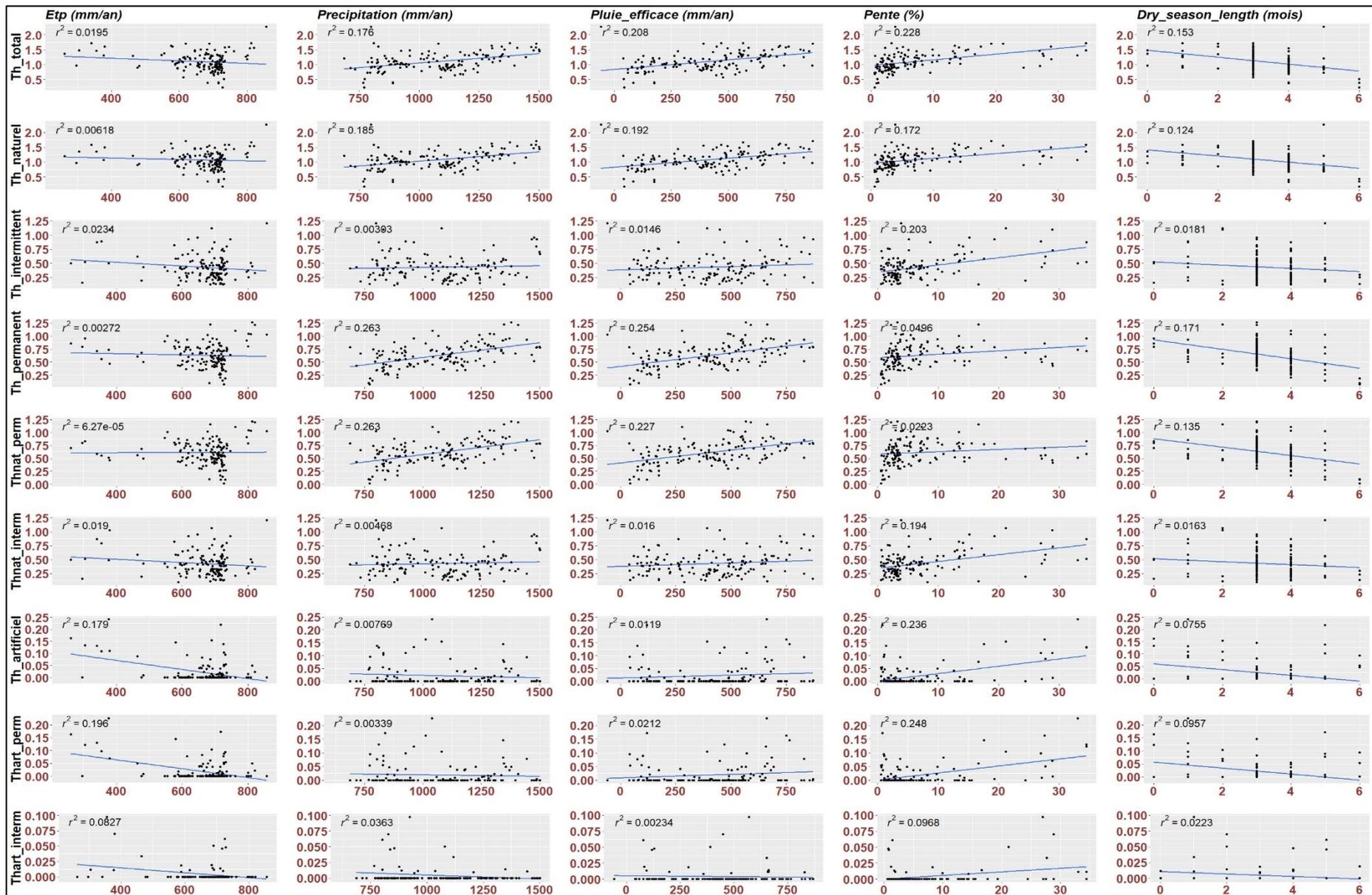


Figure 17 : Régression linéaire entre les densités de drainage des réseaux (km/km²) et certains paramètres environnementaux au niveau de la classe (Mt)

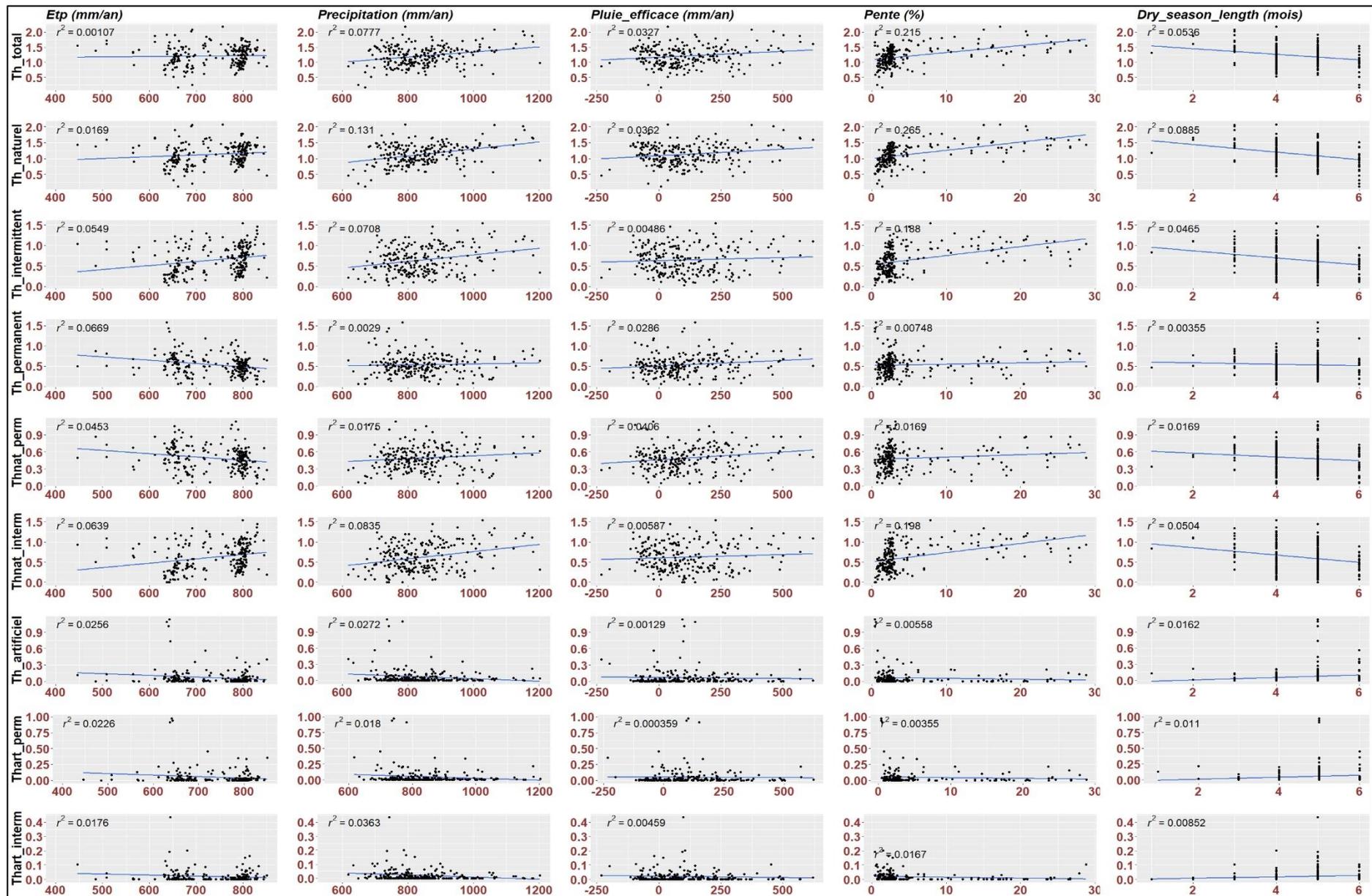


Figure 18 : Régression linéaire entre les densités de drainage des réseaux (km/km²) et certains paramètres environnementaux au niveau de la classe (Sm)

Par souci de simplicité, nous présenterons dans ce rapport les résultats pour les réseaux Th naturel, Th_total, Thnat_interm et Th_artificiel. Le **Tableau 8** résume les équations obtenues par régression linéaire multivariées pour les réseaux Th_total et Th_naturel, celles des réseaux Thnat_interm et Th_art sont répertoriés en **Annexe 3**. Sur le **Tableau 8**, les ajustements sont corrects sauf pour la couche Pa qui présente de faibles performances sur le réseau Th_naturel et Th_total et la couche Su qui présente elle de faible performance sur le réseau Th_total.

Soit $x_1 = \text{Etp (mm/an)}$

$x_2 = \text{Pluie efficace (mm/an)}$

$x_3 = \text{Pente (\%)}$

$x_4 = \text{Dry_season_length (mois)}$

y désignant la densité de drainage des différents réseaux (km^2/km^2)

Tableau 8 : Récapitulatif des régressions linéaires multiples entre les réseaux Th_naturel et Th_total et les paramètres en fonction des différentes classes de lithologies

Th_naturel	
Classe mt	
$y = 0.002 * x_1 + 0.0002 * x_2 + 0.028 * x_3 - 0.06 * x_4 - 0.4$	$r^2=0.38$
Classe pa	
$y = 0.0005 * x_1 + 0.0001 * x_2 - 0.0008 * x_3 - 0.001 * x_4 + 0.58$	$r^2=0.04$
Classe sc	
$y = 0.001 * x_1 - 0.0003 * x_2 + 0.046 * x_3 - 0.025 * x_4 - 0.258$	$r^2=0.3$
Classe sm	
$y = 0.002 * x_1 + 0.0001 * x_2 + 0.033 * x_3 - 0.006 * x_4 - 0.71$	$r^2=0.44$
Classe ss	
$y = 0.002 * x_1 + 0.0004 * x_2 + 0.028 * x_3 - 0.053 * x_4 - 1.01$	$r^2=0.39$
Classe su	
$y = -0.001 * x_1 + 0.0006 * x_2 + 0.12 * x_3 - 0.08 * x_4 + 1.42$	$r^2=0.26$
Classe va+vb	
$y = 0.002 * x_1 + 0.0003 * x_2 - 0.014 * x_3 - 0.10 * x_4 - 0.34$	$r^2=0.51$
Th_total	
Classe mt	
$y = 0.0019 * x_1 + 0.0002 * x_2 + 0.029 * x_3 - 0.056 * x_4 - 0.303$	$r^2=0.41$
Classe pa	
$y = 0.0004 * x_1 + 0.0001 * x_2 + 0.00006 * x_3 - 0.005 * x_4 + 0.6$	$r^2=0.03$
Classe sc	
$y = 0.001 * x_1 - 0.0003 * x_2 + 0.05 * x_3 - 0.01 * x_4 - 0.19$	$r^2=0.30$
Classe sm	
$y = 0.001 * x_1 + 0.0001 * x_2 + 0.03 * x_3 + 0.02 * x_4 - 0.12$	$r^2=0.30$
Classe ss	
$y = 0.002 * x_1 + 0.0004 * x_2 + 0.028 * x_3 - 0.057 * x_4 - 0.71$	$r^2=0.39$
Classe su	
$y = 0.0003 * x_1 + 0.00005 * x_1 + 0.04 * x_1 + 0.01 * x_1 + 0.7$	$r^2=0.03$
Classe va+vb	
$y = 0.002 * x_1 + 0.0004 * x_2 - 0.014 * x_3 - 0.078 * x_4 - 0.481$	$r^2=0.53$

Les densités de drainage générées en utilisant les équations de la régression linéaire multiple établies précédemment, sont une approximation des densités de drainage observées. Ce "modèle", appelé simulation 1, est comparé avec les densités de drainage observées. Les **Figure 19**, **Figure 20**, **Figure 21**, **Figure 22** montrent les différences de densité entre ces réseaux.

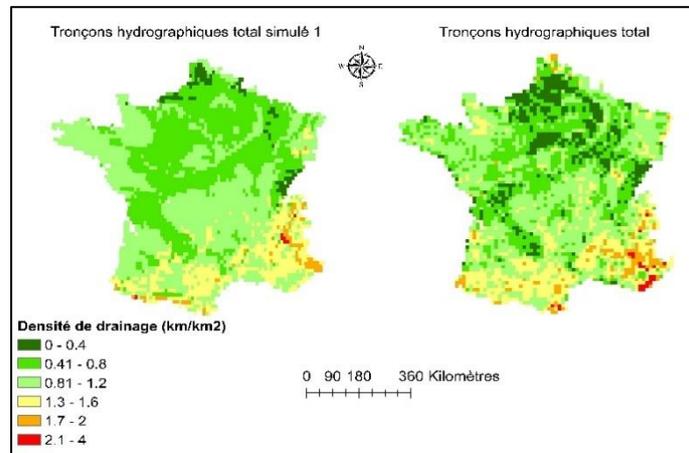


Figure 19 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques total simulé 1 avec le réseau original.

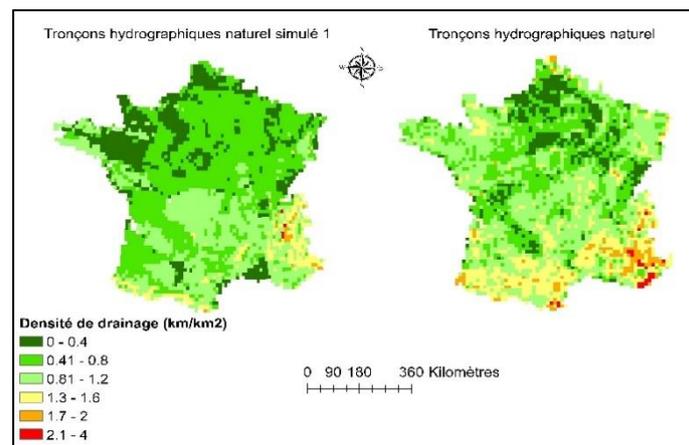


Figure 20 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques naturel simulé 1 avec le réseau original.

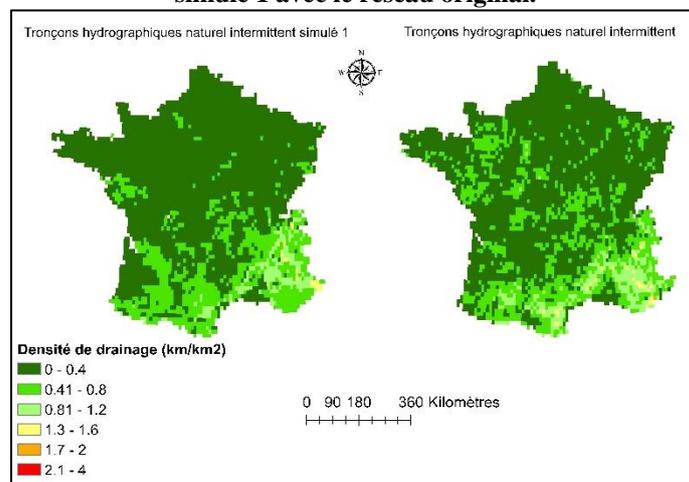


Figure 21 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques naturel intermittent simulé 1 avec le réseau original.

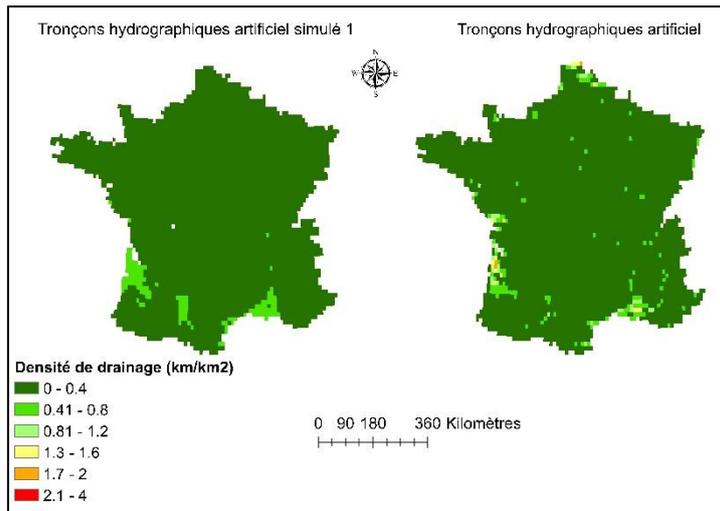


Figure 22 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques artificiel simulé 1 avec le réseau original.

Sur la **Figure 19** les patterns des densités de drainage au Sud et Sud-Est sont bien simulés tandis que sur le reste de la France cette densité de drainage est surestimée. Sur la **Figure 20**, on voit qu'à l'exception de la partie Sud-Est de la France, une grande partie de la densité de drainage naturelle est sous-estimée. La **Figure 21** nous montre que les densités sont sous-estimées principalement à l'Ouest et à l'Est de la France. Sur la **Figure 22** on voit que les densités sont en général sous-estimées.

4.2.4 Extraction des réseaux hydrographiques

A ce stade, nous avons les valeurs des densités de drainage par maille de 7.5' mais nous n'avons pas encore simulé le réseau hydrographique. Une "simulation 2" en utilisant la première équation que nous avons établie à savoir $A_{cr} = e^{-0.72} * Dd^{-1.7}$ et en prenant en compte la lithologie dominante nous permettra de dériver le réseau comme mentionné à l'étape 3 de la section 3.4. Ces aires critiques ont par la suite été utilisées pour dériver le réseau hydrographique comme l'illustre la **Figure 23** qui représente le réseaux *Thnat_interm* simulé 2 et *Thnat_interm* observés et dont l'équivalent en densité de drainage est représenté sur la **Figure 26**.

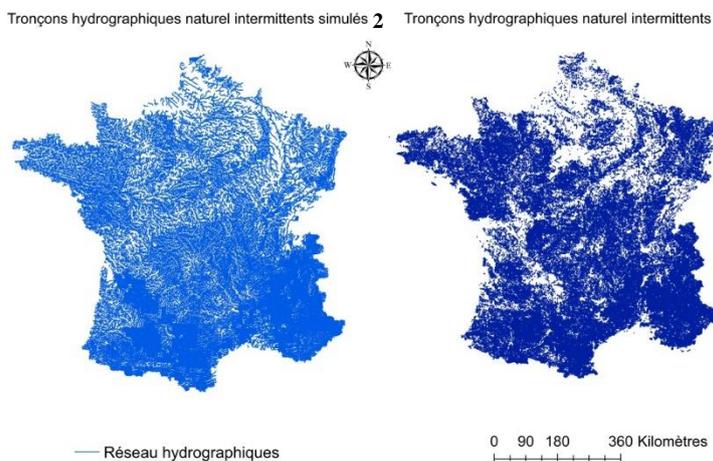


Figure 23 : Représentation des réseaux *Thnat_interm* simulés 2 et *Thnat_interm* observés

Les **Figure 24**, **Figure 25**, **Figure 26** et **Figure 27** ci-dessous représentent les résultats obtenus en termes de densité de drainage.

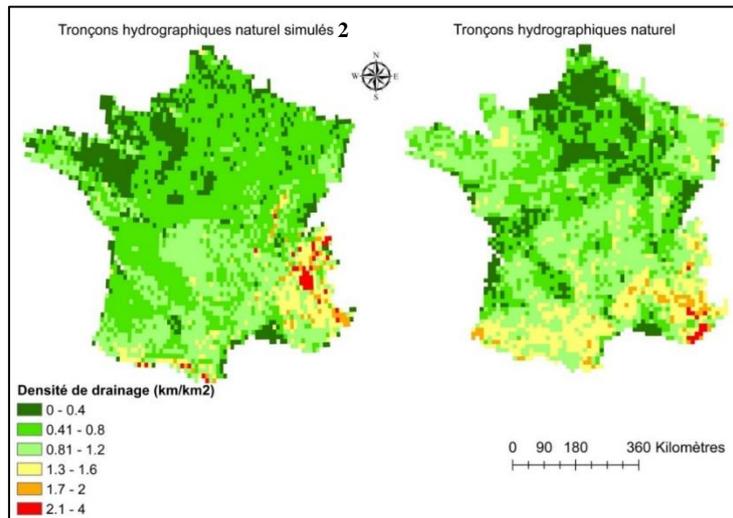


Figure 24 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques naturel simulé 2 avec le réseau original.

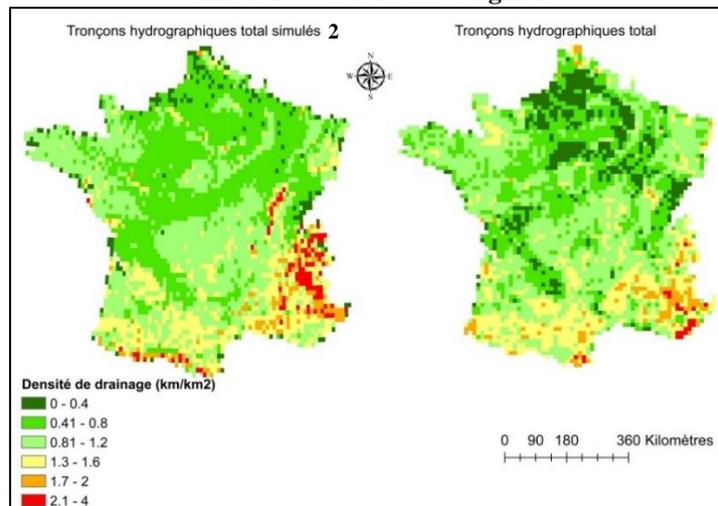


Figure 25 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques total simulé 2 avec le réseau original.

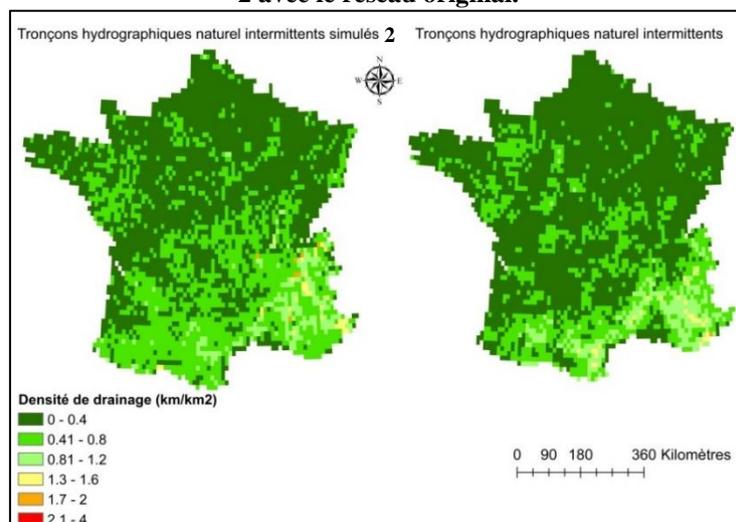


Figure 26 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques naturel intermittent simulé 2 avec le réseau original.

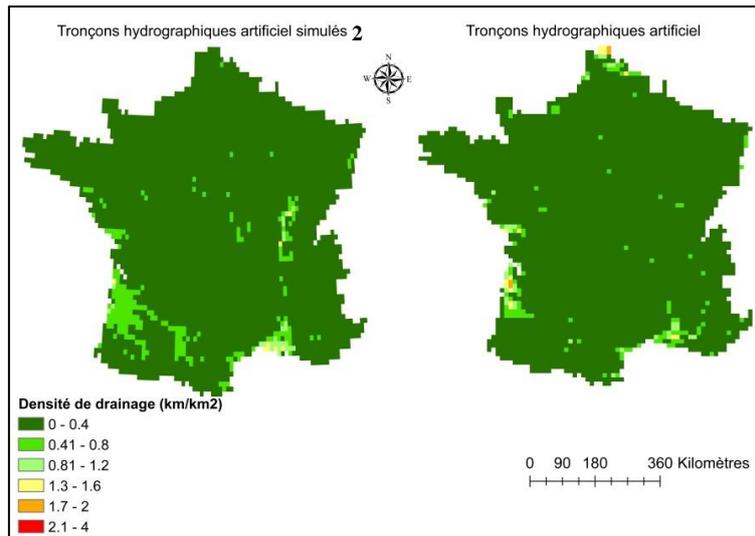


Figure 27 : Comparaison entre les densités de drainage du réseau Tronçons hydrographiques artificiel simulé 2 avec le réseau original.

Sur les figures précédentes on remarque que la répartition spatiale des densités de drainage des réseaux simulés est à certains endroits différente de ceux des réseaux observés.

Le **Tableau 9** ci-dessous présente les paramètres statistiques entre les réseaux simulés 1, simulés 2 et les réseaux observés. Les moyennes, l'écart type et le biais sont exprimés en (km/km²).

Tableau 9 : Analyses statistiques entre les réseaux simulés et les réseaux observés

Simulation 1										
	LCS	Th_total	Th_total_sim1	Th_total	Th_naturel_sim1	Th_naturel	Thnat_interm_sim1	Thnat_interm	Th_art_sim1	Th_art
Moyenne	0,71	0,90	0,87	0,90	0,66	0,82	0,35	0,33	0,05	0,08
Ecart-type	0,30	0,40	0,33	0,40	0,32	0,39	0,21	0,26	0,10	0,17
Biais	-0,21	0	-0,03	0	-0,16	0	0,02	0	-0,03	0
Corrélation	0,48	1	0,68	1	0,56	1	0,66	1	0,56	1
RMSE	0,42	0	0,30	0	0,38	0	0,20	0	0,15	0
Simulation 2										
	LCS	Th_total	Th_total_sim2	Th_total	Th_naturel_sim2	Th_naturel	Thnat_interm_sim2	Thnat_interm	Th_art_sim2	Th_art
Moyenne	0,71	0,90	0,93	0,90	0,70	0,82	0,40	0,33	0,11	0,08
Ecart-type	0,30	0,40	0,41	0,40	0,36	0,39	0,23	0,26	0,16	0,17
Biais	-0,21	0	0,03	0,00	-0,12	0,00	0,07	0,00	0,03	0,00
Corrélation	0,48	1	0,61	1,00	0,53	1,00	0,56	1,00	0,54	1,00
RMSE	0,42	0	0,36	0,00	0,38	0,00	0,24	0,00	0,16	0,00

Pour la simulation 1, la RMSE la plus faible est obtenue avec le réseau Th_art_sim1 et la plus forte avec le réseau Th_naturel_sim1. Cette bonne performance du réseau Th_art_sim1 s'explique par ses faibles valeurs de densités de drainage par rapport aux autres réseaux. Les corrélations sont bonnes dans l'ensemble avec des valeurs allant de 0.56 à 0.68. Le biais le plus important est observé avec le réseau Th_naturel.

De façon générale, pour la simulation 2 les performances sont un peu en baisse par rapport à la simulation 1. L'ordre des réseaux ayant les meilleures performances reste inchangé et les valeurs des corrélations et RMSE restent satisfaisantes. Aussi toutes les moyennes ont augmenté entre la simulation 1 et la simulation 2. Cette surestimation vient soit du changement de couches de lithologies (de dominant à 70 % à dominant simple), soit de la relation $A_{cr} = f(Dd)$, ce qui est le plus probable.

En comparant les performances obtenues pour le calibrage du réseau LCS (Schneider et al. 2017) et celui des différents réseaux qu'on a calibré, on peut dire que les performances qu'on a obtenues sont bonnes. Le coefficient de corrélation entre le modèle LCS et la couche CDO est de 0,32 tandis que celles que nous obtenons pour les réseaux Th_total_sim 1 et Th_total_sim 2 par rapport à la couche Th_total sont respectivement de 0.68 et 0.61, de même pour les valeurs de RMSE qui sont meilleurs **Tableau 10**.

Tableau 10 : Performances des réseaux LCS ,Th_total_sim 1, Th_total_sim 2

	LCS	CDO	LCS	Th_total	Th_total_sim 1	Th_total	Th_total_sim 2	Th_total
Moyenne	0,71	0,76	0,71	0,90	0,87	0,90	0,93	0,90
Ecart-type	0,30	0,36	0,30	0,40	0,33	0,40	0,41	0,40
Biais	-0,04	0	-0,21	0	-0,03	0	-0,03	0,00
Corrélation	0,32	1	0,48	1	0,68	1	0,61	1,00
RMSE	0,39	0	0,42	0	0,30	0	0,36	0,00

5 Conclusion et Perspectives

En vue de mieux caractériser le réseau hydrographique, Schneider et al., (2017) proposa plusieurs valeurs d'aires critiques en fonction de paramètres environnementaux. Ce modèle LCS (Lithologie Climate and Slope) est calibré à l'aide de la couche « Cours d'eau » de la BD Carthage et reproduit assez bien le réseau hydrographique, mais quelques erreurs subsistent.

Durant ce stage nous nous sommes attelés à caractériser ces erreurs notamment en France métropolitaine. En émettant l'hypothèse que l'artificialisation et l'intermittence des réseaux pourraient être les causes principales de ces erreurs, nous nous sommes servis de la couche « Tronçons hydrographiques » de la BD Carthage pour mieux caractériser ces réseaux. Les analyses nous ont montré que l'artificialisation et l'intermittence expliquent 40% de la part de la variance de l'erreur du modèle. Une part importante de l'erreur du modèle LCS viendrait du choix de la couche « Cours d'eau » comme couche de référence au lieu de la couche « Tronçons hydrographique » qui est plus complète.

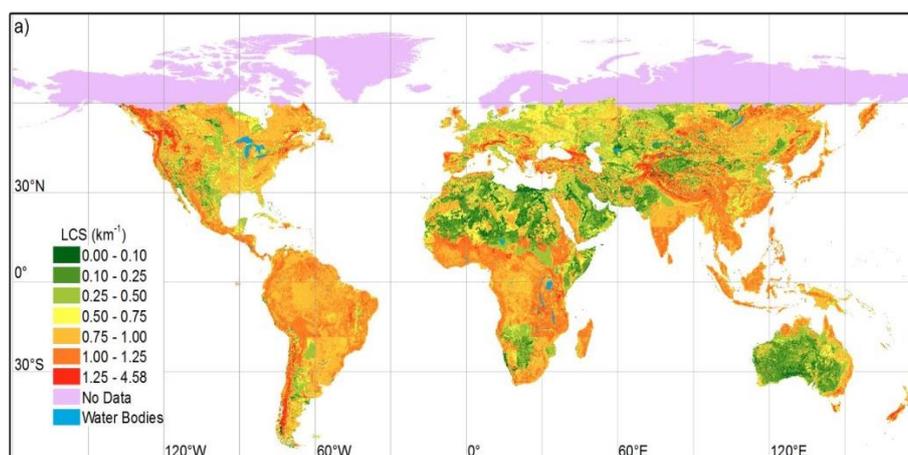
Un recalibrage en prenant en compte la couche « Tronçons hydrographique » et en sélectionnant différents types de réseaux (naturel, artificiel, permanent, intermittent) dans cette couche nous a permis d'avoir de meilleures performances. Par rapport aux réseaux de référence, les réseaux calibrés présentent de bonnes performances avec des biais en moyenne inférieur à 12 %. Pour les réseaux calibrés on a des valeurs de RMSE plus faible et des coefficients de corrélation plus fort que ceux obtenus par (Schneider et al. 2017). Ces nouvelles valeurs d' A_{cr} sont calibrées à une échelle de 1 : 50000 et pourront être utilisées pour extraire des réseaux dans des zones où l'information sur le réseau hydrographique est limitée. Dans la perspective d'améliorer les performances des réseaux calibrés par rapport aux réseaux observés, il serait possible de réévaluer les paramètres environnementaux à prendre en compte. La prise en compte de la couverture végétale et l'occupation des sols pourrait permettre d'améliorer les coefficients de détermination des relations entre les densités de drainage et les paramètres environnementaux. Les réseaux Th_{total} , $Th_{naturel}$, $Th_{artificiel}$ ont été calibrés en toute indépendance mais ces réseaux ne sont pas indépendants car Th_{total} est la somme de $Th_{naturel}$ et $Th_{artificiel}$. On s'attendrait donc à avoir de meilleures performances avec le réseau $Th_{naturel}$ qu'avec Th_{total} car $Th_{artificiel}$ est a priori dure à prévoir. Cela suggère qu'il faudrait revoir l'encodage des réseaux $Th_{naturel}$ et Th_{total} . Par la suite, en utilisant les relations que nous avons établies afin de déduire les Aires critiques des différents types de réseau en fonction de facteurs environnementaux, il serait possible de se poser la question suivante : comment le réseau hydrographique est-il susceptible de évoluer avec le changement climatique ou certains changements d'occupation des sols ?

6 Annexes

Annexe 1 : Résultats de la calibration Acr (km²) définissant le modèle LCS en France et en Australie pour tous les paramètres environnementaux sélectionnés (Schneider et al., 2017).

Lithology groups	Slope Classes (%)				Lithology groups	Slope Classes (%)				Climate classes (mm.y-1)
	S < 1	1 ≤ S < 5	5 ≤ S < 10	S ≥ 10		S < 1	1 ≤ S < 5	5 ≤ S < 10	S ≥ 10	
Mt, Pa, Pi, Py	0,6	0,6	0,6	0,4	Su	30,8	15,4	3,1	1,5	Ptot < 250
	0,6	0,6	0,6	0,4		19,3	9,6	1,9	1	250 ≤ Ptot < 500
	0,5	0,5	0,5	0,3		4,3	1,7	0,7	0,3	500 ≤ Ptot < 1000
	0,5	0,5	0,3	0,3		0,7	0,7	0,7	0,3	1000 ≤ Ptot < 1500
	0,3	0,3	0,3	0,3		0,7	0,7	0,7	0,3	Ptot ≥ 1500
Ss	1	0,6	0,6	0,4	Sc	192,6	96,3	28,9	28,9	Ptot < 250
	1	0,6	0,6	0,4		57,8	19,3	3,9	3,9	250 ≤ Ptot < 500
	0,9	0,5	0,5	0,3		5,1	3,4	1,7	1,7	500 ≤ Ptot < 1000
	0,5	0,5	0,5	0,3		2,6	1,7	1,7	1,7	1000 ≤ Ptot < 1500
	0,5	0,5	0,3	0,3		1,7	1,7	1,7	1,7	Ptot ≥ 1500
Sm, Va, Vb, Vi	2,9	0,8	0,6	0,4	Pb	1,2	1,2	1,2	1,2	Ptot < 250
	2,9	0,8	0,6	0,4		1,2	1,2	1,2	1,2	250 ≤ Ptot < 500
	1,4	0,7	0,5	0,3		1	1	1	1	500 ≤ Ptot < 1000
	0,5	0,5	0,5	0,3		1	1	1	1	1000 ≤ Ptot < 1500
	0,5	0,5	0,5	0,3		0,5	0,5	0,5	0,5	Ptot ≥ 1500

Mt = Metamorphic rocks ; Pa= Acid plutonic rocks ; Pi = Intermediate plutonic rocks; Ss = Siliciclastic sedimentary rocks; Sm = Mixed sedimentary rocks; Va = Acid volcanic rocks; Vb = Basic volcanic rocks; Vi = Intermediate volcanic rocks; Su= Unconsolidated sediments; Sc = Carbonate sedimentary rocks; Pb = Basic plutonic rocks



Annexe 2 : Carte à l'échelle globale de la densité de drainage issue du modèle LCS (Schneider et al. 2017).

Annexe 3 : Récapitulatif des régressions linéaires multiples entre les réseaux Thnat_interm et Th_artificiel et les paramètres en fonction des différentes classes de lithologies

Thnat_interm	
Classe mt	
$y = 0.0008 * x_1 - 0.000006 * x_2 + 0.02 * x_3 + 0.04 * x_4 - 0.4$	r ² =0.30
Classe pa	
$y = 0.0006 * x_1 - 0.0002 * x_2 + 0.01 * x_3 - 0.02 * x_4 + 0.01$	r ² =0.15
Classe sc	
$y = 0.0008 * x_1 - 0.0004 * x_2 + 0.03 * x_3 - 0.02 * x_4 - 0.3$	r ² =0.40
Classe sm	
$y = 0.0026 * x_1 + 0.0001 * x_2 + 0.034 * x_3 - 0.014 * x_4 - 1.46$	r ² =0.46
Classe ss	
$y = 0.001 * x_1 - 0.00006 * x_2 + 0.02 * x_3 + 0.01 * x_4 - 0.5$	r ² =0.12
Classe su	
$y = -0.0003 * x_1 + 0.0001 * x_2 + 0.07 * x_3 + 0.015 * x_4 + 0.37$	r ² =0.18
Classe va+vb	
$y = 0.001 * x_1 - 0.000006 * x_2 + 0.00003 * x_3 - 0.06 * x_4 - 0.3$	r ² =0.17
Th_artificiel	
Classe mt	
$y = -0.00009 * x_1 - 0.00002 * x_2 + 0.002 * x_3 + 0.004 * x_4 + 0.06$	r ² =0.27
Classe pa	
$y = 0.00001 * x_1 - 0.00003 * x_2 + 0.001 * x_3 - 0.006 * x_4 + 0.03$	r ² =0.04
Classe sc	
$y = -0.000001 * x_1 - 0.00003 * x_2 + 0.004 * x_3 + 0.01 * x_4 - 0.01$	r ² =0.06
Classe sm	
$y = -0.00067 * x_1 - 0.00016 * x_2 - 0.0008 * x_3 + 0.024 * x_4 + 0.47$	r ² =0.07
Classe ss	
$y = -0.0001 * x_1 - 0.00003 * x_2 + 0.003 * x_3 - 0.005 * x_4 + 0.1$	r ² =0.09
Classe su	
$y = 0.002 * x_1 - 0.00055 * x_2 - 0.07 * x_3 - 0.075 * x_4 - 0.76$	r ² =0.23
Classe va+vb	
$y = 0.00002 * x_1 + 0.0001 * x_2 + 0.0001 * x_3 + 0.03 * x_4 - 0.1$	r ² =0.40

7 Références Bibliographiques

- Barraqué, Bernard. 2001. "Les Enjeux de La Directive Cadre Sur L'eau de l'Union Européenne." *Flux*, no. 4: 70–75.
- Bishop, K, Ishi Buffam, M Erlandsson, J Fölster, Hjalmar Laudon, Jan Seibert, and Johan Temnerud. 2008. "Aqua Incognita : The Unknown Headwaters." *Hydrological Processes* 22 (8): 1239–42.
- Cacas, J., G. Degoutte, A. Dutartre, F. Gross, M. Vuillot, and G. Degoutte. 1986. *River Management : Three Case Studies*. PARIS: Ministère de l'Environnement. <http://cemadoc.irstea.fr/cemoa/PUB00026538>.
- Cazorzi, Federico, Giancarlo Dalla Fontana, Alberto De Luca, Giulia Sofia, and Paolo Tarolli. 2013. "Drainage Network Detection and Assessment of Network Storage Capacity in Agrarian Landscape." *Hydrological Processes* 27 (4): 541–53.
- Dahlgren, Per, Tomas Landelius, Per Kållberg, and Stefan Gollvik. 2016. "A High-resolution Regional Reanalysis for Europe. Part 1: Three-dimensional Reanalysis with the Regional High-Resolution Limited-Area Model (HIRLAM)." *Quarterly Journal of the Royal Meteorological Society* 142 (698): 2119–31.
- Datry, Thibault, Scott T Larned, and Klement Tockner. 2014. "Intermittent Rivers : A Challenge for Freshwater Ecology." *BioScience* 64 (3): 229–35.
- Dunn, SM, and R Mackay. 1996. "Modelling the Hydrological Impacts of Open Ditch Drainage." *Journal of Hydrology* 179 (1-4): 37–66.
- Esri. 2017. "Présentation d'ArcGIS." <http://resources.arcgis.com/fr/help/getting-started/articles/026n00000014000000.htm>.
- FAO. 2014. "AQUAMAPS Global Spatial Database on Water and Agriculture, Food and Agriculture Organization of the United Nations." <http://www.fao.org/nr/water/aquamaps/>.
- Hartmann, Jens, and Nils Moosdorf. 2012. "The New Global Lithological Map Database GLiM: A Representation of Rock Properties at the Earth Surface." *Geochemistry, Geophysics, Geosystems* 13 (12).
- Horton, Robert E. 1945. "Erosional Development of Streams and Their Drainage Basins ; Hydrophysical Approach to Quantitative Morphology." *Geological Society of America Bulletin* 56 (3): 275–370.
- IGN. 2006. "Base de Données Sur La CARTographie THématique Des AGENces de L'eau et Du Ministère Chargé de l'Institut National de L'information Géographique et Forestière." http://www.sandre.eaufrance.fr/sites/default/files/IMG/pdf/1-DC_BDCARTHAGE_3_0.pdf.
- Lagacherie, P, O Diot, N Domange, V Gouy, C Floure, C Kao, R Moussa, JM Robbez-Masson, and V Szleper. 2006. "An Indicator Approach for Describing the Spatial Variability of Artificial Stream Networks with Regard to Herbicide Pollution in Cultivated Watersheds." *Ecological Indicators* 6 (2): 265–79.
- Lehner, B, K Verdin, and A Jarvis. 2008. "New Global Hydrography Derived from Spaceborne Elevation Data." *Eos* 89 (10): 93–94.
- Marsily, Ghislain De. 2009. *L'eau, Un Trésor En Partage*. DUNOD.
- Micheaux, Pierre Lafaye, Rémy Drouilhet, and Benoît Liquet. 2011. "Statistique Descriptive." *Le Logiciel R*, 255–97.
- Oudin, Ludovic, Claude Michel, and François Anctil. 2005. "Which Potential Evapotranspiration Input for a Lumped Rainfall-Runoff Model?: Part 1—Can

- Rainfall-Runoff Models Effectively Handle Detailed Potential Evapotranspiration Inputs?" *Journal of Hydrology* 303 (1): 275–89.
- Passalacqua, Paola, Paolo Tarolli, and Efi Foufoula-Georgiou. 2010. "Testing Space-scale Methodologies for Automatic Geomorphic Feature Extraction from Lidar in a Complex Mountainous Landscape." *Water Resources Research* 46 (11): 1–17.
- Paul, Michael J, and Judy L Meyer. 2001. "Streams in the Urban Landscape." *Annual Review of Ecology and Systematics* 32 (1): 333–65.
- Pella, Hervé, André Chandesris, and Jean-Gabriel Wasson. 2004. "Constitution Dun Système D'information À Référence Spatiale Dans Le Contexte de La Directive Cadre Européenne Sur Leau." *Ingénieries-EAT*, no. 40: 1–10.
- Pella, Hervé, Jérôme Lejot, Nicolas Lamouroux, and Ton Snelder. 2012. "Le Réseau Hydrographique Théorique (RHT) Français et Ses Attributs Environnementaux." *Géomorphologie: Relief, Processus, Environnement* 18 (3): 317–36.
- Pella, Hervé, Éric Sauquet, and André Chandesris. 2006. "Construction D'un Réseau Hydrographique Simplifié À Partir de La BD Carthage®." *Ingénieries-EAT*, no. 46: 3–14.
- Raymond, Peter A, Jens Hartmann, Ronny Lauerwald, Sebastian Sobek, Cory McDonald, Mark Hoover, David Butman, Robert Striegl, Emilio Mayorga, and Christoph Humborg. 2013. "Global Carbon Dioxide Emissions from Inland Waters." *Nature* 503 (7476): 355–59.
- Salavati, Bahar, Ludovic Oudin, Carina Furusho-Percot, and Pierre Ribstein. 2016. "Modeling Approaches to Detect Land-Use Changes: Urbanization Analyzed on a Set of 43 US Catchments." *Journal of Hydrology* 538: 138–51.
- Schneider, A, A Jost, C Coulon, M Silvestre, S Théry, and A Ducharne. 2017. "Global-scale River Network Extraction Based on High-resolution Topography and Constrained by Lithology, Climate, Slope, and Observed Drainage Density." *Geophysical Research Letters* 44 (6): 2773–81.
- Snelder, Ton H, Thibault Datry, Nicolas Lamouroux, Scott T Larned, Eric Sauquet, Hervé Pella, and Clotaire Catalogne. 2013. "Regionalization of Patterns of Flow Intermittence from Gauging Station Records." *Hydrology and Earth System Sciences* 17 (7): 2685–99.
- Vogt, Jürgen V, Roberto Colombo, and Francesca Bertolo. 2003. "Deriving Drainage Networks and Catchment Boundaries: A New Methodology Combining Digital Elevation Data and Environmental Characteristics." *Geomorphology* 53 (3): 281–98.
- Weedon, Graham P, Gianpaolo Balsamo, Nicolas Bellouin, Sandra Gomes, Martin J Best, and Pedro Viterbo. 2014. "The WFDEI Meteorological Forcing Data Set : WATCH Forcing Data Methodology Applied to ERA-Interim Reanalysis Data." *Water Resources Research* 50 (9): 7505–14.

8 Résumé

La caractérisation précise des cours d'eau est un enjeu majeur pour mieux connaître le milieu et participer à la bonne gestion de la ressource en eau. Cette caractérisation se fait de plus en plus à l'aide de MNT (Modèle Numérique de Terrain). Ce stage s'appuie sur l'étude de (Schneider et al. 2017) qui propose à l'aide d'un MNT de 15 " issue de la base de données HydroSHEDS et de paramètres environnementaux (Lithology, Climate, Slope), une série de valeurs d'aires critiques (surface contributive minimum pour initier un cours d'eau) afin de mieux caractériser les cours d'eau. Ces aires critiques sont issues d'une calibration en France à l'aide de la BD Carthage (Base de Données sur la CARTographie THématique des AGences de l'eau) et en Australie à l'aide de la base de données AHGF (Australian Hydrological Geospatial Fabric). Mais le modèle de réseaux hydrographiques globale dérivés à partir de ces aires critiques (LCS) présente des erreurs . En nous focalisant sur la France continentale, nos analyses nous montrent que l'artificialisation et à l'intermittence des réseaux hydrographiques expliquerait 40% de la variance de l'erreur du modèle et qu'une part importante de l'erreur porte sur le choix de la couche de référence. Une recalibration du modèle en tenant compte de la couche « Tronçons hydrographiques » qui est mieux adapté , nous a permis d'avoir de meilleures performances.

Mots-clés : Hydrographie, ressource en eau, lithologie, climat, topographie, MNT