



**HAL**  
open science

## **CARTO ALEA (Zones inondables) 4.1.4.c : Intégration des Lames d'Eau Radar dans la méthode IRIP : IRIP\_LER**

Jérôme Pinta, Christine Poulard, Pascal Breil

► **To cite this version:**

Jérôme Pinta, Christine Poulard, Pascal Breil. CARTO ALEA (Zones inondables) 4.1.4.c : Intégration des Lames d'Eau Radar dans la méthode IRIP : IRIP\_LER. [Rapport de recherche] INRAE UR-RiverLy. 2020. hal-03787479

**HAL Id: hal-03787479**

**<https://hal.inrae.fr/hal-03787479>**

Submitted on 25 Sep 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

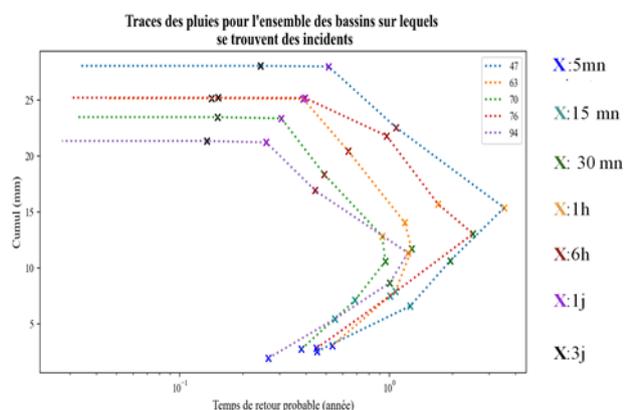
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**PROGRAMME MTES (DGPR / SRNH) - INRAE 2020**  
**Connaissance et prévention des risques naturels et hydrauliques**

Ces travaux bénéficient du soutien du Ministère chargé de l'environnement, de l'énergie et de la mer  
par décision de subvention n°21367400

## 4.1 : CARTO ALEA (Zones inondables)

### 4.1.4.c : Intégration des Lames d'Eau Radar dans la méthode IRIP : IRIP\_LER



DECEMBRE 2020

**Pinta Jérôme**  
**Poulard Christine**  
**Breil Pascal**

INRAe, Unité de recherche RIVERLY  
5, rue de la Doua  
CS 20244,  
69625 Villeurbanne cedex

[www.inrae.fr](http://www.inrae.fr)



## Préambule

Cette action répond en grande partie aux objectifs qui étaient fixés, cela malgré la situation sanitaire. Les missions de terrains n'ont pu être réalisées en totalité mais des données acquises par ailleurs ont permis de combler ce déficit. Cette action regroupe le travail d'un stagiaire de 6 mois en M2, co-encadré par C. Poulard et P. Breil, ainsi que des résultats produits par P. Breil.

L'objectif de cette action était d'étudier la variabilité spatiale des cumuls de lames d'eau, sur différentes durées, au cours d'événements orageux importants, puis de comparer ces cumuls aux zones endommagées par le ruissellement intense. Ceci afin d'évaluer si le gradient de sensibilité au ruissellement proposée par la méthode IRIP faisait sens selon les cumuls de pluie dans les zones répertoriées avec impact. Pour cela l'hypothèse formulée a été que :

**« Dans une zone endommagée par le ruissellement, plus la susceptibilité au ruissellement intense est élevée, moins le cumul de pluie nécessaire à produire un incident est élevé ».**

Cela revient à vérifier l'existence d'une relation négative entre le cumul de pluie et la susceptibilité au ruissellement intense dans les zones endommagées.

Ces résultats sont dans la continuité de la même action 4.1.4 depuis 2018. En 2019 un module de traitement des lames d'eau radar a été mis à jour et réutilisé cette année 2020, en coordination avec C.Poulard qui en développe l'architecture.

La fiche action est visualisée ci-après pour rappel.

<b>Titre complet de l'action ou s/ action (acronyme éventuel)</b>	<b>Intégration des Lames d'Eau Radar dans la méthode IRIP : IRIP_LER</b>		
<b>Action n°</b>	4.1.4c		
<b>Responsable Irstea (Unité)</b>	P. Breil (RIVERLY) <span style="color:red">coût complet: 47 € demande :20 € taux :43%</span>		
<b>Responsable SRNH (Bur.)</b>	J.M. Helmer		
<b>Contexte et enjeux</b>	Les désordres divers (coulées de boue, emportement et encombrement des réseaux de transport) attribuables à des épisodes de ruissellement intense sont présents sur tout le territoire français et cela malgré des différences de climats, de régimes de pluies fortes, de géologie et de pédologie.		
<b>Objectifs du projet</b>	Rapprocher les cartes de production IRIP et des événements survenus au cours d'une période de référence, de façon à préciser le rapport entre la susceptibilité IRIP (zones fragiles a priori), les événements dommageables survenus, et les pluies déclenchantes, exprimées tant en absolu sur le support spatial et la durée pertinente (mm x surface x durée) qu'en période de retour.		
<b>Travaux déjà réalisés et travaux prévus</b>	<p><b>Rappel des acquis si caractère pluriannuel de l'action :</b>            En 2019, la méthode d'évaluation des quantités de pluie nécessaires qui ont a priori déclenché les incidents constatés au sol a été testée sur un nouveau territoire (l'Aude) avec succès. Ces quantités ont été estimées pour des événements datés à partir des lames d'eau radar de Météo-France en calculant les cumuls de pluie sur des durées de 15, 30 et 60 minutes. L'estimation des IdF associés à partir des lames d'eau radar a fait l'objet d'une collaboration avec l'action de simulation continue portée par C. Poulard. Cela a permis la mise au point d'un algorithme avec interface utilisateur pour évaluer les IdF régionaux et le caractère multi-fréquences des lames d'eau intenses liées à des événements datés. <b><u>A ce titre la mise à jour des données de lame d'eau radar à partir de 2016 est demandée par l'entremise de la DGPR.</u></b></p> <p><b>Travaux prévus :</b></p> <ol style="list-style-type: none"> <li>1- Améliorer l'étalonnage en périodes de retour des lames d'eau radar (spécialisées, ce qui est un avantage s'agissant d'événement intéressant des zones de collecte) en les rapprochant et recalant sur les quantiles de pluies Shyreg (ponctuelles, mais qui disposent d'une meilleure stabilité statistique due à la profondeur historique des observations utilisées). Les savoir-faire à mobiliser sont listés en bibliographie.</li> <li>2- Dérouler la méthodologie présentée dans l'action 4.1.4 2018 pour évaluer le lien entre les événements générateurs de désordre, en absolu et en période de retour, et les niveaux de susceptibilité IRIP concernés. Pour cela les données d'incidents acquises sur les départements 11, 34, 66, 71 et 83 seront utilisées.</li> <li>3- Des déplacements sont prévus auprès des services de l'Etat détenteurs des données DDTM et EPTP pour présenter les résultats et les discuter.</li> </ol>		
<b>Durée si action pluriannuelle</b>	<i>Durée</i>	<i>Début de l'action</i>	<i>Fin de l'action</i>
	2 ans	Année 2020	Année 2021
<b>Produits et valorisations projetés</b>	<b>Nature</b> (publi scient, technique, guide methodo, outil, formation ...)	<b>Cibles visées</b>	<b>Echéancier</b> (au moins un rapport d'étape par an)
	1/ Publication technique (SET ou autre).	Services de l'Etat, collectivités	un rapport chaque année
<b>Equipe de travail Irstea (noms, Unités)</b>	Breil, P. ; Leblois, E; Poulard C. ; + 1 stage master 2 (Riverly); P. Arnaud (données Shyreg / thèse descente d'échelle)		
<b>Partenariats envisagés (noms, rôle)</b>	Mise à jour de la couche lame d'eau radar de Météo-France depuis Août 2016 valider avec la DGPR auprès de Météo-France et SCHAPI.		
<b>Références ou bibliographie sommaire</b>	-Poulard C., Leblois E., Faure J.B., Breil P. « Critique argumentée des pluies de bassin et crues de projet représentatives d'une période de retour ». Article de Novatech 2019 sélectionné pour numéro spécial TSM 2019. -Arnaud P. et Breil P. (2019) Cartographie du ruissellement : Application des méthodes SHYREG et IRIP sur deux cas d'étude. Rapport étude Irstea. 36 p.		
<b>Date de mise à jour de la fiche :</b>	20/11/2019		

# TABLE DES MATIERES

## Table des matières

1.	<b>Sigles</b> .....	8
2.	<b>Introduction</b> .....	1
3.	<b>Contexte de l'étude</b> .....	1
a.	Le ruissellement.....	1
b.	Cartographier le ruissellement.....	1
4.	<b>Fonctionnement d'IRIP</b> .....	2
5.	<b>Indicateurs et sources de données</b> .....	3
6.	<b>Les zones d'études</b> .....	4
7.	<b>L'interface "Pantherogation"</b> .....	5
8.	<b>Matériel et méthode</b> .....	6
a.	La donnée pluie .....	6
9.	<b>Théorie autour des lames d'eau radar</b> .....	6
a.	Le réseau ARAMIS.....	7
b.	Les bases de données disponibles.....	7
c.	Traitements appliqués à la réception des données.....	8
d.	Méthode d'exploitation.....	9
e.	Règles d'écriture des scripts.....	9
f.	Intégration des fichiers géographiques.....	9
g.	Analyse multi-durées.....	10
h.	Caractérisation par étude des évènements extrêmes .....	10
i.	La méthode SUPSEUIL .....	11
j.	Représentation graphique.....	13
k.	Exploitation des résultats IRIP .....	14
l.	Lecture des cartes.....	14
m.	Vérification des résultats.....	14
n.	Estimation des bassins versants .....	15
10.	<b>Résultats</b> .....	15
a.	Les scripts .....	16
b.	Seuil automatique .....	17
c.	Etude de cas : Le Grand Chalon .....	18
d.	Etude de cas : Les Pyrénées Orientales .....	22
e.	Résultats IRIP .....	22
f.	Les sous bassins versants .....	23

11.	<b>Discussion</b> .....	25
a.	Problèmes liés à la taille de la base PANTHERE.....	26
b.	Analyse des cas d'étude IRIP .....	26
c.	Apport de la méthode .....	27
d.	Limites .....	28
e.	Perspectives.....	28
12.	<b>Conclusion</b> .....	29
13.	Bibliographie :.....	30
14.	Sitographie :.....	30
15.	Annexes A: Script étude des bassins versant .....	31
16.	Annexes B : Script détermination automatique du seuil.....	32
17.	Annexe C : Echantillonnage .....	33
18.	Annexe D : Sélection par nuage de points.....	34
19.	Annexe F : Cartes IRIP pour les 3 bassins témoins des Pyrénées Orientales .....	36
20.	Annexe G : Relation Aléa vs Cumul max & Test de la pente .....	37



# 1. Sigles

ARAMIS : Application Radar à la Météorologie Infra-Synoptique

ESDAC : European Soil DATA Centre

GEV : loi des valeurs extrêmes généralisées

GPD : loi de Pareto généralisée

HER : HydroEcoRegion

IdF : Courbe Intensité Durée Fréquence

IGN : Institut national de l'information Géographique et forestière

PANTHERE : Projet ARAMIS Nouvelles Technologies en Hydrométéorologie Extension et RENouvellement

IRIP : Indicateur de Ruissellement Intense Pluvial

RADAR : RADio Detection And Ranging

## 2. Introduction

Les inondations sont des catastrophes naturelles récurrentes qui causent des dégâts matériels et humains importants. La Directive Inondation de 2007 de l'Union Européenne renforce la nécessité de comprendre et de représenter correctement l'aléa inondation. Parmi les processus initiaux, le débordement de cours d'eau est plutôt bien compris et des méthodes et outils existent. Au contraire, le ruissellement intense, qui lui est souvent associé, est moins facile à caractériser (chemins complexes et changeants, superposition avec des phénomènes d'érosion et de débordement de cours d'eau). Pourtant, selon la Caisse Centrale de Réassurance, il serait responsable de 45% des pertes financières dues aux inondations (Moncoulon *et al.*, 2014). Jaillet et Poulard (2012) ont montré que d'un point de vue Français comme Européen, il n'existe pas de méthode standard pour représenter le ruissellement et d'un auteur à l'autre la définition change ainsi que les variables cartographiées.

La méthode IRIP a été développée au sein de l'IRSTEA, devenu INRAE, par Dehotin et Breil, (2011a, 2011b), dans le but de proposer une méthode adaptée au ruissellement. Elle se veut simple et utilise les informations spatialisées disponibles pour en déduire des aléas de ruissellement. Des études de validation ont montré une bonne correspondance entre scores et présence d'épisodes de ruissellement (incidents), mais la précision varie encore régionalement (Breil *et al.*, 2018).

IRIP est une méthode par combinaison d'indicateurs qui sont de trois sources d'information : le sol caractérisé par ses propriétés hydrodynamiques, l'occupation du sol et la topographie. Récemment, la prise en compte des cumuls de pluie de 1 heure et de 24 heures, pour une période de retour de 30 ans, a permis d'exprimer les processus de ruissellement hortonien et par saturation du profil de sol, en comparant les cumuls avec les capacités hydrodynamiques des sols. Pour la France métropolitaine, la donnée lame d'eau radar est utilisée. Cela offre l'avantage d'exprimer une susceptibilité au ruissellement intense pour une période de retour donnée de la pluie. C'est une vision « en moyenne » qui intègre l'information de multiples événements qui se sont déroulés dans le passé. La comparaison avec des incidents géolocalisés implique alors de disposer d'une base de données de ces incidents sur la même période d'observation que celle des pluies.

Dans le cas d'un événement orageux particulier, la distribution spatiale des pluies est elle-même particulière. C'est dans cet objectif que s'intègre ce travail. L'idée est de développer des méthodes et outils qui permettront par la suite de qualifier les pluies responsables des dommages issus du ruissellement intense lors d'un événement orageux particulier.

Dans ce contexte, nous avons utilisé les lames d'eau radar de la base de donnée PANTHERE de Météo-France. Elles sont intéressantes car elles fournissent des données continues sur la structure dans l'espace, à la maille de 1 kilomètre carré, d'un événement décrit au pas de temps de 5 minutes. Pour parvenir à les exploiter rapidement et confortablement, nous avons écrit des scripts Python. Notre objectif est de les associer à une interface développée à l'INRAE par Christine Poulard. Pour tester nos méthodes nous les avons appliquées à deux cas d'étude IRIP : le Grand-Châlon sur lequel plusieurs inondations sont survenues en Juin 2016 et dans les Pyrénées Orientales pour lesquelles c'est arrivé en Novembre 2014.

Au fil de ce rapport, nous commencerons par expliquer le contexte de l'étude et l'origine des données. Nous décrirons les méthodes développées pour l'utilisation des données au service de la méthode IRIP. Nous expliquerons l'intérêt d'analyses multi-durées, afin d'estimer des relations "Intensité-durée-Fréquence" des pluies pour caractériser le régime, sur des pixels et des ensembles de pixels (IdF pour une surface donnée). Nous appliquerons cette méthode aux cas d'étude. Pour terminer, nous discuterons des résultats, des avantages de la méthode, de ses limites et nous suggérerons des pistes pour la faire évoluer.

## 3. Contexte de l'étude

Ce travail s'inscrit dans la continuité du projet IRIP développé par des membres de l'unité Riverly de l'INRAE. Nous ne travaillerons pas directement sur la méthode ici, mais plutôt sur des outils qui permettront de lier ses résultats à des données de précipitations. Avant donc de passer à l'étude en tant que telle, il nous semble important de préciser les éléments de contexte et particulièrement le fonctionnement de la méthode IRIP.

### a. Le ruissellement

Le ruissellement est l'objet d'étude de la méthode IRIP. Il est nécessaire pour la décrire de comprendre son fonctionnement.

Le ruissellement peut être défini comme la part des précipitations qui ne s'infiltrent pas dans le sol. Dans la littérature, deux processus de formation de l'écoulement de la lame ruisselante sont prédominants :

- Le ruissellement par dépassement de la capacité d'infiltration (Horton, 1933). L'intensité des précipitations est plus importante que de débit maximal qui peut transiter par la porosité. Par conséquent, l'eau va avoir tendance à s'accumuler en surface. Ce type de ruissellement est très dépendant des conditions de surface et notamment du paramètre de capacité d'infiltration.
- Le ruissellement sur surface saturée, ou ruissellement hewlettien. Le volume maximal d'eau que peut contenir le sol est atteint, peut être suite à des pluies récentes, ou encore du fait de la présence d'une nappe qui affleure.

Le concept du ruissellement est plutôt simple, il dépend en premier lieu de la nature de la surface sur laquelle la pluie tombe (Nicolas, 2010). Il est aussi relié à un ensemble de processus qui le rendent plus compliqué à étudier (convergence ou divergence d'écoulements, infiltration en chemin, érosion, colmatage, effet de la microtopographie...). Mais aussi forcément, suivant les deux définitions ci-dessus, ça genèse dépend des caractéristiques des précipitations elles-mêmes.

### b. Cartographier le ruissellement

Derrière le terme d'étude du ruissellement, il existe des objectifs et des méthodes différentes. Il est parfois compliqué de trouver celle qui permettra de répondre à un problème donné.

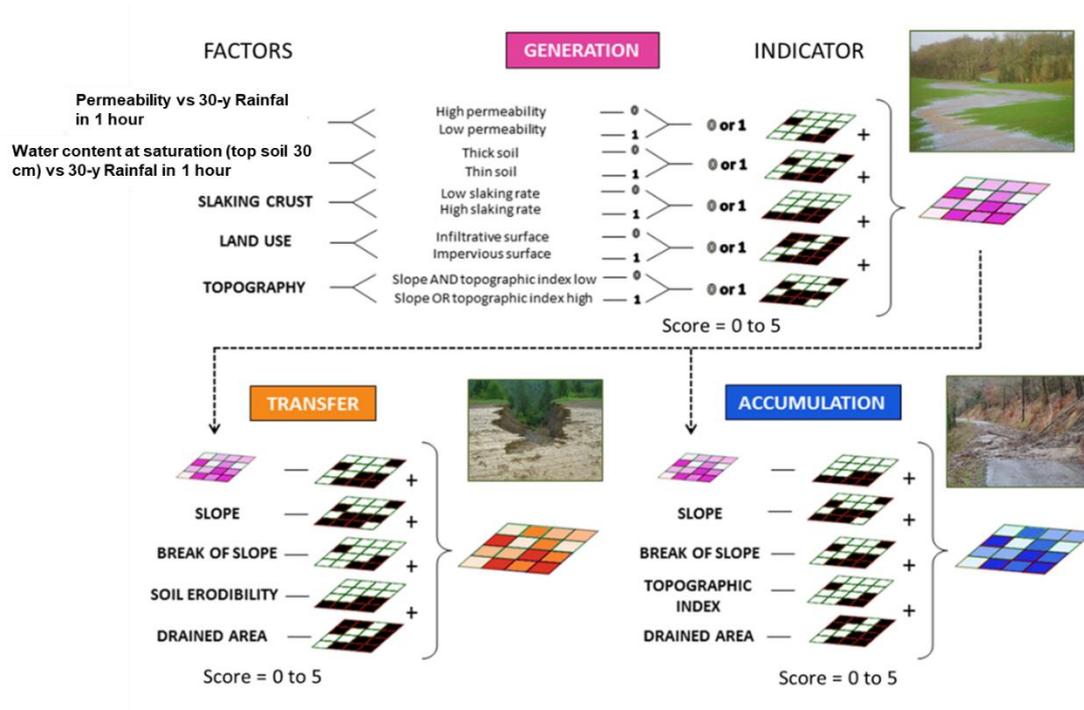
La représentation du ruissellement est nécessaire dans le contexte actuel mais est peu commune (Moncoulon *et al.*, 2014). La réalisation des cartes est plutôt concentrée sur le débordement des rivières et autres inondations par la mer (Figure 1). Un certain nombre d'auteurs ont essayé, mais les définitions, les méthodes et les objectifs diffèrent. Certaines études vont chercher à quantifier le débit qui est généré par le ruissellement sur un territoire, pour une probabilité donnée, tandis que d'autres vont cartographier les chemins de l'eau au sein du territoire (Jaillet *et al.*, 2012).

Avec l'évolution des techniques de SIG, le croisement d'informations spatialisées apporte de nouvelles possibilités pour résoudre le problème. C'est en se basant sur ces nouveaux outils qu'a été pensée la méthode IRIP. Son objectif est de proposer une approche robuste (pour être en capacité d'analyser des régions différentes) basée sur un ensemble d'indicateurs réduit mais fiables et spatialisés.

## 4. Fonctionnement d'IRIP

Le modèle IRIP est une méthode géomatique qui produit trois cartes. La première concerne la production du ruissellement (a), la deuxième son transfert (b) et la troisième son accumulation (c) (figure 2, 3) (Braud *et al.*, 2020). La méthode est adaptée pour être traitée sur QGIS.

Des facteurs explicatifs sont nécessaires pour chaque carte (5 par carte) (Figure 2). Il en existe trois types : topographique, d'occupation du sol et de type de sol (Lagadec, 2016). Récemment, des comparaisons entre les pluies et deux paramètres de sol, capacité de stockage et perméabilité ont été intégrées. Les pluies concernées sont des cumuls pour une durée de 1 heure et de 24 heures pour une période de retour 30ans (voir figure 2). De cette manière, on vérifie si une pluie (30ans) est susceptible de provoquer du ruissellement hortonien ou hewlettien.



\*Si la perméabilité d'un sol est en capacité d'infiltrer la pluie de 1 heure et de période moyenne de retour 30ans (calculée pour la zone sur laquelle se trouve le sol), la valeur binaire est de 0 sinon 1.

\*\* Si la capacité de stockage en eau du sol est suffisante pour stocker d'infiltrer la pluie de 24 heures et de période moyenne de retour 30ans, la valeur est de 0 sinon 1.

**Figure 2 :** Utilisation des indicateurs pour la construction des cartes produits du modèle IRIP. (adapté de Braud, Lagadec *et al.*, 2020, et selon P.Breil, communication personnelle)

La méthode IRIP fonctionne par combinaison d'indicateurs. Nous les obtenons à partir des facteurs explicatifs et d'un seuil. Si la valeur est supérieure au seuil, il a plutôt tendance à favoriser le processus décrit par la carte concernée, sinon c'est le contraire. Dans le premier cas l'indicateur associé prend la valeur de 1, dans le deuxième il prend la valeur de 0. Pour chaque pixel de la carte, les indicateurs sont additionnés et donnent des scores qui peuvent aller de 0 à 5.



**Figure 3 :** Illustration des cartes IRIP de gauche à droite, pour la Production (niveaux 4 et 5), le Transfert (niveaux 3,4 et5) et l'Accumulation (niveaux 4 et 5) sur le bassin versant de Givry (71) qui fait 4 km<sup>2</sup>. Les incidents liés à l'orage du juin 2016 sont reportés (points jaunes).

L'aléa est caractérisé par un score sur une échelle allant de 0 (aléa négligeable) à 5 (le plus fort). Les zones de Production fournissent l'écoulement. Les zones de Transfert et d'Accumulation sont les aléas considérés à l'origine des incidents. Les surfaces concernées sont plus réduites.

## 5. Indicateurs et sources de données

**Tableau 1 :** Les données utilisées en entrée de IRIP.

Type d'indicateur	Fichier SIG	Informations apportées	Résolution (m)	Source	Bibliographie
Type de sol	Vecteur "Mixmap"	-conductivité hydraulique (Ks) - Epaisseur - Battance - Erodibilité.	250	ESDAC	Tóth <i>et al.</i> , 2017
Occupation du sol	Vecteur OSO	23 classes d'occupation	20	OSO-CESBIO	
Topographie	Modèle numérique de terrain	Élévation	25	IGN	

La résolution de chacune des cartes en entrée montre que celle du sol est la moins précise de toute. Toutes les cartes sont ramenées à celle de la maille raster du MNT utilisé. Dans le cas de ce travail il s'agit du MNT à 25m de l'IGN. Le MNT est exploité par QGIS pour ressortir les informations topographiques. Différentes fonctions sont utilisées pendant la procédure, elles permettent par exemple d'extraire le réseau de drainage ou d'établir une carte des ruptures de pente.

L'intégration des temps de retour au modèle est nouvelle. La durée minimale de la pluie intense considérée est de 1 heure à partir des données disponibles sur la zone d'étude du Grand Châlon. Pour estimer cette valeur, une comparaison a été réalisée entre les pluies et les aléas. La durée de 15, 30 minutes et une heure présentent une tendance décroissante significative des cumuls nécessaires à produire des incidents plus les aléas IRIP sont élevés au même endroit (cette méthode sera à nouveau abordée à la fin de ce rapport) (Breil & Gally, 2018). La durée de 24h a été retenue pour prendre en compte le processus de saturation du profil supérieur du sol (horizon 0 à 30 cm). Ces calculs sont possibles du fait de la connaissance des caractéristiques hydrodynamiques du sol supérieur qui sont fournies par l'ESDAC.

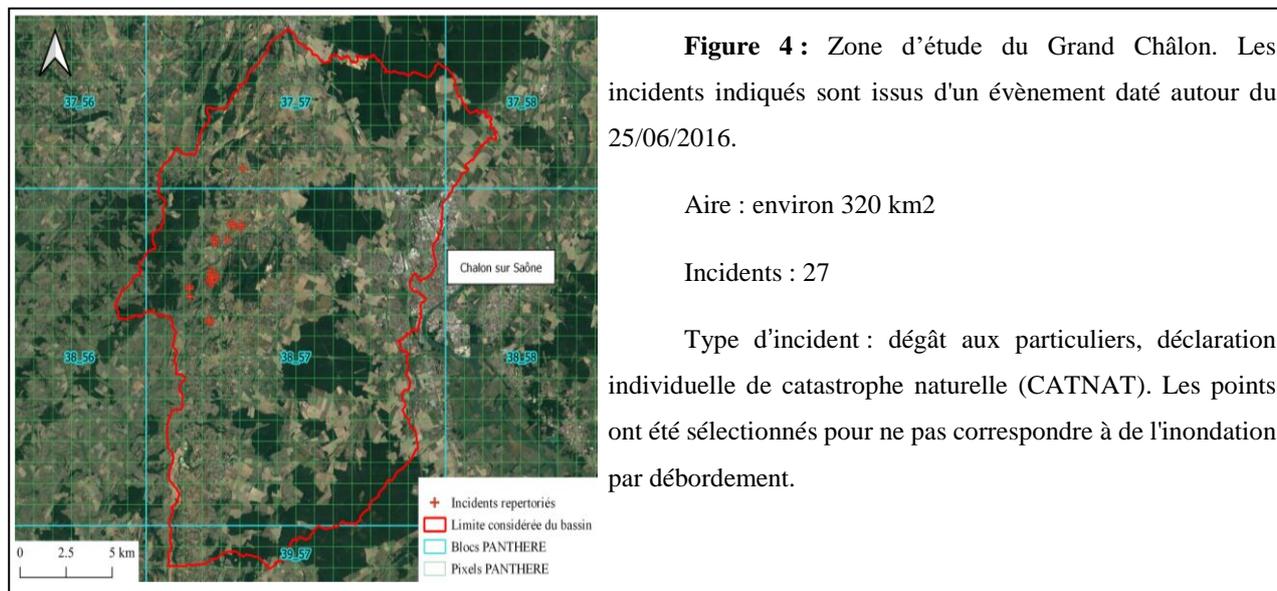
## 6. Les zones d'études

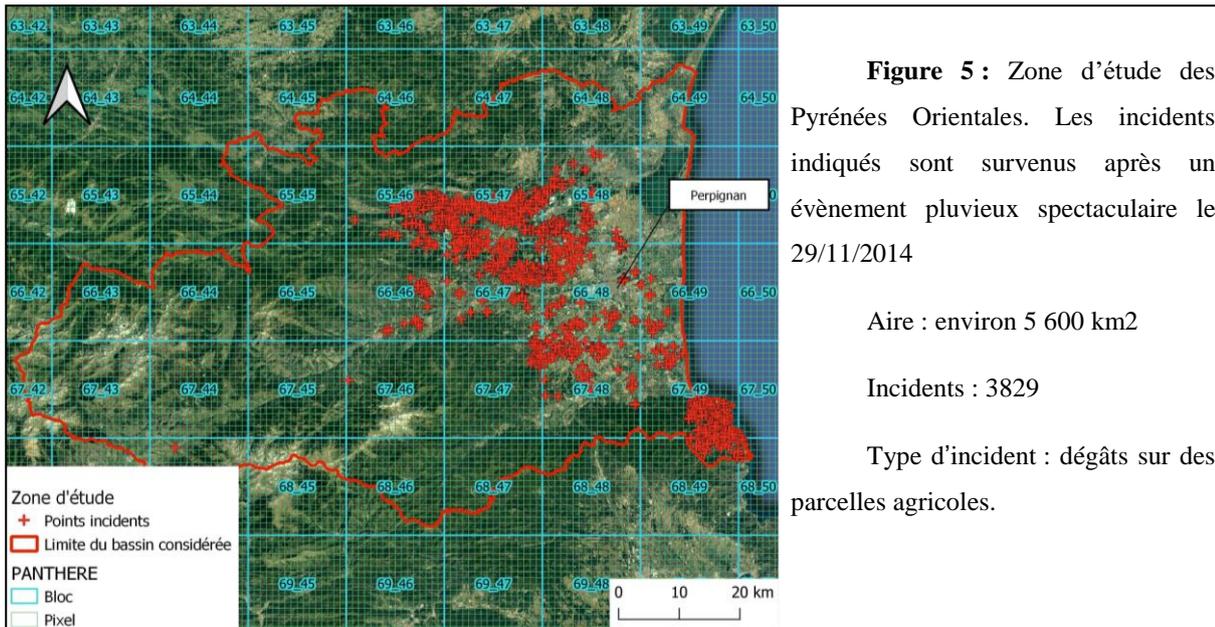
L'étude des pluies a été spécifiquement dirigée sur deux zones test IRIP (Figure 4, 5). Les deux ont été sélectionnées pour mesurer l'efficacité du modèle en fonction de la présence d'incidents passés liés à des orages pour lesquels la donnée lame d'eau radar spatialisée est disponible. Un incident est défini par un lieu et une date pour lesquels des dégâts dus au ruissellement ont été constatés. Dans ce rapport, les éléments qui ont subi des dégâts sont de trois types : des éléments urbains référencés par des coordonnées géographiques, des parcelles agricoles, des coupures de route. Ces éléments passés nous servent de point de comparaison pour tenter d'estimer l'efficacité de la méthode sur un territoire. Nous emploierons conformément aux précédents documents réalisés sur le sujet le terme "d'incident prédit" pour définir les incidents apparus sur des zones pour lesquelles l'aléa IRIP dépasse un seuil d'effet (partie "Matériel et méthode"). Ce sont donc des zones pour lesquelles le modèle indique un fort potentiel d'incident sous condition de pluie suffisante.

### Les zones d'étude suivent les HydroEcoRégions (HER) de Wasson (2004):

1- La région de Chalon (GC) (figure 4) est marquée par un relief plutôt type plaine, la géologie est essentiellement composée de roches sédimentaires détritiques. Son climat est océanique méridional (donc plutôt tempéré). Les hivers sont doux et humides et les étés moyennement chauds. La pluie y est plutôt équilibrée pendant l'année même si elle est globalement plus faible en Juillet/Août. Les plaines sont prédominantes à l'est et l'ouest est un peu plus vallonné.

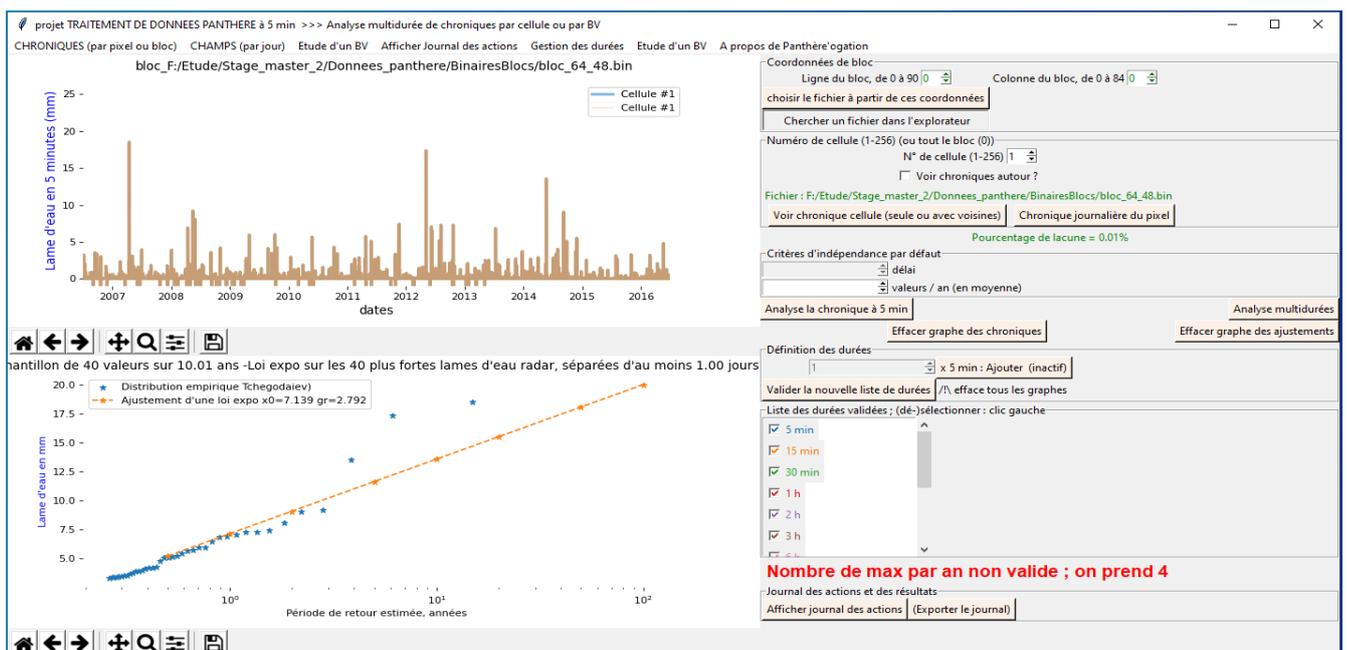
2- Le département des Pyrénées orientales (PO)(figure 5) est réparti entre deux HER : l'HER Pyrénées sur la partie Ouest. Il s'agit d'une région montagneuse avec de forts reliefs et des roches granitiques et métamorphiques. Ces caractéristiques au sol sont favorables au ruissellement. A l'Est, nous sommes sur l'HER Méditerranéen dans laquelle nous trouvons la plupart des points incidents. Le relief y est moins marqué et la géologie hétérogène. Le climat est méditerranéen avec des étés très chauds et des pluies très inégalement distribuées dans l'année, avec des événements très intenses plutôt en automne et en hiver.





## 7. L'interface "Pantherogation"

L'interface « Pantherogation » (figure 6) est développée au sein de Riverly par Christine Poulard. Elle est destinée à faciliter l'utilisation des différentes fonctions développées pour visualiser et analyser les données PANTHERE. L'interface permet au début du stage l'étude des précipitations par choix de la cellule à étudier (pixel PANTHERE). La figure 6 ci-dessous présente l'interface telle qu'elle était au début du stage. Le canevas du haut montre une chronique de pluie pour un pixel individuel pour un pas de temps de 5 minutes. Ces graphiques permettent de se faire une idée de la situation climatique de la zone. Celui du bas montre l'IDF réalisé à partir des données de la cellule, les détails de l'échantillonnage seront mentionnés dans la partie "Matériel et méthode".



**Figure 6 : Présentation de l'interface « Pantherogation » telle qu'elle était au début du stage (les menus « étude d'un BV » étaient déjà présents mais non actifs).**

En règle générale, le canevas du haut permet d'afficher les chroniques, éventuellement agrégées temporellement suivant plusieurs durées de pluies, sur l'ensemble des années disponibles ou réduites à un évènement. Le canevas du bas représente les analyses opérées sur la pluie. Il rend aussi possible l'utilisation des données par champs journaliers (Précipitations journalières sur une surface donnée), et l'étude d'un évènement compris entre deux dates. A terme, cette interface est destinée à faciliter les études d'évènements associés à IRIP. Le cahier des charges du stage impose donc de développer les scripts Python de manière à pouvoir les y intégrer.

## 8. Matériel et méthode

L'essentiel du travail a été de réaliser des scripts python pour exploiter les données PANTHERE, puis de mettre en relation nos résultats avec ceux d'IRIP. Le but est de comprendre les différents évènements survenus sur les zones d'études pour pouvoir discuter de leur différence et de l'efficacité de la méthode.

Dans cette partie, nous commencerons par expliquer l'origine des lames d'eau radar, la nature de la donnée, ses avantages mais aussi ses inconvénients. Ensuite, nous expliquons la méthode que nous avons employé pour traiter les données, l'outil employé, puis les théories statistiques utilisées. Et pour finir, nous détaillons comment nous avons utilisé la méthode IRIP dans ce contexte.

### a. La donnée pluie

L'objectif de cette étude sur les pluies est de développer des méthodes et des outils qui permettront de facilement caractériser des évènements liés à des incidents et permettre ainsi de comprendre les taux de prédiction de la méthode IRIP sous contrainte de quantité et de durée de pluie.

Les pluies peuvent être mesurées sur le terrain par des moyens physiques comme les pluviomètres. Ces mesures de la pluie, bien que précises, sont isolées et assimilables à une zone restreinte. Pour être généralisées à des zones étendues, ces données nécessitent de solides notions de géostatistiques (ex : Renard *et al.*, 2009).

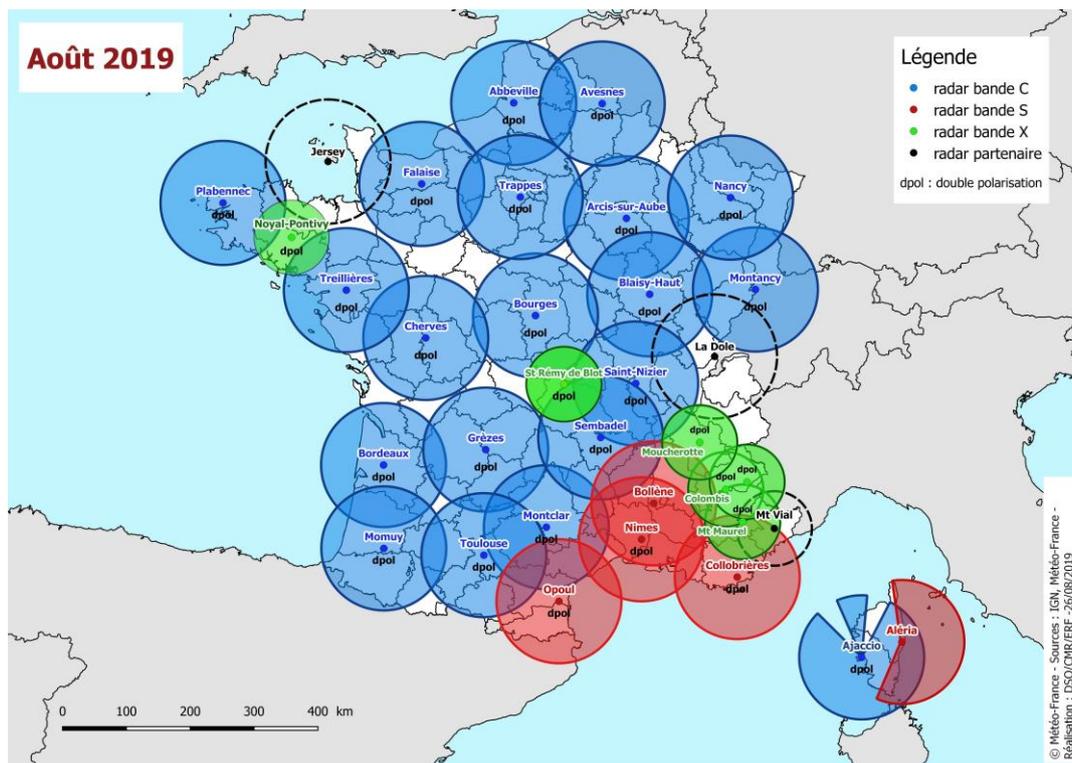
Les lames d'eau radar permettent d'obtenir un jeu de donnée continu dans l'espace et facilement exploitable. N'oublions pas qu'elles représentent cependant une information « moyennée » sur la cellule de calcul, selon des algorithmes complexes qui exploitent un écho radar. Grâce à un maillage spatial et une résolution temporelle assez précise, elles offrent la possibilité d'analyser la structure spatio-temporelle d'un évènement pluvieux. Cependant, la nature de la donnée (tirée de mesures indirectes) implique un certain nombre d'incertitudes ; les informations sur les traitements appliqués sont encore vagues et souvent la donnée nécessite d'être retravaillée. Nous avons utilisé le jeu de données PANTHERE de Météo-France qui contient une donnée par maille kilométrique à résolution temporelle de 5 minutes entre Juillet 2006 et Juillet 2016. Il permet donc d'avoir accès à 10 ans de données sur l'ensemble du territoire métropolitain, alors qu'il existe des séries de données pluviométriques pouvant contenir plusieurs dizaines d'années.

## 9. Théorie autour des lames d'eau radar

Comme leur nom l'indique les lames sont calculées à partir de radars. Ce sont des capteurs actifs, qui utilisent les propriétés des ondes électromagnétiques de se réfléchir sur tout obstacle (Darricau, 1981). Ils permettent d'en calculer notamment la distance, la vitesse, la direction. Leur utilisation peut être adaptée à la mesure des précipitations (Tabary *et al.*, 2013). La mesure s'effectue en hauteur et pas

directement au sol. L'exploitation de ces ondes est compliquée puisque la composition de l'atmosphère varie, le signal revient bruité, et faussé par des artefacts. De plus, le déplacement des gouttelettes dépend d'éléments perturbateurs comme le vent. Ce qui peut induire, une différence entre la mesure radar, et la lame d'eau qui atteint vraiment le sol à un endroit donné. Un ensemble de traitements sont nécessaires et appliqués par le distributeur sur le signal reçu (Tabary *et al.*, 2013). D'après Tabary (2013), les données estimées sont plutôt de bonne qualité. Mais elles le sont particulièrement sur les pluies intenses, au détriment des pluies faibles qui ont tendance à être surestimées.

## a. Le réseau ARAMIS



**Figure 7** : Couverture du réseau ARAMIS en Août 2019 selon Météo-France. (bleu : cercle couvert par les radars en bande C, rouge par ceux en bande S et vert en bande X)

Le réseau ARAMIS (Application Radar à la Météorologie Infra-Synoptique) est le réseau de radar de Météo-France sur le territoire de France métropolitaine. Il est composé en Août 2019, d'une trentaine d'appareils actifs sur la métropole. Ils permettent de couvrir la majeure partie du territoire (figure 7) La diversité du réseau permet d'obtenir des mesures de qualité pour à peu près l'ensemble du pays. Les données après traitements constituent plusieurs produits.

## b. Les bases de données disponibles

Les caractéristiques principales des quatre bases de données qui couvrent le territoire Français sont résumées dans le tableau 2 suivant.

**Tableau 2** : Présentation des bases de données disponibles sur le marché.

Distributeur	METEO FRANCE			RHEA
Bases de données	COMEPHORE	ANTILOPE	PANTHERE	CALAMAR
Temps réel (T) ou réanalysée (R)	R	T	T	Les deux
Début	1996	2006	<b>2006</b>	1994
Fin	2006	----	----	----
Pas de temps (min)	60	60	<b>5</b>	60 et 5

A part Météo-France, il n'existe qu'un seul autre distributeur, RHEA une entreprise privée. La source est la même mais le traitement des données est différent. Globalement, Météo-France était plus intéressant ici parce nous avons un jeu de leurs données à disposition et il est beaucoup plus simple de trouver des informations sur les traitements qu'ils utilisent sur le signal. La bibliographie ne compte pas un grand nombre d'articles qui comparent les bases de données Météo-France et RHEA. Les deux paraissent donner des résultats similaires et plutôt bons pour des pluies dont la variabilité est importante, notamment par rapport à des pluviomètres (Schoorens et al., 2013). Mais ces résultats sont tirés de petits jeux de données, ce qui limite faute d'une vérification étendue la portée des conclusions.

Ces bases de données ont toutes la même **résolution spatiale environ 1km<sup>2</sup>** (1.01337km<sup>2</sup> en réalité). PANTHERE (Projet ARAMIS Nouvelles Technologies en Hydrométéorologie Extension et RENouvellement) est la base de données « brutes » de Météo-France. Elle propose des données disponibles en temps réel. ANTILOPE a une résolution temporelle plus importante parce que la mesure est réajustée pour certaines zones. L'idée est de contrer la détérioration du signal pour les pluies faibles. Les zones en question sont identifiées et plutôt que de prendre les lames d'eau radar, les données des pluviomètres sont utilisées pour estimer les lames d'eau à l'aide d'opérations statistiques type Krigeage. Cependant, il faut se méfier des incertitudes associées à l'interpolation autour des pluviomètres et des contraintes matérielles (Tabary *et al.*, 2013).

## **c. Traitements appliqués à la réception des données**

Les données PANTHERE reçues au laboratoire sont des fichiers correspondant à des champs de lame d'eau sur la France entière pour des pas de temps de 5minutes et sous forme de matrices (raster). Ce format est favorable à l'étude "par champs", mais induit des temps d'extractions longs pour obtenir des chroniques. En conséquence, les fichiers ont été remodelés sous la forme de chroniques (pour 10ans) relatives à chaque pixel (Figure 8). Puis, ils ont été regroupés (mis à la suite) dans des fichiers "blocs" de 16X16 (256) cellules par E. Leblois, chercheur dans l'unité Riverly.

Il s'agit de données « brutes », utilisées par le SHAPI pour le déclenchement et le suivi des alertes relatives aux crues et inondations. Des traitements ont été réalisés pour réduire des artefacts, bien

reconnaissables par leurs alignements de pixels forts sur plusieurs kilométriques. Les valeurs qui paraissaient aberrantes ont été supprimées et sont à considérer comme lacunes dans le code. Ensuite, une approche géostatistique a permis de détecter et de remplacer d'autres valeurs estimées fausses mais dont la différence avec le reste du jeu de données était moins visible que pour les premières (Ekoko-Safi, 2016). Le système de coordonnées utilisé est Lambert 93.

## d. Méthode d'exploitation

Les chroniques PANTHERE sont des gros jeux de données. Ils sont compliqués à exploiter avec des outils classiques (tableurs...). Nous avons choisi d'utiliser un langage informatique, Python. Python est un langage interprété c'est-à-dire que les commandes de l'utilisateur sont traduites une à une pour l'ordinateur. Il est réputé facile à prendre en main et permet d'écrire du code rapidement et clairement (facile à écrire, facile à relire). Le langage est donc adapté pour l'utilisateur mais les calculs sont allongés par rapport à un langage compilé.

## e. Règles d'écriture des scripts

Il peut être très chronophage de lire et de comprendre un code dont on est pas le créateur. Il paraît donc important de structurer et expliquer ses scripts pour faciliter l'adaptation. Pour ce faire, nous avons utilisé la structure suivante :

1. Description du script
2. Import des packages nécessaires
3. Chemin vers les fichiers de données
4. Constantes et variables qui doivent être désignées pour les fonctions.
5. Fonctions déclarées
6. Programme (ou programme test si c'est un script destiné à l'interface) *facultatif*

A défaut d'avoir accès à une forge officielle (gitlab d'INRAE), la formation n'ayant pas été assurée en raison du contexte sanitaire, nous avons défini des "bonnes pratiques".:

- Chaque action doit être commentée et expliquée.
- Il faut que les fonctions soient assez généralisées pour pouvoir accepter des jeux de données différents (ou au moins la structure du jeu de données souhaitée doit être précisée).
- Les modifications sur le code principal sont minimales et complétées d'un indice permettant d'identifier la personne à l'origine de la modification. Ceci dans le but que le développeur principal (ici Christine Poulard) garde la main sur son code.
- Les grosses modifications ou créations sont réalisées sur des scripts à part (ce qui est aussi un gros gain de temps car il est plus facile de s'y repérer).

## f. Intégration des fichiers géographiques

### *Etude par shape*

Exploiter les données par l'intermédiaire de bassins ou sous-bassins est pratique pour étudier les incidents. Ils délimitent des zones pour lesquelles les influences hydrologiques mutuelles sont minimales et ils permettent de diviser les zones d'études pour obtenir de plus petites surfaces. La donnée de précipitation étant très variable, cela permet d'avoir une meilleure compréhension des incidents locaux tout en allégeant les temps de calcul d'un travail pixel par pixel.

L'interface ne possédait pas encore la possibilité d'étudier des chroniques de pluies par bassins versants. Nous avons dans ce but décidé d'écrire un script permettant la lecture de fichiers type "shape". Et, cela permet de faire un lien avec l'exploitation d'IRIP sur QGIS.

Le fichier va dessiner l'allure du bassin. Ensuite, sur Python grâce notamment au module "Shapely", il est possible de lui donner un sens géométrique. Sa trace va permettre de sélectionner les cellules impliquées. L'idée ensuite est de faire en sorte que la surface de chaque pixel ( $S_i$ ) à l'intérieur du bassin détermine sa contribution (et ainsi de revoir à la baisse les apports des pixels frontaliers) :

$$\frac{S_i}{Aire_{totale}} = v_i \quad \longrightarrow \quad CBV = \sum(v_i * LE_i)$$

CBV : Chronique moyenne du Bassin Versant ; LE : Lame d'eau

Ensuite, le but est de pouvoir utiliser les mêmes procédures que sur les chroniques classiques de pluie mais les données sont des lames d'eau moyennes sur une surface de 1 km<sup>2</sup>.

### *Etude par couches points*

Nous avons également créé des scripts destinés à lire des fichiers de points. En effet, la plupart des couches d'incidents sont ou peuvent être transformés en simples coordonnées géographiques. Ces scripts permettent donc d'étudier les chroniques spécifiques pour chaque incident et éventuellement de relativiser la chronique générale du bassin. Le but de ce script est d'automatiser le calcul pour un grand nombre de points sans tracer les graphiques ce qui poserait des problèmes de mémoire. Si l'utilisateur souhaite visualiser les pluies sur un seul point, il a tout intérêt à juste charger la cellule correspondante en utilisant le procédé classique.

## **g. Analyse multi-durées**

La dynamique de la pluie peut avoir un impact important sur la lame d'eau ruisselante (Fang *et al.*, 2008). Pour cette raison, prendre juste en compte des cumuls pour un événement ne suffit pas pour comprendre son processus d'apparition (Nicolas, 2010). Nous avons choisi de travailler par l'intermédiaire d'analyses multi-durées, pour avoir non seulement une information sur le cumul mais aussi sur la dynamique de la pluie. Ensuite, l'état hydrique d'un sol est conditionné par sa nature et par un contexte. La manière avec laquelle un événement va influencer l'état hydrique dépend donc aussi des caractéristiques du sol au moment de l'événement, et notamment à quel point sa capacité de stockage est disponible ou non. Si elle est pleine, la surface est déjà saturée en eau et la pluie ruisselle totalement. Il peut donc être important d'étudier en plus de l'événement, l'historique des pluies sur différentes durées.

## **h. Caractérisation par étude des événements extrêmes**

L'analyse s'est développée en hydrologie pour décrire de manière synthétique des régimes de pluies et de débits à partir de chroniques mesurées. Il s'agit de déterminer la probabilité de dépassement d'une valeur "X".

## Méthode d'échantillonnage

La démarche que nous allons employer par la suite est largement inspirée de la revue de Lang (1999). Il propose une méthode de réalisation d'un échantillonnage SUPSEUIL bien cadrée et assez simple pour pouvoir envisager une automatisation sur Python. L'étude d'évènements extrêmes sur des jeux de données aléatoires comme pour les débits ou ici les précipitations commence souvent par le choix de la méthode d'échantillonnage des valeurs. Deux méthodes principales existent : la sélection par les maxima annuels "MAXAN » et la sélection par des valeurs supérieures à un seuil "SUPSEUIL". Cette dernière favorise la sélection des plus fortes valeurs observées, ce qui peut conduire par exemple à ne pas échantillonner une année où il aurait faiblement plu. C'est la première différence avec l'échantillonnage annuel. L'autre différence consiste à pouvoir « sur-échantillonner », c'est-à-dire à sélectionner plus de valeurs que d'années disponibles. Dans ce cas l'échantillon contient plus de valeurs que celui du « MAXAN ». Ce choix apporte de l'information sur les événements fréquents, qui se produisent plus d'une fois par an en moyenne.

Nous avons choisi de travailler avec la méthode SUPSEUIL pour les raisons suivantes :

- La taille de l'échantillon est plus importante, ce qui contrebalance notre nombre d'années disponibles qui est plutôt faible (environ 10 ans)
- Les évènements extrêmes sont globalement mieux estimés, on ne contente pas d'une valeur par an qui peut être très fluctuante selon si l'année est pluvieuse ou non.
- La maîtrise du seuil rend plus flexible l'étude des valeurs et permet de s'intéresser au processus d'occurrence d'apparitions des événements exceptionnels.

Ainsi pour terminer, Lang (1999) a démontré que la modélisation par la méthode SUPSEUIL est plus intéressante que les maximums annuels si on peut déterminer plus de deux événements par an en moyenne ce qui est notre cas.

Une étape préliminaire consiste souvent à représenter graphiquement l'échantillon obtenu en plaçant les points en fonction d'une fréquence empirique et de leur rang ( $r$ ). Il existe plusieurs formules ; nous nous servons de celle de Hazen :

$$F(x_i) = \frac{r-0.5}{n}$$

Avec  $r$  le rang de  $x_i$  dans l'échantillon et  $n$  la taille de l'échantillon. Par la suite, il est possible de lui attribuer une période de retour exprimée en années. Selon la formule de Hazen, on l'exprime de la manière suivante (il s'agit ici de la période de retour liée à la fréquence au dépassement) :

$$T = \frac{1}{1-F(x_i)}$$

Pour un événement extrême caractérisé par un cumul maximal relatif à une durée donnée, la période de retour exprime l'occurrence « en moyenne » qu'un événement égal ou supérieur se produise. La période moyenne de retour s'exprime en « années ». C'est une variable fréquemment utilisée par exemple dans le dimensionnement d'ouvrage.

## i. La méthode SUPSEUIL

Les distributions de valeurs extrêmes sont censées suivre la loi d'extrémum généralisée (GEV). Dans le cas des échantillons SUPSEUIL, on leur préfère la loi de Pareto généralisée (GPD) et son cas particulier la loi exponentielle (Renard, 2014).

La première est une distribution à trois paramètres :

- le seuil «  $x_0$  », qui correspond ou est très proche du minimum de l'échantillon.
- le paramètre d'échelle ( $\lambda$ ), c'est un facteur multiplicatif, qui régit l'aplatissement de la loi de probabilité.
- le paramètre de forme ( $\xi$ ) : qui détermine la forme de la distribution.

Le choix s'est porté sur la loi exponentielle pour le code, plus simple car elle n'a que les deux premiers paramètres de la GPD et elle est plus pertinente sur des séries courtes.

Elle est définie comme suit :

$$F(x) = \begin{cases} 1 - e^{-\frac{x-x_0}{\lambda}} & \text{si } x > 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Comme la période de retour (T) est une probabilité de dépassement, les valeurs extrêmes (VE) peuvent être calculées comme des quantiles d'ordre p de la distribution de la variable étudiée. Suivant ce que nous avons dit dans la partie précédente nous avons :

$$p = 1 - \frac{1}{T}$$

Ainsi,  $F(VE) = p$  et donc :

$$VE = x_0 + \lambda \ln(T) \quad (i)$$

Si  $x_0$  est connu ce n'est pas le cas de  $\lambda$ . On peut l'estimer en comparant le seuil et la moyenne de l'échantillon :  $\lambda = \bar{x} - x_0$  (Renard, 2014).

Ainsi l'équation linéaire (i) des valeurs extrêmes n'a plus d'inconnues. L'utilisation de la GPD est complétée par une analyse de l'occurrence annuelle, dont on souhaite qu'elle soit décrite par une loi de Poisson, nous en reparlerons par la suite. En effet, la distribution de Poisson décrit l'occurrence d'évènements indépendants les uns des autres (Renard, 2014).

La détermination du seuil est une des parties sensibles de la méthode. Il doit permettre de faire ressortir un nombre suffisamment important d'échantillons pour garantir une relative homogénéité du jeu de données dans le temps. Mais d'un autre côté, s'il est trop bas, il risque de comporter des valeurs non concernées. Il y a deux manières de procéder pour s'assurer que l'échantillon final soit cohérent et composé de valeurs indépendantes :

- Utiliser des éléments statistiques pour orienter le choix du seuil.
- Retraiter les données de l'échantillonnage à l'aide de critères.

*Critère de sélection des pics :*

Nous avons utilisé deux critères pour limiter la sélection des pics et s'assurer qu'ils correspondent bien à des évènements différents. Ces critères étaient ceux déjà intégrés dans le code avant le stage :

- Laps de temps entre évènements : il permet de délimiter un temps t après lequel l'effet d'influence d'un événement sur le suivant est considérée faible.
- Coefficient de redescente ( $0 < \rho < 1$ ) : si entre deux pics,  $X_{min} \geq \rho X_{max}$  alors seul le pic le plus haut est gardé.

*Choix du seuil :*

La méthode du choix du seuil, ici est orientée en fonction de la démarche proposée dans l'article de Lang (1999) au moins pour partie, l'idée étant d'entamer une automatisation.

La moyenne du nombre d'échantillons par année ( $\mu$ ) doit être comprise dans l'intervalle :

$$[2 \leq \mu \leq 5]$$

Nous ne voulons pas trop de points dans l'échantillon pour être sûr de bien considérer les valeurs extrêmes, mais il en faut suffisamment pour que la répartition soit la plus homogène possible entre les années.

L'analyse est complétée par une analyse de la variabilité du nombre d'échantillons annuels. L'objectif est que le seuil choisi permette d'obtenir des pics dont la distribution suit une loi de Poisson. Le postulat de base est que la variance de ce type de distribution est approximativement égale à son espérance. Par conséquent son indice de dispersion (I) définie de la manière suivante est égale à 1 (en considérant m le nombre d'échantillon annuel).

$$I = \frac{Var(m)}{E(m)}$$

En citant les résultats de Cunnane (1979), Lang (1999) fait ressortir le lien entre la distribution des échantillons annuels et une distribution du khi2. Puisqu'il est possible de tester l'échantillon pour vérifier la Loi de Poisson en comparant l'indice de dispersion de l'échantillon ( $\hat{I}$ ) et l'indice suivant :

$$I(p) = \frac{\chi^2(p)}{(NY - 1)}$$

Avec, NY le nombre d'année étudiées, (NY-1) le nombre de degrés de liberté relatif à la distribution et  $\chi^2(p)$ , le quantile (q) correspondant à une probabilité p tel que  $P(X > q) = p$ . Cette dernière dépend de l'intervalle confiance ( $\alpha$ ) choisi pour l'expérience.

$$I(1 - \alpha/2) \leq \hat{I} \leq I(\alpha/2)$$

## j. Représentation graphique

La relation temps de retour/cumul est souvent tracée sous forme de courbes **Intensité-durée-Fréquence (IdF)**. On ajuste une loi exponentielle sur la chronique initiale, puis sur des chroniques agrégées sur des pas de temps plus grands. Pour cette loi, la relation entre le logarithme népérien de la période et le cumul sera linéaire. Ensuite, on lui superpose les points empiriques dont les temps de retour seront déterminés en résolvant l'équation (i). Le résultat peut éventuellement montrer certains points très éloignés de la droite théorique ce qui est normal pour les valeurs les plus extrêmes. Nous devons en revanche faire attention à ce que la répartition des points ne dessine pas une courbure.

Pour représenter un évènement par rapport aux IdFs, nous dessinerons ce que nous appelons la "**trace**" de l'incident sur l'évènement pour différentes durées. Pour chaque évènement, il est extrait le cumul maximum pour toutes les durées « d » étudiées ( $C_{max}(d)$ ), et déduit d'après la relation IdF la période de retour estimée. Les points ( $T(C_{max\_d}), C_{max\_d}$ ) sont ensuite reportés sur le graphique. Il est important de préciser que les ajustements sont effectués à partir de **10 ans** d'observations et que par conséquent les estimations des quantiles pour les périodes de retour au-delà de la décennie sont imprécises. *A noter cependant que l'étude par zone de plusieurs pixels permet d'accéder à plus de valeurs fortes et ainsi d'étendre la profondeur historique de l'analyse. Les tests menés avec le jeu de données par E. Leblois au travers d'analyses régionales ont permis une extrapolation jusqu'à la période de retour 30 ans.* Les points sont finalement reliés par ce qu'on appelle la "trace".

Ce mode de représentation est particulièrement intéressant quand il est utile d'afficher les résultats de nos analyses pour plusieurs éléments, comme plusieurs bassins versants par exemple.

## k. Exploitation des résultats IRIP

Les caractéristiques des événements pluvieux seront estimées sur les deux zones d'étude. La partie suivante va résumer sa méthode d'exploitation et les éléments nécessaires à l'interprétation des résultats.

### I. Lecture des cartes

Les cartes à partir desquelles nous estimons les aléas sont les cartes de transfert et d'accumulation (présentées dans la partie contexte et plus particulièrement dans la figure 2).

Les cartes d'aléas nécessitent d'estimer un seuil de sensibilité à partir duquel on considère que l'aléa est suffisamment fort pour que la possibilité qu'un incident apparaisse sur une zone soit non négligeable. Ce seuil peut varier en fonction des conditions climatiques des différentes régions (Breil *et al.*, 2018). Pour s'adapter à ce phénomène il a été défini que le choix des seuils serait conditionné par le découpage de la France par HER. Elles ont été définies par Wasson (2004) dans le cadre de la Directive Cadre Européenne sur l'eau. Elles sont basées sur des différences climatiques et sur leurs caractéristiques au sol, notamment la géologie et le relief.

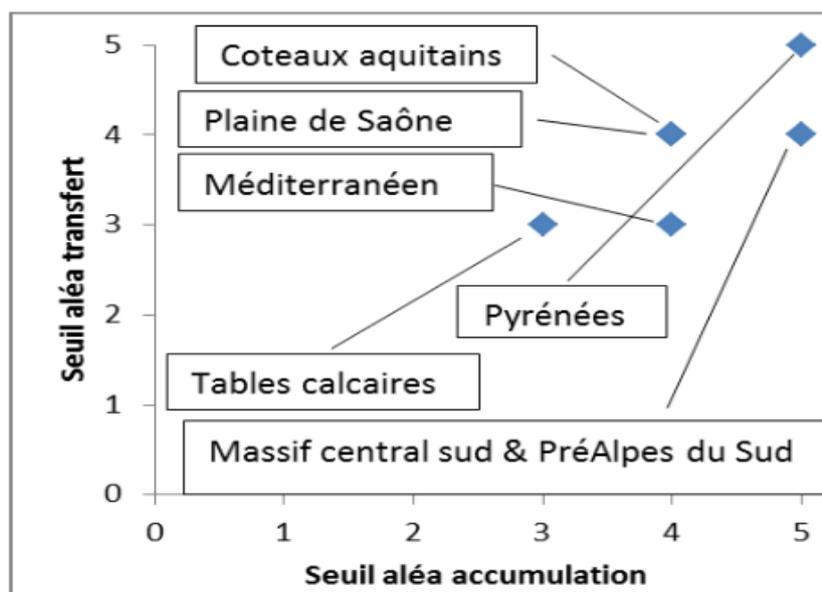


Figure 9: Seuils optimaux pour 8 HER testés (Breil *et al.*, 2018).

Ainsi, selon la figure 9, pour la zone de **Chalon**, qui correspond à la plaine de Saône, les seuils à risques seront 4 pour les deux cartes (transfert et accumulation) et pour les **Pyrénées Orientales** les seuils de risques seront de 4 pour accumulation et 3 pour le transfert.

## m. Vérification des résultats

Conformément à ce que nous avons annoncé dans la partie contexte, nous utiliserons une méthode estimant l'efficacité d'IRIP grâce au ratio suivant :

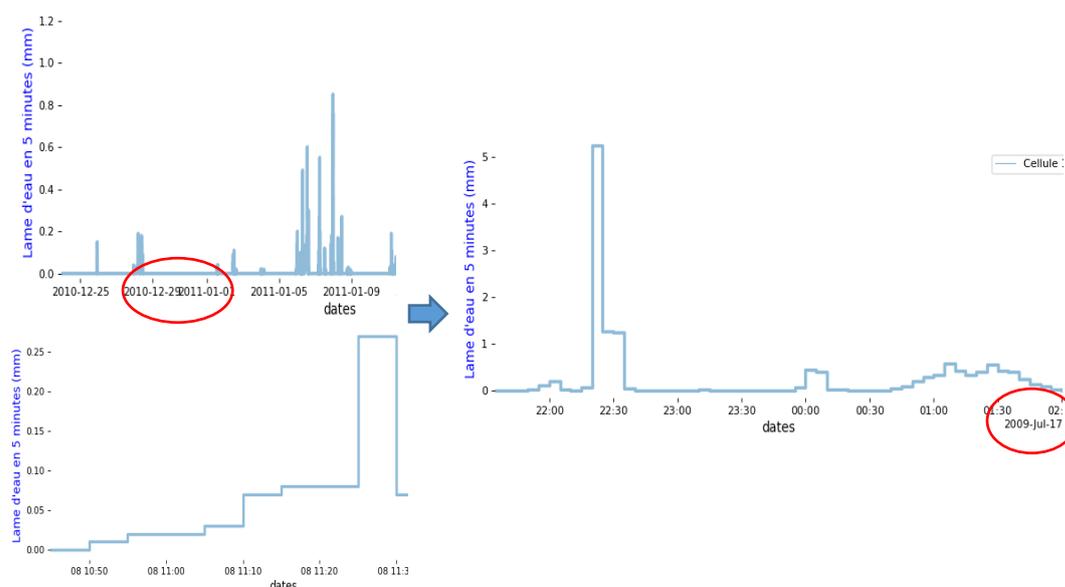
$$R = \frac{P \text{ Incident prédits}}{P \text{ Mailles en aléa fort}}$$

On estime que le modèle est efficace si plus de 80% des incidents sont prédits avec moins de 20% de la zone d'étude en aléas forts.

Les incidents sont définis pour IRIP en fonction de zones caractéristiques. Pour des couches de points, ce sont des tampons de 50 mètres de rayon autour du point. Pour les autres, nous avons utilisé directement la surface des parcelles. Un incident est considéré "prédit" si au moins 7% de la surface affiliée est couverte par des aléas fort (Breil *et al.*, 2018).

## n. Estimation des bassins versants

Pour exploiter correctement les lames d'eau il a été nécessaire de réduire la taille des zones d'études. Pour le faire, à l'intérieur des zones, les sous-bassins versants ont été définis grâce à une fonction de "GRASS" dans QGIS, "r.watershed". Elle détermine les sous bassins topographiques grâce au MNT de la zone. Ensuite, les bassins en question ont été choisis en fonction de la présence d'incidents, et en fonction des cumuls de pluie pour tester des bassins qui ont subi des incidents différents.



**Figure 10 :** A gauche le traitement automatique des dates par matplotlib, à droite géré par le package « mdates ». En haut les dates se recouvrent, en bas, l'écriture n'est pas claire.

Ils seront comparés aux chroniques de certains pixels en fonction de la présence d'incidents ou si les données de cumuls suggèrent un écart important.

## 10. Résultats

Les scripts Python sont finalement les résultats les plus importants du stage, mais par souci de clarté ils ne seront pas exposés en détail ici. Nous expliquerons dans cette partie les principes de rédaction, et ce qu'ils apportent. Ensuite, les deux cas d'études sur lesquels nous avons travaillé seront présentés. Nous avons cherché à l'aide des outils que nous avons développé à comprendre le déroulé des deux événements pluvieux. Ainsi, nous espérons trouver des pistes pour comprendre les résultats IRIP.

## a. Les scripts

Certains de nos scripts sont nouveaux, d'autres sont des améliorations de ceux qui existaient déjà. Nous aborderons plusieurs points ici. D'abord, nous résumerons l'apport des nouveaux modules dans l'optimisation du code. Ensuite, nous parlerons brièvement des différents scripts rédigés.

### *Les principaux modules*

Une partie du travail a consisté à faire des recherches sur différents modules qui pourraient permettre une optimisation des scripts et de leurs résultats. Les scripts ont été écrits en basant la structure des données sur les tableaux ("DataFrame") du module "Pandas", sauf pour quelques objets (souvent éphémères) pour lesquels nous avons préféré le format matriciel du module "Numpy". Leurs bibliothèques de fonctions, sont riches et permettent d'écrire du code à la fois concis et performant. Cette évolution est différente du code initial plutôt basé sur le module "Numpy" et d'autres types d'objets disponibles sur Python. "Pandas" a la réputation d'être particulièrement efficace sur les gros jeux de données.

Sinon, d'autres modules accessoires ont été découverts permettant d'améliorer la qualité des résultats et l'expérience utilisateur. Tous ne seront pas détaillés ici mais par exemple, le package "mdates" nous a permis d'optimiser la gestion des dates qui pouvait poser problème lors de zooms (Figure 10). "PySimpleGUI" nous a permis de rajouter des barres de progressions, qui permettent de connaître le temps d'exécution restant du code.

### *Scripts d'exploitation*

Le fonctionnement des principaux scripts écrits pour ce travail est le suivant:

- Un script optimisant l'importation et la lecture des données binaires des pixels depuis les fichiers de l'ordinateur. Le but était d'automatiser l'extraction de fichier depuis un ensemble de fichiers zip contenant les blocs vers un répertoire précis. Ensuite, le code détermine les blocs dont il aura besoin en fonction des fichiers géographiques pour ne pas avoir à les ouvrir plusieurs fois. Il lit ensuite les pixels requis à l'intérieur du bloc, et crée les chroniques (Annexe D).
- Le script d'échantillonnage, qui va sonder les maximums locaux en fonction du seuil pour vérifier s'ils correspondent aux critères cités dans la partie méthode. Ce script et celui ci-dessus utilisent les constantes résumées dans le tableau 3 (Annexe B).
- Un script de détermination automatique de seuil qui prend en compte les recommandations de Lang (1999). Pour une même durée le code va estimer différentes valeurs de seuil pour lesquelles les longueurs de l'échantillons se trouvent comprises dans l'intervalle de  $\mu$ . Enfin les indices de dispersion sont calculés et comparés aux résultats du  $\chi^2$  pour trouver le meilleur seuil possible. La routine rend un seuil, son coefficient de dispersion, et un commentaire sur sa position par rapport à l'intervalle du  $\chi^2$  (Annexe C).
- Un script permettant l'exploitation de fichiers points, pour analyser spécifiquement les données concernant les incidents. Il est par exemple utilisé pour construire les cartes de cumuls figure 11 et 17 (Annexe D).
- L'intégration à l'interface de la possibilité d'exploiter des fichiers Shape contenant un ou plusieurs polygones. Celui-ci permet de réaliser rapidement des moyennes de chroniques pour des sous bassins versants (Annexe A).
- Un dernier script contient un ensemble de fonctions graphiques pour représenter les données.

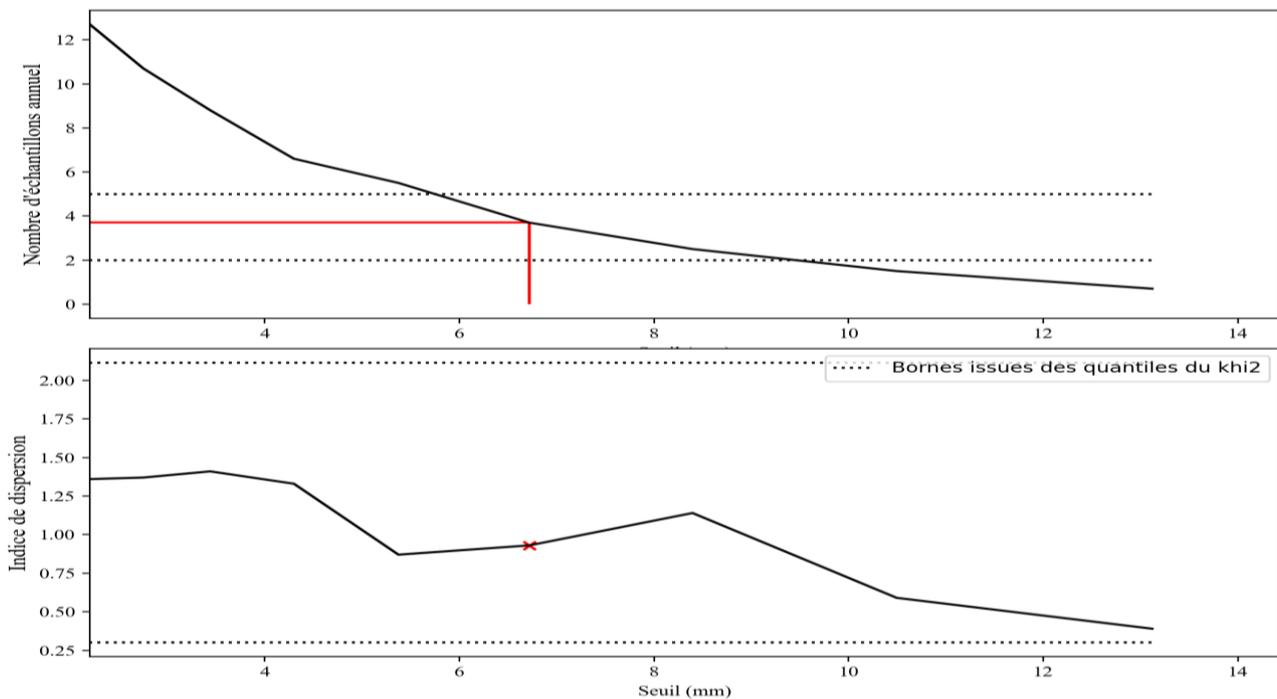
Les scripts sont assez facilement adaptables à des chroniques issues de données provenant de pluviomètres du moment que celles-ci sont à pas de temps fixes.

**Tableau 3 :** Constantes utilisées pour l'échantillonnage.

Nombre d'années (NY)	Degrés de libertés	Intervalle de confiance (%)	$\chi^2(0.975)$	$\chi^2(0.025)$	Temps entre évènements (jours)	Coefficient de redescente
10	9	5	2.7	19.02	3	0.75

## b. Seuil automatique

L'automatisation du seuil a été testée sur quelques chroniques aléatoirement pour vérifier son fonctionnement.



**Figure 11 :** Exemple des figures obtenues pour les tests de validité des seuils. En haut, le nombre moyen d'échantillon annuel en fonction du seuil. En bas, l'indice de dispersion de l'échantillon pour un seuil donné.

La détermination automatique du seuil n'a pas posé de problème pour les différentes chroniques testées. Il y a toujours un seuil qui permet de trouver un échantillon qui respecte les conditions des deux intervalles (figure 11).

Dans le cas où il y aurait un problème, un commentaire sur la validité du seuil est laissé dans les tableaux qui contiennent les résultats des codes. Si, l'indice se trouve bien dans l'intervalle, le seuil est jugé "valide", sinon, le code conserve l'information de sa position en fonction de l'intervalle. Ceci dans l'idée d'une éventuelle adaptation par les lois binomiales qui sont plus pertinentes dans ces cas là (Lang *et al.*, 1999).

## c. Etude de cas : Le Grand Chalon

L'étude portera sur des incidents survenus entre le 24 et le 26 Juin 2016, inclus. Plusieurs communes de la région ont été classées en état de catastrophe naturelle. Les incidents ont été sélectionnés de manière à ne correspondre qu'à du ruissellement (Breil *et al.*, 2018)

### *Distribution de la pluie*

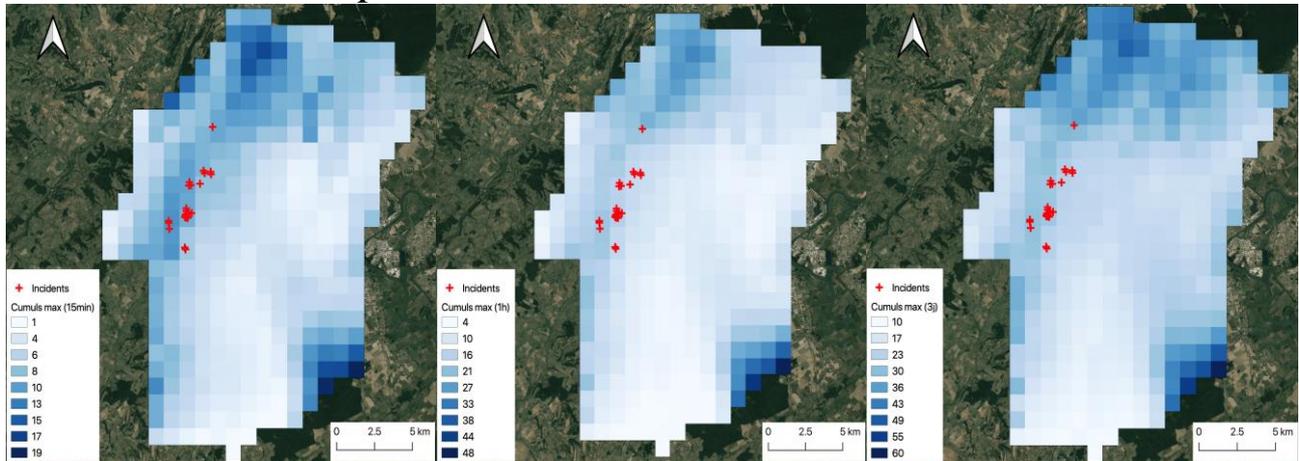


Figure 12 : Distribution des cumuls maximaux (en millimètres) pour chaque pixel sur la zone du grand Chalon. (de gauche à droite durée : 15 minutes, 1 jour, 3 jours)

A partir des cumuls glissants définis à partir de plusieurs durées, nous avons déterminé les cumuls maximaux sur chaque pixel (figure 12). Nous avons déduit que l'évènement a été plus marqué au nord-ouest et au sud-est. Pour la plupart des pixels qui se trouvent à l'emplacement des incidents, les cumuls n'atteignent pas des valeurs très importantes, et plafonnent à environ 30mm sur trois jours. Par contre, la plus grosse partie de l'eau est tombée sur une période courte, entre 15 minutes et une heure selon les deux figures de gauche. Malgré des cumuls plutôt faibles l'impact sur la zone a été considérable. Pour essayer de comprendre pourquoi, nous allons travailler sur un bassin versant témoin.

### *Le sous bassin versant*

Suivant la figure 12, ce sous bassin se situe sur une transversale pour laquelle les cumuls maximums sont importants pour chaque durée.

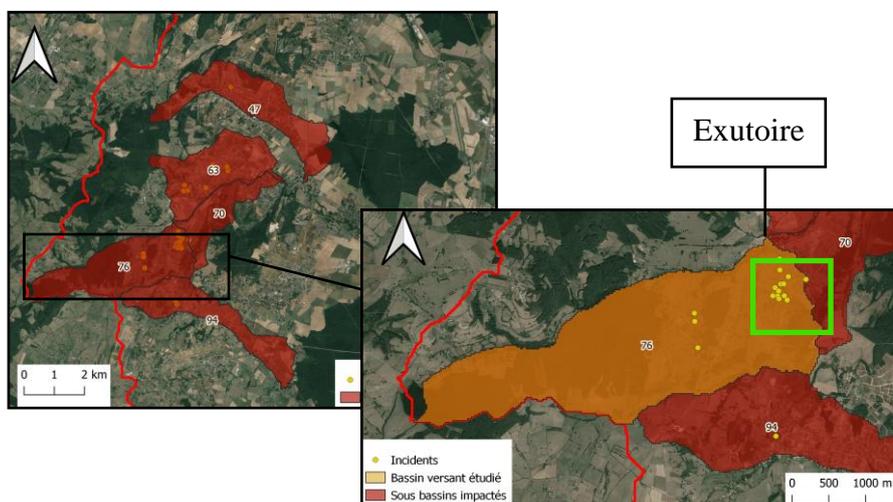


Figure 13 : Le bassin versant 76 en orange et en vert, le pixel choisit pour la comparaison.

Il compte 15 des 27 incidents répertoriés et couvre 7km<sup>2</sup>. Les outils que nous avons développés nous permettent de travailler sur le bassin ou directement sur un pixel. Or, nous avons remarqué que les incidents sur le bassin étaient pour beaucoup très proches et situés dans un seul pixel de la lame d'eau radar. Donc, nous avons choisi de comparer sa chronique moyennée (chronique de bassin) et celle du pixel (en vert sur la figure 13), où on compte 11 des 15 incidents.

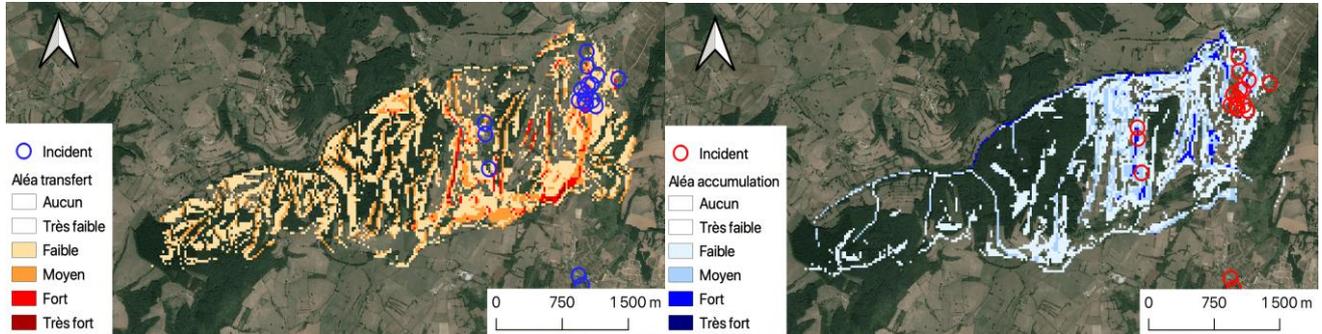


Figure 14 : Résultat IRIP sur le bassin 76.

### Résultat IRIP

Nous avons utilisé le modèle IRIP sur le bassin. Les résultats, tableau 4, montrent que le modèle est très efficace sur le grand Chalon pour la détection de l'évènement de 2016. Le résultat est encore bien plus impressionnant sur le bassin observé pour lequel la surface en aléa fort est très faible par rapport au taux de détection.

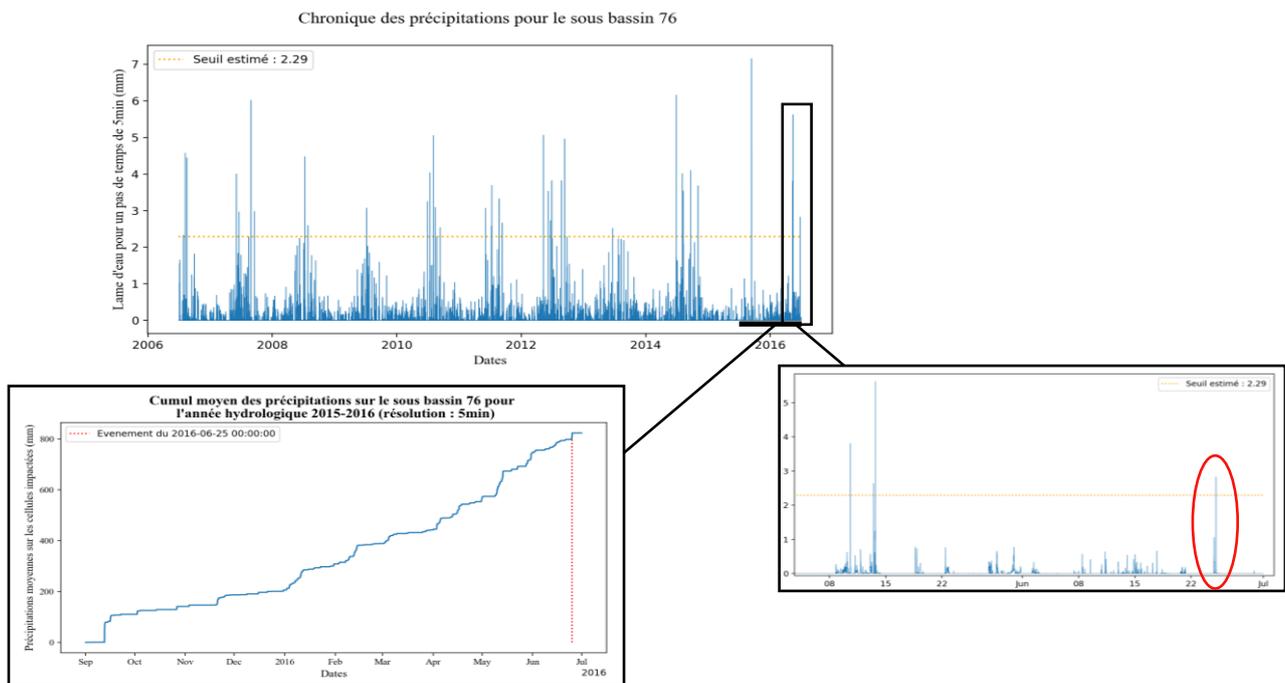
Tableau 4 : Résumé chiffré des résultats d'IRIP sur le sous bassin et le Grand Châlon

Zone considérée	Nombre d'incidents	Nombre d'incidents prédits	Incidents prédits (%)	Aléa fort dans le BV (%)	Efficacité (R)
76	15	13	87	4	19.29
GC	27	24	89	9	10.43

Les trois accidents à l'ouest semblent plutôt avoir été causés par de l'accumulation. Ceux à l'est ont subi une influence mixte mais on peut noter la tendance forte du transfert sur la partie du sud.

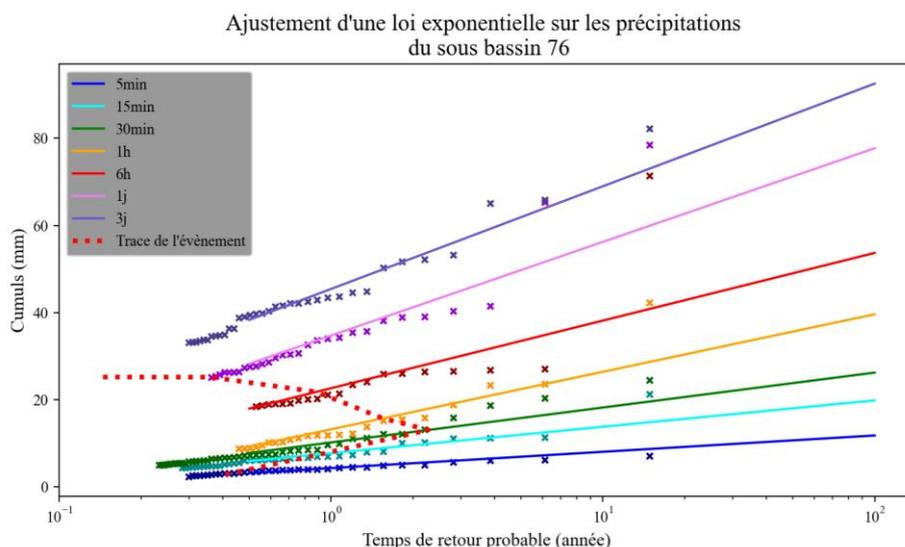
### Les précipitations

Sur la figure 15 suivante, la chronique de pluie sur le bassin versant montre bien un dépassement du seuil, mais l'évènement semble court. Il est précédé d'un évènement en Mai dont l'amplitude est plus importante (figure 15). Puis, entre les deux la chronique affiche une succession de petits évènements mineurs. La courbe des cumuls agrégés figure 15, ne montre pas de sursaut particulier qui aurait pu être synonyme de pluie très intense, mais plutôt une croissance continue depuis le mois d'Avril environ.



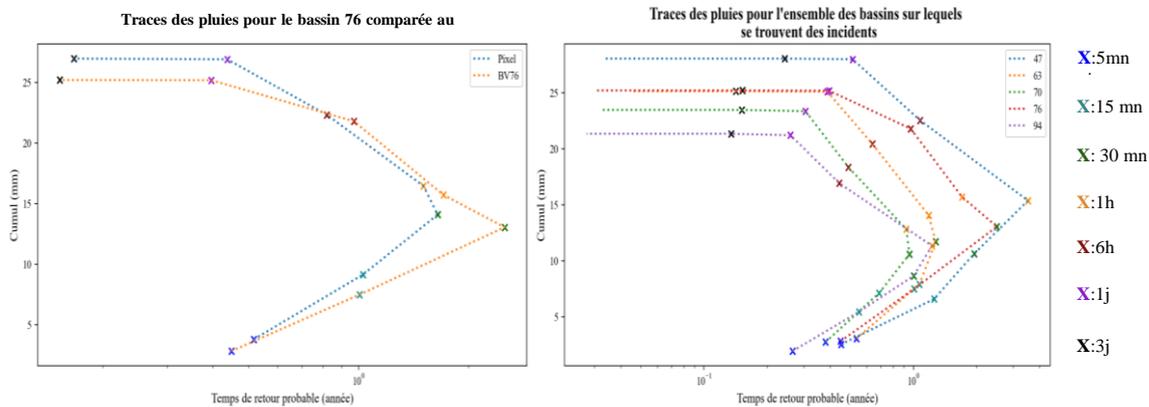
### Etude de la période de retour

Nous rappelons que l'ajustement  $P(T)$  n'est pas valable pour des périodes de retour de plus de 10ans. Au-delà, il ne faut pas les interpréter en tant que chiffre mais plutôt comme un indice qui qualifie le caractère exceptionnel de l'évènement. La trace de l'évènement dans la figure 16 suivante parait confirmer les observations réalisées grâce à la chronique. Mais, elle montre aussi une intensité forte pour la durée de 30 minutes, avec une période moyenne de retour estimée de 20 ans, et un cumul qui représente environ la moitié (# 14mm) du cumul total de pluie de l'évènement (#25mm). Cela combiné à un historique qui a pu maintenir un sol assez humide, nous avons des conditions locales favorables au ruissellement intense.



On retrouve des traces semblables pour le pixel étudié et pour les autres bassins versants qui ont subi des incidents (Figure 17). Pour le critère IDF choisi et les durées sélectionnées les temps de retour affichés sont de l'ordre de l'annuel. D'après ces figures l'évènement a été intense et rare pour ces zones

sur des durées courtes de 30 minutes à 1 heure, mais la lame d'eau cumulée d'un point de vue horaire ou journalier n'est pas exceptionnelle en comparaison à d'autres régions. Nous notons cependant qu'en dehors des zones avec incidents signalés, il a eu des cumuls horaires de plus de 40 mm (figure 12), ce qui commence à être conséquent car la vitesse d'infiltration maximale dans les 30 premiers centimètres du sol de l'ensemble des bassins est de 28mm/h et la capacité de stockage des 30 premiers centimètres du sol (non imperméable) est comprise entre 126 et 153mm. Nous pouvons donc suspecter qu'il y a eu ruissellement par excès d'infiltration dans notre cas d'étude.



**Figure 17.1 :** Comparaison de la trace du bassin avec le pixel désigné sur la figure 12 à gauche, et avec les traces des autres bassins à droite, selon les résultats de l'ajustement de leur loi exponentielle

La figure 17.2 montre les zones où la lame d'eau radar a dépassé la capacité d'infiltration pour les durées de 15, 30 et 60 mn. Cela concerne 11 points incidents sur 33 sélectionnés ici comme les points situés en dehors de l'enveloppe de débordement des talwegs (ou de l'EAIP-CE). Pour l'ensemble des points le taux de détection avec IRIP est de 31/33 et l'espace en aléas forts dans le bassin versant global est de 16%. Cela laisse entendre que les caractéristiques au sol (battance, zones imperméables) jouent aussi un rôle qui est vu apr le modèle IRIP via sa couche d'occupation du sol.

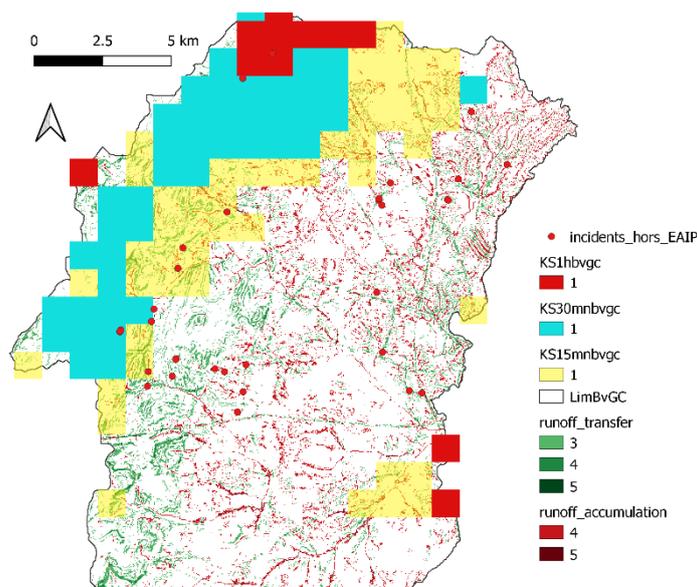
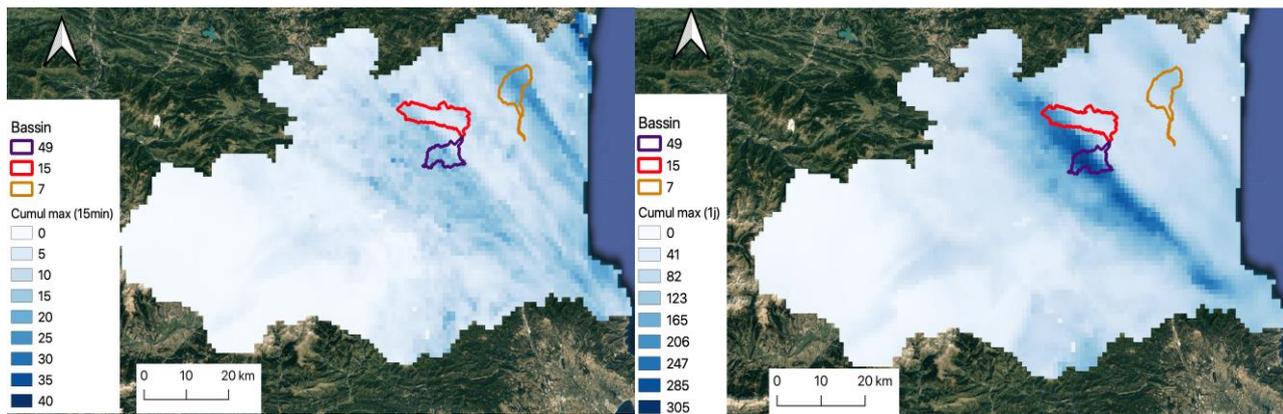


Figure 17.2 : Zones où le cumul pluviométrique a dépassé la capacité d'infiltration.

L'absence d'incidents déclarés dans les zones de plus fortes pluies peut être lié à l'absence de vulnérabilité, de non déclaration ou de déclarations relevant du système des calamités agricoles qui ne sont pas en notre possession.

## d. Etude de cas : Les Pyrénées Orientales

Ici l'étude portera sur la période entre le 28 et le 30 Novembre 2014 inclus. Nous disposons d'informations obtenues à partir des déclarations d'incidents sur parcelles agricoles. L'état de catastrophe pour les Pyrénées Orientales concerne une zone environ 10 fois plus grande que le Grand Châlon.



**Figure 18** : Cartes des cumuls maximums (en millimètre) sur les Pyrénées Orientales pour 2 durées différentes (gauche : 15 minutes, droite : 1 jour), pour l'évènement du 28/11/2014. Les bassins représentés seront ceux utilisés pour étudier les précipitations.

### *Distribution de la pluie*

Les cumuls les plus importants pour les durées faibles sont clairement dispersés sur l'est du bassin (figure 18 à gauche). Il y a un vrai contraste avec les maximums journaliers qui sont concentrés au centre du département (figure 18 à droite). Les cumuls affichés atteignent des valeurs très importantes, jusqu'à 350mm en 24h pour quelques pixels. La distribution des valeurs est très hétérogène dans le bassin versant.

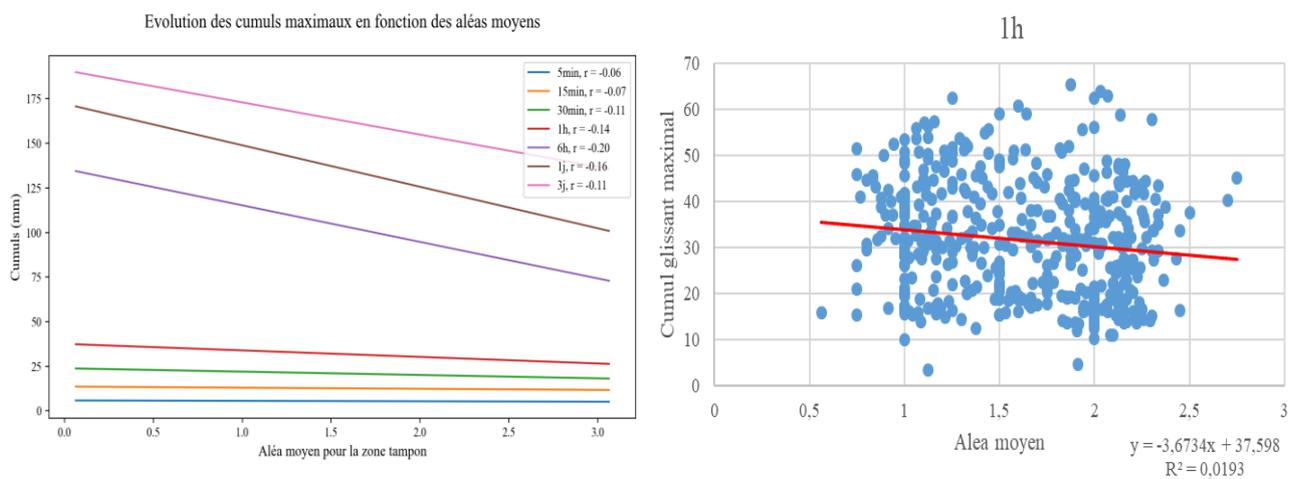
## e. Résultats IRIP

Tableau 5 : Résultat IRIP sur la région et sur les bassins versants.

Bassin versant	Nombre de parcelles touchées	Nombre d'incidents prédits	Incidents prédits (%)	Aléa fort dans le BV (%)	Efficacité (R)
7	53	48	91	15	5.93
15	357	249	70	22	3.21
49	90	55	61	14	4.30
Les PO	3829	2513	66	17	3.86

Les résultats affichés par le tableau 5 montrent une efficacité contrastée pour le modèle. Pour le bassin 7, les taux de prédictions et de surfaces couvertes par des aléas forts sont plutôt très bons. En revanche, pour les autres bassins et pour la zone totale, le taux de prédictions chute à 61% tout en gardant un % d'aléas forts inférieur ou proche de 20% (Les cartes en rapport sont disponibles annexe G).

La figure 19 affiche la relation de l'ajustement linéaire entre l'aléa moyen IRIP dans une parcelle agricole et les cumuls de pluie pour différentes durées. L'hypothèse d'une pente nulle a été testée pour chaque durée à l'aide d'un test de Student. A partir de la durée 30 minutes, les pentes sont significativement différentes de 0. On note que la pente négative augmente avec la durée (Annexe G). Ce résultat déjà observé sur d'autres zones d'études (Breil, 2020 ; Breil & Gally, 2017) montre que plus une parcelle agricole est considérée sensible aux aléas de ruissellement prédits par IRIP, moins grande est la quantité de pluie nécessaire à générer un incident sur cette parcelle. Au-delà d'un critère de performance, c'est un élément de validation du modèle.

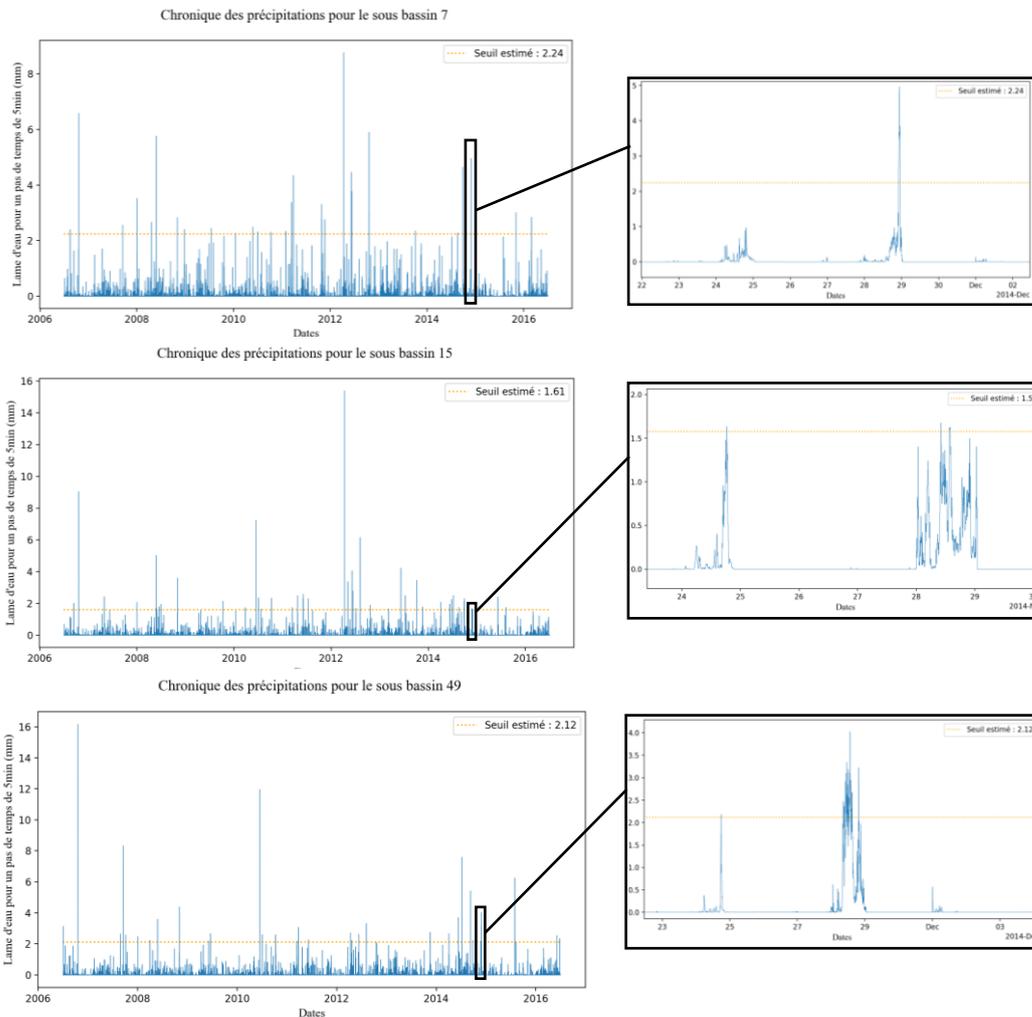


**Figure 19 :** (à gauche) Aléa moyen de chaque parcelle, comparé aux cumuls maximums pour chaque durée, (à droite) la droite de tendance et le nuage de point pour la durée de 1h. (note : si plusieurs parcelles agricoles se trouvent dans un même pixel de lame d'eau radar, seule celle dont l'aléa moyen est le plus faible est conservée)

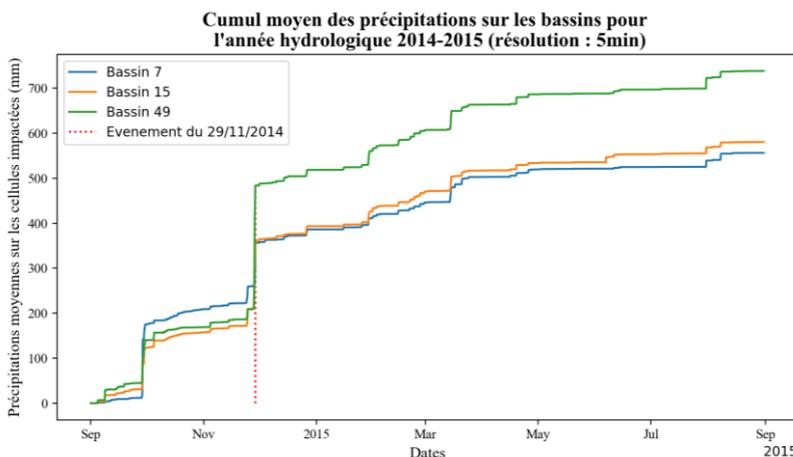
## f. Les sous bassins versants

Les cartes de cumuls mettent en évidence un fort effet de zone pour les distributions de précipitations. Pour essayer de bien comprendre l'évènement, nous allons travailler sur trois bassins versants qui ont subi des incidents sous des conditions différentes.

La figure 20 affiche les trois bassins choisis. Ils ont été sélectionnés parce qu'ils ont tous compté des incidents sur leur surface et par rapport à leur situation vis à vis de l'évènement (figure 18). Le numéro 49 a subi des cumuls importants et contient une partie des cellules les plus impactées. Le bassin 7 a subi des pluies intenses mais pour des faibles durées seulement. Enfin, le 15 est situé sur une zone contrastée, seule sa partie amont a subi des précipitations importantes sa partie aval semble avoir été épargnée pour les durées faibles comme fortes.



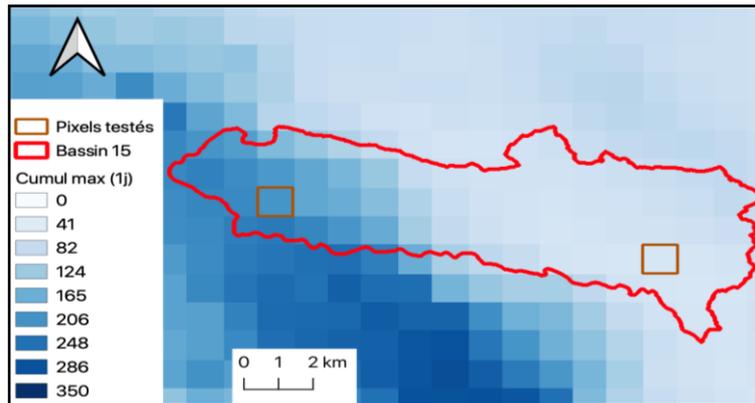
**Figure 20:** Chroniques sur 10 ans pour les trois bassins et zoom sur l'évènement de 2014.



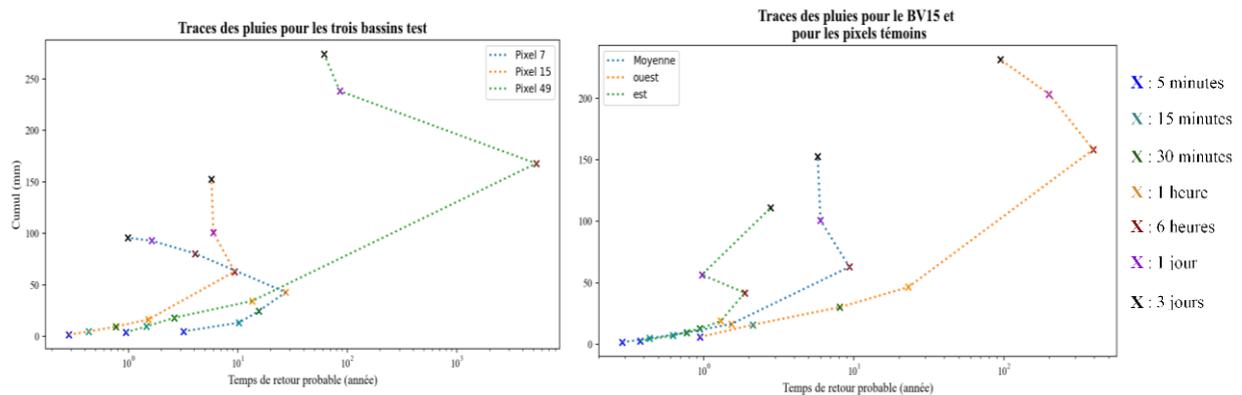
**Figure 21 :** Les valeurs de pluies cumulées pour les trois bassins versants. Deux évènements intenses en début d'année hydrologique sont responsables de plus de la moitié du cumul total sur l'année.

### Période de retour

En plus de la comparaison entre les bassins versants, nous avons décidé de comparer les résultats de l'ajustement de la loi exponentielle sur deux pixels du bassin versant 15, affichés figure 22. Ce bassin est représentatif du type de situation qu'il est possible d'observer pour un évènement aussi hétérogène. Les cellules à l'ouest ont subi jusqu'à 250mm en une journée alors qu'à l'est les valeurs sont moindre environ entre 40 et 80mm.



**Figure 22 :** Les deux pixels testés, la carte des cumuls max sur 1 jour affiche déjà une différence très importante



**Figure 23 :** Comparaison des traces entre les trois bassins versants à gauche et pour les pixels d'un même bassin versant

La comparaison des temps de retour par l'intermédiaire des traces figure 23 (gauche), montre une forte hétérogénéité des cumuls subis pour les trois bassins. Le bassin 49, affiche des valeurs très importantes qui se ressentent dans la période de retour estimée (nous rappelons ici qu'une période de retour de cette ordre est très imprécise, au vu de notre jeu de données, mais il confirme l'état remarquable de ces pluies). Les chroniques moyennées des bassins 7 et 15 affichent des résultats moins importants. Le premier a subi des précipitations intenses (de l'ordre de l'annuel) sur les petites durées, mais pas un gros cumul sur l'ensemble de l'événement, pour le deuxième c'est l'inverse. Le bassin 15 affiche une forte hétérogénéité de valeurs pour les cellules qui le composent (Figure 22, 23 (droite)). A l'ouest, il a subi des cumuls importants sur l'événement et des précipitations plus intenses même sur les petites durées. Pour l'est, les cumuls sont beaucoup plus faibles, et la moyenne est fortement impactée.

## 11. Discussion

Avant de conclure, nous allons revenir sur nos résultats, expliquer nos interprétations et les suggestions que nous proposons pour aller plus loin dans la démarche. Nous commencerons par parler des scripts, des études de cas et enfin de la méthode. Nous finirons en discutant des limites de l'approche et des perspectives que nous avons imaginées.

## a. Problèmes liés à la taille de la base PANTHERE

Au cours de ce stage nous ne disposions que des données PANTHERE de 2006 à 2016. Pour la France entière, ces dix ans représentent  $2 \cdot 10^{12}$  valeurs. Cela pose déjà des problèmes de temps de lecture et de gestion de la mémoire.

### *Temps d'exécution*

Le temps d'exécution des scripts est satisfaisant en général, si nous ne travaillons qu'avec une seule chronique (sur une cellule). Les routines qui ont besoin d'en ouvrir plusieurs, pour les comparer ou faire des moyennes spatiales, ont dû être adaptées. Pour l'importation des fichiers, le code n'ouvre qu'un fichier bloc à la fois, puis va chercher les cellules nécessaires. L'utilisation du module Pandas permet d'exploiter une structure de donnée performante (les DataFrames) et de disposer d'une bibliothèque de fonctions performantes. Ainsi, le calcul des chroniques moyennes et des cumuls est aisé et optimisé.

La détermination automatique du seuil donne les résultats attendus mais le temps de calcul est un peu long et devrait être améliorable (description Annexe X). Le gain de temps associé serait significatif sur une étude dirigée sur un grand territoire, pour laquelle le calcul serait répété.

### *Gestion de la mémoire*

La gestion de la mémoire vive de l'appareil utilisé peut donc poser un problème pour la lecture des fichiers de blocs et la manipulation de plusieurs chroniques ne serait-ce que pour dessiner les figures. Compléter la base de données avec les années récentes accentuera ces problèmes. A titre informatif les scripts ont bien tourné sur un MacBook pro de 2015 (avec 8Go de mémoire vive). Il fonctionne moins bien sur les machines INRAE. Nos comparaisons nous laissent à penser que l'anti-virus ralentit les traitements (en lecture), et consomme une partie de la mémoire instantanée.

## b. Analyse des cas d'étude IRIP

Les deux cas d'étude ont été choisis en fonction de la disponibilité et de la pertinence de leurs données. Ces deux cas de figures sont très différents. Ils ont permis de tester nos outils et d'imaginer des pistes pour compléter la méthode IRIP.

### *Le Grand Châlon*

Les résultats IRIP sur la région de Chalon sur Saône montrent une bonne détection des incidents par le modèle sur cet événement (tableau 4). L'étude des précipitations en rapport avec l'évènement a permis d'établir des cartes de cumuls maximums qui montrent une relative homogénéité de l'évènement pour les zones concernées par les incidents (figure 12). Entre le 24 et le 26 Juin 2016, les cumuls déduits des chroniques moyennes des bassins versants atteignent les 30 - 40mm (Figure 17, droite). L'ajustement IdF (loi exponentielle) a montré que cet événement n'était pas particulièrement exceptionnel au point de vue des cumuls pour le jeu de durées testé (de 5 minutes à trois jours) (Figures 16, 17).

Les données des mois passés ont montré que la période qui a précédé l'incident (depuis Avril environ, figure 15), était le théâtre d'une succession d'évènements pluvieux. Ils ont fourni entre 300 et 400mm. Les bassins versants avaient déjà subi au mois de Mai un événement important. La succession de petits événements entre les deux a probablement empêché une vidange suffisante du sol, qui n'aurait pas pu infiltrer les pluies. L'élargissement de la fenêtre temporelle pour une période de 60 jours

(Annexe F) nous a montré qu'on atteignait en effet des valeurs de cumul (cumuls glissants sur les 30 et 60 jours qui précèdent la valeur) importantes sur la période notamment par rapport aux autres années. Même si les cumuls n'ont pas été exceptionnels au cours de la journée. Ils sont tombés au cours d'une fenêtre temporelle réduite (30 minutes - 1 heure figures 16, 17) qui a probablement favorisé le ruissellement.

### *Les Pyrénées Orientales*

Les résultats d'IRIP se sont montrés en moyenne moins satisfaisants sur la région, pour l'événement qui concerne les jours entre le 28 et le 30 Novembre 2014 (tableau 5, Annexe F). Malgré tout, d'un bassin à l'autre l'efficacité de prédiction d'IRIP varie fortement.

Les cartes de cumuls ont montré une forte hétérogénéité de l'événement qui en l'espace d'un kilomètre pouvait varier du simple au double (figures 18, 22). Elles nous ont permis de sélectionner quelques bassins versants pour constater des différences.

Pour chaque bassin versant l'événement paraît plutôt isolé dans le temps (figure 20, 21). Quelques chutes d'eau mineures le séparent d'un autre intense mais sans incidence qui a eu lieu en Septembre. Les résultats ont montré qu'une tendance significative existe entre les aléas IRIP et les cumuls de pluies. Il semblerait que des zones les moins sensibles (au sens aléa IRIP faible) qui ont été impactées, aient subi de plus fortes précipitations (au moins à partir de 30 minutes - 1 heure, figure 19, Annexe G). A partir de ce constat, nous faisons l'hypothèse que des cumuls importants peuvent avoir un impact sur les résultats IRIP, puisque le score élevé ne suffit plus pour définir les zones endommagées. Cela reste cohérent avec la physique du processus qui pour des pluies exceptionnelles entraîne un ruissellement généralisé.

L'étude comparative menée sur les bassins (Figure 23) souligne l'hétérogénéité de l'événement sur le département. Les bassin 49 par exemple, a subi presque deux fois plus de précipitations que les deux autres sur les trois jours.

Le test effectué sur le bassin 15 (Figure 23), montre le contraste de valeurs possibles entre deux pixels d'un même bassin versant. En fonction des points soulignés précédemment, l'étude des incidents sur cette surface ne devrait pas donner des résultats équivalents si la parcelle se trouve à l'ouest ou à l'est de la zone.

## **c. Apport de la méthode**

Les deux situations étudiées ont montré des cas de figures très différents. Aucune conclusions définitives ne seront énoncées, mais, nos analyses ont permis de mettre en lumière quelques points intéressants :

- L'étude par bassin versant est intéressante pour nous. La création des chroniques moyennes et leur étude prend moins de temps que l'étude de chaque cellule et elle a un sens hydrologique. Si les événements sont très hétérogènes l'idéal serait de définir un bassin par point incident, pour mettre en relation la cause et l'effet.
- Si les résultats IRIP sont bien dépendants des cumuls, les cartes de cumuls maximaux (figure 12, 18) peuvent aider à comprendre la structuration de l'événement et à adapter l'angle d'approche d'étude pour une zone. Mais ce point souligne aussi le fait qu'il n'est pas anormal qu'IRIP ne "prédise" pas les incidents provoqués par des cumuls exceptionnels : en effet, ce constat implique que du ruissellement peut apparaître même sur les surfaces non propices a priori (aléas faibles IRIP).
- Les traces (figure 16, 17, 23) soulignent l'importance du multi-durées pour étudier un événement extrême. Les temps de retour peuvent fortement varier pour un même événement. Elles donnent aussi la possibilité de se faire une idée de la dynamique de la pluie, qui comme le cumul à son importance dans la genèse du ruissellement hortonien (Horton, 1933).
- Nous avons mentionné que des notions de temps de retour ont déjà été intégrées à IRIP et permettent d'estimer le ruissellement sur surface saturée ou par dépassement de la capacité

d'infiltration. Le temps de retour 30 ans a été choisi par expertise et la durée 1h parce que c'est la durée pour laquelle la première pente a été significativement différente de 0 sur la région de Chalon. Sur les Pyrénées Orientales, la première pente significativement différente de 0 pour un seuil de confiance 5% est celle de 30 minutes et pour un seuil de 1% c'est celle d'une heure (Annexe G). Ces premiers résultats sur les Pyrénées vont dans le sens de l'utilisation faite d'IRIP jusqu'alors.

## d. Limites

- Les résultats d'IRIP sont compliqués à interpréter tels quels. L'étude du modèle dépend fortement de la donnée disponible. Le relevé d'incident n'est pas systématique et non exhaustif car limité aux zones à enjeux. De plus, les jeux de données peuvent être très hétérogènes (nature, précision de la localisation, indication de la gravité, description précise des phénomènes, disponibilité).
- La donnée PANTHERE est issue de mesures indirectes qui ont subi de nombreux post-traitements, et qui comportent un certain nombre d'incertitudes. Les données disponibles s'étalent pour le moment sur une trop faible profondeur temporelle (10 ans pour nous 15 ans bientôt).
- Le reconditionnement des fichiers champs PANTHERE en fichiers binaires mentionné plus haut, a entraîné pour certaines dates l'écrasement de certains fichiers champs, et l'apparition de doublons sur certaines journées dans les chroniques (figure 23). Le cas s'est présenté pour les Pyrénées mais la journée a été supprimée. Le problème ayant été identifié l'algorithme sera corrigé.

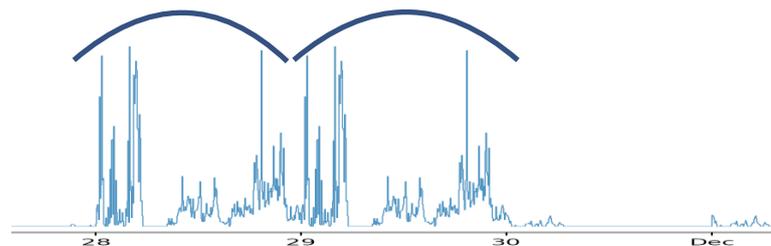


Figure 24 : Doublons pour la date du 29 Novembre 2014.

## e. Perspectives

Pour aller plus loin, nous proposons les pistes de réflexions suivantes :

- Des problèmes informatiques se sont matérialisés sous la forme de limitation de la mémoire et de temps d'exécution trop important. Ces problèmes s'aggraveront quand les bases de données seront élargies. Des solutions sont possibles, par exemple, aller plus loin dans la réflexion sur le code, paralléliser les calculs ou encore l'utilisation de la mémoire virtuelle.
- L'utilisation des données ANTILOPE devrait être envisagée, car elles sont réputées plus fiables que PANTHERE sur les valeurs faibles d'après Tabary (2013). Cependant, ANTILOPE est agrégée à l'heure et nos travaux n'ont pas tranché sur les durées pertinentes.
- Les ajustements IdF au pixel pourraient être consolidés par régionalisation (méthode Hosking-Wallis in Ekoko-Safi, 2018).
- La figure 19 associée à l'annexe G montrent que la différence entre aléa faible et aléa fort est la plus importante pour la durée de 6h. Il serait intéressant de tester les résultats IRIP en le prenant en compte pour éventuellement aller vers une adaptation par région de cette durée. Mais cela reste très dépendant de la manière de caractériser l'aléa d'une parcelle car sa susceptibilité peut être forte sur

une partie seulement. Le choix de la moyenne de l'aléa dans cette étude était exploratoire et suivait la méthode utilisée dans d'autres régions étudiées par ailleurs.

- De nouvelles données pourraient être intéressantes pour IRIP. Les mesures satellitaires d'humidité du sol par exemple, permettraient à la fois d'apporter une explication précise aux problèmes comme celui de Chalon et de tenir à jour les prédictions d'IRIP en fonction de l'humidité évolutive du sol. La connaissance des lames d'eau radar en temps réel permettrait aussi de calculer un indice d'humidité des sols qui servirait à la mise en vigilance en cas de pluies intenses annoncées.

- Nous avons proposé une méthodologie destinée à comprendre les répartitions spatiales et temporelle d'un événement de manière à pouvoir préciser les résultats IRIP. Les constats que nous avons faits sont eux-mêmes voués à être réétudiés sur des jeux de données plus importants.

## 12. Conclusion

Ce travail a permis d'avancer dans la conception d'un outil pratique d'exploitation des lames d'eau RADAR. Les scripts proposés sont fonctionnels et prêts à être intégrés au code. Ils ajouteront la possibilité de travailler sur des nuages de points et des fichiers shapes (par moyenne spatiale). Ils permettent ainsi de faire un lien avec QGIS sur lequel fonctionne IRIP.

L'échantillonnage par le biais de la méthode SUPSEUIL s'est montré efficace, et permet de construire des courbes IdF plutôt fiables et d'estimer des temps de retour crédibles, sans que l'exécution du code soit chronophage au moins pour l'analyse d'une chronique. L'analyse multi-durées a montré son caractère indispensable pour qualifier un événement extrême. Le jeu de données PANTHERE, trop court (seulement 10ans), ne peut en revanche pas nous donner des idées précises des temps de retour probable les plus rares.

Les études des événements de Chalon et des Pyrénées Orientales ont permis de tester le fonctionnement des scripts et de mettre à jour les problèmes et les conditions de bon fonctionnement. Les précipitations qui ont impacté les Pyrénées ont été bien plus intenses que pour Chalon, avec 300mm en quelques heures pour le coeur de l'événement. D'après nos résultats, il apparaît que les cumuls importants sont susceptibles de déclencher l'apparition de ruissellement sur des zones à aléas faibles, ce qui impliquerait de considérer des seuils IRIP plus faibles que 3 ou 4. Dans ce cas la surface de bassin versant concernée par l'aléa ruissellement devient importante et perd de son intérêt opérationnel en terme de mesures préventives ciblées. Au-delà des cumuls pluviométriques, il faut bien noter la différence de nature des données d'incident entre Chalon et les P.O. Dans le cas des calamités agricoles nous traitons des surfaces de parcelles parfois de plusieurs hectares et les impacts peuvent être internes ou sur les bordures (routes, fossés). Une réflexion est à mener sur la manière d'évaluer la détection par IRIP des parcelles endommagées.

L'utilisation des scripts a permis de mettre en évidence des méthodes qui pourraient s'avérer utiles pour adapter les prochaines études sur IRIP. Notamment du point de vue de l'utilisation de la structure spatiale de la pluie pour déterminer des zones d'études choisies. Ils permettent une étude des périodes de dépassement plus poussée qui sera sûrement nécessaire pour qualifier les territoires sur lesquels le modèle est déployé.

## 13. Bibliographie :

- Braud I., Lagadec L.R., Moulin L., Chazelle B., & Breil P. 2020. A method to use proxy data of runoff-related impacts for the evaluation of a model mapping intense storm runoff hazard: application to the railway context. *Natural Hazards and Earth System Sciences*. 20(4), p. 947-966.
- Breil P. & Gally T. 2018. *CARTO ALEA (Zones inondables) : Améliorations de la méthode IRIP au niveau national*. Villeurbanne : IRSTEA, 25 p.
- Cunnane C. 1979. A note on the Poisson assumption in partial duration series models. *Water Resources Research*. 15(2), p. 489-494.
- Darricau J. 1981. *Physique et théorie du RADAR, Description générale*. SODIPE, Tome 1. Disponible sur : <https://radars-darricau.fr/livre/1-Pages/page-01.html> (Consulté le 10 Juillet 2020)
- Ekoko-Safi A. 2016. *Analyse des données de lame d'eau radar pour l'étude des pluies fortes en France métropolitaine. Une contribution à la méthode IRIP de détermination de l'aléa de ruissellement intense* (Master, Sciences des Environnements Continentaux et Côtiers). Rouen : Université de Rouen, 55 p. Disponible sur : <https://hal.inrae.fr/hal-02606539> (Consulté le 12 Août 2020).
- Fang H.Y., Cai Q.G., Chen H., & Li Q.Y. 2008. Effect of Rainfall Regime and Slope on Runoff in a Gullied Loess Region on the Loess Plateau in China. *Environmental Management*. 42(3), p. 402- 411.
- Jaillet C., Poulard C., & Breil P. 2012. *Méthodes de cartographie du ruissellement (au sens D.I.) : Synthèse des méthodes de cartographie des inondations par ruissellement en France*. Villeurbanne : IRSTEA, 33 p.
- Horton R.E. 1933. The role of infiltration in the hydrologic cycle. *American Geophysical Union Transactions*. 14, p. 445-460.
- Lagadec L.R. 2016. *Evaluation et développement de la méthode IRIP de cartographie du ruissellement. Application au contexte ferroviaire* (Thèse, spécialité "Océan, Atmosphère, Hydrologie"), Grenoble : Université Grenoble Alpes, 336 p.
- Lang M., Ouarda T.B.M.J., & Bobée B. 1999. Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*. 225(34), p. 103-117.
- Moncoulon D., Labat D., Ardon J., Leblois E., Onfroy T., Poulard C., Aji S., Rémy A., & Quantin A. 2014. Analysis of the French insurance market exposure to floods: a stochastic model combining river overflow and surface runoff. *Natural Hazards and Earth System Sciences*. 14(9), p. 2469- 2485.
- Nicolas M. 2010. *Étude expérimentale et numérique du ruissellement de surface: effets des variations d'intensité de la pluie. Application à une parcelle de vigne en Cévennes-Vivarais*. (Thèse, spécialité "Terre, Univers, Environnement"), Grenoble : Université de Grenoble, 217 p.
- Renard, B. 2014. Probabilités et Statistiques appliquées à l'Hydrologie. Présenté à : Ecole Nationale des Travaux Publics de l'Etat (ENTPE, 3<sup>e</sup> année), Vaulx-en-Velin.
- Renard B., Leblois E., Kuczera G., Kavetski D., Thyer M., & Franks S. 2009. Characterizing errors in areal rainfall estimates: application to uncertainty quantification and decomposition in hydrologic modelling. Présenté à : *32nd Hydrology and Water Resources Symposium*, Newcastle.
- Schoorens J., Emmanuel I., & Balier V.L. 2013. Influence de l'entrée de pluie en modélisation des systèmes d'assainissement : cas du bassin versant de Boulogne-Billancourt. Présenté à : *Novatech 2013*, Lyon-Villeurbanne.
- Tabary P., Augros C., Champeaux J.-L., Chèze J.-L., Faure D., Idziorek D., Lorandell R., Urban B., & Vogt V. 2013. Le réseau et les produits radars de Météo-France. *La Météorologie*. 8(83), p. 15.
- Tóth B., Weynants M., Pásztor L., & Hengl T. 2017. 3D soil hydraulic database of Europe at 250 m resolution. *Hydrological Processes*. 31(14), p. 2662-2666.
- Wasson J.G., Chandresris A., Pella H., & Blanc L. 2004. Les hydro-écotémoins: une approche fonctionnelle de la typologie des rivières pour la Directive cadre européenne sur l'eau. *Ingénieries eau-agriculture-territoires*, 40. Disponible sur : <https://hal.archives-ouvertes.fr/hal-00475620> (Consulté le 8 Septembre 2020).

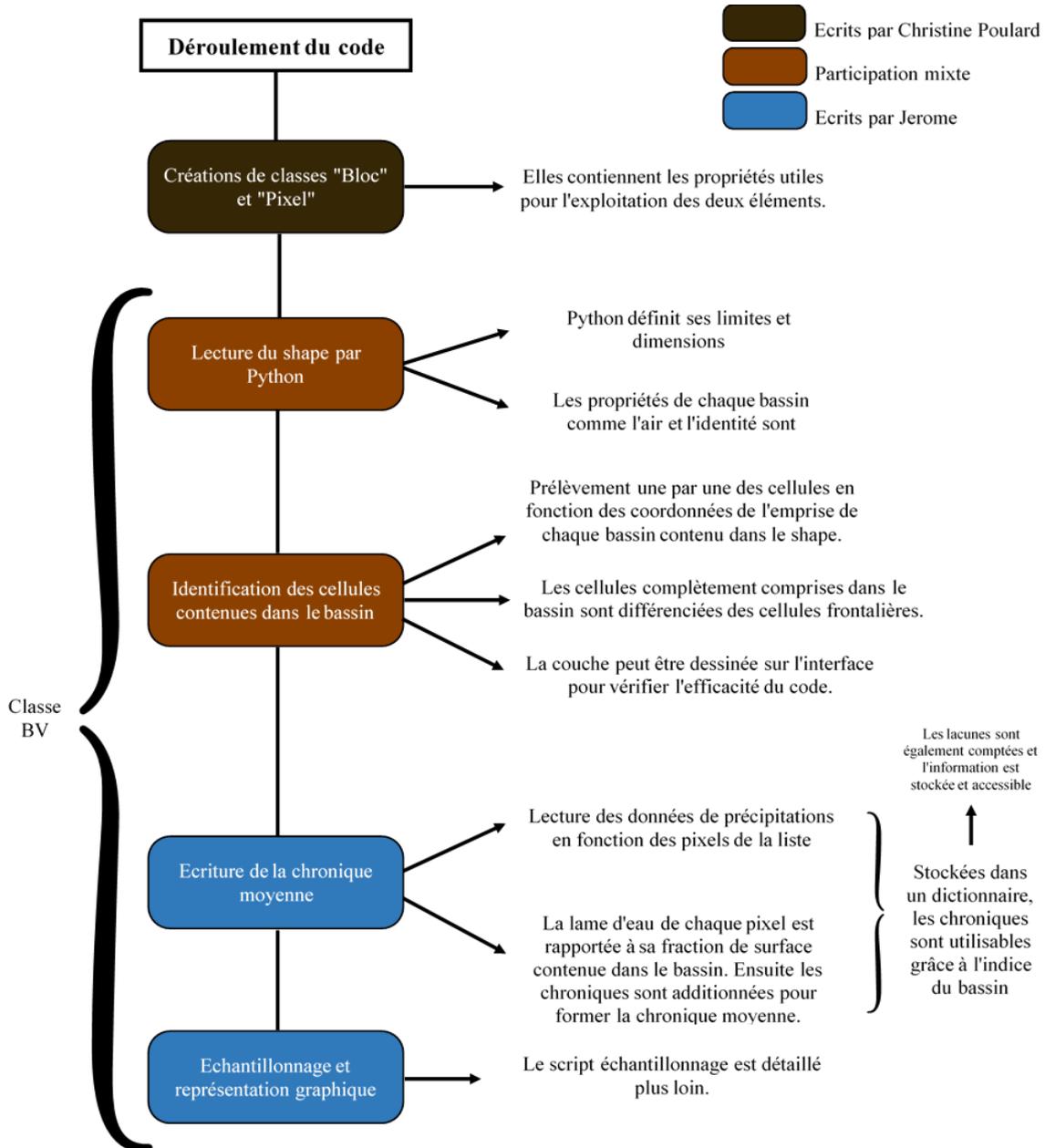
## 14. Sitographie :

Carte d'inondation du gouvernement Anglais : <https://flood-warning-information.service.gov.uk/long-term-flood-risk/map>

Pour aller plus loin sur les données radar : <http://wikhydro.developpement-durable.gouv.fr/index.php/B.04> -

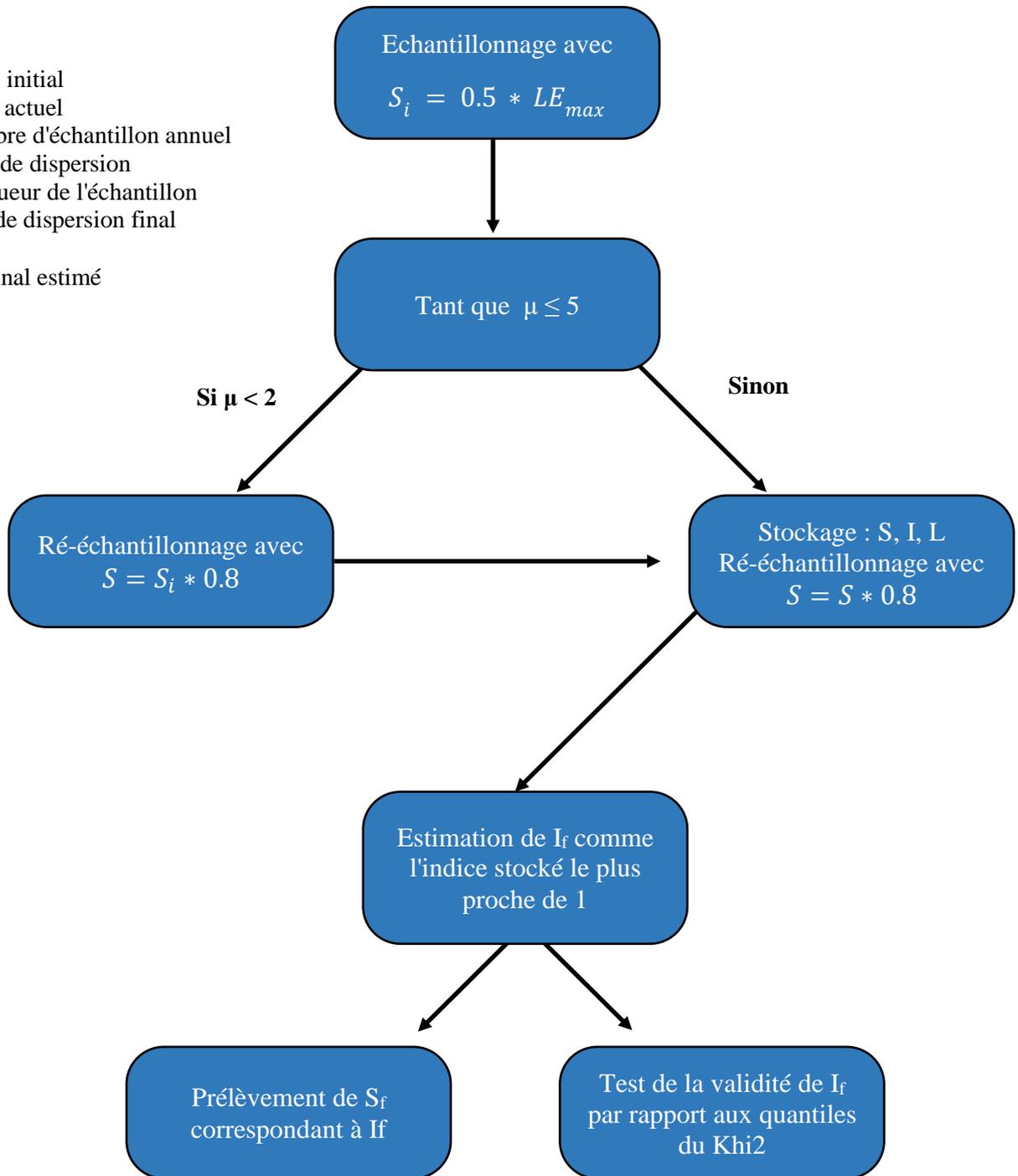
[Estimation d'une pluie de bassin par observation RADAR](#)

# 15. Annexes A: Script étude des bassins versant

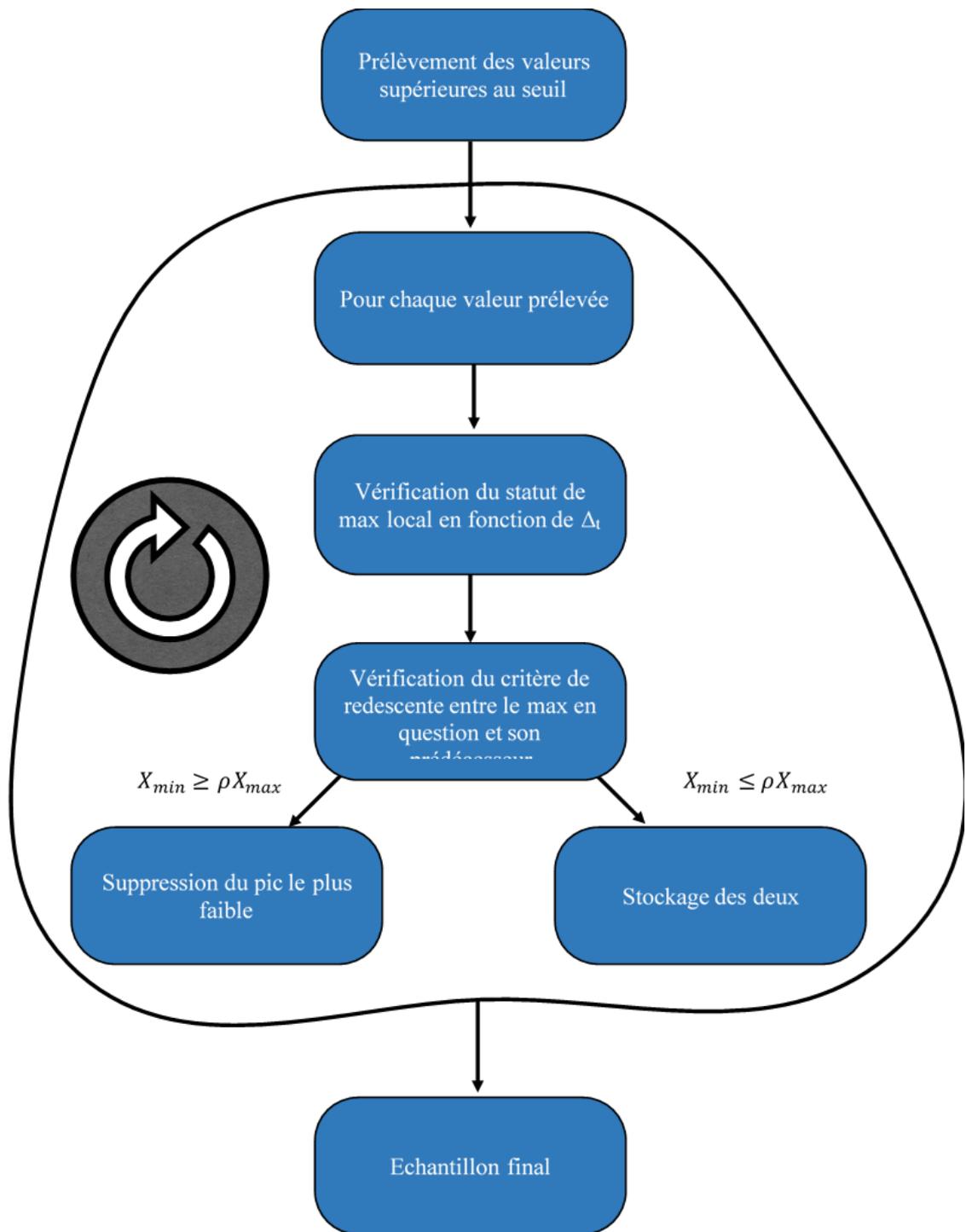


# 16. Annexes B : Script détermination automatique du seuil

Avec :  
 $S_i$  : le seuil initial  
 $S$  : le seuil actuel  
 $\mu$  : le nombre d'échantillon annuel  
 $I$  : l'indice de dispersion  
 $L$  : la longueur de l'échantillon  
 $I_f$  : indice de dispersion final estimé  
 $S_f$  : seuil final estimé

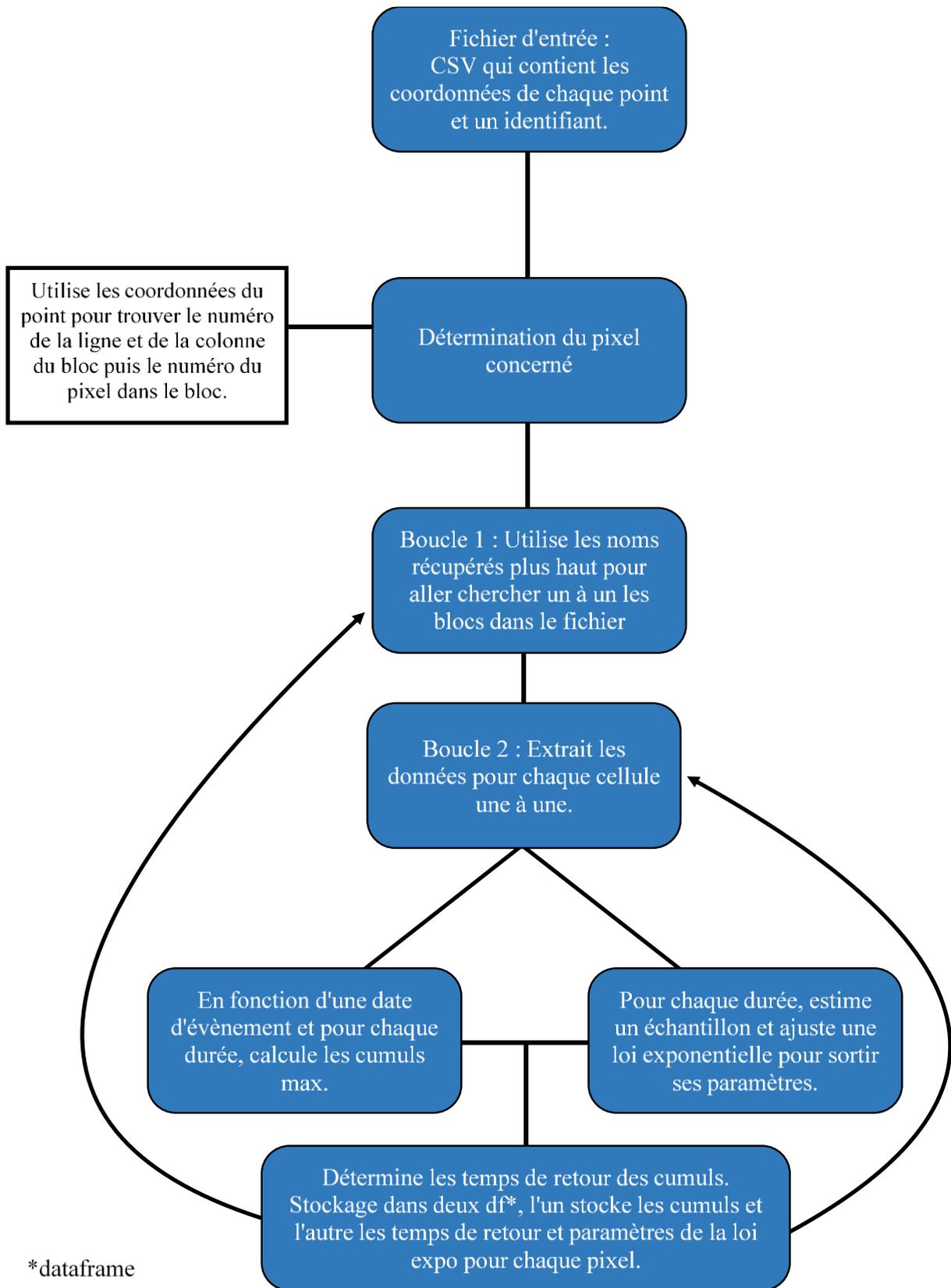


# 17. Annexe C : Echantillonnage



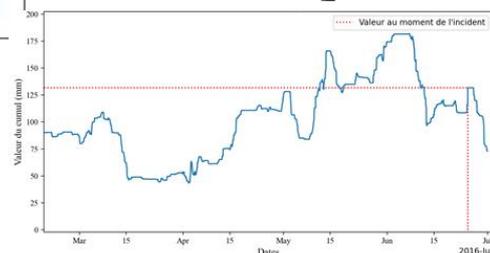
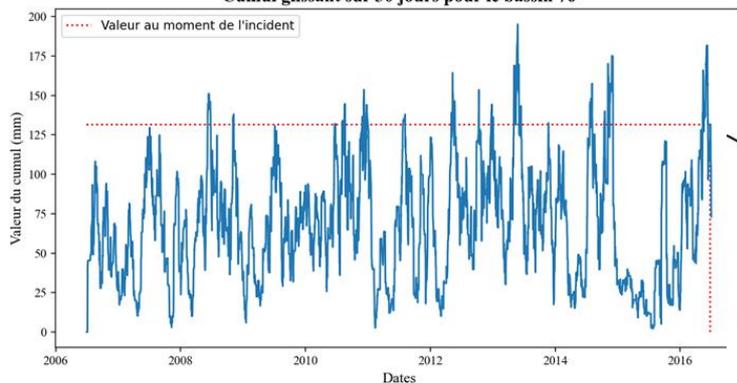
$\rho$  : coefficient de redescente

# 18. Annexe D : Sélection par nuage de points

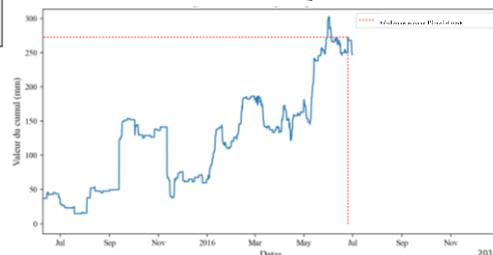
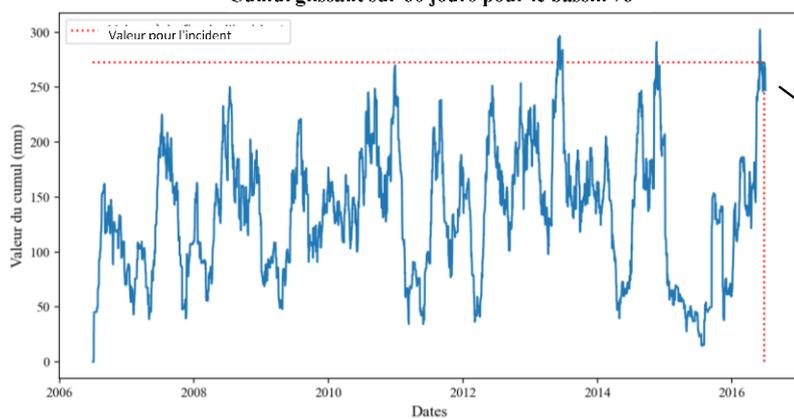


# Annexe E : Cumuls glissants sur des périodes longues pour le bassin test de Châlon

Cumul glissant sur 30 jours pour le bassin 76

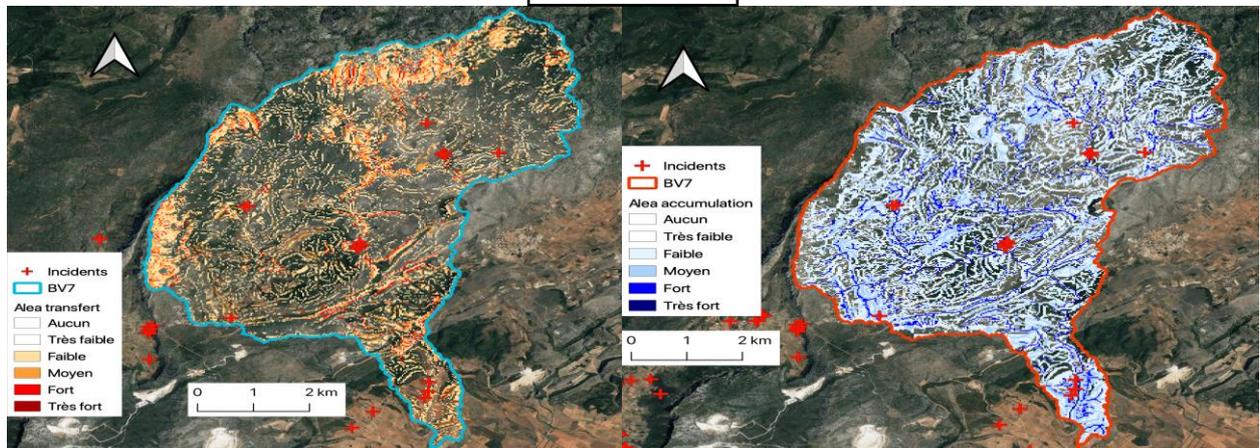


Cumul glissant sur 60 jours pour le bassin 76

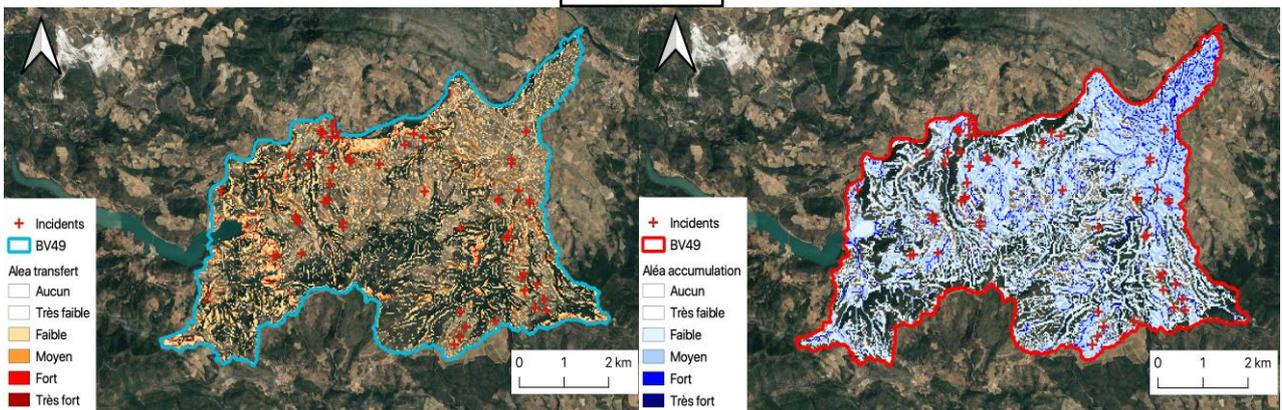


# 19. Annexe F : Cartes IRIP pour les 3 bassins témoins des Pyrénées Orientales

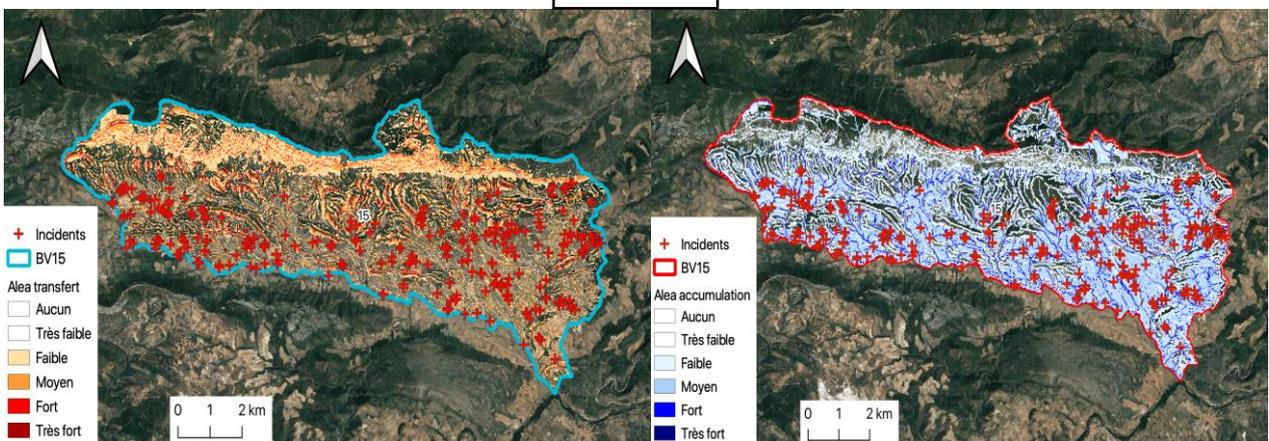
**BV7**



**BV49**



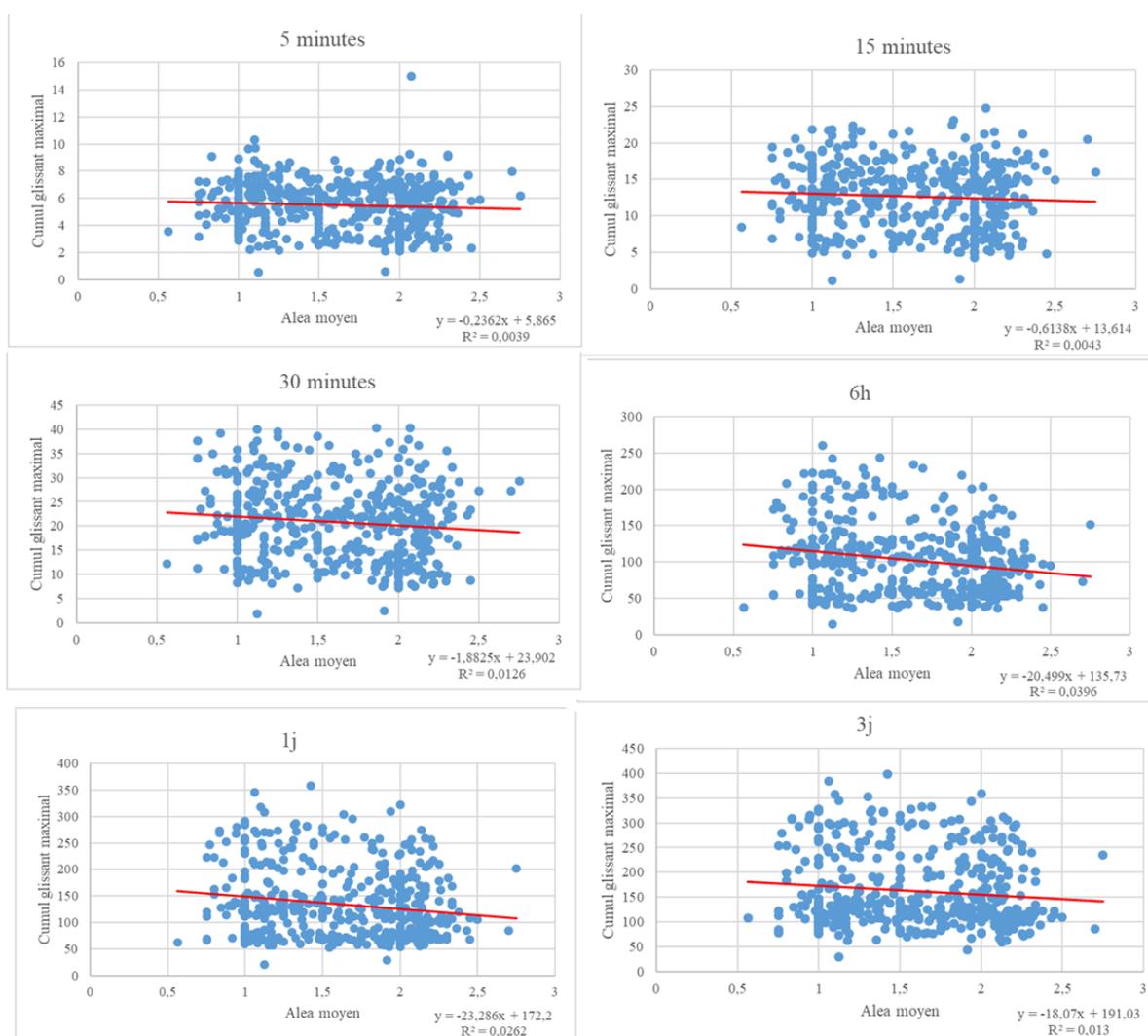
**BV15**



# 20. Annexe G : Relation Aléa vs Cumul max & Test de la pente

Résultat du test de Student sur les pentes pour les différentes durées

Durée (min)	5	15	30	60	360	1440	4320
p	0.17	0.15	0.016	0.0028	0.000017	0.000482	0.014364







## Résumé

Dans cette action, la première étape a été de contribuer à l'élaboration d'outils pour visualiser et analyser des données de pluies issues des lames d'eau radar (produit PANTHERE de MétéoFrance). Leur nature spatialisée nous permet de calculer des chroniques de lame d'eau moyennées à l'échelle d'un sous-bassin, ou de tout territoire pertinent pour étudier la genèse du ruissellement. Nous avons développé des outils d'analyse multi-durées, après échantillonnage SUPSEUIL, qui estiment une relation locale Intensité-durée-Fréquence-Surface (IdF). Nous proposons de représenter les événements par leur "trace" obtenue grâce aux différentes courbes Intensité-durée-Fréquence.

Nos outils nous ont permis de mettre au point une méthodologie d'étude qui complète l'étude IRIP. Certains événements avec des cumuls exceptionnels peuvent survenir dans des zones où le score IRIP est faible mais où des incidents sont constatés. De même des intensités plus faible peuvent expliquer des incidents où les scores IRIP sont élevés. Il est alors intéressant de qualifier les incidents par les périodes de retour des pluies. Il est cependant difficile de définir un événement par une durée unique. De même quelle surface considérer si l'on convient d'un transfert amont – aval du ruissellement ou encore d'un décalage entre la zone de pluie au sol et sa représentation par maille radar du fait de vents violents qui peuvent accompagner les pluies intenses. La méthodologie mise au point dans ce travail devra être testée sur un plus grand nombre d'échantillons pour tenter de répondre aux questions soulevées.

Les scripts Python produit dans ce travail peut-être utilisé soit indépendamment soit intégré dans une interface graphique préexistante (C. Poulard et al., 2019) et avec d'autres sources de données à pas de temps fixe. Ils pourraient à terme être intégrés au plugin IRIP pour QGIS.

**Mots clé :** Ruissellement intense, Lame d'eau radar, Intensité-durée-fréquence, période de retour