



**HAL**  
open science

# Développement d'un outil opérationnel de recherche automatique d'analogues d'évènements de crues

Charlotte Raynal

► **To cite this version:**

Charlotte Raynal. Développement d'un outil opérationnel de recherche automatique d'analogues d'évènements de crues. Sciences de l'ingénieur [physics]. 2023. dumas-04462962

**HAL Id: dumas-04462962**

**<https://dumas.ccsd.cnrs.fr/dumas-04462962>**

Submitted on 16 Feb 2024

**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.

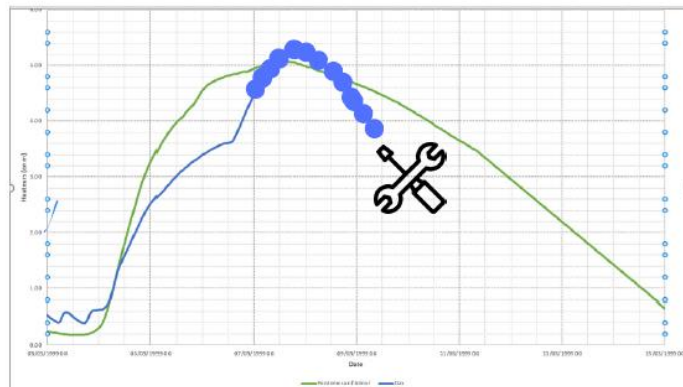
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Travail de fin d'études

pour le diplôme d'ingénieur de l'École nationale des travaux publics de l'État

## Développement d'un outil opérationnel de recherche automatique d'analogues d'évènements de crues



Réalisé par :

Charlotte RAYNAL

3A Risques nuisances pollutions (ENTPE, Année 2022-2023)

Soutenu le 28/08/2023

Devant le jury composé de :

- Président du Jury : Laurence Volatier
- Maître de TFE [ou Tuteur] : Yan Lacaze (DREAL Nouvelle Aquitaine)
- Expert : Elsa Laganier (DREAL AuRA)

# Notice analytique

AUTEUR			
Nom	Raynal		
Prénom	Charlotte		
ORGANISME D'ACCUEIL			
Nom de l'organisme et Localité	Direction Régionale de l'Environnement, de l'Aménagement et du Logement Nouvelle Aquitaine, Bordeaux		
Nom du Tuteur	Yan Lacaze		
ANALYSE DU TFE			
Titre (français)	Développement d'un outil opérationnel de recherche automatique d'analogues d'évènements de crues		
Titre (anglais)	Development of an operational automatic tool for the search of flood event analogue.		
Résumé (français)	<p>Le risque inondation est le premier risque naturel en France, possédant des conséquences économiques et humaines catastrophiques, notamment dans le contexte de changement climatique. Il est donc important de développer de nouveaux outils et d'améliorer ceux existants afin de prévoir les phénomènes de débordement des cours d'eau. L'objectif de ce TFE est de développer un outil de recherche d'analogues des évènements de crue en s'appuyant sur un outil existant : le catalogue de crues. Ce travail développe dans un premier temps les recherches nécessaires pour déterminer les paramètres permettant la description d'un évènement. Puis, il présente une méthode permettant de prévoir les crues futures à partir de celles passées, sur le tronçon Adour moyen-Adour des Barthes.</p>		
Résumé (anglais)	<p>The flood risk is the first natural hazard in France, with catastrophic economical and social consequences, especially in a context of global warming. That is why, it is important to develop new tools and improve existing ones, to predict flooding of rivers. The purpose of this TFE is to develop a tool for the search of flood event analogue using an existing one : the flood catalogue. In a first time, this work develop the necessary research to find parameters that describe a flood event. Then, it describe a methodology that can help to predict future flood event from the past ones. This methodology is applied on the Adour moyen-Adour des barthes section.</p>		
Mots-clés (français, 5 maxi)	Catalogue de crues, prévision 24h, critère d'évaluation, critère de choix		
Mots-clés (anglais, 5 maxi)			
Termes géo-graphique (français)	Bassin Gironde-Adour-Dordogne (GAD), Adour moyen, Adour des Barthes, Dax		
COLLATION			
	Nb de pages	Nb d'annexes (nb de pages)	Nb de réf. biblio.
	53	7	10

# Déclaration de travail personnel

---

Je déclare que ce rapport constitue l'aboutissement d'un travail personnel et ne peut être suspecté de plagiat.

Le travail présenté distingue explicitement ce que j'ai produit de ce que j'ai emprunté à d'autres. A ce titre, les citations sont clairement identifiables et les sources (écrits, images) qui ont alimenté ma réflexion sont référencées.

# Sommaire

---

Notice analytique.....	2
Déclaration de travail personnel.....	3
Sommaire.....	4
Liste des illustrations.....	6
Liste des tableaux.....	7
Sigles & Acronymes.....	8
1. Cadre de travail du Service Prévision des crues, structure d'accueil du travail de fin d'étude.....	9
1.1 La prévision des crues en France.....	9
1.2 Le SPC Gironde Adour Dordogne (SPC GAD).....	10
1.2.1 Territoire.....	10
1.2.2 Organisation et missions du SPC.....	11
1.3 Contexte, moyen et outils.....	12
1.3.1 Contexte du SPC GAD.....	12
1.3.2 Moyens et outils disponibles.....	13
1.4 Zoom sur un outil indispensable : le catalogue de crue.....	16
1.4.1 Les données du catalogue.....	16
1.4.2 Son utilisation actuelle.....	17
1.5 Enjeux et objectifs du TFE.....	18
1.5.1 Les enjeux.....	18
1.5.2 Objectifs.....	18
2. Secteur global de l'étude : les tronçons Adour moyen, Midouze et Adour des Barthes.....	19
2.1 Choix d'une station pertinente pour réaliser une méthode de recherche d'analogues : la ville de Dax.....	19
2.1.1 Difficultés rencontrées actuellement en prévision des crues à Dax.....	19
2.1.2 Le fonctionnement hydraulique du secteur.....	20
2.1.3 Les enjeux dépendants d'une bonne prévision à Dax.....	20
2.2 Données utilisées.....	22
2.2.1 L'Adour des Barthes.....	22
2.2.2 La Midouze.....	23
2.2.3 L'Adour moyen.....	24
3. Méthodologie.....	26
3.1 Principes de fonctionnement de la méthode de recherche d'analogues.....	26
3.1.1 La méthode d'analogues en météorologie.....	26
3.1.2 Application à la prévision des crues.....	26
3.2 Les variables d'analogie en prévision des crues.....	28
3.2.1 Hauteur max.....	28
3.2.2 Volume.....	28

3.2.3	Durée.....	29
3.2.4	Gradient de montée .....	29
3.3	Critère d'analogie en prévision des crues .....	30
4.	Procédure d'obtention du jeu de variables optimal .....	32
4.1	1ere méthode de recherche : la méthode « intégratrice » .....	32
4.2	2 <sup>ème</sup> méthode de recherche : la méthode de calage .....	32
4.3	3 <sup>ème</sup> méthode de recherche : la méthode finale .....	34
5.	Proposer une prévision à partir des analogues .....	40
5.1	Combinaison linéaire : la méthode « intégratrice » et calage .....	40
5.2	Prise en compte de plusieurs analogues pour proposer une prévision : méthode finale.....	42
6.	Résultats .....	43
6.1	Résultats de la première méthode .....	43
6.2	Résultat de la deuxième méthode .....	46
7.	Conclusion.....	51
8.	Perspectives.....	52
	Bibliographie.....	53
	Annexes .....	54

# Liste des illustrations

Figure 1.1.1: Organisation des services de prévision des crues (SPC) Bilan 2022 (réseau Vigicrues).....	9
Figure 1.2.1: Bassins versants du SPC GAD (RIC SPC GAD).....	10
Figure 1.2.2: Organigramme simplifié du DHPC GAD .....	11
Figure 1.3.1 : Processus de fonctionnement d'un modèle pluie-débit (Institut des Sciences Technologiques de l'Environnement (ISTE), Ecole Polytechnique Fédérale (EPFL)) .....	13
Figure 1.3.2 : Cartes des outils disponibles pour la prévision des crues sur les stations réglementaires .....	14
Figure 1.3.3: Schéma de construction de la lame d'eau ANTILOPE (météo France) .....	14
Figure 1.3.4: Processus de production des courbes de tarage.....	15
Figure 1.3.5: Processus de production de données hydrométriques .....	16
Figure 1.4.1: Extrait de la première feuille du catalogue de crues Adour moyen-Adour des Barthes .....	17
Figure 2.1.1: Limnigramme de la station de Dax lors de la crue de mars 1999.....	19
Figure 2.1.2 : Carte de Dax et des barthes de l'Adour.....	20
Figure 2.1.3: Les niveaux de protection pour chaque tronçon du système d'endiguement de Dax .....	21
Figure 2.1.4 : Les zones concernées par le risque d'inondation à Dax (PPRI Ville de Dax) .....	22
Figure 2.2.1 : Seuils de vigilance et évènement importants sur le tronçon Adour des Barthes .....	23
Figure 2.2.2: Seuils de vigilance et évènement importants sur le tronçon Midouze.....	24
Figure 2.2.3: Seuils de vigilance et évènement importants sur le tronçon Adour moyen .....	25
Figure 3.1.1 : 4 évènement observé à la station Pontonx-sur-l'Adour.....	26
Figure 3.2.1: Schéma représentant le paramètre « hauteur max » .....	28
Figure 3.2.2: Schéma représentant le paramètre « volume » .....	28
Figure 3.2.3: Schéma représentant le paramètre « durée ».....	29
Figure 3.2.4: Schéma représentant le paramètre « gradient de montée » .....	29
Figure 3.3.1: Représentation du $t_0$ , $t_0+24h$ et $t_0+48h$ .....	30
Figure 3.3.2 : Calcul du critère de Nash-Sutcliffe en fonction des hauteurs.....	31
Figure 4.1.1: Schéma explicatif du calcul des 3 meilleurs évènements pour chaque évènement du catalogue .....	32
Figure 4.2.1: Schéma de différents évènements de crues ayant lieu sur la station de Pontonx sur l'Adour .....	33
Figure 4.3.1: Graphique de l'évènement de mars 1999 à Dax et Pontonx-sur-l'Adour.....	34
Figure 4.3.2: Schéma explicatif Hmaxabs (a) et Hmaxrebond (b) pour un évènement de catalogue.....	35
Figure 4.3.3: Logigramme de sélection des Hmaxabs ou Hmaxrebond .....	36
Figure 4.3.4: Pente dans le cas $H_{max}=H_{maxrebond}$ (b) Pente dans le cas $H_{max}=H_{maxabs}$ .....	36
Figure 4.3.5: Schéma explicatif des volumes calculés pour les évènements du catalogue en fonction de l'évènement observé.....	37
Figure 4.3.6: Nuage de points théorique des valeurs de Nash24(ou Nash48) en fonction de la distance avec courbe de régression linéaire .....	38
Figure 4.3.7: Logigramme de la méthode finale de recherche d'analogie d'évènement de crue.....	39
Figure 5.1.1: Logigramme de la première méthode de recherche .....	40
Figure 5.1.2: Schéma explicatif de la méthode d'optimisation de la combinaison linéaire dans le cas où l'évènement observé possède 3 « meilleurs » évènements .....	41
Figure 5.2.1 : Repère temporel de l'évolution des différentes méthodes testées pendant mon stage.....	42
Figure 6.1.1: Image de l'évènement 8 à Pontonx et de son meilleur évènement dont le Nash est inférieur à 0.5 .....	43
Figure 6.1.2 : Graphe de l'évènement observé à Dax ainsi que de la combinaison linéaire des trois meilleurs évènements. ....	44
Figure 6.1.3: Images du calage de l'évènement 1 avec l'évènement 2 .....	45
Figure 6.2.1: Nuages de points avec le meilleur $R^2$ .....	47
Figure 6.2.2 : Graphe des hauteurs à Dax entre l'évènement observé et celui prévu selon le critère de distance .....	48
Figure 6.2.3: Nuages de points avec les meilleurs $R^2$ .....	49
Figure 6.2.4 : Graphique de 2 évènements observés avec leur évènement prévu et la valeur de leur critère de Nash en titre des graphes.....	50

# Liste des tableaux

---

Tableau 1.2.1: Tableau récapitulatif des missions du RRI.....	12
Tableau 2.1.1: Hauteurs d'eau auxquelles les zones de Dax ne sont plus protégées.....	21
Tableau 3.1.1: Tableau descriptif des différents groupe de score pour les modèle de prévision (Guide d'utilisation scores V1.3.3, réseau Vigicrues).....	27
Tableau 6.1.1: Tableau récapitulatif du nombre d'évènements possédant un, deux trois ou aucun « sosie ».....	43
Tableau 6.1.2: Tableau récapitulatif des coefficients optimaux pour la combinaison linéaire des évènements « prévus ».....	44
Tableau 6.1.3 : Tableau récapitulatif du nombre d'évènements possédant un, deux trois ou aucun « sosie » après calage du Nash à Pontonx .....	45
Tableau 6.1.4: coefficient optimaux pour une combinaison linéaire des évènements « prévus » .....	46
Tableau 6.1.5: Tableau détaillant les valeurs des moyennes de Nash à Dax en fonction de la catégorie à laquelle appartiennent les évènements observés .....	46
Tableau 6.2.1: Résultats obtenus par station ( $a > 0$ signifie que le coefficient directeur de la courbe de régression linéaire est supérieur à 0).....	47
Tableau 6.2.2: Résultats obtenus par paramètres ( $a > 0$ signifie que le coefficient directeur de la courbe de régression linéaire est supérieur à 0).....	47
Tableau 6.2.3 : Résultats obtenus par station, après coïncidence des rebonds.....	48
Tableau 6.2.4: Résultats obtenus par paramètres, après coïncidence des rebonds .....	48



# Sigles & Acronymes

---

**ANTILOPE** Analyse par spatialisatiOn hOraire des prEcipitations.

**AROME** Application de la Recherche Opérationnelle à MEsoéchelle

**ARPEGE** Action de Recherche Petite Echelle Grande Echelle

**CVH** Cellule de Veille Hydraulique

**DHPC** Département Hydrométrie et PrÉvision des Crues

**PPRI** Plan de PrÉvention Risque Inondation

**RIC** Règlements d'Information sur les crues

**RR3** Rainfall Rate 3heures

**SCHAPI** Service Central d'Hydrométéorologie d'Appui à la PrÉvision des Inondations

**SDPC** Schéma Directeur de PrÉvision des Crues

**SPC** Service PrÉvision des Crues

**SRNH** Service des Risques Naturels et Hydrauliques

**UH** Unité Hydrométrique

# 1. Cadre de travail du Service Prédiction des crues, structure d'accueil du travail de fin d'étude.

## 1.1 La prédiction des crues en France

En France le risque inondation est le premier risque naturel, menaçant tous les territoires, avec 17,1 millions d'habitants exposés aux conséquences des débordements de cours d'eau (dont 16,8 millions en France métropolitaine).

La surveillance, la prédiction et la transmission de l'information sur les crues sont mises en œuvre par une organisation de l'Etat, conformément au code de l'environnement. Les conditions de cette prise en charge par l'Etat sont définies par les Schémas Directeurs de Prédiction des Crues (SPDC) et complétés par les Règlements d'Information sur les Crues (RIC). La vigilance « crues » est un dispositif de surveillance des cours d'eau mis en place en juillet 2006 et opérée par le réseau Vigicrues. Ce réseau est constitué du Service Central d'Hydrométéorologie d'Appui à la Prédiction des Inondations (SCHAPI), des services de terrain, 17 Services de Prédiction des Crues (SPC) (figure 1.1.1) et 20 Unités Hydrométriques (UH) en France métropolitaine ainsi que 5 Cellules de Veille Hydrologique (CVH) en outre-mer.



Figure 1.1.1: Organisation des services de prédiction des crues (SPC) Bilan 2022 (réseau Vigicrues)

La vigilance « crues » est basée sur l'analyse des SPC, des scénarios d'hydrométéorologie, qui permet de proposer au SCHAPI une couleur de vigilance sur les tronçons des cours d'eau surveillés (328 tronçons). Les tronçons intégrés au dispositif sont ceux sur lesquels « ont été identifiés des enjeux significatifs exposés au risque d'inondation » et ceux dont le « fonctionnement hydrologique est connu et pour lesquels il est techniquement possible de produire des informations d'anticipation à une échéance suffisante et à un coût proportionné aux enjeux à protéger ». Le choix de la couleur de vigilance se fait sur une échelle de quatre couleurs (vert, jaune, orange, rouge) permettant d'indiquer le niveau de danger maximal prévu sur la période de prédiction, avec une durée de validité des informations de 24 heures.

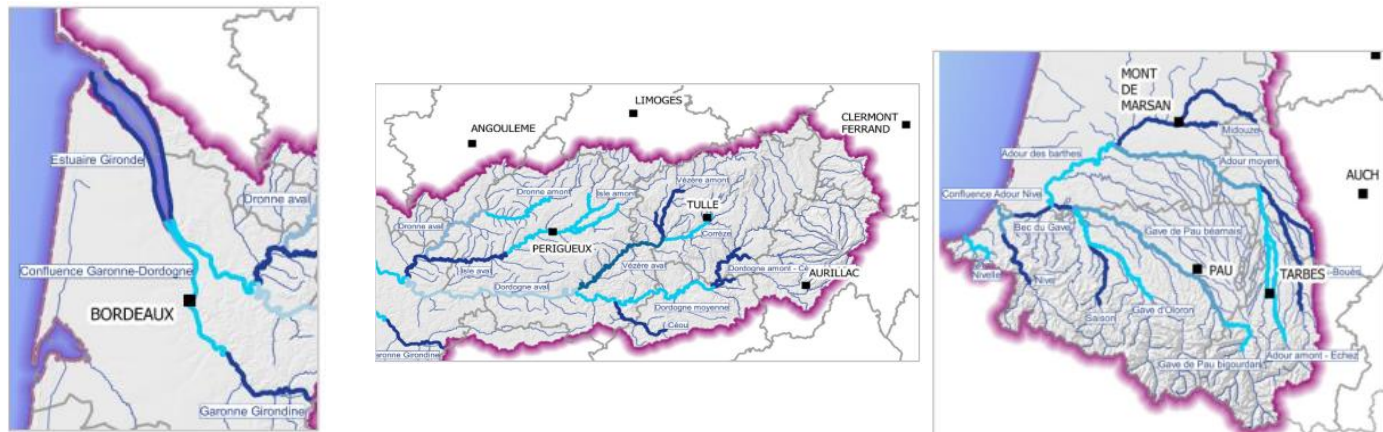
Ce travail se traduit par une carte de vigilance « crues » avec un bulletin d'information accessible en permanence sur le site internet [vigicrues.gouv.fr](http://vigicrues.gouv.fr), actualisés au minimum deux fois par jour (à 10 heures et à 16 heures) et plus fréquemment si nécessaire.

## 1.2 Le SPC Gironde Adour Dordogne (SPC GAD)

### 1.2.1 Territoire

Le 19 décembre 2012, le SDPC du bassin Adour-Garonne est arrêté par le préfet coordonnateur de bassin, et est réactualisé le 29 décembre 2015.

Le territoire du SPC GAD s'étend sur trois régions : Nouvelle Aquitaine, Occitanie, et Auvergne Rhône-Alpes. Sur ces trois régions, onze départements sont concernés : Cantal, Charente, Charente-Maritime, Corrèze, Dordogne, Gers, Gironde, Landes, Lot, Pyrénées-Atlantiques et Hautes-Pyrénées. La délimitation de son territoire est définie sur les limites de trois bassins versants : l'Adour, la Nivelle, la Dordogne, auquel s'ajoute l'estuaire de la Gironde et la Garonne.



(a) Estuaire de la Gironde

(b) Bassin de la Dordogne

(c) Bassin de l'Adour et de la Nivelle

Figure 1.2.1: Bassins versants du SPC GAD (RIC SPC GAD)

De nombreux phénomènes hydrométéorologiques peuvent avoir lieu sur ces bassins. Les principaux étant des averses orageuses d'intensité élevée sur les reliefs, des pluies d'intensités moyennes mais durables sur les vallées, l'éventuelle fonte printanière du manteau neigeux, des orages violents pouvant générer de fortes crues. A cela s'ajoute l'influence fluvio-maritime sur le bassin de l'Adour et de la Nivelle et l'estuaire de la Gironde, ainsi que l'influence des nombreux barrages sur le bassin de la Dordogne. Ainsi le territoire du SPC GAD fait face à une diversité d'événements de crues importantes, crues lentes, crues rapides, concentration de la pluie ou propagation d'une crue formée en amont, influence fluvio-maritime, neiges, orages.

De plus, l'augmentation de l'urbanisation dans les dernières années a facilité l'installation des populations dans les plaines inondables des différents bassins. Il y a donc trois types d'enjeux à signaler sur le territoire : des enjeux particulièrement exposés dans les zones de crue les plus fréquentes, des enjeux dans les lits majeurs des cours d'eau soumis à des crues rapides dont la difficulté est le déclenchement de l'alerte, la fiabilité de la prévision des crues et le temps de réaction des secours sur les secteurs habités susceptibles de se retrouver noyés ou isolés, et des enjeux dans les plaines inondables plus vastes, avec la présence d'agglomérations importantes telles que Toulouse, Bordeaux, ou des zones fortement touristiques. [SPC ADOUR GARONNE]

## 1.2.2 Organisation et missions du SPC

Le SPC GAD appartient au service risques naturels et hydrauliques de la DREAL Nouvelle-Aquitaine, l'organisation du département Hydrométrie et Préviation des Crues Gironde Adour Dordogne est schématisée sur la figure 1.2.4 ci-dessous :

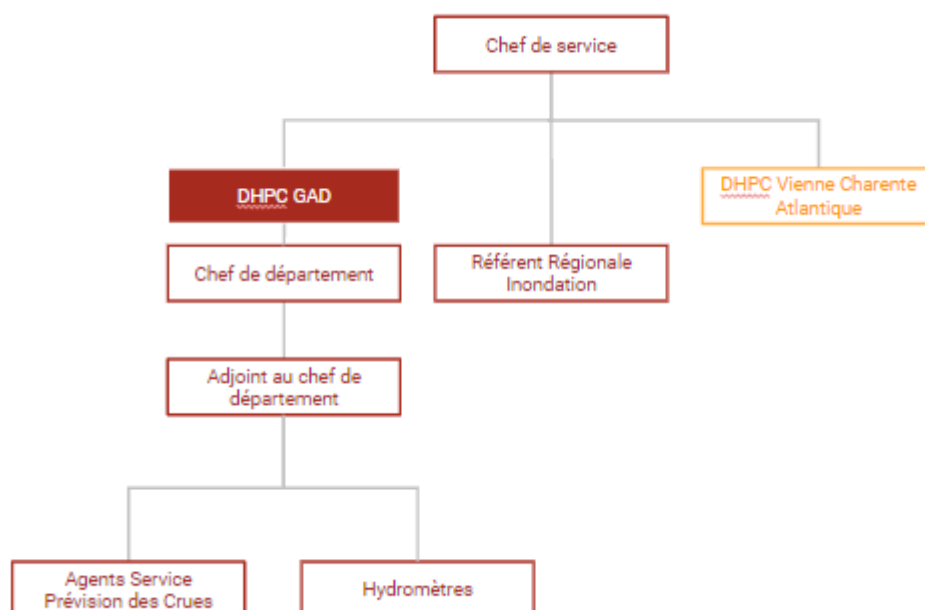


Figure 1.2.2: Organigramme simplifié du DHPC GAD

Ainsi pour assurer la mission principale, la prévision des crues sur le territoire, les missions au sein du SPC GAD sont nombreuses et réparties entre les agents du service.

Les missions peuvent être réparties en différentes catégories :

- *La modélisation et le développement des outils* (3 agents)
- *La mise en opérationnelle des outils* : mise en application des outils développés sur le territoire (1 agent)
- *Le pôle informatique* (2 agents) : Gérer le réseau du SPC (nécessité d'un accès au réseau 24h/24h) et le réseau de secours en cas de problème, s'assurer que la récupération des données hydrométriques fonctionne (récupération des données via IP ou radio cf partie 1.2.3), assurer que les bureaux des prévisionnistes soient fonctionnels
- *La vigilance opérationnelle* (3 agents) : Permet de rendre la vigilance opérationnelle. En effet la vigilance « crue » ne repose pas uniquement sur les outils utilisés mais aussi sur les agents effectuant cette prévision. Ainsi, les agents en charge de la vigilance opérationnelle doivent faire en sorte que la vigilance fonctionne au mieux. Cela se traduit par un travail interne : la prise en charge de formations, assurer la cohésion de l'entièreté de l'équipe, gérer l'évolution réglementaire de la prévision (RIC, SDPC, ...), effectuer les retours d'expériences (REX) si nécessaire et mettre en place des actions suite à ces REX (exemple : développement de fiches réflexes sur les différents bassins du territoire).

Mais aussi un travail externe : formation des acteurs liés à la vigilance (DDT, cadre d'astreinte, ...), fournir des outils indispensables à la vigilance (cartes de zones d'inondation potentielle (ZIP)).

- *Le référent régional inondation* : appartient au service des risques naturels mais travaille en parallèle avec le SPC GAD : ses missions peuvent se diviser en 3 parties

« SPC GAD »	Référent Régional Inondation	Missions transverses
- assurer la prévision - apporter son expérience via des formations, l'organisation d'exercices	Animation régionale des RDI (référent départemental inondation), personne ou équipe possédant une culture technique de la prévision et une culture du terrain (les enjeux , ...) afin de traduire la vigilance en terme local et aider le préfet dans ses décisions	Appui au 2 SPC (GAD et VCA)

Tableau 1.2.1: Tableau récapitulatif des missions du RRI

- *Les Hydromètres* (18agents) : Déplacement sur le terrain pour mesurer les hauteurs et les débits afin d'assurer la fiabilité des courbes de tarage, et de vérifier le bon fonctionnement des stations hydrométriques (maintenance préventive et curative). Expertise des données : tous les 2 mois, vérification de leur cohérence, et suivi d'étiage, afin de les contrôler pour la rédaction des arrêtés préfectoraux de sécheresse.

## 1.3 Contexte, moyen et outils

### 1.3.1 Contexte du SPC GAD

Le SPC GAD, qui a été créé il y a 10ans, est cadré par un plan quadriennal défini par le SCHAPI. Dans ce plan, il y a les objectifs suivants :

- Assurer la cohérence entre les différentes informations de vigilance
- Publier des prévisions en crue et les mettre à jour
- Elaborer des prévisions avec des échéances temporelles suffisantes
- Quantifier et afficher les incertitudes de prévision
- Visualiser la dynamique de la crue
- Evaluer les prévisions réalisées

Le SPC est donc dans une période de changements de ses pratiques puisque des prévisions graphiques sur l'ensemble de ses stations seront attendues pour l'automne 2023 (alors qu'actuellement, le SPC fait des prévisions expertisées et chiffrées sur l'ensemble des stations auxquelles s'ajoutent des prévisions graphiques uniquement sur ses stations sous influence maritime). Ces prévisions graphiques devront être publiées à 6h, 10h, 16h et 22h dès la vigilance jaune, ce qui nécessitera d'adapter la production individuelle. Le SPC cherche donc à gagner en

efficacité sur les outils des différents secteurs et à multiplier les outils pour consolider les prévisions par une approche multi-modèles.

De plus, le service Vigicrues est également en train d'être repensé à l'échelle de la France avec une volonté du SCHAPI de faire des prévisions sur tous les cours d'eau d'ici 2030. En parallèle des mutations en cours, les SPC vont donc devoir équiper de nouveaux secteurs non surveillés actuellement, construire de nouveaux outils, et tout ceci à moyen humain constant.

### 1.3.2 Moyens et outils disponibles

Pour assurer les prévisions des crues du territoire Gironde Adour Dordogne, 22 agents du Services Risques Naturels et Hydrauliques font partie de l'équipe prévisionniste. Pour exercer leurs missions, ceux-ci ont à leur disposition de nombreux outils, qui peuvent être regroupés dans les catégories suivantes :

- Les modèles hydrologique, transformation pluie-débit (type GRP ou Plathynes) ;

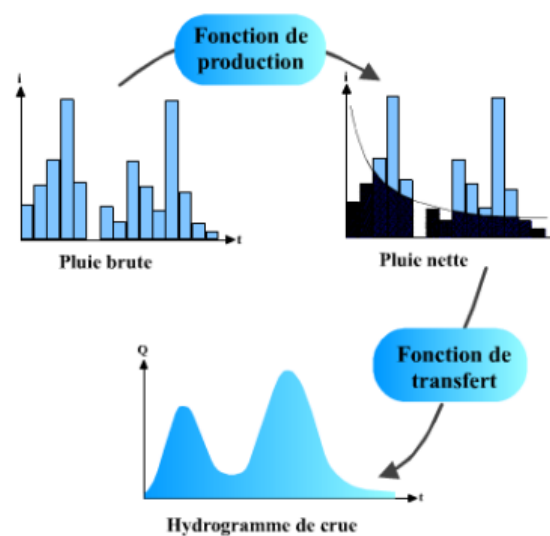


Figure 1.3.1 : Processus de fonctionnement d'un modèle pluie-débit (Institut des Sciences Technologiques de l'Environnement (ISTE), Ecole POLYTECHNIQUE FÉDÉRALE (EPFL))

- Les modèles hydrauliques, transformation débit-débit (Type Mascaret ou Télémac) ;
- Les abaques (pluie/débit – débit/débit – hauteur/hauteur) ;
- Les relations linéaires ;
- Les catalogues de crues.

Certains de ces outils ont été construits en essayant de trouver des relations entre les stations grâce aux données observées en crue (abaques, catalogues, etc).

D'autres outils essaient de modéliser le comportement des cours d'eau grâce à des données complémentaires comme la topographie et la morphologie du territoire (GRP, Plathynes, Mascaret, Télémac). Toutefois, ces outils ont également besoin de données observées pour être calés et avoir des résultats proches de la réalité.

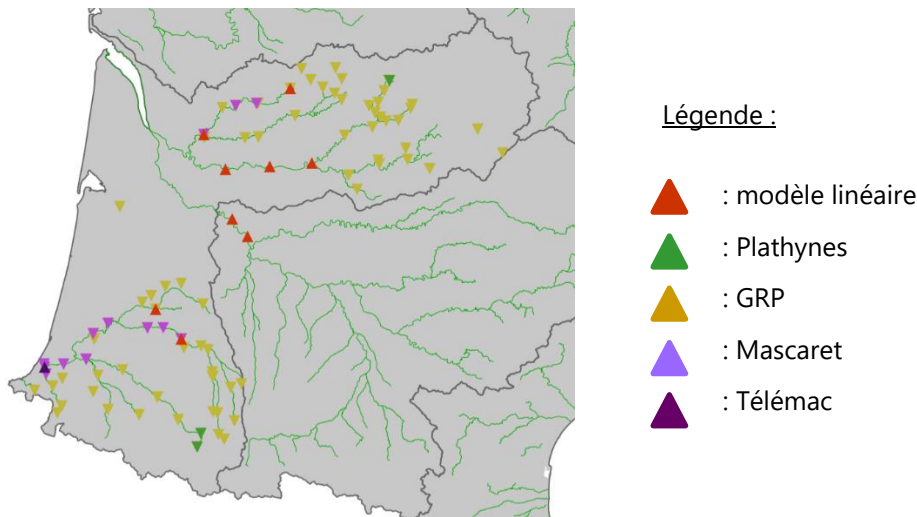


Figure 1.3.2 : Cartes des outils disponibles pour la prévision des crues sur les stations réglementaires

Afin de permettre un fonctionnement optimal de tous ces outils, les prévisionnistes du SPC ont besoin de données d'observation. Pour cela, ils ont à disposition des outils développés par Météo-France : **la lame d'eau** (estimations des pluies observées sur des grilles de mailles carrées) **ANTILOPE** dont le fonctionnement est schématisé en figure 1.3.3.

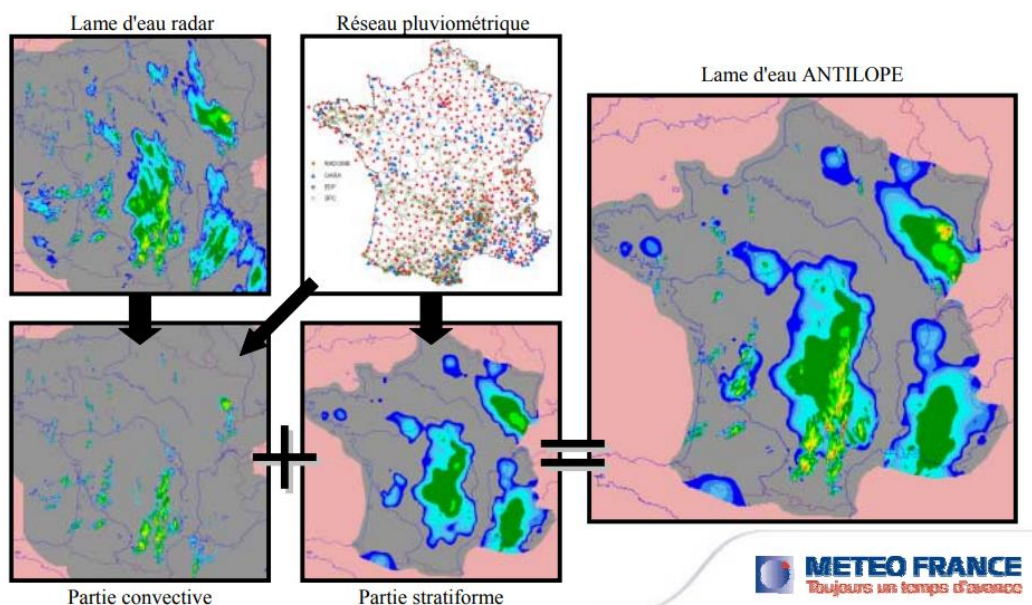


Figure 1.3.3:Schéma de construction de la lame d'eau ANTILOPE (météo France)

Le SPC dispose également de prévisions RR3, c'est-à-dire de la lame d'eau prévue au pas de temps 3h donnée par deux modèles de prévisions de météo-France, Arôme et Arpege. En plus de cet outil Météo-France fournit un **bulletin AP-BP** (Alerte Précipitation-Bulletin Prévision). Ce bulletin est fourni 2 fois par jour (avant 8h et avant 13h30) afin de permettre la vigilance « crue ». Il est découpé en 3 parties :

La première est composée des hauteurs de précipitations en mm à J, J+1 et J+2 sur l'ensemble du territoire Sud-Ouest.

La deuxième est la période AP, dans le cas où une alerte précipitation est mise en place par Météo-France, le début et la fin de l'évènement est indiqué sur le bulletin.

La troisième partie du bulletin est la description de la situation actuelle et de l'évolution prévue ainsi que le degré de confiance et les sources d'incertitudes.

Ainsi la prévision des crues ne pourrait se faire sans la prévision des précipitations.

Cependant, les prévisionnistes ont aussi besoins des **données hydrométriques**. Sur le territoire du SPC GAD il y a 3 antennes : Gironde (3 personnes), Adour (8 personnes) et Dordogne (7 personnes), ainsi 18 personnes vont effectuer des jaugeages et des mesures de hauteurs afin d'obtenir un outil indispensable pour les prévisionnistes : **les courbes de tarage** (La courbe de tarage est une loi de correspondance monotone, croissante et univoque entre le débit et la hauteur pour une station hydrométrique donnée. Elle n'est pas transposable dans l'espace). Une station peut posséder plusieurs courbes de tarage, en fonction des périodes de temps. Une courbe de tarage est utilisée pour une mesure effectuée à une certaine date si cette date est comprise dans la période d'utilisation de la courbe.

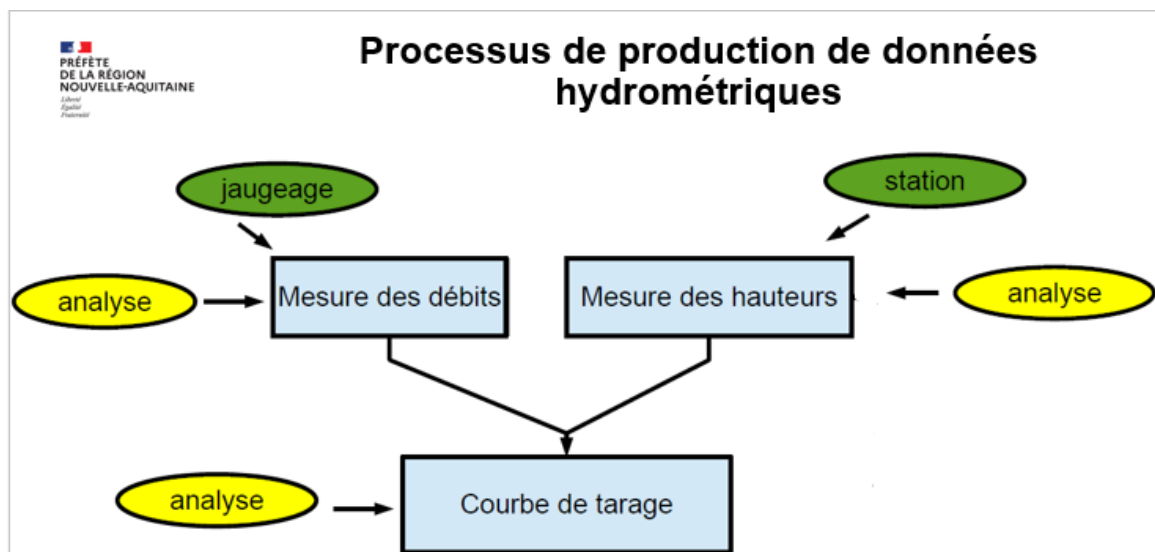


Figure 1.3.4:Processus de production des courbes de tarage

En plus de ces courbes de tarage, les hydromètres assurent le fonctionnement des stations hydrométriques permettant le bon fonctionnement du processus de production et de diffusion des données hydrométriques (les hauteurs d'eau des stations devant être disponibles sur le site hydroportail en toutes circonstances).



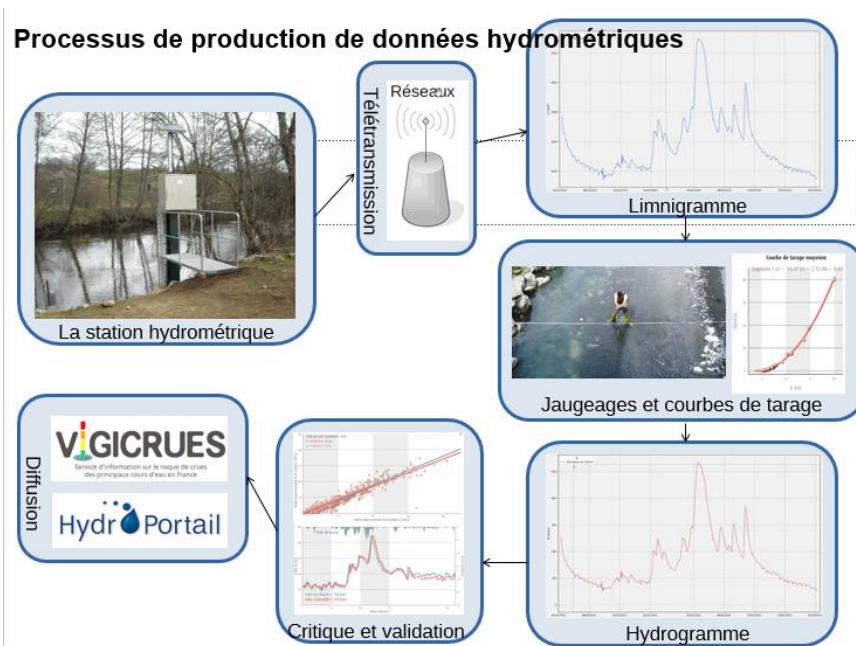


Figure 1.3.5: Processus de production de données hydrométriques

En addition de tous ces outils, les prévisionnistes ont accès au **superviseur**, outil qui leur permet d'obtenir les données instantanées (pluie, débits, hauteurs), les seuils de vigilance, la conversion des hauteurs/débits, et de visualiser les résultats des modélisations Mascaret et Plathynes.

## 1.4 Zoom sur un outil indispensable : le catalogue de crue

### 1.4.1 Les données du catalogue

Le catalogue de crues est composé des données en hauteurs et débits des stations réglementaires auxquelles on peut ajouter la hauteur, le débit et le taux d'humidité initiale sur l'une des stations du tronçon. Ainsi que le cumul total de pluie, les cumul max pour des pas de temps différents (allant de 3h à 48h), la hauteur au pas de temps 1h. Prenons l'exemple du catalogue de crue de l'Adour moyen et l'Adour des Barthes, qui contient les hauteurs et débits des stations réglementaires du tronçon, la station de Tartas pour la Midouze et quelques stations d'observations.

Ce catalogue regroupe 43 évènements allant de la vigilance verte à la vigilance orange des tronçons. Le premier évènement date de mars 1999. En effet, l'automatisation des données sur le linéaire ne permet pas d'obtenir suffisamment d'informations pour remonter au-delà de l'année 1999. Ces valeurs ont été récupérées sur la banque hydro et sont disponibles pour le public. La mise à jour de ce catalogue a été faite pour la dernière fois en décembre 2022. Afin de permettre aux prévisionnistes une vision optimale des évènements ayant eu lieu sur ces cours d'eau, une mise à jour est prévu tous les ans et lorsque qu'un évènement survient.

Les données présentes dans le catalogue de crues ne sont pas les données brutes ; elles ont été travaillées pour la cohérence. En effet, les stations hydrométriques envoient des données toutes les 5 minutes, il est donc nécessaire de les modifier pour obtenir des valeurs avec un pas de temps de 1heure. Pour cela, on effectue une moyenne des données récupérées durant l'heure. Pour nos tronçons cela ne pose pas de problème car le temps de propagation est lent, il y a peu de risque de perte d'information en effectuant cette moyenne.

De plus ce travail sur les données a pu mettre en évidence un problème de cohérence de débit en période de forte crue, dû au manque de jaugeages à fortes hauteurs d'eau (en raison du faible nombre d'évènements), ce qui rend l'appréhension de la relation hauteur/débit difficile. C'est pour cela que la prévision via ce catalogue se fait en hauteur.

N° évènement	Date début	Date fin	Risque H max (en m)	Lannux H max (en m)	Bernède H max (en m)	Aire-sur-l'Adour H max (en m)	Grenade-sur- l'Adour H max (en m)	Saint-Sever H max (en m)	Audon [Onard] H max (en m)	Tartas H max (en m)	Pontonx-sur- l'Adour H max (en m)	Garde H max (en m)	Dax H max (en m)
1999_03	03/03/1999 00	15/03/1999 00	1.92	4.33		2.95	3.30		5.26	3.02	5.09		5.54
2000_02	13/02/2000 00	27/02/2000 00	2.12	Si H max Lannux > 3.50 m, cellule en jaune clair		2.28	1.99	1.47	4.07	2.20	3.87		3.40
2000_06	09/06/2000 00	19/06/2000 00	3.17			2.96	3.39	2.01		1.73	4.48		3.52
2000_11	03/11/2000 00	17/11/2000 00	0.79			1.93	1.20	1.37	4.53	3.26	4.46		4.35
2001_02	27/02/2001 00	19/03/2001 00	1.64	4.15		2.52	2.26	1.84	4.99	3.28	4.69		4.78
2001_04	06/04/2001 00	16/04/2001 00	1.67			2.18	1.45	1.31	3.77	2.22	3.64		3.06
2002_11	28/11/2002 00	12/12/2002 00	1.99	3.48		2.44	2.05	1.48	4.49	1.48	3.90		3.42
2003_01	27/01/2003 00	14/02/2003 00	2.74	4.30		2.91	3.06	2.01	5.19	2.73	4.81		4.84
2003_03	02/03/2003 00	13/03/2003 00	1.79			2.18	1.61	1.38	3.69	1.87	3.33		2.90
2004_01	22/01/2004 00	06/02/2004 00	2.13	3.63		2.42	2.01	1.57	4.47	2.37	4.33		3.64
2004_05	04/05/2004 00	17/05/2004 00	1.31	3.06		2.13	1.76	1.36	4.30	2.36	4.15		3.46
2006_03	09/03/2006 00	19/03/2006 00	2.69	4.45		3.04	3.66	1.98	5.12	2.35	4.71		4.37
2008_01	16/01/2008 00	25/01/2008 00	1.89	3.93		2.45	2.63	1.52	4.66	1.78	4.15		3.30
2009_01	19/01/2009 00	05/02/2009 00	2.11	4.19		2.68	3.30	1.82	5.16	3.12	4.82		5.25
2009_02	10/02/2009 00	21/02/2009 00	2.31	4.09		2.51	3.01	1.68	4.68	2.26	4.45		3.77
2010_01	12/01/2010 00	26/01/2010 00	1.97	3.88		2.28	2.46	1.37	4.23	1.93	3.75		3.08
2013_01	14/01/2013 00	20/02/2013 12	2.92	4.27		3.15	4.11			2.93	5.09		5.40
2013_03	29/03/2013 00	10/04/2013 00	1.82	3.82		2.31	2.45		4.24	1.98	3.72		3.15

Les couleurs correspondent au seuil de vigilance ainsi qu'au zone de transition entre les seuils

Figure 1.4.1:Extrait de la première feuille du catalogue de crues Adour moyen-Adour des Barthes

#### 1.4.2 Son utilisation actuelle

Actuellement l'utilisation des catalogues se fait par analyse propre du prévisionniste. En effet, pour sélectionner les crues analogues, le prévisionniste va observer les données en temps réel, et parcourir le catalogue pour essayer de déterminer quels évènements ressemblent le plus. Que doit-il regarder ? Où doit-il regarder ? Doit-il se positionner sur la station juste avant la station étudiée, ou doit-il se placer sur celle la plus en amont du cours d'eau ? Quels paramètres regarde-t-il ? La hauteur ? Le volume ? Le gradient de montée ? Les trois en même temps ? Celle-ci est similaire en terme de hauteur mais celle-là en terme de volume, quelle est la meilleure ? Cet évènement est analogue pour la première station amont mais celui-ci l'est encore plus pour la station la plus en amont du cours d'eau, comment décider ?

De plus une fois que le prévisionniste a décidé sur quelles variables il allait filtrer sa recherche, il doit ensuite analyser les évènements restants. Par exemple pour effectuer une prévision sur Dax, la hauteur maximum atteinte à Saint-Sever pourra être un critère de sélection. Cependant cela peut ne pas conduire à une crue ressemblante car des apports intermédiaires peuvent avoir lieu entre Dax et Saint-Sever.

A cela s'ajoute encore des questions, combien d'analogues choisir pour la prévision ? Comment à partir de ces analogues prévoit-on le comportement sur la station aval ?

Malgré la formation des prévisionnistes à l'utilisation du catalogue de crues, la sélection d'analogues ne sera pas similaires d'un prévisionniste à l'autre.

Une fois la pré-sélection effectuée, le prévisionniste va parcourir les évènements restants en exportant les données en temps réel afin d'effectuer un ajustement graphique entre les hauteurs des analogues et de l'évènement observé. Il va aussi effectuer une analyse critique de chaque évènement conduisant à modifier les hydrogrammes sélectionnés pour établir la prévision et prendre en compte les écarts des évènements passés à la situation en cours.

Ainsi, on remarque que cette méthode repose énormément sur les choix du prévisionniste et donc de son expérience.

## 1.5 Enjeux et objectifs du TFE

### 1.5.1 Les enjeux

Ainsi, comme expliqué précédemment, la prévision des crues par analogue dépend énormément du choix ainsi que de la vision critique du prévisionniste. En effet, le prévisionniste se doit d'avoir un regard critiques sur les choix des analogues pour juger du niveau de prise en compte de chaque évènement dans la prévision finale. Cette dépendance vis-à-vis du prévisionniste cela soulève certains enjeux.

En effet, la question de la ressemblance entre deux évènements possède pour le moment une réponse subjective en fonction de la personne qui effectue la recherche d'analogue. Qu'est-ce qui fait qu'une crue ressemble à celle observée ?

Combien d'évènements ressemblants doit-on choisir pour permettre une prévision optimale ?

Doit-on donner le même poids à tous les analogues sélectionnés ?

Si non, comment décider quel évènement est « meilleur » que l'autre ?

Comment peut-on objectiver la recherche de crues analogues ?

Ainsi comme expliqué précédemment de nombreux outils permettent la prévision des crues sur un territoire présentant de forts enjeux. Cependant l'utilisation de certains de ces outils, tel que le catalogue de crues, n'est pas encore optimal. C'est pourquoi le premier objectif de ce stage a été de développer une méthode de recherche automatique d'analogues d'évènements de crues « passés ».

### 1.5.2 Objectifs

Nous avons donc un outil dont l'utilisation soulève énormément de questions. L'objectif de mon stage est d'essayer d'y répondre. Ces questions se posent pour tous les catalogues de crues. Néanmoins mon stage porte sur un secteur de prévision précis : Dax.

Les lignes directrices du travail sont donc :

- Détermination des critères permettant d'affirmer la ressemblance entre 2 évènements sur le tronçons Adour moyen-Adour des barthes
- Détermination du nombre optimal d'analogue pour la prévision
- Développement d'un outil de recherche et de sélection automatique des évènements analogues sur Python

## 2. Secteur global de l'étude : les tronçons Adour moyen, Midouze et Adour des Barthes.

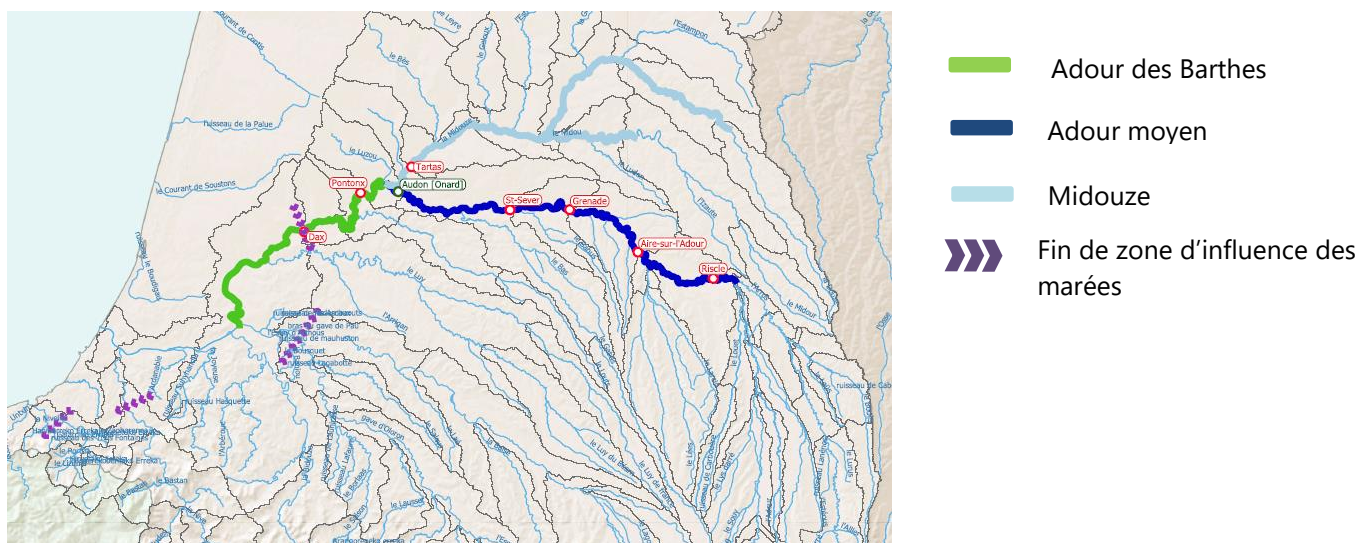


Figure 2.1 ; Carte des tronçons Adour moyen, Adour des Barthes et de la Midouze

### 2.1 Choix d'une station pertinente pour réaliser une méthode de recherche d'analogues : la ville de Dax.

#### 2.1.1 Difficultés rencontrées actuellement en prévision des crues à Dax

Durant mon stage, je me suis concentrée sur la recherche d'évènements analogues pour aider à la prévision sur la station réglementaire de Dax. Le choix de cette localité de travail s'est fait car c'est l'un des endroits du territoire le plus critique en terme de prévision, c'est-à-dire une station où la prévision est difficile, les outils de modélisation ne permettant une bonne anticipation de l'évènement avec énormément d'enjeux à prendre en compte. C'est aussi le tronçon sur lequel la prévision via le catalogue de crues est utile. En effet, les différents modèles développés sur la station de Dax n'arrivaient pas à reproduire le rebond ci-dessous :

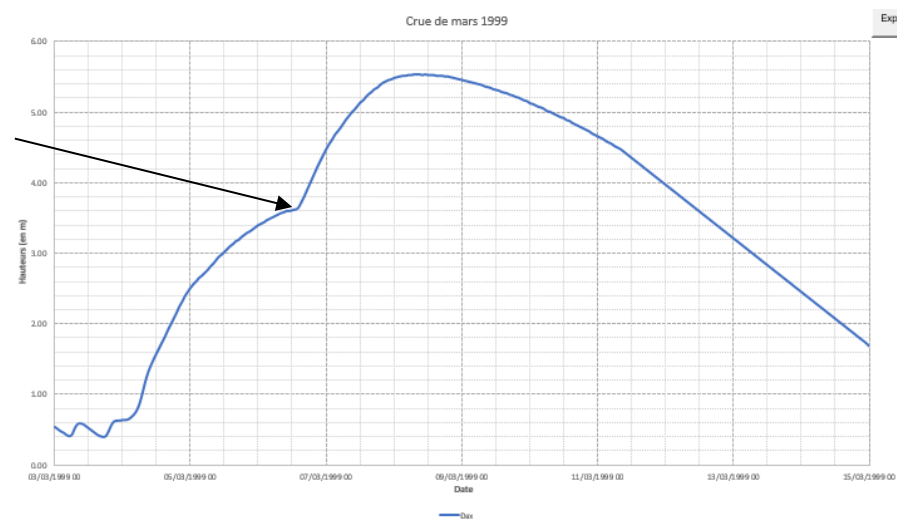


Figure 2.1.1: Limnigramme de la station de Dax lors de la crue de mars 1999

Ce « rebond » systématique pour les crues importantes, correspond à un phénomène hydraulique décrit en partie 2.1.2

### 2.1.2 Le fonctionnement hydraulique du secteur

La ville de Dax est entourée de zones boisées, chênaies et de prairies inondables, zone que l'on appelle les barthes de l'Adour (cf figure 2.1.2). Les barthes de l'Adour constituent un écosystème adapté à la présence de l'eau durant une grande partie de l'année, et jouent un rôle d'écroulement des débits de l'Adour. Lorsque la crue arrive à Dax, les barthes étant une zone appartenant au lit majeur de l'Adour, elles vont se remplir d'eau ce qui va permettre à la hauteur de l'Adour de se stabiliser. Cependant, si la hauteur atteinte par l'Adour dépasse ~3.50m alors les barthes ne seront plus efficace et cela va entraîner une augmentation de la hauteur mesurée par la station, ce qui explique le rebond observé en figure 2.1.1. Ce phénomène hydraulique est inévitable en cas de crue importante, ce qui en fait un instant de prévision crucial.

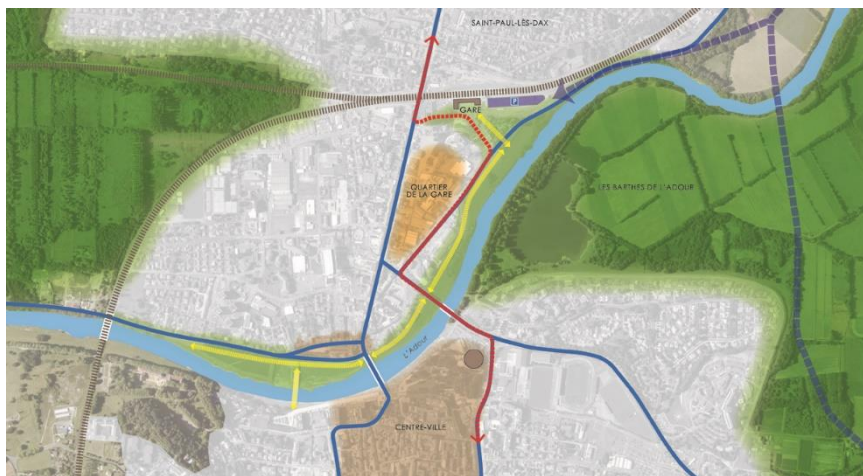


Figure 2.1.2 : Carte de Dax et des barthes de l'Adour

### 2.1.3 Les enjeux dépendants d'une bonne prévision à Dax

La prévision à Dax en utilisant le catalogue de crues ne se fait qu'à l'appréciation du prévisionniste pour la sélection des épisodes les plus « ressemblant », et le calage entre les épisodes « passés » et l'épisode observé. Ainsi, la prévision étant très liée au choix et à l'expérience du prévisionniste, il est difficile d'obtenir une bonne prévision sur ce secteur à enjeux.

De plus, la population de Dax est d'environ 21 000 habitants dont environ 14 000 (chiffre maximisant) sont protégés par des systèmes d'endigements très hétérogènes. Par conséquent le niveau de protection des digues, indiqué par la performance du système d'endiguement, est un facteur important à prendre en compte car cette performance impacte les seuils de vigilance donnés par la figure 2.1.3. Or à Dax les digues sont en mauvais états, ce qui entraîne une performance faible, et donc des seuils de vigilance bas. Actuellement le niveau de protection du système d'endiguement est compris pour des événements de période de retour comprise entre Q2 et Q10. Il est prévu de gros travaux de confortements de ces ouvrages d'ici 2029, qui devraient permettre d'améliorer le niveau de protection jusqu'à une période de retour de Q50 maximum.

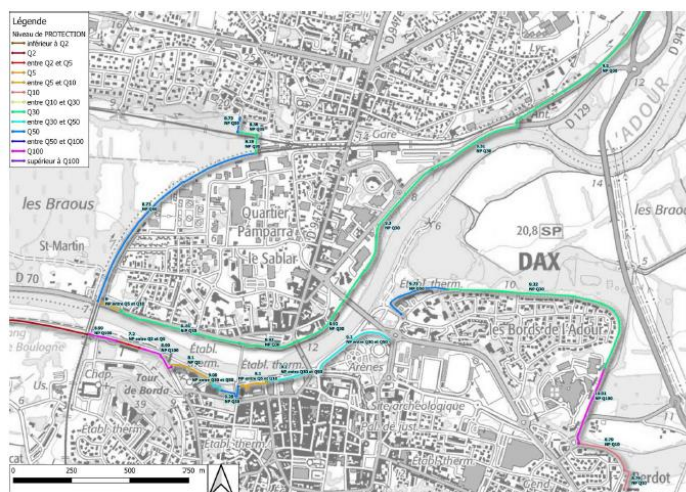


Figure 2.1.3: Les niveaux de protection pour chaque tronçon du système d'endiguement de Dax

Il est nécessaire d'appréhender le rebond à Dax car au-delà de 3.80m d'eau, de nombreux enjeux sociaux et économiques apparaissent. En effet, au-delà de certaines hauteurs (cf. tableau 2.1.1), le système d'endiguement de la ville de Dax ne la protège plus contre les inondations.

Zone protégée	Niveau de protection	Hauteur d'eau au point de référence sur l'échelle
Rive-Droite-Sablar	Crue de retour 5 à 10ans	5.42m (correspondant à 8.26m NGF)
Rive Gauche Arènes et ZAC Adour.	Crue de retour 5 à 10ans	5.47m (correspondant à 8.31m NGF)
Rive Gauche quartier des Baignots et quartier Thermal	Crue de retour 2 à 5 ans	4.57m (correspondant à 7.41m)
Rive Gauche -Saubagnac	Crue de retour 2ans	3.83m (correspondant à 6.67m NGF)

Tableau 2.1.1: Hauteurs d'eau auxquelles les zones de Dax ne sont plus protégées

Cette partie de Dax endiguée présente de nombreux enjeux, en effet en cas d'inondation de nombreux secteurs sont touchés, tels que la gare SNCF, le centre urbain. Mais aussi le quartier du Sablar dans lequel on trouve des résidences collectives ou individuelles abritant des milliers de personnes, des activités artisanales ainsi que deux unités d'enseignement et des installations sportives. Ces différentes zones ne sont plus accessibles en cas de crue centennale car les accès aux deux ponts seront coupés.

Sur la rive gauche, les zones touchées sont un quartier résidentiel mais surtout le centre hospitalier, car bien que l'hôpital soit construit hors zone inondable il est nécessaire de prendre en compte le fait que la rive droite ne peut avoir accès à ce service en cas d'inondation.

De plus, de nombreuses routes départementales seront submergées en cas d'inondation ce qui entraîne une difficulté voire impossibilité d'accès à certaines zones de la ville de Dax.

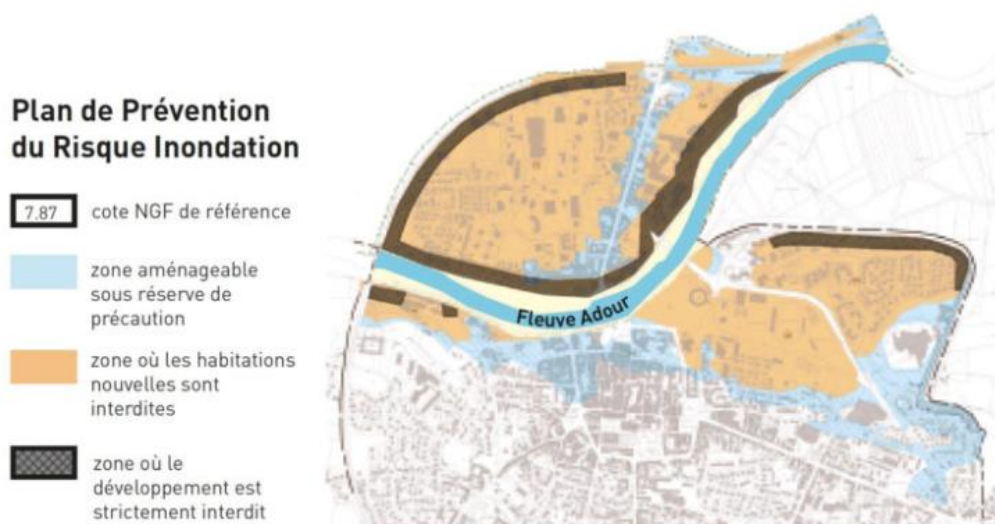


Figure 2.1.4 : Les zones concernées par le risque d'inondation à Dax (PPRI Ville de Dax)

En conclusion, la ville de Dax est une ville à fort enjeux avec des outils de modélisation peu efficaces et dont la prévision via catalogue de crues peut différer en fonction des prévisionnistes en astreinte durant les évènements, c'est pourquoi développer un outil opérationnel de recherche automatique d'analogues d'évènements de crue sur la base de catalogues de crues sur ce tronçon est un choix pertinent.

## 2.2 Données utilisées

### 2.2.1 L'Adour des Barthes

L'Adour des Barthes prend son origine dans la confluence entre la Midouze et l'Adour moyen. C'est sur ce tronçon que se trouve la station de Dax. Les grandes crues observables sur l'Adour des Barthes se forment de la même manière que les grandes crues de l'Adour moyen, par apports généralisés des affluents suite à des évènements pluvieux étalés dans le temps avec des cumuls importants. Elles se forment par propagation des crues formées en amont sur l'Adour moyen et la Midouze (la contribution de la Midouze en débit à Pontonx-sur-l'Adour est de 25% à 30%).

Ainsi la crue peut toucher le territoire en décalage par rapport aux moments forts des précipitations. Ce contexte favorise les crues largement débordantes et dommageables avec de faibles vitesses d'écoulement. De plus, le lit majeur de l'Adour des Barthes est organisé pour ralentir l'écoulement par le remplissage des Barthes (parcelles largement inondables). Une fois arrivé à saturation, l'écoulement reprend entraînant la vidange des Barthes.

Comme le montre la figure 2.1, on remarque que ce tronçon est soumis à l'influence des marées jusque Dax.

Cinq communes sont des secteurs à enjeux, Saint-Jean-de-Lier, Gousse, Téthieu, Yzosse et Dax, dont les enjeux sont liés à l'importante population saisonnière.

Les crues significatives les plus récente connues de l'Adour des Barthes datent de décembre 2020 et février 2021.

**FICHE TRONCON DE VIGILANCE – Version destinée aux partenaires opérationnels (Préfecture, RDI...)**

Tronçon Adour des Barthes Rivière : Adour			STATIONS DE REFERENCE DU TRONCON			
Vigilance	Définition et conséquences attendues		Pontonx		Dax	
			Crues historiques	Hauteur / Débit	Crues historiques	Hauteur / Débit
ROUGE	Niveau 4 : ROUGE Risque de crue majeure. Menace directe et généralisée de la sécurité des personnes et des biens.	Crue rare et catastrophique, nombreuses vies humaines menacées, débordements généralisés, évacuations généralisées et concomitantes, paralysie à grande échelle du tissu urbain, agricole et industriel.	Février 1952	5,75 m	Avril 1770 Février 1879 Février 1952	7,10 m 6,54 m 6,52 m
			Février 1979 – 5,60 m	5,70 m 5,60 m	Décembre 1981 6,03 m	6,25 m 6,00 m
ORANGE	Niveau 3 : ORANGE Risque de crue génératrice de débordements importants susceptibles d'avoir un impact significatif sur la vie collective et la sécurité des biens et des personnes.	Crue majeure et dommageable, vies humaines menacées, nombreuses évacuations, paralysie d'une partie de la vie sociale, agricole et économique.	Décembre 1981	5,48 m	Janvier 2014 Décembre 2020 Février 2021 Mars 1999	5,97 m 5,89 m 5,86 m 5,54 m
			Décembre 2020 Janvier 2014 Décembre 1976 Février 2021 Mars 1999 Février 2013	5,37 m 5,32 m 5,30 m 5,23 m 5,10 m 5,08 m	Février 2013 Janvier 2009 Février 2015 Novembre 2019 Juin 2018 Décembre 2021 Mai 2020 Janvier 2022	5,40 m 5,26 m 5,14 m 5,13 m 4,84 m 4,75 m 4,56 m 4,55 m
JAUNE	Niveau 2 : JAUNE Risque de crue ou de montée rapide des eaux n'entraînant pas de dommages significatifs, mais nécessitant une vigilance particulière dans le cas d'activités saisonnières et/ou exposées.	Perturbation des activités liées au cours d'eau, premiers débordements localisés, coupures ponctuelles de routes secondaires, maisons isolées touchées, caves inondées, activité agricole perturbée.	Novembre 2019 – 4,92 m	5,00 m 4,90 m		4,50 m 4,40 m
			Décembre 2021 Juin 2018 Mai 2020 Mars 2006 Février 2016 Juin 2000 Décembre 2019 Janvier 2018 Janvier 2022 Février 2003	4,87 m 4,86 m 4,75 m 4,71 m 4,66 m 4,48 m 3,97 m 3,87 m 3,74 m 3,55 m	Novembre 2006 Novembre 1992 Juin 2000 Décembre 2019 Janvier 2018	4,37 m 4,02 m 3,52 m 3,36 m 3,13 m
VERT	Niveau 1 : VERT Pas de vigilance particulière requise	Situation normale.		3,50 m 3,40 m		2,90 m 2,80 m

Figure 2.2.1 :Seuils de vigilance et évènement importants sur le tronçon Adour des Barthes

### 2.2.2 La Midouze

La Midouze trouve son origine à Mont-de-Marsan, à la confluence de deux rivières : la Douze et le Midou, quant à l'aval, elle rejoint l'Adour au niveau de trois communes dans les Landes. Elle contribue à 25-30% au débit de l'Adour des Barthes.

Le bassin peut être divisé en deux parties :

- Les coteaux armagnacais à l'amont, caractérisé par des fortes pentes et un ruissellement assez rapide
- Le plateau landais à l'aval, plateau sableux, perméable et plat.

En hiver, des crues rapides, liés au fort ruissellement pendant des épisodes pluvieux importants, peuvent être observées sur la partie des coteaux.

Quant au plateau, les nappes superficielles effectuent un travail de tampon et absorbent une partie des débits lors des périodes pluvieuses. Ainsi, cela entraîne une formation de crue lente jusqu'à saturation complète des terrains. Une fois cette saturation atteinte, la montée est rapide et on observe des crues avec un retard important par rapport aux précipitations.

Les enjeux sur le bassin sont assez rares sauf le centre-ville de Tartas qui est une zone inondable.

Les crues les plus fortes connues de la Midouze sont celle de décembre 1976 et décembre 1981.



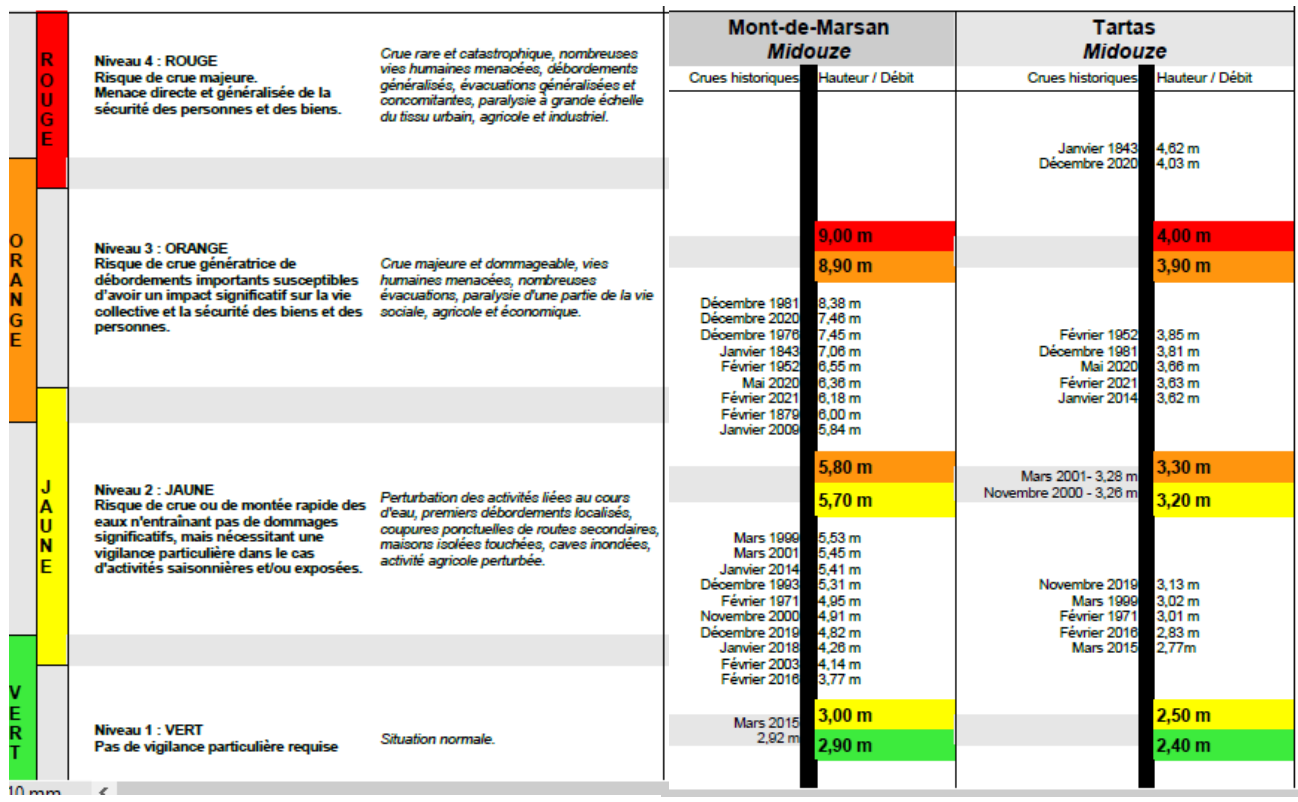


Figure 2.2.2:Seuils de vigilance et évènement importants sur le tronçon Midouze

### 2.2.3 L'Adour moyen

L'Adour moyen contribue à 70-75% au débit de l'Adour des Barthes. Sur l'Adour moyen les grandes crues se forment par un apport généralisé de tous les affluents, principalement suite à des pluies étalées dans le temps (durée supérieure à 48h) et entraînant des cumuls importants (supérieurs à 100mm).

Les sous-bassins de la rive gauche de l'Adour sont collés aux premiers contreforts des Pyrénées et subissent les premiers effets orographiques (élévation d'un flux d'air lorsqu'il se heurte à une montagne ou une colline) des perturbations venant principalement du Nord-Ouest. De plus, les sols sont imperméables ou semi-perméables. Ainsi on obtient des conditions favorables pour des crues largement débordantes et dommageables.

Les enjeux se concentrent sur six secteurs différents, la commune de Riscle, Barcelone du Gers, Aire-sur-l'Adour, Grenade-sur-l'Adour, Larrivière-Saint-Savin et Onard.

Les crues les plus fortes de l'Adour moyen datent d'avril 1770 et février 1952.

Tronçon Adour Moyen Rivière : Adour		STATIONS DE REFERENCE DU TRONCON							
Vigilance	Définition et conséquences attendues	Riscle		Aire-sur-l'Adour		Grenade-sur-l'Adour		Saint-Sever	
		Crues historiques	Hauteur / Débit	Crues historiques	Hauteur / Débit	Crues historiques	Hauteur / Débit	Crues historiques	Hauteur / Débit
ROUGE	Niveau 4 : ROUGE Risque de crue majeure. Menace directe et généralisée de la sécurité des personnes et des biens.  Crue rare et catastrophique, nombreuses vies humaines menacées, débordements généralisés, évacuations généralisées et concomitantes, paralysie à grande échelle du tissu urbain, agricole et industriel.	Février 1952	4,69 m	Février 1952	5,45 m			Jun 1885	5,43 m
			4,60 m		5,00 m			Février 1952	5,17 m
ORANGE	Niveau 3 : ORANGE Risque de crue généralisée de débordements importants susceptibles d'avoir un impact significatif sur la vie collective et la sécurité des biens et des personnes.  Crue majeure et dommageable, vies humaines menacées, nombreuses évacuations, paralysie d'une partie de la vie sociale, agricole et économique.		4,50 m		4,80 m		4,60 m		3,80 m
		Décembre 2019	3,73 m	Jun 1855	4,71 m	Janvier 2014	4,56 m	Décembre 1930	3,40 m
				Mars 1930	4,40 m	Décembre 2020	4,14 m	Mars 1999	3,35 m
JAUNE	Niveau 2 : JAUNE Risque de crue ou de montée rapide des eaux n'entraînant pas de dommages significatifs, mais nécessitant une vigilance particulière dans le cas d'activités saisonnières et/ou exposées.  Perturbation des activités liées au cours d'eau, premiers débordements localisés, coupures ponctuelles de routes secondaires, maisons isolées touchées, caves inondées, activité agricole perturbée.	Février 2015 - 3,43 m Jun 2018 - 3,43 m	3,60 m 3,40 m	Jun 2000 - 2,96 m Mars 1999 - 2,96 m Décembre 2020 - 2,91 m	3,00 m 2,90 m		3,65 m 3,55 m		2,80 m 2,70 m
		Jun 2000	3,17 m	Décembre 2021	2,82 m			Décembre 1976	2,25 m
		Décembre 2021	3,07 m	Février 2015	2,86 m	Janvier 2022	3,44 m	Janvier 2014	2,34 m
VERT	Niveau 1 : VERT Pas de vigilance particulière requise  Situation normale.	Janvier 2022	2,89 m	Décembre 1962	2,80 m	Février 2018	3,33 m	Février 2019	2,15 m
		Février 2003	2,74 m	Janvier 2022	2,76 m	Mai 2020	3,25 m	Décembre 2018	2,13 m
		Mars 2006	2,69 m	Janvier 2009	2,69 m	Novembre 2019	3,13 m	Jun 2018	2,08 m
	Février 2021	2,59 m	Mai 2020	2,64 m	Décembre 2019	2,33 m	Février 2021	2,06 m	
	Décembre 2020	2,37 m	Février 2016	2,52 m			Décembre 2020	2,02 m	
	Janvier 2009	2,11 m	Novembre 2019	2,34 m			Mars 2006	1,98 m	
			Janvier 2018	2,10 m			Février 2019	1,77 m	
							Décembre 2021	1,66 m	
							Février 2015	1,71 m	
							Janvier 2022	1,70 m	
							Mai 2020	1,68 m	
							Février 2018	1,61 m	
			2 m		2,00 m		2,50 m		1,50 m
			1,85 m		1,90 m		2,30 m		1,40 m

Figure 2.2.3: Seuils de vigilance et évènement importants sur le tronçon Adour moyen

### 3. Méthodologie

#### 3.1 Principes de fonctionnement de la méthode de recherche d'analogues

##### 3.1.1 La méthode d'analogues en météorologie

Selon le Larousse une analogie est un « rapport existant entre des choses ou entre des personnes qui présentent des caractères communs ; ressemblance ; similitude. ». Dans notre domaine, le principe d'analogie a d'abord été étudié en météorologie. Le principe est d'associer une situation météorologique en cours à des situations passées similaires et ce, en vue de prévoir l'évolution attendue.

Cependant, comment cela s'applique dans la prévision de précipitations ? De nombreuses questions apparaissent : « Quels paramètres descriptifs d'une situation météorologique allons-nous choisir ? Quel est le critère qui va nous permettre d'affirmer qu'une situation est plus analogue à la situation prévue qu'une autre ? »

La recherche d'une situation analogue repose sur plusieurs paramètres. Il faut avant tout préciser ce qu'on cherche à prévoir. Ensuite il est nécessaire de définir les variables permettant de décrire de la meilleure façon la situation prévue, c'est ce qu'on appelle les **variables d'analogie**. Afin de pouvoir utiliser le principe d'analogie il faut posséder une **archive météorologique** suffisamment longue. Et enfin, le principe d'analogie ne peut s'appliquer sans un critère numérique permettant l'identification de la meilleure situation analogue, le **critère d'analogie**. [Thèse : *Améliorations et développements d'une méthode de prévision probabiliste des pluies par analogie*, Ben Daoud, 2010]

##### 3.1.2 Application à la prévision des crues

Un parallèle peut se faire entre l'application du principe d'analogie en météorologie et en prévision des crues. En effet, de nombreuses questions apparaissent dans l'application de ce principe à la prévision des crues : Quels sont les paramètres décrivant le mieux l'évènement observé ? Comment choisir le(s) meilleur(s) évènement(s) analogue(s) ?

Prenons un exemple pour visualiser un peu ces problématiques :

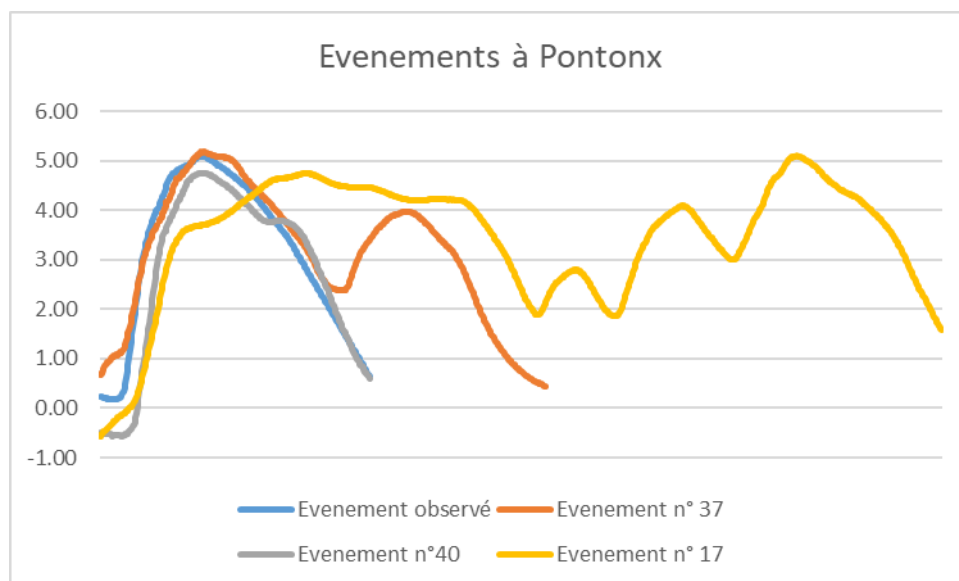


Figure 3.1.1 : 4 évènements observés à la station Pontonx-sur-l'Adour

Sur la figure 3.1.1 on observe 4 évènements, l'évènement observé ainsi que 3 autres évènements. Lequel ressemble le plus à l'évènement observé ? Pourquoi l'un plus que l'autre ? Sur quels critères le choix a été fait ? Est-ce que le choix effectué est vraiment le meilleur ? Comment le vérifier ?

Ainsi, tout comme en météorologie il est nécessaire de définir ce que l'on cherche à prévoir. Dans notre cas nous cherchons à prévoir l'évènement sur la station réglementaire de Dax en utilisant les informations sur les stations amont. Nous devons aussi définir nos variables d'analogie, effectivement de nombreux paramètres permettent de décrire les évènements de crues : la hauteur maximum atteinte, le volume, la durée de l'évènement, le gradient de montée. Dans la suite, il sera expliqué comment le choix des variables est effectué et quelles sont celles qui décrivent le mieux la situation observée.

Toutes ces informations sont stockées dans le catalogue des crues, équivalent à l'archive météorologique, qui décrit les évènements du tronçon depuis 1999.

En prévision des crues, il est utile d'effectuer le calcul de scores de performances afin d'évaluer la performance de modèles de prévisions. Il existe de nombreuses façon de calculer le score d'une prévision, réparti en plusieurs groupes, Groupe 1 : Scores d'erreurs simples, Groupe 2 : Score de persistance, Groupe 3 : Critère de Nash-Sutcliffe, Groupe 4 : Erreurs sur les volumes, Groupe 5 : Erreurs sur les pointes, Groupe 7 : Scores de franchissement de seuil.

Numéro de groupe	Définition du groupe
1	Groupe qui comprend les calculs d'erreur les plus « classiques »(erreurs moyennes, relatives, ...)
2	Score qui revient à comparer le modèle étudié au « modèle naïf » (modèle qui propose comme prévision à $t+\Delta t$ la valeur observé à $t$ )
3	Critère d'efficacité imaginé pour les modèle de prévision hydrologiques, il peut varier en $-\infty$ et 1.
4	Calculer les erreurs en volume sur les débits permet d'apporter un éclairage hydrologique supplémentaire sur le comportement du modèle
5	Met en évidence les performances du modèle sur le maximum de crue. Permet de donner un ordre de grandeur de fourchette d'incertitude sur les maxima
7	Permet de déterminer la qualité de la prévision en terme de franchissement de seuil.

Tableau 3.1.1: Tableau descriptif des différents groupe de score pour les modèle de prévision (Guide d'utilisation scores V1.3.3, réseau Vigicrues)

Au sein de ces différents groupes sont présents des critères numériques permettant de juger de la qualité d'une prévision vis-à-vis de l'évènement observé. Ainsi, j'ai décidé de choisir le critère de Nash-Sutcliff pour permettre l'identification du meilleur évènement analogue dans ces groupes de score.

## 3.2 Les variables d'analogie en prévision des crues

Afin de prévoir un évènement par la méthode d'analogie il est nécessaire de trouver les paramètres permettant de décrire les évènements. Après une phase de recherche (cf. partie 4.1 & 4.2) j'ai isolé les paramètres suivants comme les meilleures variables d'analogie.

### 3.2.1 Hauteur max

Le premier paramètre est la hauteur maximale atteinte lors de l'évènement. C'est un paramètre important car c'est le plus visible pour un prévisionniste lors qu'il effectue une recherche manuelle dans le catalogue de crue.

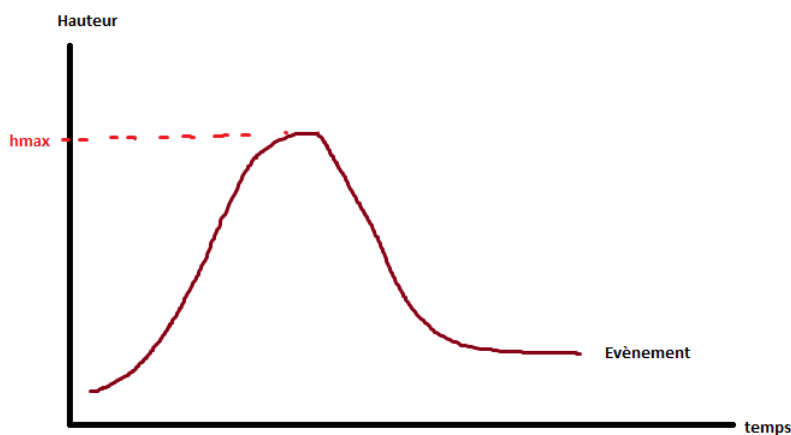


Figure 3.2.1:Schéma représentant le paramètre « hauteur max »

### 3.2.2 Volume

Le volume est l'aire sous la courbe de l'hydrogramme (ne nous intéressant qu'aux évènements ayant dépassé le seuil de vigilance jaune, le calcul ne se fait que pour les valeurs supérieures au seuil jaune). C'est un paramètre important pour décrire un évènement car le volume représente la quantité d'eau qui s'est écoulé au niveau de la station sur laquelle le volume est calculé.



Figure 3.2.2:Schéma représentant le paramètre « volume »

### 3.2.3 Durée

La durée de l'évènement se calcule de la façon suivante,  $\text{durée} = t_{\text{fin}} - t_{\text{début}}$ . La durée permet de décrire le temps durant lequel le cours d'eau a eu une hauteur supérieur au seuil jaune. C'est un paramètre important dans la description d'un évènement. Ce paramètre n'est pas calculé sur toutes les stations. Par exemple, lorsque l'évènement est en cours sur la station de Pontonx-sur-l'Adour, la durée de l'évènement n'est pas un paramètre optimal car il n'est pas fixe. Cependant, la crue étant le résultat d'un phénomène de propagation, il est possible que l'évènement soit en train de se finir sur les stations amont à Pontonx-sur-l'Adour. C'est sur ces stations que le calcul de la durée est un paramètre intéressant.

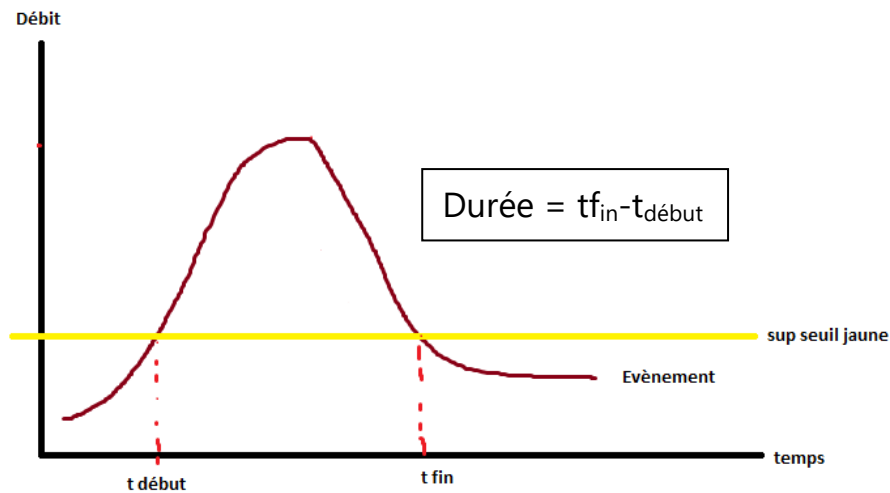


Figure 3.2.3:Schéma représentant le paramètre « durée »

### 3.2.4 Gradient de montée

Le gradient de montée représente la pente entre la hauteur max et le début de l'évènement (cf figure 3.2.4). Ce paramètre est aussi un paramètre visuel permettant au prévisionniste un nouveau filtre des évènements ressemblants lors de la recherche manuelle dans le catalogue de crue.

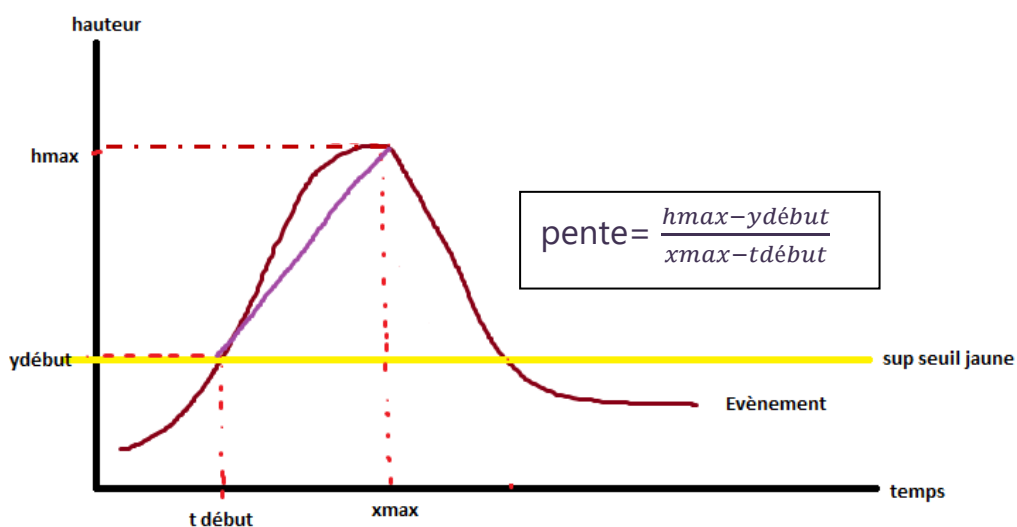


Figure 3.2.4:Schéma représentant le paramètre « gradient de montée »

Ainsi ces différents paramètres forment les variables d'analogie permettant de décrire de façon optimale l'évènement passé sur une station. Cependant ces paramètres ne nous permettent pas de

vérifier que le choix d'analogie effectué est le bon, c'est pour cela qu'il nous a fallu définir un critère d'analogie.

### 3.3 Critère d'analogie en prévision des crues

Le critère d'analogie est un critère numérique qui permet de comparer les évènements et d'identifier le meilleur analogue.

Pour se faire nous avons défini deux critères différents :

**Le critère de choix** : critère mathématiques calculant la « distance » entre les deux évènements en fonction des variables d'analogie.

Le calcul de la distance est le suivant :

$$D = \sqrt{(h_{maxobs} - h_{maxprev})^2 + (volume_{obs} - volume_{prev})^2 + (pente_{obs} - pente_{prev})^2}$$

(ou  $D = \sqrt{\frac{(h_{maxobs} - h_{maxprev})^2 + (volume_{obs} - volume_{prev})^2 + (pente_{obs} - pente_{prev})^2}{(durée_{obs} - durée_{prev})^2}}$ )

Où *obs*=évènement en cours

*prev*=évènement du catalogue

Ainsi cette distance permet de définir mathématiquement quel est l'évènement le plus proche de l'évènement observé.

Cependant cela ne suffisait pas pour permettre d'affirmer que cet évènement permettait une bonne prévision, nous avons donc ajouter le critère d'évaluation. **Le critère d'évaluation** est le calcul du critère de Nash-Sutcliff (que l'on nommera critère de Nash dans la suite de ce rapport) au niveau de la station de Dax à  $t_0+24h$  et  $t_0+48h$ .

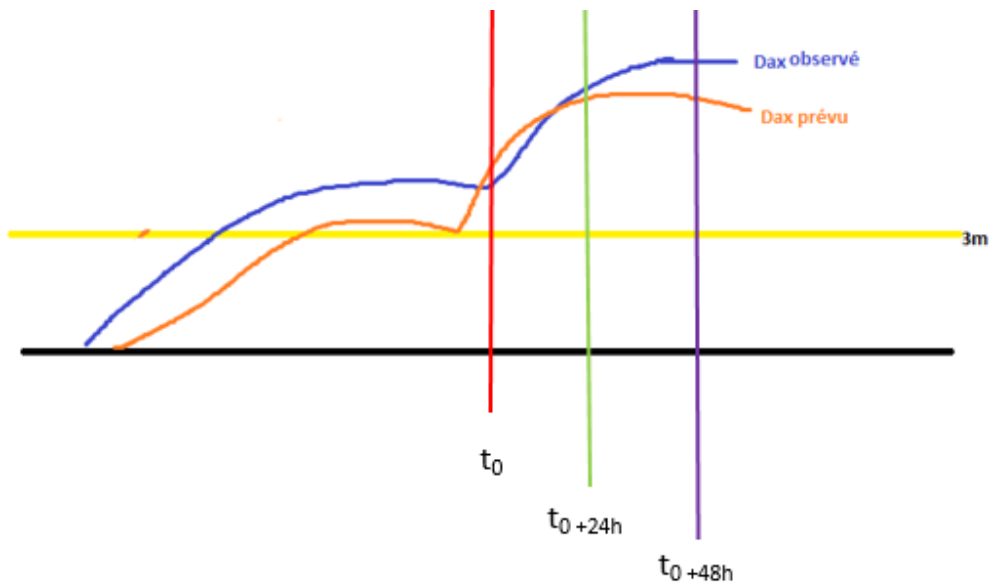


Figure 3.3.1: Représentation du  $t_0$ ,  $t_0+24h$  et  $t_0+48h$

En effet, le critère de Nash est un critère mathématique permettant de calculer la performance d'une prévision. Le calcul est le suivant :

$$N = 1 - \frac{\sum_1^n [H_{obs}(t) - H_{modèle}(t)]^2}{\sum_1^n [H_{obs}(t) - \overline{H_{obs}(t)}]^2}$$

Figure 3.3.2 : Calcul du critère de Nash-Sutcliffe en fonction des hauteurs

Ce critère est un critère d'efficacité pour les modèles de prévision hydrologiques. Ce critère peut être utilisé en débit et en hauteur. De plus ce critère est compris dans l'intervalle  $]-\infty ; 1]$  le résultat pouvant s'interpréter de la façon suivante :

- Une valeur nulle signifie que le modèle ne donne pas de meilleur résultat qu'un modèle basique donnant à chaque pas de temps un débit constant égal à la moyenne des débits observés ;
- Une valeur positive signifie que le modèle donne de meilleurs résultats et qu'il est d'autant meilleur que le critère se rapproche de 1 ;
- Une valeur négative est le signe d'une inadéquation du modèle.



## 4. Procédure d'obtention du jeu de variables optimal

### 4.1 1ere méthode de recherche : la méthode « intégratrice »

Le premier objectif de ce stage était de trouver le critère de choix optimal qui permettrait la sélection de l'évènement le plus ressemblant à l'évènement observé.

La définition des paramètres ainsi que du critère d'analogie a été le résultat d'un premier travail de recherche sur le catalogue de crues.

Dans un premier temps, la vision de la recherche d'analogues était différente de la vision finale (détaillée dans la suite de ce rapport). En effet dans un premier temps, le critère de choix et d'évaluation n'était pas réellement défini, l'idée était la suivante :

Dans un premier temps, on enlève un évènement du catalogue de crue, qui devient *l'évènement observé*. Cet évènement observé sera comparé aux évènements restants du catalogue de crues. La comparaison s'est fait selon le critère de Nash sur les évènements de la station Pontonx-sur-l'Adour. Cette station a été choisi car c'est la première station en amont de Dax et elle se situe après la confluence entre l'Adour et la Midouze. De plus, le temps de propagation entre Pontonx-sur-l'Adour et Dax est estimé entre 4h et 34h, ce qui permet d'affiner la prévision à Dax lorsqu'on observe le maximum de la crue à Pontonx.

L'entièreté des évènements étaient comparé à l'évènement observé. De l'ensemble des valeurs de Nash obtenues on ne gardait que les 3 meilleures. On recommençait cette méthodologie avec un autre évènement de crue du catalogue, et ainsi de suite jusqu'à que chaque évènement possède ses 3 meilleures analogues.

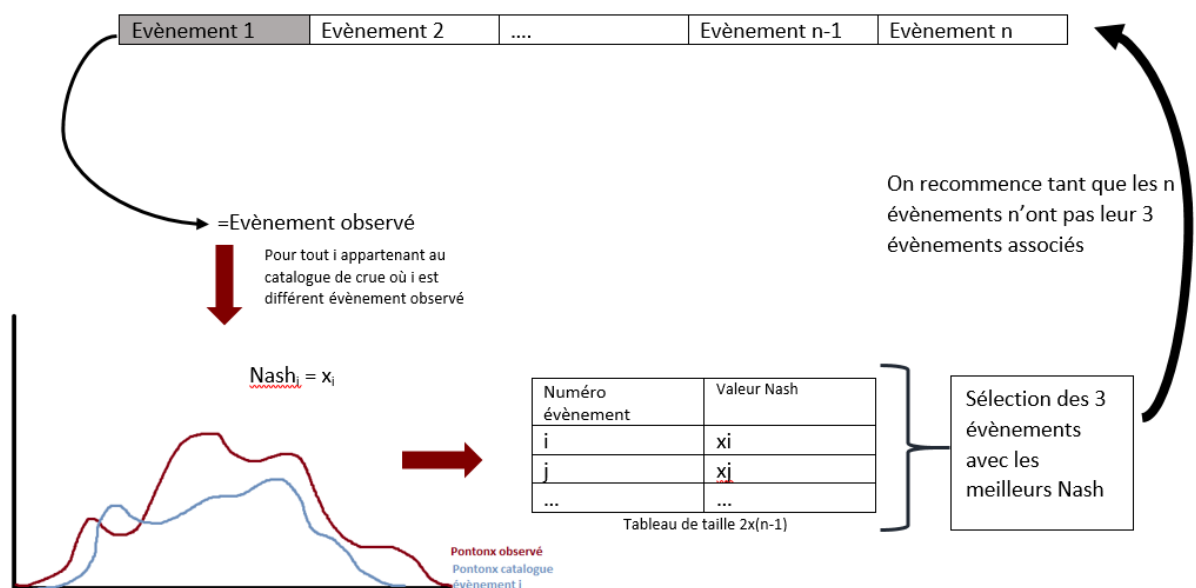


Figure 4.1.1: Schéma explicatif du calcul des 3 meilleurs évènements pour chaque évènement du catalogue

### 4.2 2ème méthode de recherche : la méthode de calage

Les résultats n'étant pas convaincants (cf. partie 6.1), j'ai essayé une nouvelle méthode de sélection. Dans la première méthode la sélection des meilleurs évènements se faisait de façon « brut », c'est-à-dire que le seul ajustement effectué était la sélection des valeurs de l'évènement

supérieur au sup seuil jaune de la station de Pontonx-sur-l'Adour. Ainsi, certains évènements pouvaient présenter un décalage temporel qui entraîne une valeur de Nash mauvaise alors que c'était l'un des évènements le plus ressemblant à l'œil par rapport à l'évènement observé.

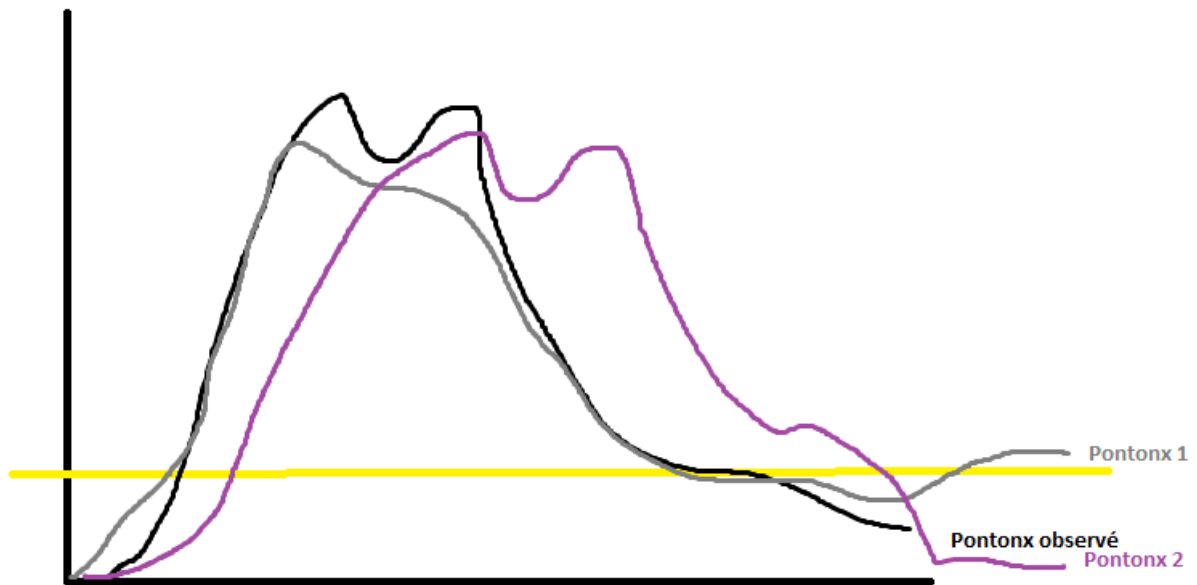


Figure 4.2.1: Schéma de différents évènements de crues ayant lieu sur la station de Pontonx sur l'Adour

La figure 4.2.1 illustre bien ce problème, en effet, l'œil humain remarque ici que Pontonx 2 est l'évènement ressemblant le plus à Pontonx observé, or le calcul de Nash ne nous donnera pas la même conclusion car la partie montante de l'évènement de Pontonx 1 est en adéquation avec la partie montante de Pontonx observé.

Ainsi, pour résoudre ce problème on a d'abord eu l'idée de caler les hauteurs maximales, c'est-à-dire de décaler les évènements du catalogue de telle façon que la hauteur max de l'évènement observé et de l'évènement du catalogue se trouvait sur le même point d'abscisse. Cependant, si la ressemblance était basée sur un décalage à Pontonx alors il fallait aussi décaler à Dax, ainsi la fonction de calage des Hmax prenait comme argument la liste des valeurs de l'évènement observé à Pontonx, la liste des valeurs de l'évènement du catalogue qui allait être comparé (que l'on appellera évènement comparé dans la suite) et retourné la liste des valeurs de l'évènement comparé décalé, ainsi que l'indice de décalage.

Une autre idée a été proposée. Celle de décaler la courbe comparée jusqu'à obtenir la valeur de Nash optimale. Pour cela j'ai écrit une fonction calage qui prend en argument la liste des valeurs de l'évènement observé à Pontonx-sur-l'Adour et celle de l'évènement comparé, et qui décale l'évènement comparé d'un pas de temps 1h (on décale donc valeur par valeur), et calcule la valeur de Nash à chaque décalage. Une fois le décalage effectué la fonction trouve dans le tableau récapitulatif quel est le meilleur indice de calage pour l'évènement comparé. Ceci est fait pour tous les évènements du catalogue et la méthodologie se répète encore une fois mais cette fois-ci avec un « calage Nash ».

Cependant, les résultats n'étaient toujours pas concluants, en effet le critère de Nash ne peut se calculer que sur des événements de même longueur, ainsi pour contourner ce problème, je coupais la liste de valeur de l'évènement le plus long et perdais donc en informations sur l'évènement coupé. J'ai donc pu déduire que le critère de Nash n'était pas un paramètre suffisant pour permettre de décrire correctement un évènement.

### 4.3 3<sup>ème</sup> méthode de recherche : la méthode finale

Ainsi, avec mon tuteur nous avons donc mis au point une nouvelle méthodologie de recherche. Pour cette nouvelle méthode on part d'un évènement du catalogue de crue, l'évènement observé, que l'on coupe à un instant  $t_0$ . Dans un premier temps nous avons défini le  $t_0$  au niveau du rebond de Dax. En effet, le rebond étant un évènement difficile à modéliser et un point clé de l'évènement, cela paraissait pertinent de commencer avec ce moment.

Les valeurs de l'évènement ont donc été coupées à  $t_0$  et l'étude de ressemblance se fait sur les valeurs comprises entre le dépassement du seuil jaune de la station et la valeur  $t_0$ . Sur la figure 4.3.1 l'étude de ressemblance se fera à Pontonx-sur-l'Adour sur les valeurs représentées en orange.

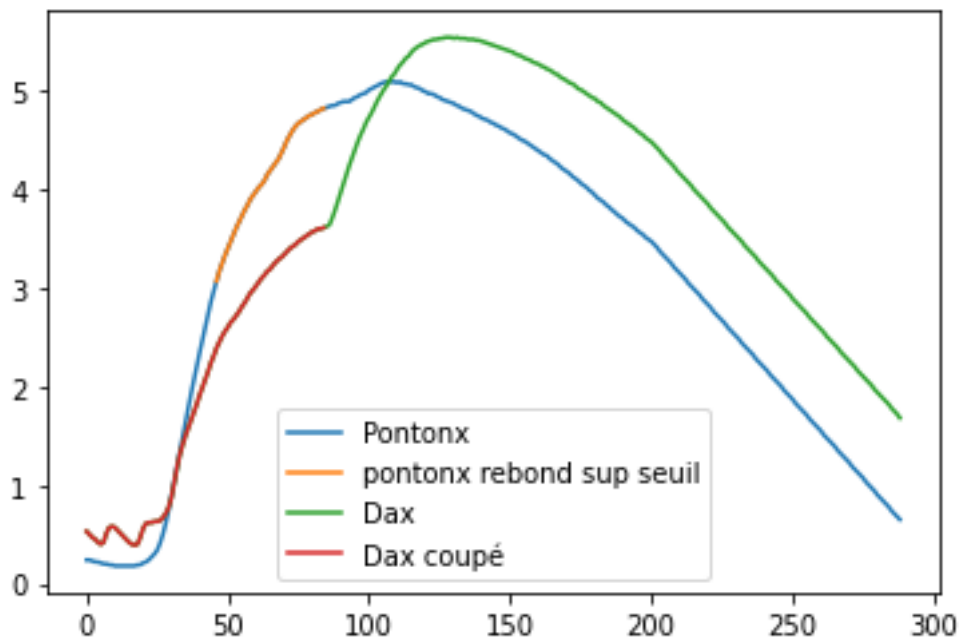


Figure 4.3.1: Graphique de l'évènement de mars 1999 à Dax et Pontonx-sur-l'Adour

Afin de permettre cette étude à partir du rebond de Dax, il faut permettre une reconnaissance facile et rapide de la présence du rebond ou non. Pour cela, j'ai ajouté une colonne au catalogue de crues avec la valeur 0 dans le cas où aucun rebond n'était observé à Dax et 1 sinon.

La caractérisation informatique du rebond a été faite via les pentes de la courbe des hauteurs à Dax. En effet comme on peut le voir sur la figure 4.3.1 le rebond est l'endroit où la différence des pentes est la plus importante.

Ainsi pour se faire j'ai pris les valeurs de l'évènement observé à Dax et j'ai coupé toutes les valeurs inférieures au seuil jaune (on coupe ces valeurs car si elles sont toutes inférieures au sup seuil jaune alors la vigilance est au vert et donc la prévision n'est pas utile). Des valeurs restantes j'ai créé une liste(=daxobs) de valeurs comprises entre le sup seuil jaune et la hauteur max (la suite n'est pas utile car le rebond apparait avant la hauteur maximum). Une fois la liste daxobs complétée je calcule la pente point à point de la courbe avec la formule donnée sur la figure 3.2.4. Ces valeurs de pente sont conservées dans un tableau de la taille de la liste.

Ensuite j'ai construit un nouveau tableau dans lequel je notais la valeur de la différence des pentes calculées précédemment. Une fois ce tableau rempli, j'ai cherché le maximum(=m) car cela me donnait l'endroit où la différence des pentes est maximum. Je crée ensuite une liste de toutes les valeurs à Dax comprise entre le sup seuil jaune et t0. Cette caractérisation du rebond se fait par une fonction rebond qui prend en argument la liste de Dax et retourne la liste comprise entre 0 et le rebond, ainsi que l'indice d'apparition du rebond. Cette fonction a posé un problème sur certains évènements. En effet, quelques-uns des évènements observés se coupaient avant l'apparition du rebond, cela était dû à des petites imperfections sur la courbes, ce problème a été réglé via une fonction de lissage, permettant de supprimer ces imperfections qui entraînaient des différences de pentes sur la courbe plus importante que celle au niveau du rebond.

Une fois l'indice de t0 obtenu, on passe à l'étude sur les stations amont. On commence par Pontonx-sur-l'Adour, on coupe « Pontonx observé » au niveau de t0 et on calcule les paramètres descriptifs de l'évènement. Dans le cas de Pontonx-sur-l'Adour, on calcule la pente, le volume et Hmax. Cependant, après avoir observé les évènements coupés on remarque que le Hmax peut être divisé en deux Hmax différents : le Hmaxabs et le Hmaxrebond.

Le Hmaxrebond signifie que l'on n'a pas encore observé de max à Pontonx au moment du rebond, tandis que le Hmaxabs signifie qu'au moment de la coupure les valeurs à Pontonx-sur-l'Adour sont en train de redescendre et qu'un max a été observé avant le rebond. Cette distinction est importante car elle va modifier la façon dont la recherche d'analogue va s'effectuer.

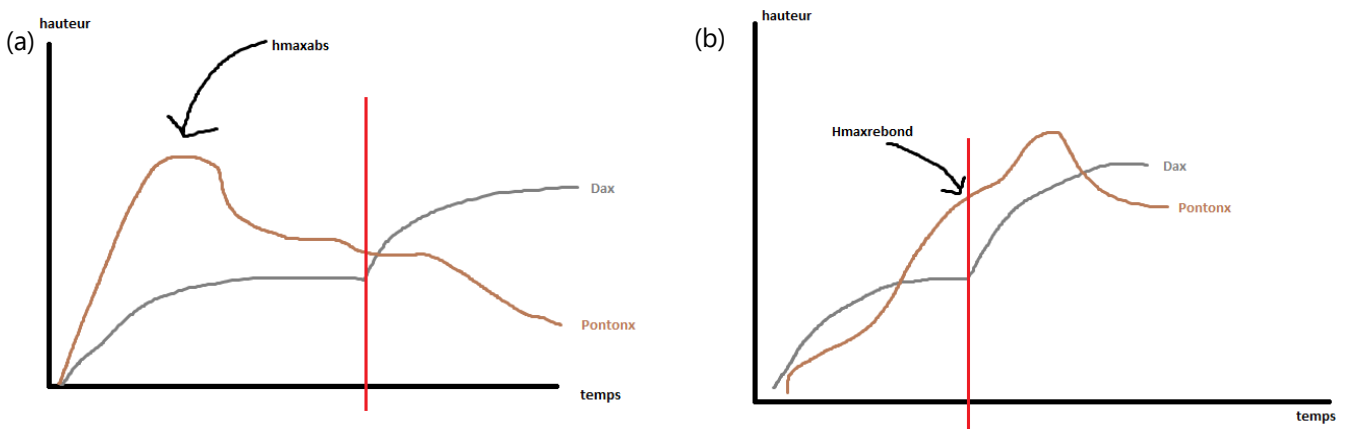


Figure 4.3.2: Schéma explicatif Hmaxabs (a) et Hmaxrebond (b) pour un évènement de catalogue

En fonction du type de hauteur max atteinte la recherche va être différente. Dans le cas (b), le Hmax atteint est celui au moment du rebond, ainsi lorsqu'on parcourra les autres évènements du catalogue on cherchera la valeur du Hmax lors du rebond. Prenons un exemple, notre évènement observé atteint un Hmaxrebond, ainsi on parcourt le reste du catalogue, on sélectionne l'évènement j et on calcule les paramètres pour permettre le calcul de D. La hauteur max de l'évènement observé étant un Hmaxrebond il est nécessaire de couper l'évènement j pour déterminer la valeur de son Hmax mais avant la coupure. Attention cependant, la coupure de l'évènement j se fait en fonction du rebond de Dax<sub>j</sub> et pas Dax observé.

Dans le cas (a) on ne coupe pas l'évènement j et on cherche juste le Hmax dans l'entièreté des valeurs.

Afin de déterminer quel type de hauteur max est atteinte une fonction a été écrite. La détermination informatique du type de Hmax se fait via les pentes, comme pour le rebond. En effet

cette fonction prend en arguments le tableau de données totales (ensemble des évènements, ensemble des valeurs de chaque station, en hauteur et débit), le nom de la station étudiée (cela permet de facilement changer l'étude du Hmax en fonction de la station) et la valeur du sup seuil jaune (différentes pour chaque station comme le montre les figures 2.2.1, 2.2.2 et 2.2.3). Cette fonction va couper les valeurs de hauteur de la station étudiée au niveau du rebond, lisser la courbe restante (via la fonction lissage mentionnée dans le calcul de rebond), et calculer la pente point par point. Si le max atteint est de type Hmaxabs alors en amont du max les pentes sont positives ou nulles (sur le point maximum de la courbe) tandis qu'une fois le point culminant atteint, les pentes deviennent négatives. Ainsi, j'ai donc conservé les valeurs des pentes dans un tableau et lorsque qu'une valeur devenait négative cela signifiait que le max atteint était de type Hmaxabs. Le lissage de la courbe permet d'assurer une continuité des pentes, c'est-à-dire que les pentes soient bien positives avant l'atteinte du max et négative après.

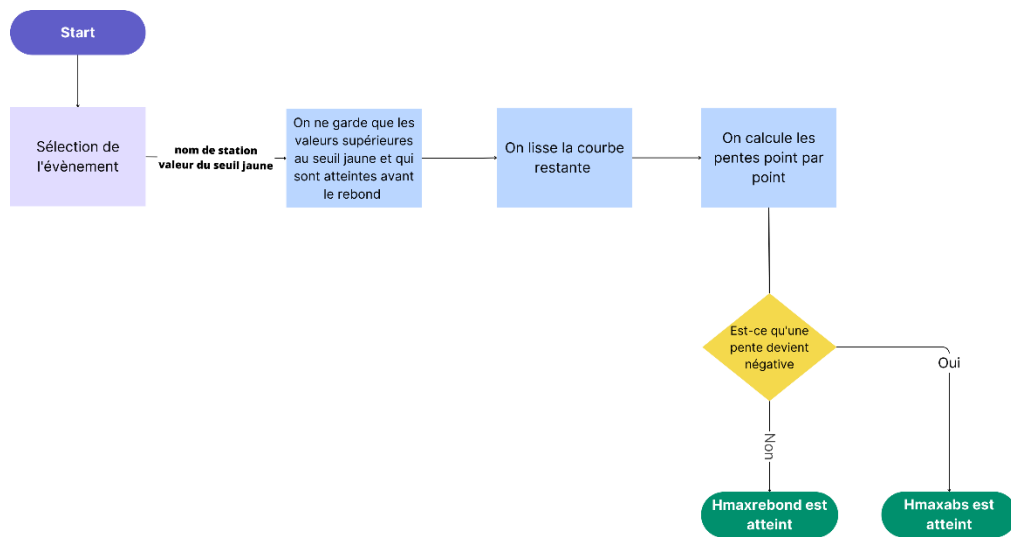


Figure 4.3.3: Logigramme de sélection des Hmaxabs ou Hmaxrebond

Le calcul de la pente se fait via une fonction prenant en arguments 4 nombres :  $x_a, y_a, x_b, y_b$ , et retourne la valeur de la pente  $(y_b - y_a) / (x_b - x_a)$  si  $x_b \neq x_a$  (cf. équation figure 3.2.4). Le calcul de la pente diffère aussi en fonction du type de Hmax. En effet comme le montre la figure 3.2.4, le calcul de la pente se fait en fonction du Hmax, ainsi la méthode de calcul de la pente diffère. Le schéma ci-dessous explique les 2 méthodes de calcul en fonction du Hmax :

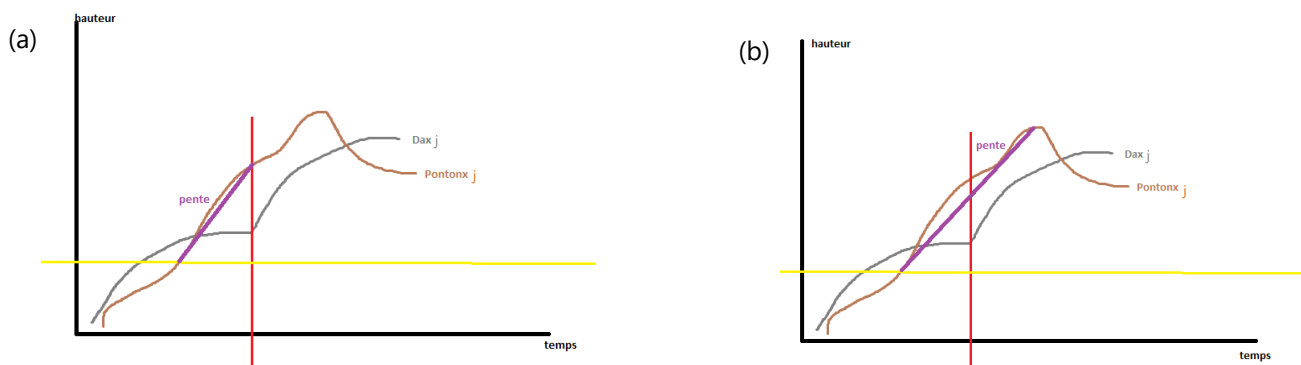


Figure 4.3.4: Pente dans le cas Hmax=Hmaxrebond (a) Pente dans le cas Hmax=Hmaxabs (b)

Le volume est calculé via une fonction qui prend en argument un tableau de données composé des valeurs de hauteur, de débit et les dates (DD/MM/YYYY) de la station de notre choix, la valeur du rebond de l'évènement étudié, ainsi que le seuil jaune. De ces deux arguments la fonction coupe le tableau de valeurs au niveau du rebond. De cette nouvelle table de valeurs elle supprime celles

inférieures au sup seuil jaune et calcule ensuite l'aire sous la courbe de l'hydrogramme via la méthode python d'intégration des trapèzes.

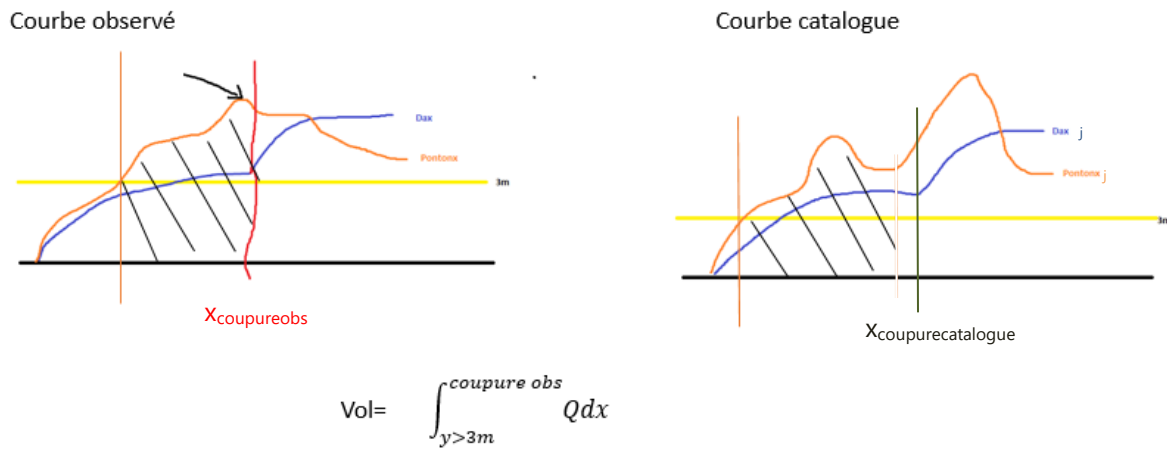


Figure 4.3.5: Schéma explicatif des volumes calculés pour les événements du catalogue en fonction de l'évènement observé

Ainsi on calcule l'ensemble de ces paramètres pour l'évènement observé et un évènement du catalogue. On calcule ensuite la distance

$$D = \sqrt{(h_{maxobs} - h_{maxprev})^2 + (volumeobs - volumeprev)^2 + (penteobs - penteprev)^2}$$

Ce choix de « distance » a été un choix fait par intuition, inspiré par la distance que l'on calcule entre 2 points dans l'espace. La recherche de la meilleure distance est une composante à part entière de la recherche d'analogie. Elle pourrait être une combinaison linéaire des différentes variables pour ne pas affecter le même poids d'importance à chaque variable. Qui plus est, les poids affectés pourraient aussi être eux-mêmes fonctions de l'instant de prévision. Ainsi notre composante « distance » est une hypothèse faite pour permettre de continuer d'avancer durant le stage.

Cette valeur est conservée dans un tableau associé au numéro de l'évènement du catalogue.

On effectue ces calculs pour tous les évènements restant du catalogue. On obtient donc un tableau (Di) de la taille 2 x (nombre d'évènement-1), récapitulant la distance de tous les évènements du catalogue à l'évènement observé. De ce tableau on choisit la distance minimale ainsi que le numéro de l'évènement du catalogue associé, évènement que l'on appellera évènement prévu dans la suite. On effectue ces étapes pour chaque évènement du catalogue, on obtient donc un tableau (=N3) de taille 3 x (nb d'évènements) contenant le numéro de l'évènement observé, de l'évènement prévu, ainsi que la valeur de la distance associée. Ainsi on a obtenu un critère de choix.

Il faut maintenant vérifier que ce critère de choix fonctionne. Pour le vérifier, on calcule le critère d'évaluation. Pour calculer les valeurs de Nash entre t0 et t0+24h, que l'on appellera Nash24 dans la suite, et les valeurs de Nash48 (entre t0 et t0+48h) trois fonctions ont été utilisées. La fonction Nash mentionnée précédemment, mais aussi une fonction `previ_24h` (resp. `previ_48h`) qui prend en argument le tableau de données totales, N3, et R une liste récapitulatif des indices de rebond de

chaque évènement. Cette fonction récupère les valeurs de Dax de l'évènement observé au niveau du rebond, coupe l'évènement prévu au même endroit et calcule le Nash à t0+24( resp. t0+48). Elle retourne un tableau final de taille 3 x (nb d'évènements) contenant le numéro de l'évènement observé, la valeur du Nash et la valeur de la distance.

En effet afin de vérifier que notre critère de choix fonctionne on va utiliser la corrélation entre la valeur du critère de choix et la valeur du critère d'évaluation. Pour se faire, on trace le nuage des valeurs de Nash en fonction de la distance. L'idée est que le critère de choix est validé si en effectuant une régression linéaire on obtient un nuage de cette forme :

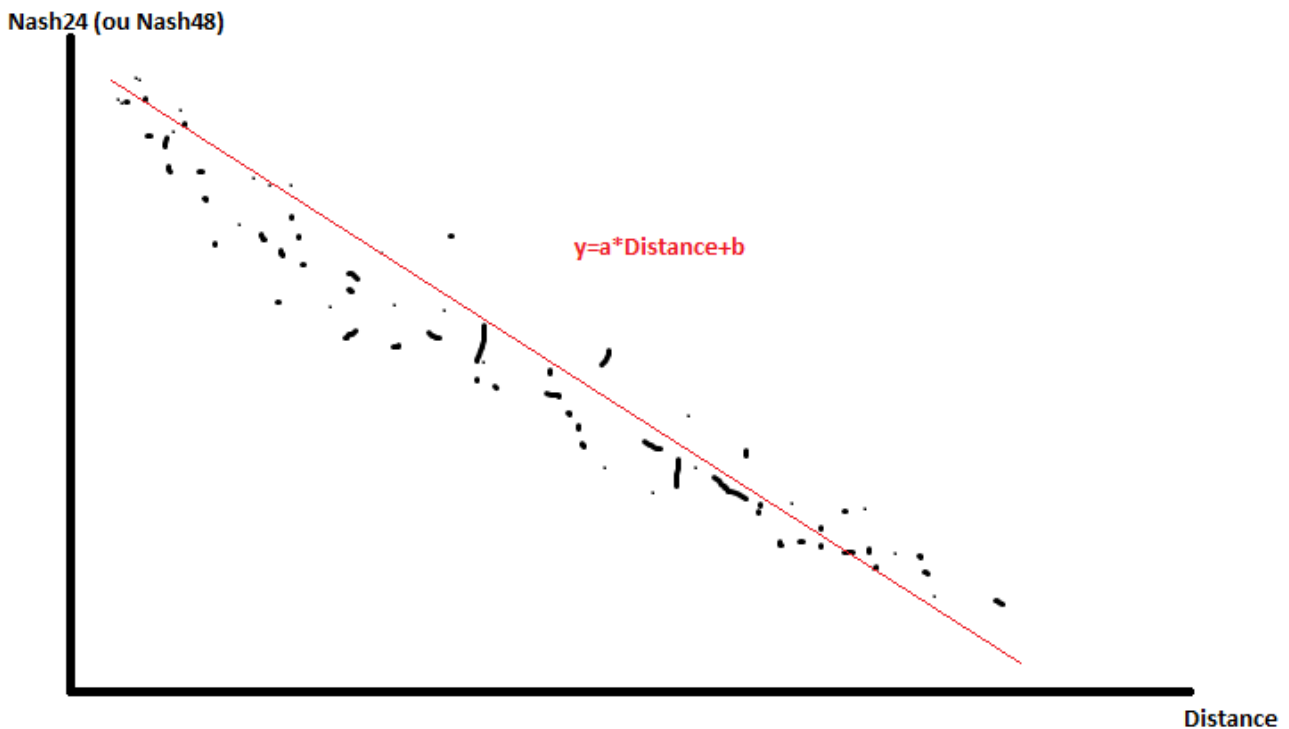


Figure 4.3.6: Nuage de points théorique des valeurs de Nash24(ou Nash48) en fonction de la distance avec courbe de régression linéaire

Si la régression linéaire n'est pas bonne on joue sur les coefficients du calcul de distance afin d'améliorer les résultats.

Le logigramme ci-dessous résume le fonctionnement de notre méthodologie finale :

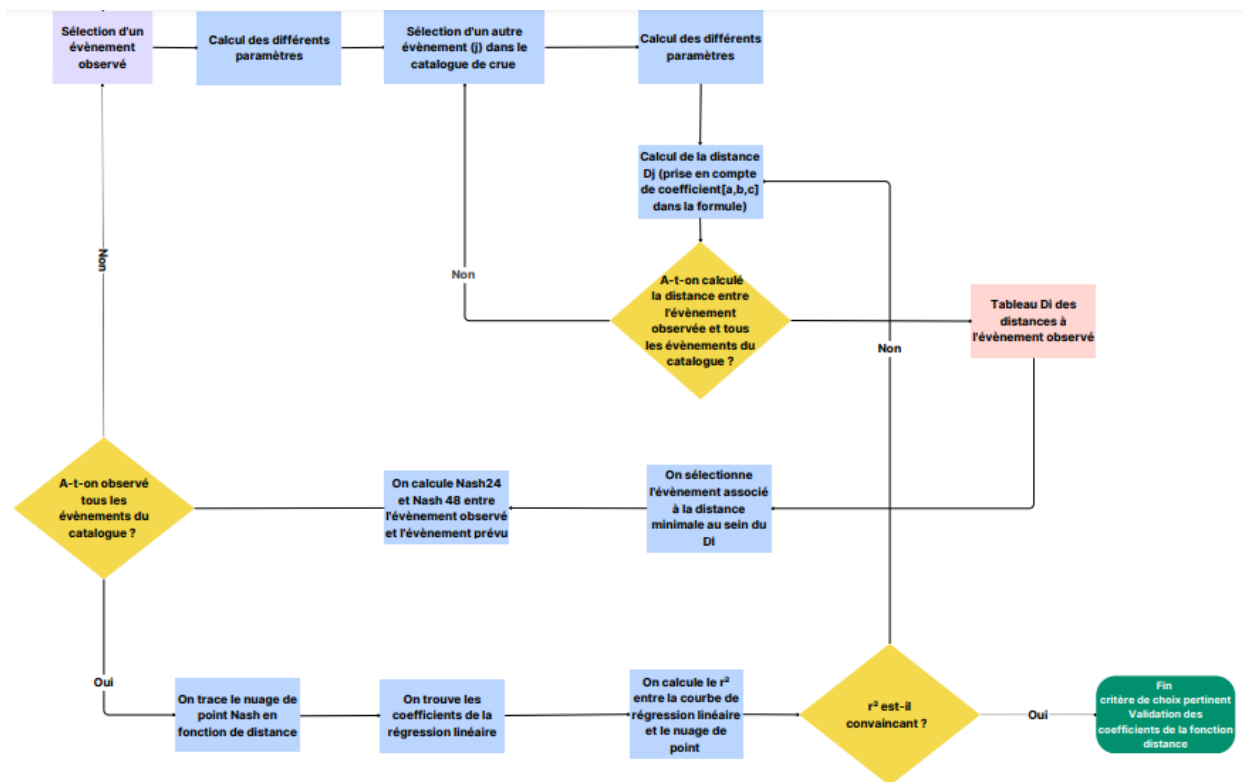


Figure 4.3.7: Logigramme de la méthode finale de recherche d'analoge d'évènement de crue

L'intérêt de cette méthode est de permettre une automatisation de la recherche d'analoge en fonction du moment de découpe, ce qui permet de l'appliquer à n'importe quel moment de la prévision.



## 5. Proposer une prévision à partir des analogues

### 5.1 Combinaison linéaire : la méthode « intégratrice » et calage

Dans les 2 premières méthodes, une fois les évènements sélectionnés grâce aux valeurs sur la station Pontonx-sur-l'Adour, je me suis concentrée sur les valeurs à Dax de ces « meilleures » évènements. L'objectif était de trouver une combinaison linéaire des évènements sélectionnés à Dax qui permettrait d'approcher au mieux l'évènement observé. Cette combinaison linéaire doit être la même pour tous les évènements afin d'obtenir une méthode s'appliquant aussi aux évènements futurs non catalogués.

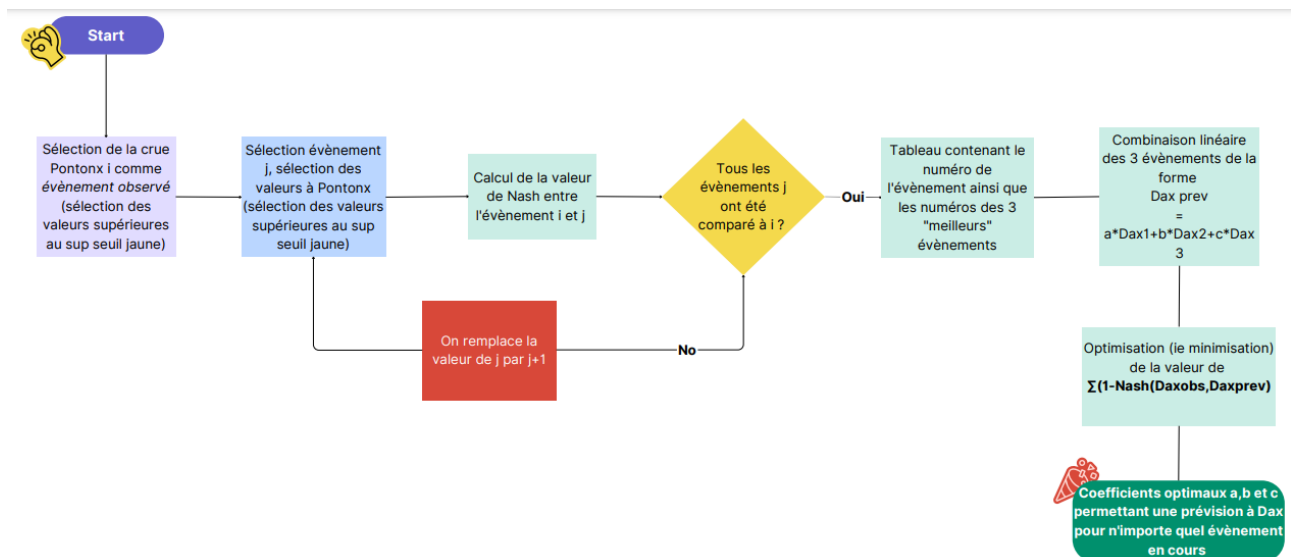


Figure 5.1.1: Logigramme de la première méthode de recherche

Pour pouvoir mettre en place cette méthode il a fallu écrire la fonction calculant le Nash, cette fonction prend en argument deux listes, la liste des valeurs observées et la liste des valeurs prévues et renvoie la valeur du Nash. Attention cependant, l'ordre des listes est important, en effet si par erreur vous demandez au script de calculer  $\text{Nash}(\text{liste}_{\text{prévue}}, \text{liste}_{\text{observée}})$  le résultat sera différent du résultat recherché. De plus, il est possible que la liste des valeurs ne soit pas complète, qu'il manque quelques valeurs sur l'ensemble de l'évènement. Ce manque traduit par des NaN (Not a Number) sur Python qui empêche la fonction Nash de tourner correctement. Ainsi il a fallu créer une fonction qui permet de régler ce problème, fonction qui prend en argument la liste à problème et retourne la liste complète sans les NaN.

Une fois le problème des valeurs réglé, on crée une boucle for dans laquelle on va parcourir tous les évènements du catalogue, cette première boucle va permettre de sélectionner l'évènement observé. Une deuxième boucle est ensuite créée pour parcourir le reste des évènements, calculer les Nash et sélectionner les 3 meilleures valeurs de Nash et donc les 3 « meilleurs » évènements.

Il est important de noter que le nombre d'analogues à sélectionner est aussi une composante à part entière de la recherche par analogie. Combien d'évènements sont nécessaire pour permettre une bonne prévision ? Ainsi afin de pouvoir continuer à travailler j'ai fait l'hypothèse que 3 était un chiffre optimal pour la prévision

Une fois ces évènements sélectionnés, j'ai fait face à un problème auquel nous n'avions pas pensé lors de l'élaboration de la première méthode de recherche, les valeurs de Nash négatives. En effet, nous avons décidé que pour pouvoir considérer que le choix des 3 « meilleurs » évènements était pertinent la valeur de Nash devait être supérieur ou égale à 0.5. Or, pour certains évènements le 3<sup>ème</sup>, voir 2 ou même les 3 « meilleurs » évènements sélectionnés possédaient un critère de Nash inférieur

à 0.5. Il a donc fallu pallier ce problème, pour cela j'ai écrit une fonction qui prend en argument le tableau récapitulatif contenant l'évènement observé, les 3 « meilleurs » évènements ainsi que la valeur du critère de Nash associée, et renvoie 4 tableaux différents, celui contenant les évènements dont aucun Nash n'était supérieur ou égal à 0.5 (catégorie sans Nash), celui avec les évènements possédant un « meilleur » évènement (catégorie 1 Nash), puis un récapitulatif les évènements avec 2 « meilleurs » (catégorie 2 Nash) et enfin un dernier avec les évènements et ses 3 « meilleurs » (catégorie 3 Nash).

Le nombre de « meilleurs » évènements n'étant pas identiques pour tous (certains en posséder 1, d'autres 2,...), la combinaison linéaire ne pouvait pas être communes aux 43 évènements du catalogue. Il a donc fallu différencier les fonctions d'optimisation de la combinaison linéaire différente en fonction de l'appartenance de l'évènement observé à sa catégorie (la catégorie sans Nash n'étant pas incluse).

Le principe d'optimisation se décompose en trois temps :

1<sup>er</sup> temps : En fonction de sa catégorie, pour chaque évènement observé on calcule l'évènement prévu ( $a \cdot \text{Dax1} + b \cdot \text{Dax2} + c \cdot \text{Dax3}$  pour un évènement observé appartenant à la catégorie 3 Nash)

2<sup>ème</sup> temps : On définit l'erreur qu'il nous faudra minimiser dans notre processus d'optimisation. Dans notre cas  $\text{erreur} = \sum (1 - \text{Nash}_{\text{évènementobs}})$  où  $\text{Nash}_{\text{évènementobs}} = \text{Nash}$  entre l'évènement observé et l'évènement prévu dans le 1<sup>er</sup> temps

3<sup>ème</sup> temps : on minimise l'erreur via une fonction minimize, qui va chercher de façon automatique les coefficients a,b,c qui vont minimiser l'erreur définit dans le 2<sup>ème</sup> temps.

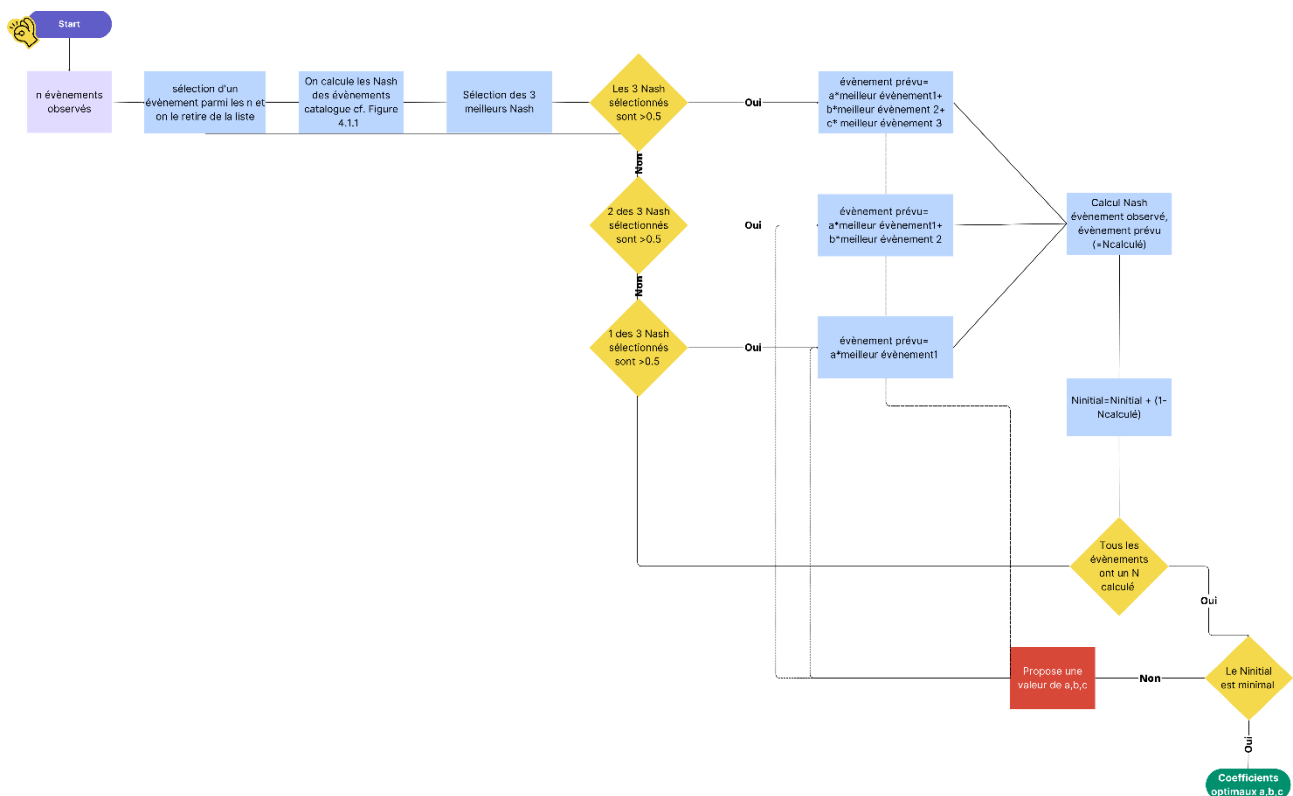


Figure 5.1.2: Schéma explicatif de la méthode d'optimisation de la combinaison linéaire dans le cas où l'évènement observé possède 3 « meilleurs » évènements

La méthode est la même pour le cas 2 « meilleurs évènements) (resp. 1), on recherche les coefficients a et b (resp. a) optimaux.

Une fois les coefficients optimaux déterminés, on obtient  $Da_{x_{prévu}} = A_{final} * Da_{x1} + B_{final} * Da_{x2} + C_{final} * Da_{x3}$ .

Ensuite on calcule le Nash( $Da_{x_{observé}}$ ,  $Da_{x_{prévu}}$ ) pour valider la méthode.

## 5.2 Prise en compte de plusieurs analogues pour proposer une prévision : méthode finale

Le critère de choix nous permet de trouver le meilleur analogue à l'évènement observé. Cependant ce critère étant validé par la régression linéaire il ne garantit pas la meilleure prévision au niveau de Dax. Ainsi, afin d'optimiser la prévision, l'idée n'est pas de sélectionner LE meilleur analogue mais LES meilleurs analogues.

L'idée est la suivante, considérons un évènement en cours E pour lequel nous aimerions émettre une prévision chiffrée et graphique sur la station de Dax. L'évènement E a atteint le rebond à Dax, ainsi le critère de choix permet

- i) de sélectionner un certain nombre d'évènements analogues  $\{P_{ana}\}$  à l'évènement E
- ii) d'établir une prévision sous la forme d'une combinaison linéaire ou statistique des situations analogues

La sélection des situations analogues se fait grâce à une fonction qui prend en argument le tableau des distances  $D_E$  de l'évènement E, ainsi que k le nombre de situation voulu et retourne les k distances les plus petites et permettre ainsi la sélection de k évènements analogues.

L'idée est de rechercher ensuite quelle est la meilleure combinaison possible de ces k analogues pour prévoir la situation à Dax entre  $t_0 : t_0+24h$  et  $t_0 : t_0+48h$ .

Dans un premier temps la sélection se portera sur maximum trois situations. De ces trois sélections nous vérifierons leur pertinence en fonction de leur valeur de N24 et N48. Une fois cette vérification effectuée, on pourra lancer une optimisation comme présenté dans la première phase de recherche pour l'obtention d'un jeu de variables d'analogie optimal.



Figure 5.2.1 : Repère temporel de l'évolution des différentes méthodes testées pendant mon stage

## 6. Résultats

### 6.1 Résultats de la première méthode

Dans un premier temps, l'étude s'est portée sur la ressemblance des évènements via le critère de Nash. Une fois l'évènement observé sélectionné, on calculait la valeur du critère de Nash entre l'évènement observé à Pontonx-sur-l'Adour et celui du catalogue. Dans un premier temps, l'analyse était brute, l'entièreté des données (pas seulement les valeurs supérieures au seuil jaune) était analysée. De ces valeurs, les trois meilleures étaient retenues.

Le script écrit pour réaliser cet objectif permet d'obtenir un tableau de taille 43x7, contenant l'évènement observé, les trois meilleurs évènements ainsi que la valeur du critère de Nash associé. Ces valeurs varient entre 0.96 et -3.54. Cependant, un nash inférieur à 0.5 indique une non-ressemblance des deux évènements comparés. En effet, comme le montre la figure 6.1.1, un nash inférieur à 0.5 entraîne des différences qui font qu'en cas de sélection « manuel » cet évènement n'aurait pas été choisi.

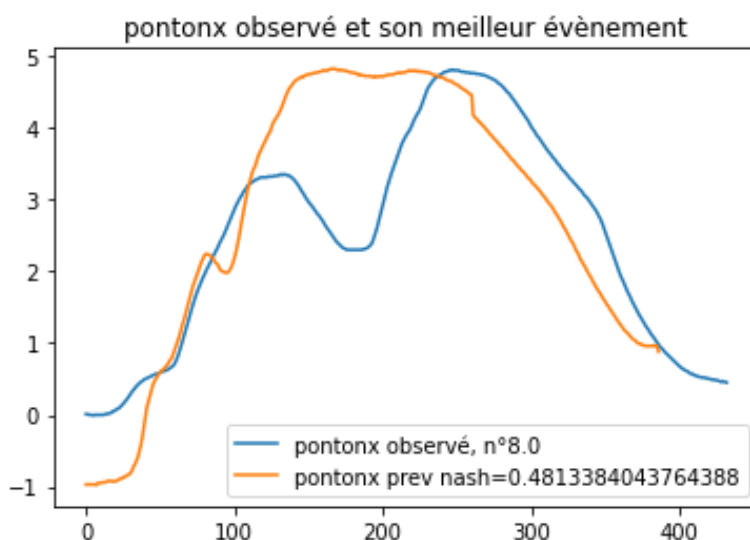


Figure 6.1.1: Image de l'évènement 8 à Pontonx et de son meilleur évènement dont le Nash est inférieur à 0.5

Ainsi il a fallu trier les évènements observés en 4 catégories différentes :

Catégorie	Nombre d'évènements observés dans cette catégorie
Trois Nash (évènement dont 3 évènements ont été trouvés dans le catalogue avec un Nash >0.5)	30
Deux Nash (évènement dont 2 évènements ont été trouvés dans le catalogue avec un Nash >0.5)	5
Un Nash (évènement dont 1 évènements a été trouvé dans le catalogue avec un Nash >0.5)	1
Zéro Nash (évènement dont aucun évènement a été trouvé dans le catalogue avec un Nash >0.5)	7

Tableau 6.1.1: Tableau récapitulatif du nombre d'évènements possédant un, deux trois ou aucun « sosie »

Une fois ces résultats enregistrés, j'ai cherché à effectuer une recherche de combinaison linéaire optimisant les Nash (à Dax) entre l'évènement observé et la combinaison linéaire des évènements prévus. Cette combinaison était différente en fonction de la catégorie dans laquelle les évènements observés appartiennent.

Catégorie	Coefficients de la combinaison linéaire		
Trois Nash	a=0.37	b=0.28	c=0.35
Deux Nash	a=0.4		b=0.57
Un Nash	a=0.82		

Tableau 6.1.2: Tableau récapitulatif des coefficients optimaux pour la combinaison linéaire des évènements « prévus »

Ces différentes combinaisons linéaires amènent à une moyenne de Nash pour tous les évènements à 0.83. Ce qui est un très bon score, cependant, en observant les courbes de Dax observé et la combinaison linéaire des Dax prévus, on remarque que pour certains évènements le Nash est très bon et pourtant certains aspects importants de l'évènement ne sont pas représentés. Prenons l'exemple de l'évènement 25 :

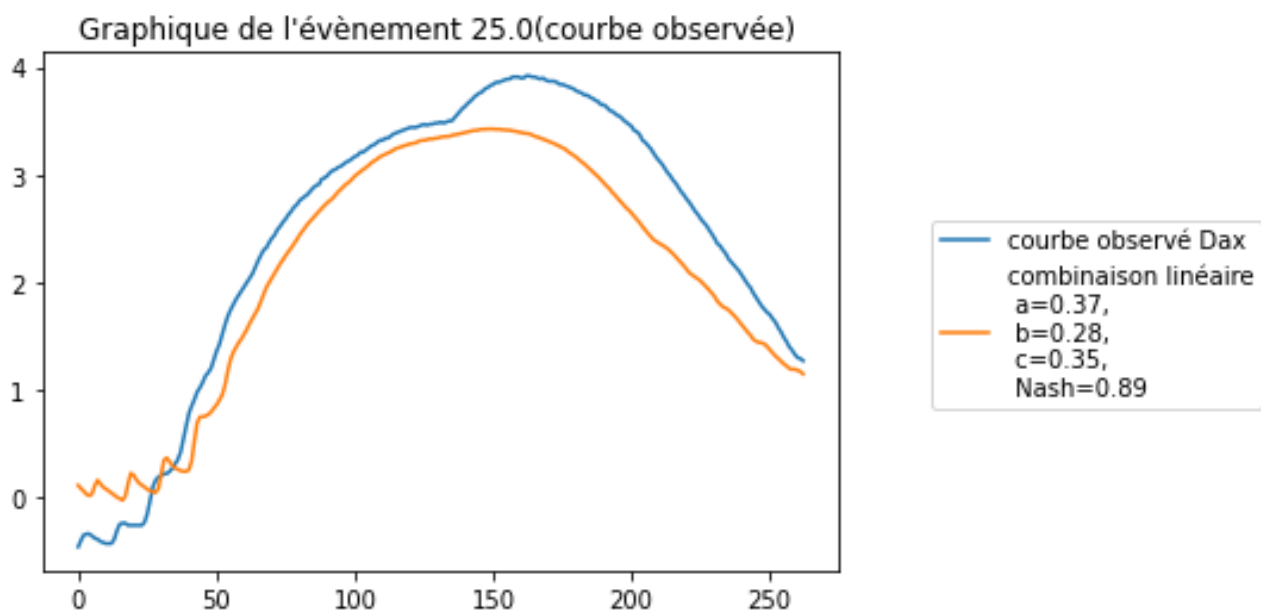


Figure 6.1.2 : Graphe de l'évènement observé à Dax ainsi que de la combinaison linéaire des trois meilleurs évènements.

On remarque sur la figure 6.1.2 que la courbe observée présente un rebond qui n'apparaît pas sur la combinaison linéaire. De plus la hauteur maximum atteinte par la combinaison linéaire est inférieure de plus de 0.5m par rapport au maximum de la courbe observée, et pourtant la valeur de Nash vaut 0.89 ce qui est un très bon score. Cette valeur peut s'expliquer par les vaguelettes que l'on observe au début de l'évènement. On remarque que la hauteur de ces vaguelettes est inférieure au sup seuil jaune.

Ainsi prendre les valeurs brutes comme ceci n'est pas concluant car la prévision se fait même pour les valeurs inférieures au sup seuil jaune, ce qui signifie une prévision pour des valeurs correspondant à une vigilance verte, ce qui ne nous intéresse pas.

Dans un second temps, je me suis donc intéressés aux évènements « coupés », c'est-à-dire, on étudie seulement les valeurs supérieures au sup seuil jaune des stations. L'étude de ressemblance était toujours faite à Pontonx, mais je n'ai pas seulement coupé les valeurs inférieures à 3m, j'ai aussi effectué un calage afin de sélectionner les évènements dont le décalage permettait d'obtenir le meilleur Nash.

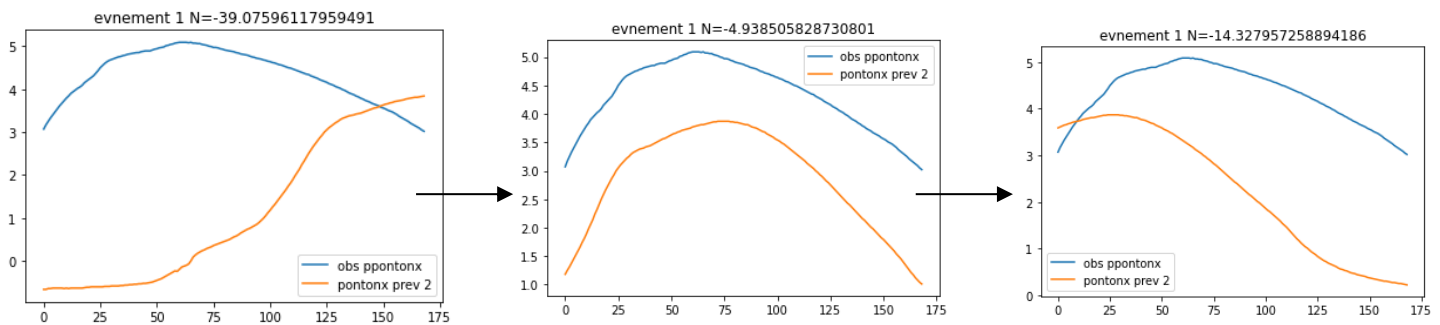


Figure 6.1.3: Images du calage de l'évènement 1 avec l'évènement 2

Du calage on obtenait les trois meilleurs évènements associés à l'évènement observé. Chaque évènement observé est de nouveau trié en fonction de son nombre de Nash :

Catégorie	Nombre d'évènements observés dans cette catégorie
Trois Nash (évènement dont 3 évènements ont été trouvés dans le catalogue avec un Nash >0.5)	29
Deux Nash (évènement dont 2 évènements ont été trouvés dans le catalogue avec un Nash >0.5)	5
Un Nash (évènement dont 1 évènements a été trouvé dans le catalogue avec un Nash >0.5)	2
Zéro Nash (évènement dont aucun évènement a été trouvé dans le catalogue avec un Nash >0.5)	5

Tableau 6.1.3 : Tableau récapitulatif du nombre d'évènements possédant un, deux trois ou aucun « sosie » après calage du Nash à Pontonx

On remarque que le nombre total d'évènements triés est différent de 43. Cela est dû au fait que deux évènements du catalogue n'ont aucune valeur supérieure au seuil jaune donc ne constituent pas un évènement amenant à effectuer une prévision.

On effectue de nouveau une combinaison linéaire pour obtenir le jeu de coefficient optimal pour le critère de Nash à Dax :

Catégorie	Coefficients de la combinaison linéaire		
Trois Nash	a=0.25	b=0.37	c=0.32
Deux Nash	a=0.39		b=0.61
Un Nash	a=0.95		

Tableau 6.1.4:coefficient optimaux pour une combinaison linéaire des évènements « prévus »

On obtient donc une moyenne de Nash qui vaut 0.2 avec en détails :

Catégorie	Moyenne Nash à Dax
Trois Nash	0.12
Deux Nash	0.46
Un Nash	0.7

Tableau 6.1.5:Tableau détaillant les valeurs des moyennes de Nash à Dax en fonction de la catégorie à laquelle appartiennent les évènements observés

On remarque que la valeur moyenne pour la catégorie Un Nash est élevée, ceci s'explique par le nombre d'évènement appartenant à cette catégorie. En effet, seulement deux évènements ne possèdent qu'un Nash supérieur à 0.5 ce qui rend aisé l'optimisation linéaire. Il y a seulement un coefficient à optimiser et il faut prendre en compte seulement deux évènements, ce qui donne une optimisation plutôt bonne. Tandis que l'optimisation de trois coefficients est plus compliquée car elle doit être optimale pour 29 évènements qui peuvent être complètement différents.

Ainsi, on peut déduire de ces différents résultats, que le critère de Nash à Pontonx n'est pas un critère suffisant pour sélectionner les meilleurs évènements. De plus, l'objectif de ce stage est d'aider à la prévision, or une prévision ne se fait pas à plus de 48h du jour-même de la prévision, ainsi nous avons revu la méthodologie afin de trouver une méthode de recherche adaptée aux conditions réelles de prévision (cf partie 4.3)

## 6.2 Résultat de la deuxième méthode

Les différents résultats de la première méthodologie ont donc aidé à définir cette méthodologie. Dans un premier temps, l'objectif de cette méthode est de trouver le jeu de variables optimal pour le calcul de distance qui donnerai la meilleure prévision à 24h et 48h.

Pour chaque station les paramètres calculés sont les suivants :

- Hmaxabs ou hmaxrebond à Pontonx-sur-l'Adour
- Hmaxrebond sur les autres stations
- Volume
- Pente
- Durée pour les stations Aire-sur-l'Adour et Riscle

J'ai d'abord porté l'étude sur les stations (coefficient de 1 devant les paramètres pour la station étudiée) :

Paramètres du critère distance	R <sup>2</sup> pour P24 & P48 Hmax rebond	R <sup>2</sup> pour P24 Hmax abs	R <sup>2</sup> pour P48 Hmax abs
Pontonx	AC>0	AC>0	AC>0
Saint-Sever	0.06	0.31	0.42
Audon	0.18	AC>0	AC>0
Grenade	0.002	0.09	0.02
Aire sur l'Adour	0.002	0.006	AC>0
Riscle	0.08	0.23	0.24

Tableau 6.2.1: Résultats obtenus par station (ac>0 signifie que le coefficient directeur de la courbe de régression linéaire est supérieur à 0)

Puis par paramètres :

Paramètre du critère distance	R <sup>2</sup> pour P24 & P48 Hmax rebond	R <sup>2</sup> pour P24 & P48 Hmax abs
Que les hauteurs (odg Différence : 0-5[0.1]) (m)	0.11	0.04
Que les pentes (odg : 0.01-1)	ac>0	0.0005
Que les volumes (odg : 10 <sup>7</sup> -10 <sup>8</sup> ) (m <sup>3</sup> )	0.32	0.009
Que les durées (0-250 000) (s)	0.17	0.03

Tableau 6.2.2: Résultats obtenus par paramètres (ac>0 signifie que le coefficient directeur de la courbe de régression linéaire est supérieur à 0)

On remarque que les résultats des régressions ne sont pas bon, tous les R<sup>2</sup><0.8. Pour essayer de comprendre pourquoi ces résultats étaient si mauvais j'ai fait afficher les régressions linéaires qui donnaient le meilleur R<sup>2</sup> pour chaque catégorie.

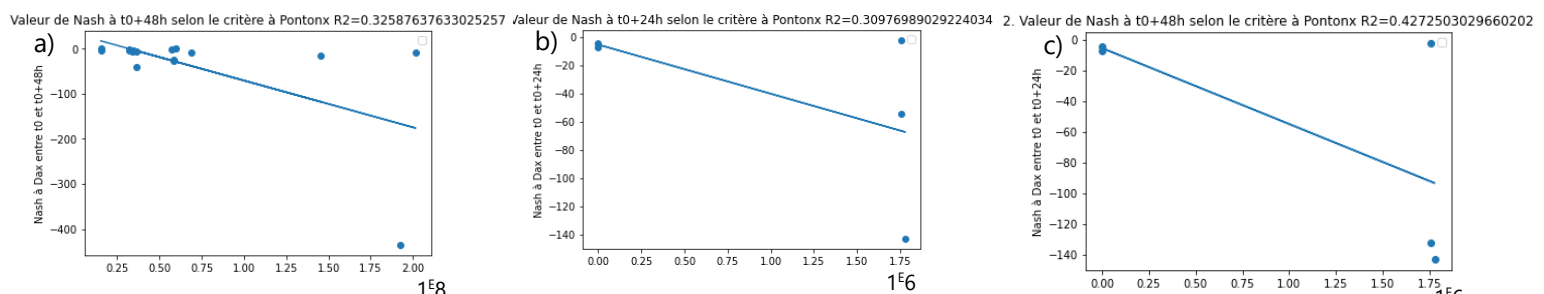


Figure 6.2.1: Nuages de points avec le meilleur R<sup>2</sup>

- Critère de distance : volume de chaque station, prévisions P24h et P48h (Hmaxrebond à Pontonx)
- Critère de distance : volume, pente et hauteur max Saint-Sever, prévisions P24h (Hmaxabs à Pontonx)
- Critère de distance : volume, pente et hauteur max Saint-Sever, prévisions P48h (Hmaxabs à Pontonx)

On remarque qu'en plus de valeurs de R<sup>2</sup> moyenne, le Nash calculé peut atteindre -400, ce qui n'est pas une valeur espérée.



Dax observé en 2 parties et dax prévu entre t0 et t0+48h, Nash=-436.4007070468996

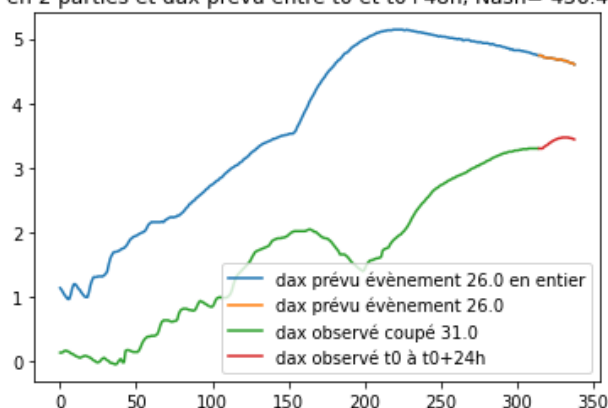


Figure 6.2.2 : Graphe des hauteurs à Dax entre l'évènement observé et celui prévu selon le critère de distance

On remarque qu'en plus d'un Nash mauvais, les rebonds apparaissent à deux moments différents. Le choix de l'évènement se faisant à partir des caractéristiques de chaque évènement avant son rebond, il paraît nécessaire de faire coïncider les rebonds des évènements observés et prévus.

J'ai refait les mêmes calculs sur les stations :

Paramètres du critère distance	R <sup>2</sup> pour P24 Hmax rebond	R <sup>2</sup> pour P24 Hmax abs	R <sup>2</sup> pour P48 Hmax rebond	R <sup>2</sup> pour P48 Hmax abs
Pontonx	ac>0	ac>0	ac>0	ac>0
Saint-Sever	0.07	0.12	0.07	0.14
Audon	0.08	0.97	0.09	0.32
Grenade	0.025	0.39	0.07	0.13
Aire sur l'Adour	ac>0	ac>0	ac>0	ac>0
Riscle	0.9	Ac>0	0.91	Ac>0

Tableau 6.2.3 : Résultats obtenus par station, après coïncidence des rebonds

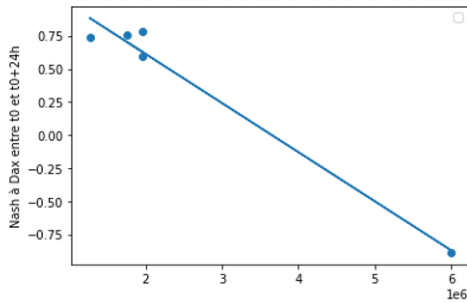
Et pour chaque paramètre :

Paramètre du critère distance	R <sup>2</sup> pour P24 Hmax rebond	R <sup>2</sup> pour P24 Hmax abs	P48 Hmax rebond	R <sup>2</sup> pour P48 Hmax abs
Que les hauteurs	0.003	0.02	0.02	0.003
Que les pentes	ac>0	0.02	ac>0	0.05
Que les volumes	0.31	ac>0	0.28	ac>0
Que les durées	0.22	0.36	0.25	0.05

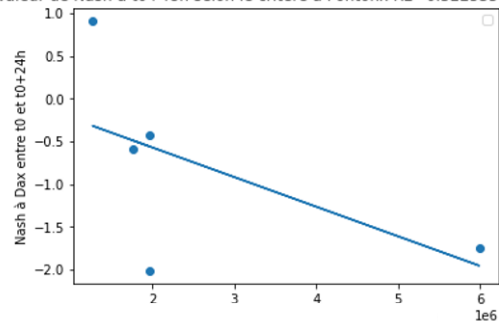
Tableau 6.2.4:Résultats obtenus par paramètres, après coïncidence des rebonds

On observe une nette amélioration du  $R^2$  pour les évènements possédant un Hmaxrebond et la prévision 24h des évènements atteignant un Hmaxabs.

Valeur de Nash à t0+24h selon le critère à Pontonx  $R^2=0.9765255549723084$

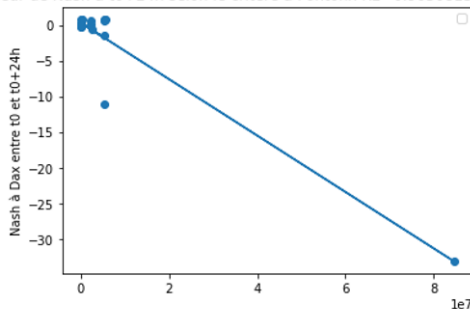


2. Valeur de Nash à t0+48h selon le critère à Pontonx  $R^2=0.3225357728725551$

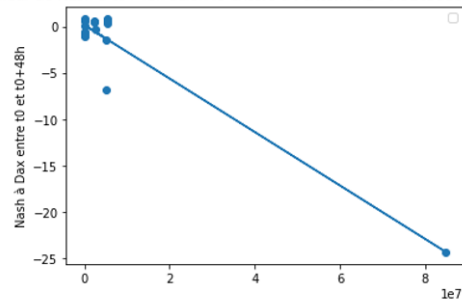


Nuage de points Hmaxabs (paramètres : Audon)

Valeur de Nash à t0+24h selon le critère à Pontonx  $R^2=0.9030813808666696$



Valeur de Nash à t0+48h selon le critère à Pontonx  $R^2=0.9163550360639533$



Nuage de points : Hmaxrebond (paramètres : Risclé)

Figure 6.2.3: Nuages de points avec les meilleurs  $R^2$

- Prévision 24h, Hmaxabs à Pontonx, critère : Hauteur, volume, pente Audon
- Prévision 48h, Hmaxabs à Pontonx, critère : Hauteur, volume, pente Audon
- Prévision 24h, Hmaxrebond à Pontonx, critère : Hauteur, volume, pente, durée Risclé
- Prévision 48h, Hmaxrebond à Pontonx, critère : Hauteur, volume, pente, durée Risclé

Les valeurs de Nash sont largement supérieures à celles observées sans la coïncidence des rebonds. On remarque cependant que les points à -25 possède un poids trop important dans la régression, il faut donc vérifier qu'un Nash à -25 entraîne une prévision vraiment mauvaise afin de pouvoir retirer ce point de la régression et voir le résultat sans. En effectuant cette vérification, nous avons rencontré un problème non négligeable.

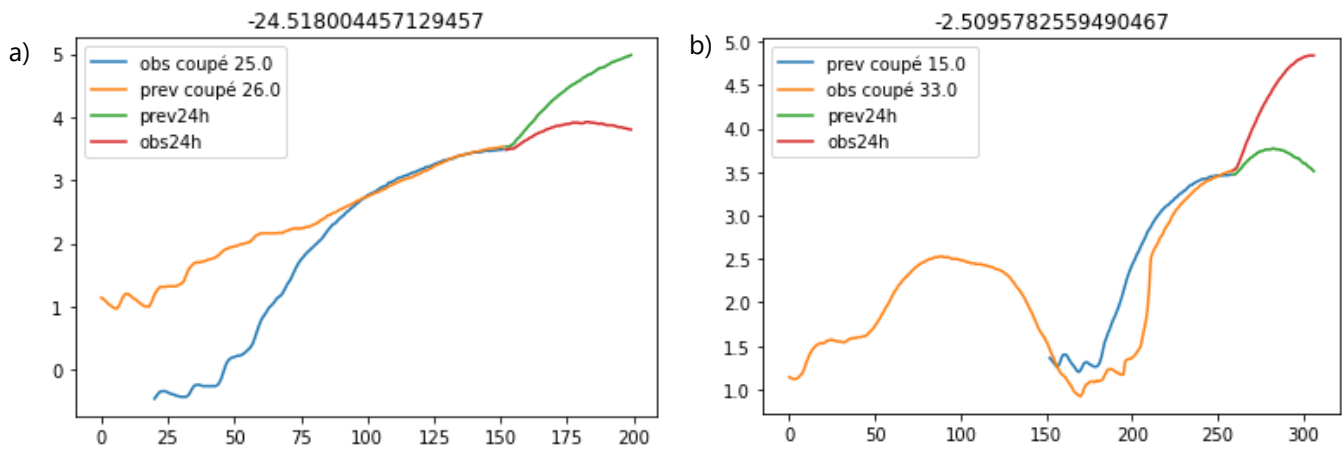


Figure 6.2.4 :Graphique de 2 évènements observés avec leur évènement prévu et la valeur de leur critère de Nash en titre des graphes

En effet, sur la figure 6.2.4 on remarque que selon les valeurs de Nash (en titre des graphes), la prévision de la figure 6.2.4.b est meilleure que celle de la a. Or, à l'œil, ce serait le contraire que l'on aurait affirmé. Cela s'explique par la variance des valeurs de l'évènement observé qui influe sur la valeur du Nash calculé. Cela montre que le choix du critère d'analogie est lui aussi une composante à part entière de la recherche d'analogie et que ce critère peut lui aussi être amélioré.

## 7. Conclusion

L'objectif de ce Travail de Fin d'Études était de développer une méthode de recherche automatique d'analogues de crues qui permettrait en situation de prévision une sélection d'évènements de crues pour aider à la prévision. Pour cela une démarche a été construite et appliquée sur le tronçon Adour Moyen Adour des Barthes, et notamment Dax, une station réglementaire surveillée par le SPC GAD. Cette ville a été choisie pour l'importance des enjeux liés à la prévision : la réaction de rebond dû aux remplissages des barthes rendent indispensable la prévision par catalogue.

Une réflexion sur la caractérisation mathématique de la ressemblance des courbes a permis d'identifier des choix de critères permettant la sélection des analogues mais aussi la vérification de ces choix. La sélection s'est fait selon plusieurs critères (hauteur max, gradient de montée, volume, durée) sur les différentes stations en amont de Dax. Le critère d'évaluation du choix d'analogues était le critère de Nash-Sutcliffe sur Dax.

Cette réflexion est suivie par une méthode mise au point pour la sélection d'évènements analogues. Dans un premier temps, on coupe l'évènement « observé » et on sélectionne l'évènement le plus ressemblant selon le critère de distance (combinaison des critères de hauteurs, pente, volume, durée). Dans un second temps on calcule le critère de Nash entre l'évènement « observé » et l'évènement sélectionné 24h et 48h après l'endroit de coupure. Une fois ces évènements calculés on effectue une régression linéaire des valeurs de Nash calculées en fonction de la distance.

L'ensemble de cette démarche a été appliquée sur la station de Dax, avec une découpe au niveau du rebond. Les différents essais de distance ont permis d'obtenir des premiers éléments de réponses quant aux critères de choix d'évènements permettant une bonne prévision. On peut notamment citer des sélections amenant à des Nash  $>0.9$ . Cependant la fiabilité du choix du critère d'évaluation n'est pas parfaite. En effet, la variance de l'évènement observé ayant un impact sur le résultat du critère d'évaluation, certains évènements sont considérés comme ressemblants pour l'œil humain mais présente un Nash faible. Cependant certains résultats sont tout de même encourageants au regard de l'utilité d'un outil automatique de recherche d'analogues de crue. Ce stage permet de proposer des évènements au prévisionniste qui peut être complété par sa propre recherche. De plus ce stage a mis en évidence certains points importants pour la prévision des crues : la sensibilité et la difficulté à traduire la ressemblance entre deux évènements, ainsi que la difficulté à produire une prévision  $y$  compris à partir d'évènement passés. Le travail d'expertise des prévisionnistes est difficile à mathématiser et automatiser.

Il n'existe pas d'antécédents sur le sujet, étant donné que le catalogue de crue n'est pas un outil national développé dans tous les SPC. Cette étude apporte une contribution aux réflexions qui peuvent être menées sur la façon de décrire la ressemblance d'évènements de crue. Elle a abouti à une méthodologie qui se veut générale et peut être appliquée aux catalogues de crues des autres tronçons du territoire en adaptant les critères de sélection et d'évaluation aux stations étudiées. Ce travail présente de nombreuses perspectives. En effet, la prévision des crues est un travail nécessitant des outils en constante évolution pour permettre aux agents des prédictions de plus en plus précises au fil du temps. Ainsi l'outil développé durant mon stage est un outil qui sera susceptible d'évoluer en fonction des évènements futurs. En effet, le nombre d'évènements présents dans le catalogue de crue peut être considéré comme une limite (difficile de comparer correctement des crues quand l'échantillon est petit), et l'ajout d'évènements permettrait une meilleure comparaison et donc une meilleure sélection d'analogues. De plus comme évoqué dans les résultats, le critère d'évaluation n'est pas totalement fiable, ce qui montre une possible évolution de la définition de ce critère.

## 8. Perspectives

A l'issue de cette étude, plusieurs pistes de travail ont été ouvertes.

- Les variables d'analogie, il est possible de tester d'autres variables permettant de décrire au mieux un événement de crue. Le paragraphe 3.2 décrit les variables d'analogie sur lesquelles je me suis concentrée durant cette étude, cependant il n'est pas impossible que le critère de ressemblance de deux événements nécessite d'autres variables afin d'améliorer la sélection.
- On peut aussi également envisager une approche différente qui se concentrerait sur la composante « distance ». En effet, comme expliqué dans le paragraphe 4.3, le choix de la distance fait durant cette étude est une hypothèse afin de pouvoir avancer dans le travail. Ainsi dans le futur, il pourrait être intéressant d'étudier cette composante afin de l'optimiser.
- Une autre hypothèse a été faite dans cette étude et mériterait une étude plus approfondie, le nombre d'analogues nécessaire à la prévision. Comme écrit dans le paragraphe 5.1, j'ai fait le choix de sélectionner seulement 3 événements analogues pour tenter d'optimiser la prévision. Or est-ce le nombre optimal ?
- On peut aussi envisager de faire évoluer cette méthode. Durant cette étude, j'ai étudié l'événement au moment du rebond à Dax, il est possible de changer de  $t_0$  et d'appliquer de nouveau cette méthode de recherche. Mais cette méthode peut aussi s'appliquer aux autres catalogues de crues

# Bibliographie

---

Ben Daoud, A. B. (2010). *Améliorations et développements d'une méthode de prévision probabiliste des pluies par analogie*. [Thèse de doctorat]. Université de Grenoble.340p

Chesneau, S. C., & Danné, A. D. (s. d.). *L'hydrométrie pour les prévis* [Diapositives ; PDF].

*Comment mettre en œuvre une régression linéaire avec Python ?* (2018, 21 janvier). Germain Salvato Vallverdu. <https://gsalvatovallverdu.gitlab.io/python/moindres-carres/#%C3%A9tape-2-cr%C3%A9er-une-fonction>

Dourdet, V. D., & Ollivier, D. O. (s. d.). *Formation modélisation* [Diapositives ; PDF].

Ministère de la transition écologique et de la cohésion des territoires. (2023). *Note technique du 18 janvier 2023 relative à la production opérationnelle de la vigilance crues : NOR : TREP2301553N*.

PERON, F. P. (2022). Règlement de surveillance, de prévision et de transmission de l'Information sur les Crues (RIC) : SPC Gironde-Adour-Dordogne. Dans *vigicrues.gouv.fr*.

Réseau VIGICRUES. (s. d.). *Scores V1.3.3 : Guide d'utilisation* [Diapositives ; PDF].

Réseau Vigicrues. (2023). La vigilance « crues » : Bilan 2022. Dans *vigicrues.gouv.fr*.

SCHAPI. (2021). *Plan d'Action Quadriennal des Départements Hydrométrie et Prévision des Crues Gironde-Adour-Dordogne et Vienne Charente Atlantique*.

StackLima. (s. d.). *Python – Coefficient de détermination-R2 Score – StackLima*. <https://stacklima.com/python-coefficient-de-determination-du-score-r2/>

# Annexes

---

## Annexe A :

### Ensemble des fonctions python écrites et utiles pour l'étude

1. *Fonction de calcul du Nash, prend en 1<sup>er</sup> argument la liste des valeurs observées et en 2<sup>ème</sup> la liste des valeurs prévues (l'ordre à son importance !)*

```
def Nash(x,y):  
    #fonction qui prend en argument 2 listes de valeurs et retourne le critère Nash  
    V=np.var(x) #calculé la variance de la liste des valeurs observées  
    diff=np.subtract(x,y) #calculé la différence entre la valeur observé (ici les valeurs x)  
    square=np.square(diff) #calculé le carré de la diff  
    MSE=square.mean()  
    N=1-(MSE/V)  
    return(N)
```

2. *Fonction `supp_nan_liste` qui supprime les valeurs Nan de la liste*

```
def supp_nan_liste(x): #fonction qui prend en argument une liste et renvoie la liste sans les nans  
    newlist=[y for y in x if np.isnan(y)==False]  
    return (newlist)
```

3. *Fonction `choix_valeurs` qui prend en argument une liste et un nombre et supprime toutes les valeurs inférieures à ce nombre*

```
def choix_valeurs(x,s): #fonction qui prend en argument une liste x  
    newlist=[]  
  
    for i in range (len(x)):  
        if x[i]>=s:  
            newlist.append(x[i])  
  
    return (newlist)
```

4. Fonction `calcul_max_n` qui prend en argument une liste et un nombre  $n$  et retourne les  $n$  premiers max de la liste

```
def calcul_max_n(liste,n) :
    for i in range (len(liste)-1) :
        if liste[i]==400 :
            liste=np.delete(liste,i,0)
    liste=liste.tolist()
    it=liste[:]
    return([it.pop(it.index(max(it))) for i in range(3) if len(it)])
```

5. Fonction `calcul_pente`, qui prend en arguments 4 nombres représentant les coordonnées de 2 points et calcule la pente entre ces 2 points

```
def calcul_pente(xa,ya,xb,yb) :
    if xa!=xb :
        p=(yb-ya)/(xb-xa)
        return(p)
    else :
        return(0)
```

6. Fonction `rebond`, qui va prendre en argument une liste et va retourner la liste coupée ainsi que l'indice de découpe

```
def rebond(L):#liste qui prend en argument une liste (celle de l'évènement de Dax) et renvoie la liste coupée au niveau du
#rebond ainsi que l'indice de coupure
    daxobs=[]
    indicearrive=0
    m=max(L)
    for j in range (len(L)):
        if m<4:
            if L[j]>=3.2 and L[j]<m: #on ne garde que les valeurs comprise entre 3,2 et la hauteur max
                daxobs.append(L[j])
            if L[j]==m :
                break
        else :
            if L[j]>=3.2 and L[j]<=4:
                daxobs.append(L[j])
            if L[j]>4 :
                break
    indicearrive=j
    l=len(daxobs)
    daxobs=lissage([s for s in range(l)], daxobs,2)[1]
    P=np.zeros((len(daxobs)))
    for i in range (1,len(daxobs)) :
        xa=i-1
        xb=i
        ya=daxobs[i-1]
        yb=daxobs[i]
        p=calcul_pente(xa,ya,xb,yb) #on calcule la pente point par point
        P[i-1]=p #on met les pentes point par point dans un tableau
    F=np.zeros((len(P)))
    for i in range (1,len(P)):
        F[i]=P[i]-P[i-1] #on calcule la différence des pentes
    m=np.argmax(F) #indique l'indice du max, ie l'indice pour laquelle la différence entre 2 pentes est la
#plus importante est représentée souvent l'endroit du rebond
    t=len(L[0:indicearrive]) #calcule la taille de la liste entre le départ et la fin de la liste daxobs
    p=t-l #calcul l'indice de départ de la liste daxobs dans la liste obsdax
    obsdaxcoupe=L[0:p+m] #donne la liste de obsdax avant le rebond
    return(F,obsdaxcoupe, p+m)
```



7. Fonction `duree_volume_événementpart` qui prend en argument un tableau, le rebond ainsi que le seuil jaune et va retourner les volumes et durée des événements sur la station choisie de tous les événements du catalogue

```
def duree_volume_événementpart(T,r,s):
    #evenementobs=donnees_totales[donnees_totales["N°"]==i] #on isole un événement qui devient l'évènement observé
    #donnees_totales=donnees_totales.fillna(0)
    # evenementobs=evenementobs.reset_index()
    #obsponontx=evenementobs.loc[:,["Pontonx-sur-l'Adour","Date / Hauteurs(m)","Pontonx-sur-l'Adour.1"]] #on récupère les hauteurs,
    #la date et les débits de Pontonx
    obsponontx=T.iloc[0:r]
    obsponontx=obsponontx.reset_index()
    L=[]
    vr=0
    ttot=0
    for j in range (len(obsponontx.iloc[:,2])):
        if obsponontx.iloc[j,1]<s : #on enlève les valeurs inférieures à 3m
            L.append(j)
    obsponontx.drop(L,axis=0, inplace=True)
    if len(obsponontx)!=0 :
        dfy=obsponontx.iloc[:,3] #onrécupère les valeurs de débits
        dfx=np.array([])
        for l in range (len(obsponontx)) :
            t1=dt.datetime.strptime(obsponontx.iloc[0,2], '%d/%m/%Y %H')
            t2=dt.datetime.strptime(obsponontx.iloc[l,2], '%d/%m/%Y %H')
            t=(t2-t1) #on calcule le temps pas à pas
            dfx=np.append(dfx,t.total_seconds()) #on créer une liste avec le pas de temps appliqué à l'évènement
        # for j in range (len(dfy)):
        # if dfy.iloc[j,1]==0:
        #     dfy[j]=(dfy[j-1]+dfy[j+1])/2
        vr=sp.integrate.trapz(dfy,dfx) #on calcule l'intégrale de la courbe des débits pour obtenir le volume
        ttot=t.total_seconds()
        if np.isnan(vr)==True : |
            vr=0
    return(vr,ttot)
```

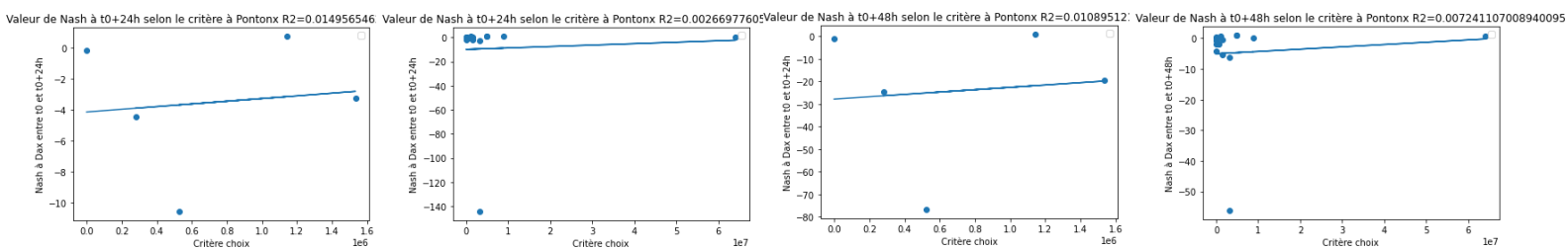
8. Fonction `Hmaxabs` qui prend en argument le tableau de données totales, la station voulue et le seuil jaune de cette station, et va retourner les événements pour lesquels le Hmax est absolue sur la station voulue

```
def Hmaxabs(donnees_totales,station, seuil) :
    L=[]
    for i in range (0,43):
        evenement=donnees_totales[donnees_totales["N°"]==i+1] # on récupère l'évènement observé
        evenement.reset_index()
        if evenement.iloc[0,25]==1 :
            Dax=evenement["Dax"]
            Pontonx=evenement[station]
            LDax=supp_nan_liste(Dax.tolist())
            LPontonx=supp_nan_liste((Pontonx.tolist()))
            r=rebond(LDax)
            pontonxc=LPontonx[0:round(r[2])]
            pontonxc2=choix_valeurs(pontonxc, seuil)
            Lx=[l for l in range(len(pontonxc))]
            Lise=lissage(Lx,pontonxc,2)
            # plt.plot([l for l in range (len(r[1]))], r[1], label="Dax")
            # plt.plot([l for l in range (len(pontonxc))], pontonxc, label="Pontonx")
            # plt.plot(Lise[0],Lise[1], label="pontonx lissé")
            # plt.legend()
            # plt.show()
            pontonxc2=choix_valeurs(Lise[1], seuil)
            if len(pontonxc2)!=0 :
                # plt.plot([l for l in range(len(pontonxc2))], pontonxc2)
                # plt.title(f"Evenement {i+1} pontonx lissé sup à 3")
                # plt.show()
                D=np.zeros((len(pontonxc2)))
                for j in range (len(pontonxc2)-1):
                    ya=pontonxc2[j]
                    yb=pontonxc2[j+1]
                    xa=j
                    xb=j+1
                    D[j]=calcul_pente(xa, ya, xb, yb)
                I=pontonxc2.index(max(pontonxc2))
                for j in range(I,len(D)):
                    if D[j]<0 :
                        L.append(i+1)
                        break
    return(L)
```

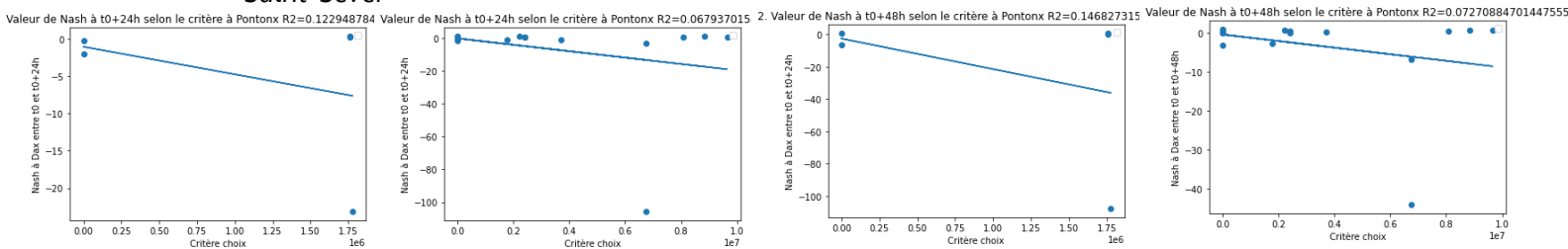
## Annexe B :

# Ensemble des nuages de points des résultats des tableaux 6.2.3 & 6.2.4

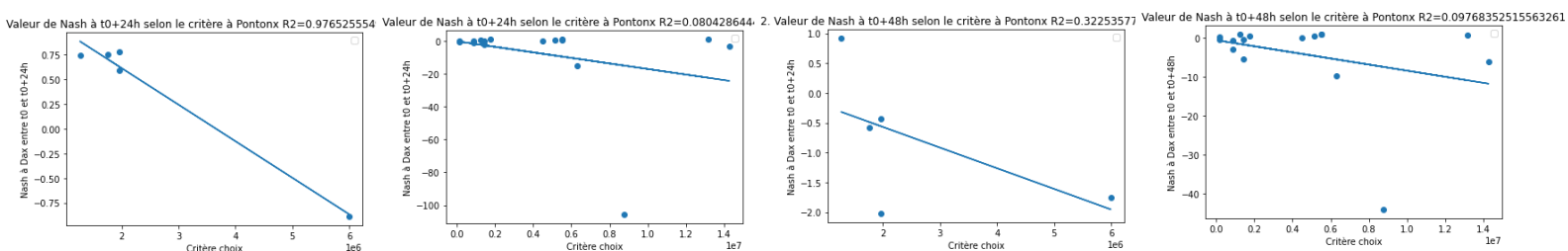
### 1. Nuage de points $t_0+24h$ Hmaxabs et Hmaxrebond, $t_0+48h$ Hmaxabs et Hmaxrebond, critère : Pontonx-sur-l'Adour



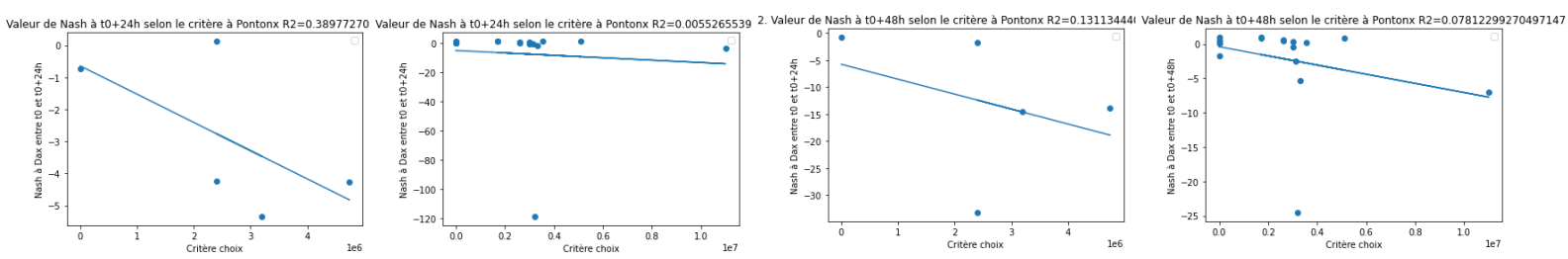
### 2. Nuage de points $t_0+24h$ Hmaxabs et Hmaxrebond, $t_0+48h$ Hmaxabs et Hmaxrebond, critère : Saint-Sever



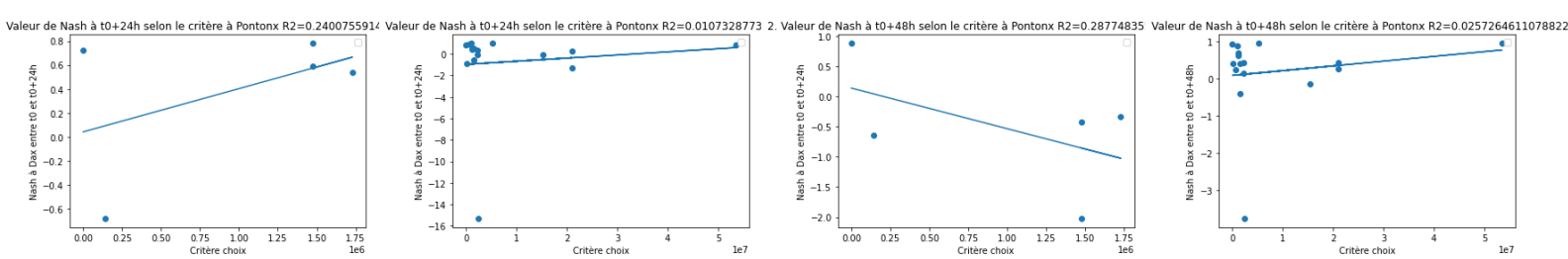
### 3. Nuage de points $t_0+24h$ Hmaxabs et Hmaxrebond, $t_0+48h$ Hmaxabs et Hmaxrebond, critère : Audon



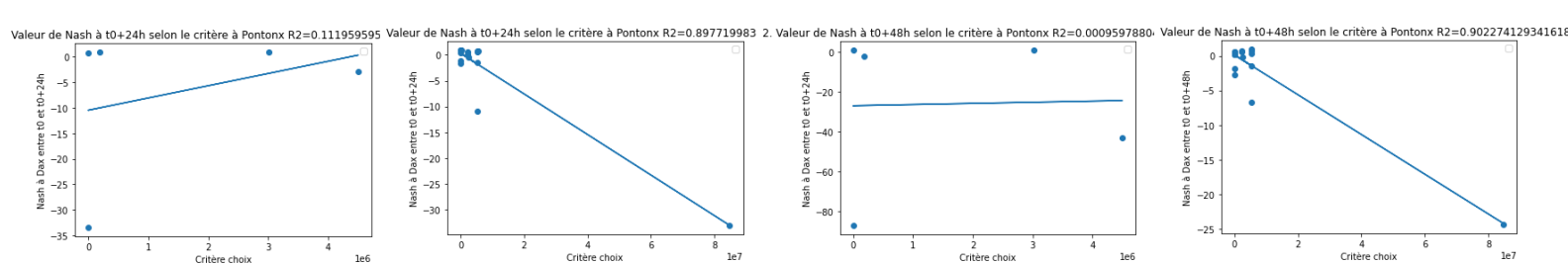
4. Nuage de points  $t_0+24h$  Hmaxabs et Hmaxrebond,  $t_0+48h$  Hmaxabs et Hmaxrebond, critère : Grenade sur l'Adour



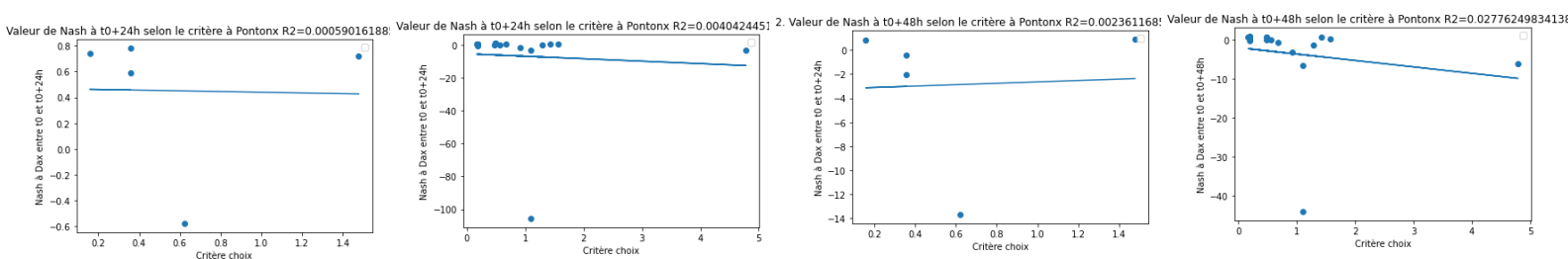
5. Nuage de points  $t_0+24h$  Hmaxabs et Hmaxrebond,  $t_0+48h$  Hmaxabs et Hmaxrebond, critère : Aire-sur-l'Adour



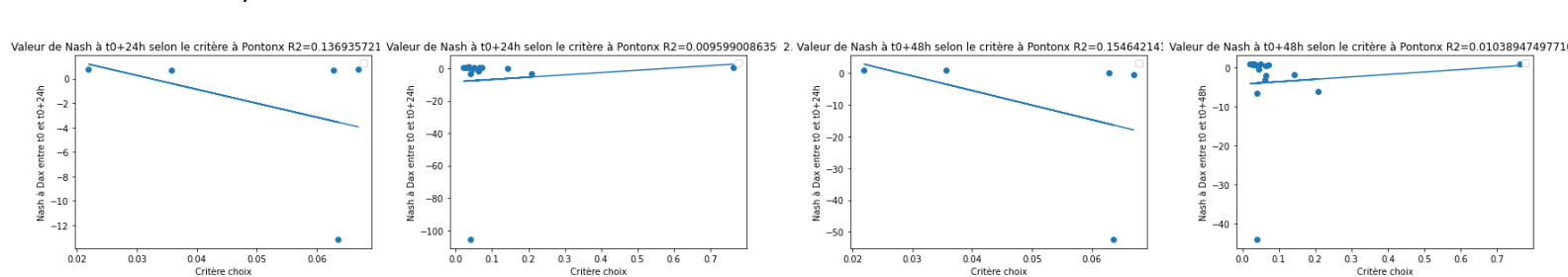
6. Nuage de points  $t_0+24h$  Hmaxabs et Hmaxrebond,  $t_0+48h$  Hmaxabs et Hmaxrebond, critère : Risle



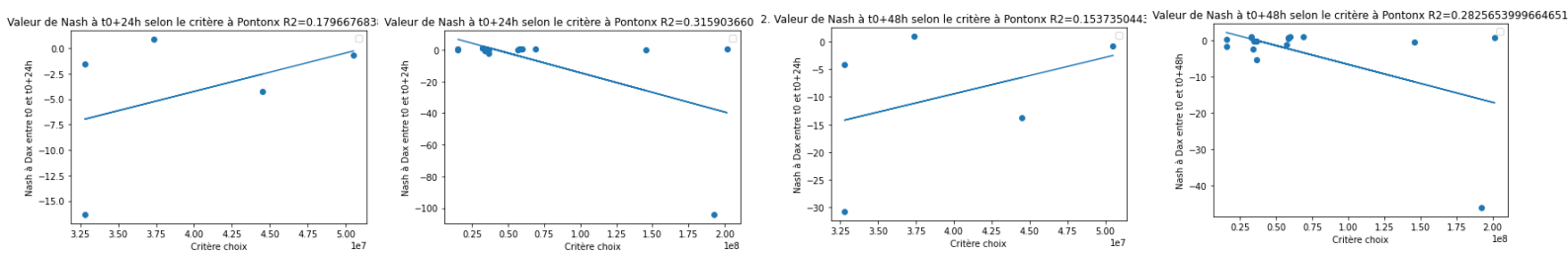
7. Nuage de points  $t_0+24h$  Hmaxabs et Hmaxrebond,  $t_0+48h$  Hmaxabs et Hmaxrebond, critère : hauteur max



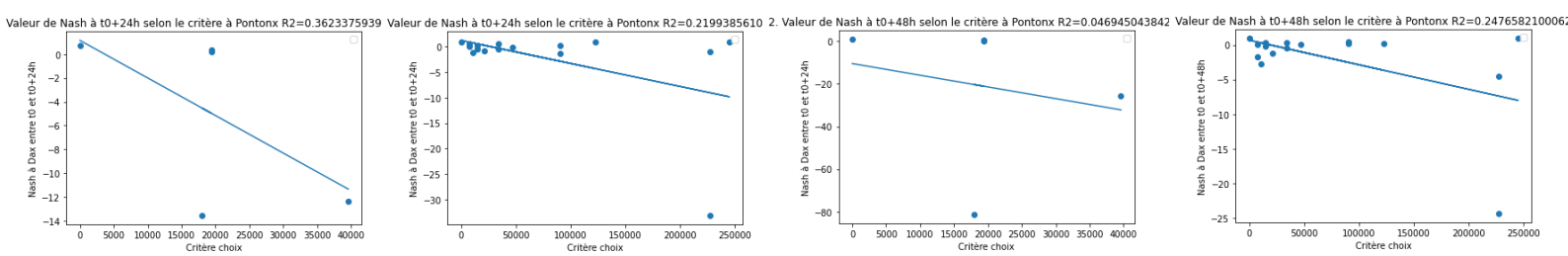
8. Nuage de points  $t_0+24h$  Hmaxabs et Hmaxrebond,  $t_0+48h$  Hmaxabs et Hmaxrebond, critère : pente



9. Nuage de points  $t_0+24h$  Hmaxabs et Hmaxrebond,  $t_0+48h$  Hmaxabs et Hmaxrebond, critère : volume



10. Nuage de points  $t_0+24h$  Hmaxabs et Hmaxrebond,  $t_0+48h$  Hmaxabs et Hmaxrebond, critère : durée



## Annexe C :

# Dispositif de vigilance crues

