HAL open science

Modélisation de la conjonction pluie-niveau marin et prise en compte des incertitudes et de l'impact du changement climatique : application au site du Havre

Amine Ben Daoued

▶ **To cite this version:**

Amine Ben Daoued. Modélisation de la conjonction pluie-niveau marin et prise en compte des incertitudes et de l'impact du changement climatique : application au site du Havre. Autre. Université de Technologie de Compiègne, 2019. Français. NNT : 2019COMP2528 . tel-02952803

HAL Id: tel-02952803

https://theses.hal.science/tel-02952803v1

Submitted on 29 Sep 2020

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.

Par **Amine BEN DAOUED**

Modélisation de la conjonction pluie-niveau marin et prise en compte des incertitudes et de l'impact du changement climatique : application au site du Havre

Thèse présentée pour l'obtention du grade de Docteur de l'UTC

Soutenue le 16 décembre 2019
Spécialité : Modélisation des Systèmes Urbains : Laboratoire Avenues - GSU (EA-7284)

D2528

UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE

Thèse de Doctorat

pour obtenir le titre de

Docteur de l'Université de Technologie de Compiègne
Discipline : Modélisation des Systèmes Urbains

par

Amine BEN DAOUED

Modélisation de la conjonction pluie-niveau marin et prise en compte des incertitudes et de l'impact du changement climatique : application au site du Havre

préparée à l'UTC, équipe AVENUES (EA 7284)
soutenue le 16 décembre 2019

Jury :

— Gaume Éric	Chargé de recherche, IFSTTAR
— Gourbesville Philippe	Professeur des universités, Polytech Nice Sophia
— Benoit Michel	Professeur des universités, IRPHE
— Ducharne Agnès	Directrice de recherche, Sorbonne Université
— Hamdi Yasser	Chercheur, IRSN
— Lefrançois Emmanuel	Professeur des universités, UTC
— Voyneau Nassima	Maître de conférences, UTC
— Sergent Philippe	Ingénieur de recherche, Cerema

À ma grand-mère...

Remerciements

Je souhaite dans un premier temps remercier les membres de mon jury de thèse d'avoir accepté de faire partie de ce comité, d'avoir accordé de leur temps pour relire mon manuscrit de thèse et de leurs avis constructifs sur mes travaux. Je remercie particulièrement Éric Gaume de m'avoir guidé afin de perfectionner ces travaux. J'ai bien apprécié tous les échanges techniques avec lui et j'espère que j'aurai de futures opportunités de travailler avec lui.

Je suis très reconnaissant à mes directeurs de thèse, Nassima Voyneau et Philippe Sergent, de m'avoir supporté pendant ces trois années de thèse. Je les remercie de leur soutien scientifique considérable et de leurs précieux conseils tout au long de cette thèse. Ils m'ont appris le bon sens de la recherche scientifique et m'ont aidé à atteindre mon but de devenir un jeune chercheur. Merci particulièrement à ma codirectrice de thèse Nassima pour sa confiance, sa gentillesse, ses encouragements et son support pendant des moments assez difficiles de ma thèse.

Je remercie mon enseignant, mon encadrant de stage d'ingénieur et mon encadrant de thèse de doctorat, Yasser Hamdi, de son soutien pendant plusieurs années, de m'avoir offert cette une opportunité de mener une thèse et de sa contribution majeure dans cette thèse. Mes remerciements vont également à tous les membres du Bureau d'Expertise en Hydrogéologie, Risques Inondation, météorologiques et Géotechniques et à tous les collègues de l'Institut de Radioprotection et de Sûreté Nucléaire ayant contribué de près ou de loin à la réussite de ce projet de thèse.

J'adresse également mes remerciements au Service Hydrographique et Océanographique de la Marine, à Météo-France et à la Communauté de l'agglomération havraise pour avoir fourni des données nécessaires à l'application de cette thèse.

Merci à tous mes amis et collègues du département Génie des Systèmes Urbains pour leur compréhension et convivialité. Merci à Raphael, Jamila, Wenshai, Dian, Ghada et Solarso.

Mes sincères remerciements vont également à mes parents, mon frère et ma soeur, pour leurs encouragements et soutien durant toute ma vie. Cette réussite vous est dédiée.

Table des matières

.1 Introduction **2**

I Données et variables aléatoires **10**

I.1 Présentation des données et des variables aléatoires de la thèse **12**
I.1.1 Site d'étude 14
I.1.2 Sélection des phénomènes d'inondation 16
I.1.3 Coïncidence des phénomènes d'inondation 17
I.1.4 Définition des variables aléatoires de la démarche 19
I.1.5 Analyse de données de pluie et de niveau marin 20
I.1.5.1 Statistiques descriptives des données 23
I.1.5.2 Corrélation des données 24
I.1.6 Conclusion 25

II Échantillonnage - analyses univariées et multivariées **28**

II.1 Échantillonnage sous l'hypothèse de non-coïncidence **32**
II.1.1 Méthodes d'échantillonnage univariées 35
II.1.1.1 La méthode des Maxima Annuels (MA) 35
II.1.1.2 La méthode POT 36
II.1.1.3 Hypothèses de base 37
II.1.1.4 Adéquation des fonctions de distributions 37
II.1.2 Échantillonnage de l'intensité de pluie et du niveau marin maximal 38
II.1.3 Conclusion 41

II.2 Échantillonnage sous l'hypothèse de coïncidence **43**
II.2.1 Échantillonnage de la combinaison intensité de pluie-niveau marin maximal 45
II.2.1.1 Construction des marginales univariées 45
II.2.1.2 Calculs des probabilités conjointes 48
II.2.2 Conclusion 60

III Approche probabiliste **64**

III.1 Scénarios d'inondation **68**
III.1.1 Construction des scénarios 70
III.1.1.1 Définition d'un scénario 70
III.1.1.2 Cas de non-coïncidence 71
III.1.1.3 Cas de coïncidence 72
III.1.2 Propagation de la source aux points d'intérêt sur le site - modèle hydraulique 72
III.1.2.1 Modèle hydraulique 73
III.1.2.2 Propagation des scénarios 75
III.1.3 Conclusion 79

III.2 Agrégation - courbe d'aléa finale **81**
III.2.1 Méthode d'agrégation 83
III.2.1.1 Cas de non-coïncidence 83
III.2.1.2 Cas de coïncidence 87
III.2.2 Conclusion 95

IV Incertitudes et changement climatique **99**

IV.1 Evaluation et intégration des incertitudes et de l'impact du changement climatique **101**
IV.1.1 Théorie des fonctions de croyance 103
IV.1.1.1 Fondements théoriques des fonctions de croyance 104
IV.1.1.2 Modélisation des informations statistiques 106
IV.1.1.3 Modélisation d'autres informations : avis d'experts 106
IV.1.1.4 Combinaison de fonctions de croyance 106
IV.1.1.5 Propagation d'incertitudes 107
IV.1.1.6 Prise de décision dans l'incertain en utilisant les fonctions de croyance 108
IV.1.2 Quantification et intégration des incertitudes 109
IV.1.2.1 Incertitudes statistiques 109
IV.1.2.2 Incertitudes liées au choix du seuil POT 111
IV.1.2.3 Incertitudes liées au changement climatique 112
IV.1.2.4 Combinaison des incertitudes 116
IV.1.2.5 Propagation des incertitudes et courbes d'aléa 119
IV.1.3 Conclusion 120

Annexes **142**

A Ajustement et choix des copules pour différents déphasages **144**

Table des figures

.1.1 Tempête Xynthia à La Rochelle - Dégâts au Port de Plaisance des Minimes (source : wikimedia.org) 4
.1.2 Inondations à la suite du passage de la tempête Xynthia à Sainte-Marie-de-Ré sur l'île de Ré (Charente-Maritime) (source : wikimedia.org) 6

I.1.1 Site du Havre 14
I.1.2 Vue depuis le haut de l'escalier funiculaire sud au Havre (source : wikimedia.org) 15
I.1.3 Tempête Eleanor touchant la digue Nord du Havre, provoquant des débordements de bassins dans les quartiers de Saint-François et du Commerce de la ville (source : paris-normandie.fr) 16
I.1.4 Débordements à différents endroits du réseau d'évacuation des eaux pluviales au Havre le 10 août 2014 (source : youtube.com/driverhavre) 16
I.1.5 Variables aléatoires de la pluie et du niveau marin ($V1$ et $V2$) sous les hypothèses de non-coïncidence (a) et de coïncidence (b) . 20
I.1.6 Illustration des entrées-sorties d'un modèle hydraulique sous l'hypothèse de non-coïncidence 21
I.1.7 Illustration des entrées-sorties d'un modèle hydraulique sous l'hypothèse de coïncidence 21
I.1.8 Disponibilité des données de pluie et de niveau marin 22
I.1.9 Validation croisée des chroniques de pluies $D1 = 1\ h$ de Cap-De-La-Hève et de Havre-Perret pour la période du 07/12/1994 au 09/12/1994 ; vérification des évènements de hauteur de pluie en commun (a) et première bissectrice (b) 22
I.1.10 Scatter plot des observations de pluie ($D1 = 1\ h$) et de niveau marin maximal (probabilités empiriques calculées à partir des rangs des observations de pluie et de niveau marin à la même date) . . . 25
I.1.11 Chi-plot d'un échantillon d'observations de pluie-niveau marin (à la même date) (a) et de pluie-surcote (à la même date) (b) ; les points représentent les mesures de dépendance obtenues avec la statistique de Chi (χ) en fonction de la distance (λ) des couples (pluie-niveau marin ou pluie-surcote) par rapport à leur médiane . 26

II.1.1 Principales étapes d'une analyse fréquentielle univariée 35
II.1.2 Construction des échantillons d'intensité de pluie $V1$ (gauche) et de niveau marin maximal $V2$ (droite) 38

II.1.3 Graphes de variation du quantile centennal (courbes noires), du paramètre de forme ξ (courbes rouges) et du nombre d'évènements (chiffres en bleu). (a) pour l'intensité de pluie 1h ; (b) pour le niveau marin maximal 39
II.1.4 Ajustement de la GPD à l'intensité de pluie 1 h (a) et au niveau marin maximal (b) 40

II.2.1 Extraction des couples d'intensité de pluie $V1$ et de niveau marin maximal $V2$ afin de construire les marginales univariées (POT) . 45
II.2.2 Principales étapes d'une analyse fréquentielle multivariée 49
II.2.3 Extraction des couples d'intensité de pluie $V1$ et de niveau marin maximal $V2$ afin de construire les marginales univariées (POT) . 57

III.1.1 Scénarios simples de pluie (a) et de niveau marin (b) 71
III.1.2 Étendue de la zone 2D du modèle du Havre 74
III.1.3 Murets de protection autour du quartier Saint-François et point d'intérêt (étoile rouge) 75
III.1.4 Exemple de propagation d'un scénario (ici un scénario simple de niveau marin est présenté) 75
III.1.5 Hauteurs d'eau issues de la propagation des scénarios simples de pluie (a) et de niveau marin (b) 76
III.1.6 Ligne de côte du quartier Saint-François - scénarios de niveau marin 77
III.1.7 Hauteurs d'eau issues de la propagation des scénarios composés de pluie et de niveau marin déphasés de $\delta = -6\ h$ 78

III.2.1 Courbes d'aléa $h = f(T)$ liées à la pluie (a) et au niveau marin (b) 84
III.2.2 Sens d'écoulement de l'eau dans le réseau 1D du quartier Saint-François 84
III.2.3 Courbes d'aléa de pluie (bleue), de niveau marin (verte) et agrégée (noire) pour le cas de non-coïncidence 86
III.2.4 Couples d'intensité de pluie-niveau marin 88
III.2.5 Contours d'iso-hauteurs d'inondation h 88
III.2.6 Calcul de la probabilité totale de dépasser h avec la méthode de l'intégrale 89
III.2.7 Méthode graphique de construction d'une courbe d'aléa à partir des contours d'iso-hauteurs d'eau : les points C_j représentent les couples pluie-niveau marin simulés à travers le modèle hydraulique et donnant lieu à des hauteurs d'eau maximales Z 89
III.2.8 Surfaces liées aux probabilités de non-coïncidence de pluie et de niveau marin 90
III.2.9 Courbe d'aléa pour $\delta = -6\ h$ (même courbe pour le cas cas d'indépendance et de dépendance) 91
III.2.10 Courbes d'aléa pour tous les déphasages (mêmes courbes pour les cas de dépendance et d'indépendance) 91
III.2.11 Agrandissements des courbes d'aléa aux périodes de retour fréquentes (a) et rares (b) 92

III.2.12 Courbe d'aléa finale (noire continue) et courbes d'aléa d'aléa liées aux différents déphasage (superposées) 94
III.2.13 Comparaison des courbes d'aléa finales pour les cas de non-coïncidence (en noir) et de coïncidence (en bleu) ; courbes superposées 94

IV.1.1 Principales théories de modélisation des incertitudes 103
IV.1.2 Bornes d'une famille de fonctions de distributions de probabilités 104
IV.1.3 Courbes de plausibilité pour un quantile centennal de niveau marin (a) et de pluie (b) 110
IV.1.4 Illustration de la méthode de calcul de l'incertitudes liées au seuil POT sur un quantile d'intérêt 111
IV.1.5 Plausibilité de la montée du niveau marin portant sur l'avis du GIEC (bleu) de la NOAA (vert) et leur combinaison (rouge) . . . 115
IV.1.6 Plausibilité (bleu) et croyance (rouge) cumulées pour un niveau marin centennal (a) ; incertitudes statistiques uniquement (courbes discontinues) ; incertitudes statistiques combinées avec les incertitudes liées au seuil POT (courbes en pointillé) ; toutes incertitudes combinées (courbes continues) 117
IV.1.7 Ajustements de la fonction de distribution GPD aux données de niveau marin (a) et de pluie (b) ; ajustements initiaux sans incertitudes (en bleu) ; nouveaux ajustements en intégrant les incertitudes statistiques (discontinu) avec les incertitudes sur le choix du seuil POT (pointillé) et avec les incertitudes sur l'impact du changement climatique à l'horizon de 2100 (noire continue) dans le cadre de la théorie des fonctions de croyance (en choisissant les quantiles médians correspondant à 50% de croyance et 50% de plausibilité) 118
IV.1.8 Courbes d'aléa du niveau marin (vert continue), de la pluie (bleue continue) ; les bornes inférieure et supérieure des intervalles de confiance obtenues à partir de 50% de plausibilité et de 50% de croyance (en bleu et vert transparents) ; la courbe d'aléa agrégée en trait noir continu et les intervalles de confiances agrégés en trait noir discontinu 120

Résumé

La modélisation des combinaisons de phénomènes d'inondation est une problématique d'actualité pour la communauté scientifique qui s'intéresse en priorité aux sites urbains et nucléaires. En effet, il est fort probable que l'approche déterministe explorant un certain nombre de scénarios possède certaines limites car ces scénarios déterministes assurent un conservatisme souvent excessif.
Les approches probabilistes apportent une précision supplémentaire en s'appuyant sur les statistiques et les probabilités pour compléter les approches déterministes. Ces approches probabilistes visent à identifier et à combiner plusieurs scénarios d'aléa possibles pour couvrir plusieurs sources possibles du risque.

L'approche probabiliste d'évaluation de l'aléa inondation (Probabilistic Flood Hazard Assessment ou PFHA) proposée dans cette thèse permet de caractériser une (des) quantité(s) d'intérêt (niveau d'eau, volume, durée d'immersion, etc.) à différents points d'un site en se basant sur les distributions des différents phénomènes de l'aléa inondation ainsi que les caractéristiques du site. Les principales étapes du PFHA sont : i) identification des phénomènes possibles (pluies, niveau marin, vagues, etc.), ii) identification et probabilisation des paramètres associés aux phénomènes d'inondation sélectionnés, iii) propagation de ces phénomènes depuis les sources jusqu'aux point d'intérêt sur le site, iv) construction de courbes d'aléa en agrégeant les contributions des phénomènes d'inondation. Les incertitudes sont un point important de la thèse dans la mesure où elles seront prises en compte dans toutes les étapes de l'approche probabiliste.

Les travaux de cette thèse reposent sur l'étude de la conjonction de la pluie et du niveau marin et apportent une nouvelle méthode de prise en compte du déphasage temporel entre les phénomènes (coïncidence). Un modèle d'agrégation a été développé afin de combiner les contributions des différents phénomènes d'inondation. La question des incertitudes a été étudiée et une méthode reposant sur la théorie des fonctions de croyance a été utilisée car elle présente des avantages divers par rapport aux autres concepts (modélisation fidèle dans les cas d'ignorance totale et de manque d'informations, possibilité de combiner des informations d'origines et de natures différentes, etc.). La méthodologie proposée est appliquée au site du Havre, en France.

Mots-clés — inondation, probabilisation, propagation, agrégation, dépendance, coïncidence, incertitudes, fonctions de croyance

Abstract

The modeling of the combinations of flood hazard phenomena is a current issue for the scientific community which is primarily interested in urban and nuclear sites. Indeed, it is very likely that the deterministic approach exploring several scenarios has certain limits because these deterministic scenarios ensure an often excessive conservatism.
Probabilistic approaches provide additional precision by relying on statistics and probabilities to complement deterministic approaches. These probabilistic approaches aim to identify and combine many possible hazard scenarios to cover many possible sources of risk.

The Probabilistic Flood Hazard Assessment (PFHA) proposed in this thesis allows to characterize a quantity(ies) of interest (water level, volume, duration of immersion, ect.) at different points of interest of a site based on the distributions of the different phenomena of the flood hazard as well as the characteristics of the site. The main steps of the PFHA are: i) screening of the possible phenomena (rainfall, sea level, waves, ect.), ii) identification and probabilization of the parameters representative of the selected flood phenomena, iii) propagation of these phenomena from their sources to the point of interest on the site, iv) construction of hazard curves by aggregating the contributions of the flood phenomena. Uncertainties are an important topic of the thesis insofar as they will be taken into account in all the steps of the probabilistic approach.

The work of this thesis is based on the study of the conjunction of rain and sea level and provide a new method for taking into account the temporal phase shift between the phenomena (coincidence). An aggregation model has been developed to combine the contributions of different flood phenomena. The question of uncertainties has been studied and a method based on the theory of belief functions has been used because it has various advantages (faithful modeling in cases of total ignorance and lack of information, possibility to combine information of different origins and natures, ect.). The proposed methodology is applied on the site of Le Havre in France.

Keywords— flood, probabilization, propagation, aggregation, dependence, coincidence, uncertainties, belief functions

Chapitre .1

Introduction

Contexte et motivation

De nombreux évènements d'inondations relativement importants ont eu lieu dans le passé dans plusieurs pays causant des dommages matériels et humains importants. Les inondations de 1984, de 2000 et de 2003 ont impacté la côte Manche de la France et particulièrement la ville basse du Havre. Une grande inondation a touché l'Europe en 1953 faisant 2500 morts et des dégâts matériels importants. La survenue de la tempête Xynthia en 2010 (figure .1.1), provoquant 47 victimes et 2 milliards d'euros de dommages, est classée parmi les plus meurtrières en France [28, 93, 69]. Ces évènements exceptionnels ont provoqué un débat sur la prise en compte de la combinaison de différents phénomènes d'inondation dans l'évaluation du risque inondation. L'inondation partielle du site du Blayais suite à la tempête Martin en 1999 a été l'élément déclencheur de plusieurs études statistiques portant sur l'aléa inondation (REX Blayais, [81]). Le but a été de compléter les actions à mener sur les sites d'exploitation en termes de protection. Ces études ont été considérablement accentuées suite à Fukushima (Japon, 2011)[8].

Figure .1.1: Tempête Xynthia à La Rochelle - Dégâts au Port de Plaisance des Minimes (source : wikimedia.org)

La probabilité de dépassement et/ou la fréquence d'apparition d'un évènement extrême sont des paramètres clés dans le dimensionnement des structures de protection des sites urbains et nucléaires contre les inondations. En effet, les approches traditionnellement utilisées sont purement déterministes, essentiellement parce qu'elles impliquent un seul scénario (même si ce scénario est basé sur une analyse de valeurs extrêmes). Les modèles déterministes, basés sur des scénarios indépendants des spécificités des sites, peuvent avoir certaines limites lorsqu'il s'agit d'estimer le risque d'une façon complète au niveau des sites. Pour compléter les connaissances acquises à partir de l'analyse traditionnelle, l'approche probabiliste a été identifiée comme étant un outil efficace d'évaluation du risque associé aux aléas et de l'estimation des incertitudes. C'est dans cette optique qu'une approche probabiliste doit être introduite pour compléter les modèles déterministes.

Des approches probabilistes ont déjà été développées dans d'autres domaines tels que la modélisation probabiliste des séismes (Probabilistic Seismic Hazard Assessment ou "PSHA") et des tsunamis (Probabilistic Tsunami Hazard Assessment

ou "PTHA"). Une méthode typique de PSHA est présentée dans [99]. De nombreux programmes de recherche utilisent cette méthode, en particulier les travaux de l'Agence Internationale de l'Énergie Atomique (AIEA) sur la prise en compte du risque sismique dans les études probabilistes de sûreté (EPS) [78], la thèse [12] réalisée à l'Institut de Radioprotection et de Sûreté Nucléaire (IRSN) sur les incertitudes dans une estimation probabiliste de l'aléa sismique et d'autres travaux dans le même contexte peuvent être trouvés dans [65, 66, 23]. Certains travaux ont abordé le PTHA, tels que le projet Tsunamis en Atlantique et MaNche (TANDEM) [75].
En revanche, l'expérience en matière d'évaluation probabiliste de l'aléa inondation (PFHA) est limitée. En 2013, la US Nuclear Regulatory Commission (NRC) a proposé une approche probabiliste de l'aléa inondation mais sans la décliner [2, 1]. Dans un article publié en 2015 [19], Bensi et Kanney ont présenté le cadre d'un PFHA réalisé spécifiquement pour l'aléa submersion marine, appelé Probabilistic Storm Surge Hazard Assessment (PSSHA) pour les centrales nucléaires des États-Unis. La méthode des probabilités conjointes (JPM) a été discutée parmi d'autres méthodes existantes pour une caractérisation probabiliste des tempêtes. D'autres bureaux de recherche se réfèrent également à la JPM pour la caractérisation probabiliste des tempêtes, y compris l'USACE [106]. Le projet européen $ASAMPSA - E$ (2013-2016) [112] visait à définir des bonnes pratiques pour le développement et l'application des EPS. L'étude de l'aléa inondation et son intégration dans un cadre d'évaluation de risque faisaient partie des objectifs de ce projet qui introduit également quelques notions liées à la dépendance entre les phénomènes. Klügel [87] a mis en oeuvre une méthodologie de caractérisation de l'aléa inondation (PFHA) dans le cas d'un site nucléaire fluvial en Suisse, ainsi que le lien entre la modélisation de l'aléa et l'EPS. Certains auteurs utilisent une méthode "Source-Pathway-Receptor-Consequence" (SPRC) pour évaluer le risque d'inondation (par exemple, [101, 77]). Ce modèle SPRC présente un focus sur le lien entre l'aléa et le risque. Une méthode similaire appelée MADS (Méthode d'analyse du dysfonctionnement du système) a été utilisée dans le projet "SAO POLO" pour évaluer le risque de submersion au Havre avec une approche multi-échelles et prise en compte du changement climatique [125]. Un travail de recherche a été effectué à l'IRSN (2014-2016) et a permis d'explorer les approches probabilistes et leur applicabilité à l'aléa inondation [15]. Malgré toutes ces recherches, les travaux dans la littérature sont rares en matière d'études probabilistes bien établies et largement acceptées pour caractériser l'aléa inondation en intégrant tous les phénomènes d'inondation (pluies locales, crue fluviale, submersions marines, etc.), tout en considérant la combinaison et la dépendance de ces phénomènes. Cette thèse de doctorat conduite à l'Université de Technologie de Compiègne (UTC), en collaboration avec l'IRSN, a comme objectif de développer une méthode PFHA tenant compte de la combinaison et de la dépendance des phénomènes d'inondation ([16, 17, 18, 71]). Le PFHA devrait fournir, pour les périodes de retour élevées (1000 ans par exemple), une probabilité totale (ou fréquence) de dépasser une certaine hauteur d'eau sur le site, quels que soient les phénomènes déclenchant l'inondation.

Les phénomènes d'inondation peuvent être caractérisés par plusieurs variables

aléatoires dont certaines peuvent être corrélées. Par exemple, les pluies intenses sont décrites par leur intensité et leur durée dont la corrélation n'est pas négligeable généralement, ou encore, les crues qui sont décrites par les trois variables débit de pointe, durée et volume dont la corrélation n'est pas nulle. Par ailleurs, de nombreuses études ont démontré que l'analyse fréquentielle univariée ne permet pas d'estimer d'une manière complète la probabilité des valeurs extrêmes dans le cas d'un aléa caractérisé par plusieurs variables [29]. Selon Salvadori et De Mechele [119], la modélisation de la dépendance permet une meilleure compréhension de l'aléa et évite de sous/sur-estimer le risque.
D'autre part, certains couples d'aléas comme la pluie et la surcote, peuvent être indépendants mais se produisent en même temps ou avec un décalage temporel. Ce contexte de combinaison des phénomènes d'inondation est rarement étudiée dans la littérature. Seuls quelques travaux utilisent le terme "coïncidence" (e.g., [5]) pour désigner la probabilité que deux phénomènes arrivent en même temps. La notion de décalage entre les deux phénomènes n'est pas considérée. Dans ce manuscrit de thèse, on introduit une nouvelle méthode de prise en compte de la concomitance dans l'étude des aléas tout en conservant le terme anglais de "coïncidence". Cette nouvelle méthode consiste à prendre en compte le décalage de survenue des évènements des phénomènes d'inondation dans le calcul des courbes d'aléa finales. Ce décalage peut avoir un impact non négligeable sur l'ampleur de l'inondation et sur sa probabilité. En effet, la survenue des conjonctions d'aléas naturels peut s'avérer très destructive et peut avoir un impact important parfois même lorsque les deux phénomènes sont de faibles intensité, e.g., l'inondation du site du Blayais a été la conséquence d'une conjonction d'une forte marée et d'une surcote importante causée par des vents extrêmes [81], les départements littoraux de la France on été touchés en 2010 par des inondations meurtrières suite au passage de la tempête Xynthia qui a causé également une conjonction de vents violents avec de fortes marées (figure .1.2).

Figure .1.2: Inondations à la suite du passage de la tempête Xynthia à Sainte-Marie-de-Ré sur l'île de Ré (Charente-Maritime) (source : wikimedia.org)

Ces évènements d'inondation sont de plus en plus fréquents dans le monde. Cela nécessite dans un premier temps la mise en place de stratégies d'aménagement efficaces (e.g., aménagement des terres basses, construction et renforcement

des ouvrages de protection). Outre ce besoin de stratégies, les décideurs ont aussi besoin de connaître et de prendre en compte les impacts du changement climatique afin de pouvoir adapter les mesures d'aménagement. En effet, le réchauffement climatique a un effet direct sur les phénomènes d'inondation. Il induit avant tout une forte remontée du niveau marin et une évolution des caractéristiques des tempêtes (changement de l'intensité et de la fréquence des tempêtes). Cet impact du réchauffement climatique fera que des évènements extrêmes des phénomènes d'inondation (e.g., niveau marin, pluie) deviennent de plus en plus fréquents. La coïncidence de ces phénomènes et leur dépendance sont aussi susceptibles d'évoluer sous l'impact du changement climatique. L'évaluation des incertitudes s'impose également comme un élément essentiel dans l'évaluation de l'aléa inondation.

Objectif

Le but de cette thèse est de mettre en œuvre une approche probabiliste d'évaluation de l'aléa inondation (PFHA). Cette problématique est d'actualité pour la communauté scientifique dans les domaines de l'urbain et du nucléaire car l'approche déterministe explorant un certain nombre de scénario présente certaines limites. En effet, les analyses fréquentielles univariées et multivariées sont des méthodes utilisées pour estimer les probabilités (ou périodes de retour) de dépasser (ou non) les niveaux de retour (intensité de pluie, niveau de la mer, etc.). Ces approches sont très souvent déployées dans des contextes déterministes par les concepteurs et les décideurs, généralement en prenant en compte un seul scénario et un seul phénomène (par exemple, la conception d'un réseau de drainage se fait avec une pluie de période de retour de 10 ans). Cependant, les travaux menés dans cette thèse permettront d'exploiter le savoir-faire utilisé dans les approches déterministes pour la mise en œuvre d'une nouvelle approche probabiliste. Cette approche probabiliste vise à identifier et à combiner tous les scénarios possibles pour couvrir toutes les sources possibles de l'aléa inondation. Ces scénarios sont ensuite propagés à l'aide d'un modèle hydraulique, puis leurs contributions sont agrégées dans le but d'obtenir des courbes d'aléa. Le PFHA permet de caractériser une variable d'intérêt (hauteur d'eau maximale avec la probabilité associée, volume, durée, etc.) à différents endroits d'un site en fonction de la contribution des différents phénomènes d'inondation ainsi que des caractéristiques du site. Les incertitudes forment un point important de la thèse dans la mesure où elles seront identifiées, quantifiées et intégrées dans les courbes d'aléa finales.

Plan de la thèse

Ce mémoire est structuré en quatre parties principales et un total de sept chapitres. Les données et variables aléatoires qui seront utilisées dans ce travail de thèse seront introduits dans le chapitre I.1. D'abord, le site d'étude du Havre et les phénomènes d'inondation (pluie et niveau) auxquels ce site est exposé seront présentés. Puis, nous introduirons les différentes hypothèses de coïncidence et de

dépendance ainsi que les variables aléatoires associées aux phénomènes d'inondation retenus. Ensuite, nous présenterons les données de pluie et de niveau marin ainsi que leurs caractéristiques statistiques et nous discuterons la corrélation de ces données. Nous montrerons en particulier que cette corrélation est très faible.

Les chapitres II.1 et II.2 seront dédiées aux différentes approches d'échantillonnage utilisées sous les hypothèses de non-coïncidence et de coïncidence des variables aléatoires. Dans le chapitre II.1, nous présenterons le concept d'une analyse fréquentielle univariée. D'abord, les fondements mathématiques des différentes méthodes d'échantillonnage univarié (Peaks-Over-Threshold ou "POT" et Maxima Annuels ou "MA") seront détaillés. Nous retiendrons la méthode POT car elle permet d'améliorer la taille de l'échantillon. Ensuite, nous montrerons les échantillons obtenus avant de présenter le test d'adéquation utilisé ainsi que les tests des hypothèses de base (homogénïté, indépendance et stationnarité).
Le chapitre II.2 détaillera la méthode d'échantillonnage sous l'hypothèse de coïncidence tout en montrant les deux étapes de modélisation des marginales univariées et de calculs des probabilités conjointes en cas d'indépendance et de dépendance des variables aléatoires.

Les chapitres III.1 et III.2 seront consacrés à la méthode probabiliste (PFHA) développée dans ce travail de thèse. Le chapitre III.1 présentera la méthode de construction des scénarios de pluie et de niveau marin, leur propagation de la source au point d'intérêt en utilisant le modèle hydraulique de la ville du Havre et la construction des courbes d'aléa. Le chapitre III.2 présentera la méthode d'agrégation de ces courbes d'aléa dans le cas de non-coïncidence et de coïncidence. Nous verrons à travers les résultats l'importance de la prise en compte de la notion de coïncidence surtout pour les fortes périodes de retour.

Le chapitre IV.1 sera dédié aux incertitudes et à l'impact du changement climatique. D'abord, le cadre théorique des fonctions de croyance sera présenté vu ses avantages divers (modélisation fidèle des informations incertaines, possibilité de combiner des informations, etc.). Ensuite, cette théorie sera appliquée en combinant des incertitudes d'origines et de natures différentes (incertitudes aléatoires d'origine statistiques, incertitudes épistémiques sur un paramètre de la fonction de distribution Général Pareto ou "GPD" et incertitudes épistémiques liées à l'impact du changement climatique).

Première partie

Données et variables aléatoires

Chapitre I.1

Présentation des données et des variables aléatoires de la thèse

Introduction

La connaissance des phénomènes initiateurs permet de mieux comprendre la physique des phénomènes d'inondation. Un phénomène initiateur peut être représenté par l'ensemble des conditions météorologiques (pression atmosphérique, vent, température, etc.) à l'origine des phénomènes d'inondation (niveau marin, pluie, etc.). Dans le présent chapitre, on s'intéresse à des phénomènes d'inondation dont les variables sont mesurables (e.g. hauteur de pluie, niveau marin). La sélection de ces phénomènes d'inondation et leur nombre dépend principalement des caractéristiques géographiques et morphologiques du site ainsi que des risques auxquels il est soumis. C'est pour cette raison que le choix de présenter le site d'étude en premier a été fait.

Une première section de ce chapitre sera dédiée à la présentation du site d'étude. La sélection des phénomènes d'inondation et la présentation des hypothèses de non-coïncidence et de coïncidence feront l'objet des sections 2 et 3. La définition des variables aléatoires et l'analyse des données seront présentées en section 4 et 5.

I.1.1 Site d'étude

La ville du Havre est une ville urbaine du département de la Seine-Maritime, sur la côte Française de la Manche en Normandie. C'est une grande ville Française située au nord-ouest de la France (17 communes – environ 250 000 habitants). La carte de la figure I.1.1 montre la situation de la ville du Havre. Le nom de la ville du Havre signifie "le port". Ce dernier est un des plus grands ports de France. Pour ces raisons, la ville du Havre reste profondément influencée par ses traditions maritimes.

Figure I.1.1: Site du Havre

En raison de sa situation sur la côte de la Manche, le climat du Havre est tempéré océanique. Les jours sans vent sont rares et la ville subit des influences

maritimes tout au long de l'année. D'après les relevés météorologiques, les précipitations sont réparties tout au long de l'année, avec un maximum en automne et en hiver. Les mois de juin et juillet sont marqués par quelques tempêtes relativement extrêmes en moyenne 2 jours par mois. Une des caractéristiques de la région est la grande variabilité de la température, même pendant la journée. Les vents dominants sont du nord-nord/est pour les brises et du secteur sud-ouest pour les vents forts.

Figure I.1.2: Vue depuis le haut de l'escalier funiculaire sud au Havre (source : wikimedia.org)

La ville du Havre est traversée de l'Est à l'Ouest par une falaise "morte" qui marque une séparation ancienne entre la ville haute et la ville basse (figure I.1.2). La ville basse est donc construite dans un ancien espace intertidal, c'est-à-dire un espace de propagation des ondes de marées, où le niveau marin évolue entre marée haute et marée basse. Cette ville dite "basse" a été le théâtre de développements divers et variés : canalisation du cours d'eau, remblais, construction de bassins, etc. Cependant, cette ville basse et son port restent naturellement soumis au risque de submersion marine [42, 125]. Plusieurs évènements d'inondations relativement importants ont touché la côte Manche de la France. Certains de ces évènements sont indiqués dans l'archive municipale de la ville du Havre. D'autres inondations de la ville basse du Havre ont eu lieu en juin 1982 suite à des précipitations torrentielles, provoquant des sinistres importants dans plusieurs quartiers et un important glissement de terrain sur une falaise de la ville. D'importantes inondations ont également été provoquées par les précipitations torrentielles d'un ouragan qui a dévasté le nord de la France en décembre 1999. D'autres évènements d'inondation ont eu lieu suite à des tempêtes, notamment la tempête Xynthia en 2010 et Eleanor en 2018 qui ont provoqué beaucoup de dommages dans la ville du Havre (e.g., figures I.1.3 et I.1.4). De ce fait, la ville du Havre a été choisie pour servir de cas d'étude dans le cadre de cette thèse. D'autant plus qu'elle est soumise à des submersions marines et des instabilités des falaises côtières [53, 54, 96]. En particulier, la partie basse de la ville (Quartier Saint-François) est susceptible

d'être inondée par des inondations marines et pluviales.

Figure I.1.3: Tempête Eleanor touchant la digue Nord du Havre, provoquant des débordements de bassins dans les quartiers de Saint-François et du Commerce de la ville (source : paris-normandie.fr)

Figure I.1.4: Débordements à différents endroits du réseau d'évacuation des eaux pluviales au Havre le 10 août 2014 (source : youtube.com/driverhavre)

I.1.2 Sélection des phénomènes d'inondation

Une étape préliminaire de la démarche consiste à identifier et sélectionner des phénomènes pressentis pour l'étude de l'aléa inondation. Cela repose sur un processus de "screening" qui vise à identifier les phénomènes qui ont une contribution non-négligeable au niveau de l'aléa.

Dans la suite de ce mémoire et pour des raisons de généralisation de cette méthode, on retiendra deux phénomènes notés ϕ_1 et ϕ_2. Dans l'application de cette méthode, ϕ_1 et ϕ_2 représentent respectivement la pluie et le niveau marin.

I.1.3 Coïncidence des phénomènes d'inondation

La coïncidence est définie par la chance d'occurrence de deux phénomènes en même temps ou avec un décalage temporel. Cette coïncidence n'implique aucune dépendance entre les phénomènes. Les travaux sur ce sujet dans la littérature sont rares.
Apel et al. [5] distinguent la dépendance et la coïncidence dans la combinaison des crues fluviales et pluviales. La dépendance implique une relation fonctionnelle ou de causalité entre les deux phénomènes alors que la coïncidence ne fait intervenir aucune relation de dépendance. Si l'indépendance entre les crues fluviales et pluviales est supposée, la coïncidence des deux phénomènes est prise en compte dans le calcul des probabilités de coïncidence. La probabilité conjointe d'occurrence des évènements de crues fluviales et pluviales est calculée par le produit des probabilités individuelles multiplié par la probabilité de coïncidence i.e., dans ce contexte, la probabilité de coïncidence est la chance que deux évènements de crue fluviale et pluviale se produisent à la même saison e.g., dans [5] la probabilité de coïncidence des évènements de crues fluviales et pluviales est calculée à partir de la longueur de la saison d'inondation et des durées des évènements de crues fluviales.

Quelques travaux dans la littérature ont traité cette question de coïncidence dans le cas de dépendance et dans un contexte de sélection des combinaisons (couples) en particulier. Chebana and Ouarda [29] ont proposé la décomposition d'un contour d'isoprobabilité en partie "naïve" (queue de distribution) et partie "propre" (centrale), ils ont ensuite proposé de sélectionner uniquement les couples formant les extrémités de la partie propre. Salvadori et al. [120] ont proposé d'introduire une fonction liée au dimensionnement (i.e., courbe de fragilité d'un ouvrage) et de sélectionner un seul couple du contour bivarié (combinaison) permettant de minimiser cette fonction, identiquement à la méthode appliquée dans le projet "SAO POLO" [125]. Dung et al [52] ont utilisé le concept de la densité de coïncidence dans le cas de dépendance introduit par Volpi and Fiori [146]. Les couples (débits, volumes) sur les contours bivariés possèdent la même probabilité de non dépassement ou de dépassement p. Mais, la probabilité de coïncidence des couples sur un même contour n'est pas la même i.e., la vraisemblance qu'une valeur de débit extrême survienne en même temps qu'un petit volume est faible (vice-versa). La méthode d'estimation de cette densité de coïncidence proposée par Volpi and Fiori [146] consiste à définir les probabilités d'occurrence de couples bivariés en coordonnées curvilignes. Les points d'intersection des contours bivariés avec la droite de dépendance totale (première bissectrice), appelés sommets, définissent l'origine de ce système de coordonnées et constituent les combinaisons les plus probables des contours d'isoprobabilité. Le calcul des probabilités de coïncidence est réalisé par la résolution numérique d'équations implicites décrivant le contour d'isoprobabilités. Dans d'autres travaux la distinction entre la probabilité conjointe de dépassement (ou de non-dépassement) et la probabilité de coïncidence n'a pas été faite, e.g., Baowei et Lu [158] définissent la probabilité de coïncidence des précipitations comme étant la probabilité de combinaison calculée à travers une analyse fréquentielle trivariée permettant d'estimer la distribution de probabilités conjointes des variables étudiées.

Dans cette thèse, une nouvelle méthode est proposée pour prendre en compte la coïncidence des phénomènes d'inondation indépendants ou dépendants dans un contexte de PFHA.
Dans un premier temps, la corrélation des VAs $V1$ et $V2$ est évaluée. Cette corrélation peut être quantifiée par la technique de Chi-plot [3] ou par des estimateurs non-paramétriques (par exemple, le coefficient de dépendance de queue) [126]. Cela permettra ensuite de décider de modéliser la dépendance des VAs $V1$ et $V2$ en utilisant la théorie des copules [129] (si $V1$ et $V2$ sont dépendants) et de ne considérer que les fonctions de distribution univariés en cas d'indépendance. L'étape suivante dans la prise en compte de la coïncidence consiste à considérer n déphasages δ équiprobables, définis entre les pics des VAs $V1$ et $V2$. Cela suppose de considérer l'une des VAs comme référence, disons $V2$. Pour chaque déphasage, plusieurs combinaisons $(V1, V2)$ sont considérées. Les probabilités de dépassement conjointes de ces combinaisons sont estimées selon que les deux variables sont indépendantes ou dépendantes. Ceci sera détaillé dans les sections suivantes.

Le cas de non-coïncidence est présenté dans ce mémoire à titre de référence permettant d'interpréter les résultats obtenus avec le cas de coïncidence. Comme on le verra plus loin, le cas de non-coïncidence est déjà inclus dans le cas de coïncidence. Les sections suivantes dans ce chapitre montrent comment ces probabilités conjointes sont calculées dans les cas d'indépendance et de dépendance. En outre, plus de détails sur la méthode de construction des scénarios seront présentés plus loin dans le chapitre III.1.

Trois ensembles d'hypothèses de complexités croissantes ont été considérés pour évaluer les effets combinés de deux variables explicatives (e.g., intensité de pluie, niveau main maximal) sur la distribution statistique de la quantité d'intérêt (e.g., hauteur d'eau) :

- 1) Les niveaux de mer et les intensité de pluies susceptibles de générer une inondation ne se produisent jamais ensemble (hypothèse de non-coïncidence) ;
- 2) Ils sont susceptibles de se produire au cours du même cycle de marée - la probabilité de cette coïncidence ainsi que le déphasage temporel entre le moment où le niveau maximal est observé et l'heure de pluie intense devront être considérés pour évaluer l'effet combiné des deux phénomènes. Les deux variables aléatoires sont indépendantes ;
- 3) Idem cas 2, mais la dépendance éventuelle entre les deux variables aléatoires est prise en compte.

L'inondation du site peut être provoquée par l'un des deux phénomènes ϕ_1 ou ϕ_2 indépendamment et séparément (Cas 1) ou par ces deux phénomènes lorsqu'ils sont indépendants mais arrivent en même temps (ou avec un déphasage temporel) (Cas 2) ou lorsqu'ils sont dépendants et surviennent en même temps (Cas 3). La démarche considérant à la fois ces trois cas de figures est un apport important de cette thèse, car cela vise à prendre en compte toutes les origines d'inondation tout en considérant l'éventuelle dépendance et coïncidence pour estimer des probabilités totales d'inondation.
Il est important de noter que, pour un site donnée, deux phénomènes d'inondation ne peuvent pas être dépendants et indépendants à la fois. Il est donc nécessaire de réaliser au préalable une étude fine de la dépendance entre les différents phéno-

mènes retenus. Toutefois, le cas de non-coïncidence peut-être inclus dans le cas de coïncidence (une pluie nulle peut bien survenir avec un niveau marin ne donnant pas lieu à des déversements, et vice-versa). En effet, le premier cas est tout de même traité à part dans ce rapport car il servira de repère pour les deux autres cas.
Chacune des hypothèses de non-coïncidence et de coïncidence nécessite une étape d'échantillonnage différente qui sera détaillée les deux prochains chapitres.

I.1.4 Définition des variables aléatoires de la démarche

On définit une variable aléatoire (VA) comme étant une entité décrivant un phénomène aléatoire. Une VA X est une fonction, souvent à valeurs réelles, définie sur un espace échantillonal. Cet espace correspond à l'ensemble des valeurs possibles d'un phénomène aléatoire. En hydrologie, une VA peut représenter l'intensité d'une pluie, la durée d'une pluie, la hauteur maximale du niveau marin, les hauteurs d'eau produites en un point d'intérêt sur un site, la durée d'inondation de ce point d'intérêt, etc.

Par convention, les VAs sont désignées par des lettres majuscules (X, Y, Z, etc.) et les réalisations de ces VAs sont désignées par des lettres minuscules (x, y, z, etc.). On désigne par F_X et f_X les fonctions de répartition et de densité d'une VA X, respectivement.

$$\begin{aligned} F_X(x) &= Pr(X \leq x) \\ &= Pr(\text{VA X prenne une valeur inférieure ou égale à x}) \\ &= \int_{-\infty}^{x} f_X(y)dy \end{aligned} \tag{I.1.1}$$

Dans la suite, on considère que les phénomènes d'inondation ϕ_1 et ϕ_2 sont décrits chacun par une VA d'intensité ($V1$ et $V2$) et une VA caractérisant la durée du phénomène ($D1$ et $D2$).

Dans le cas de non-coïncidence, les VAs $V1$ et $V2$ représentent les intensités de pluies horaires et les niveaux marins maximums, respectivement (figure I.1.5 (a)). Tandis que dans le cas de coïncidence, $V2$ est le niveau marin maximal et $V1$ est l'intensité de pluie qui peut être déphasée par rapport au moment de la maximale (figure I.1.5 (b)). Plusieurs échantillons de pluie peuvent alors être constitués pour des déphasages différents.

$D1$ et $D2$ représentent les durées de ces deux phénomènes (pluie et niveau marin respectivement). La variable δ caractérise le déphasage temporel entre les deux pics de la pluie et du niveau marin.
Dans la suite de cette méthode, les VAs $D1$, $D2$ et δ seront considérées déterministes ; $D1 = 1\ h$ (durée très proche du temps de concentration du réseau d'assainissement du site du Havre, estimé entre 30 et 45 minutes à travers des simulations effectuées dans cette thèse), $D2 = 12\ h$ (à peu près la durée d'un cycle de marée) et $\delta = \{-6\ h, -5\ h, -4\ h, -3\ h, -2\ h, -1\ h, 0\ h, +1\ h, +2\ h,$

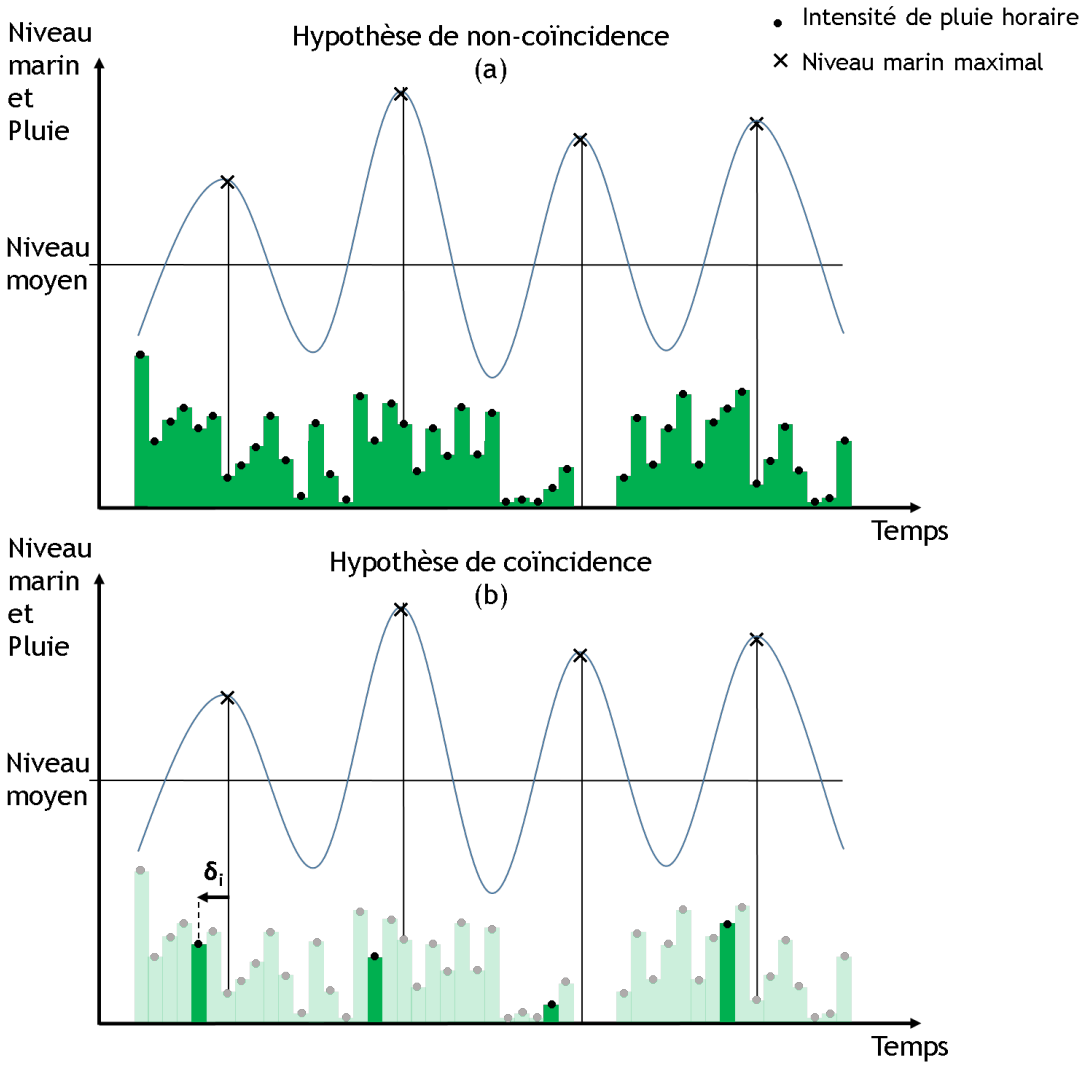


Figure I.1.5: Variables aléatoires de la pluie et du niveau marin ($V1$ et $V2$) sous les hypothèses de non-coïncidence (a) et de coïncidence (b)

$+3\ h, +4\ h, +5\ h, +6\ h\ \}$) et ne seront donc pas probabilisées.
Les VAs $X1$ et $X2$ décrivent les hauteurs d'eau maximales obtenues en propageant séparément des scénarios simples définis par les VAs d'intensités $V1$ et $V2$ ayant pour durées $D1$ et $D2$ (figure I.1.6). La VA Z décrit quant à elle les hauteurs d'eau maximales provenant de la propagation de scénarios combinés et formés par des couples $(V1, V2)$ de durées $D1$ et $D2$ (figure I.1.7). La notion de scénario sera présentée plus loin dans ce rapport.

I.1.5 Analyse de données de pluie et de niveau marin

Les niveaux marins observés enregistrés entre 1971 et 2015 au port du Havre sont fournis par le Service Océanographique Français (SHOM - Service Hydro-

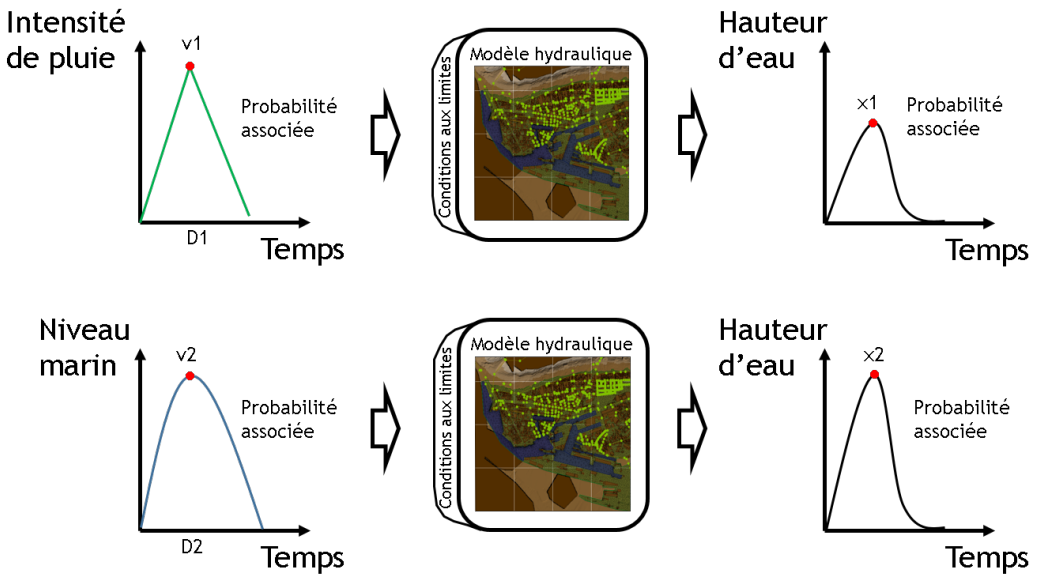


Figure I.1.6: Illustration des entrées-sorties d'un modèle hydraulique sous l'hypothèse de non-coïncidence

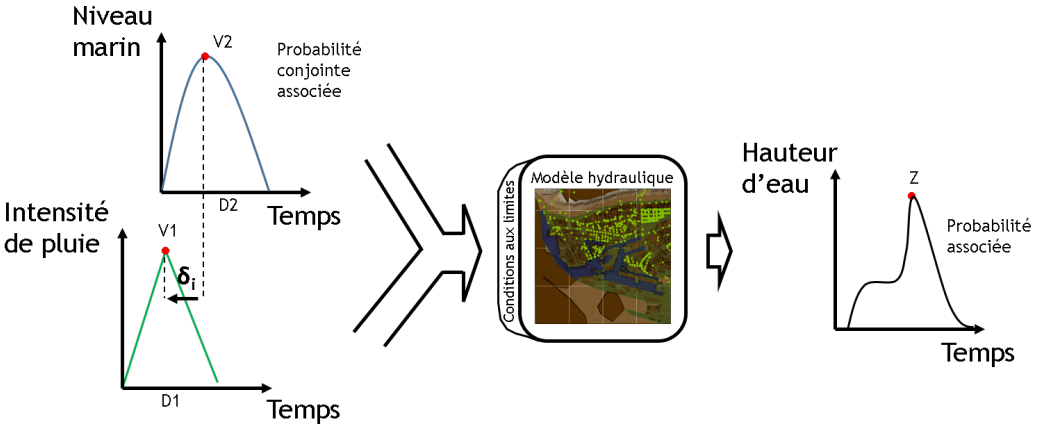


Figure I.1.7: Illustration des entrées-sorties d'un modèle hydraulique sous l'hypothèse de coïncidence

graphique et Océanographique de la Marine). Les différentes séries de données de pluies proviennent de Météo-France. Les caractéristiques des données de pluies ($D1 = 1\ h$) et les observations de niveau marin sont présentées en figure I.1.8. Ces séries contiennent des données manquantes, le tableau I.1.1 récapitule les différentes durées d'oservation.

	Période d'observation totale (années)		Période d'observation commune (années)	
	Taille initiale	Durée effective	Taille initiale	Durée effective
Niveau marin	79.03	46.80	10.67	10.67
Pluie	10.67	10.54	10.67	10.54

Tableau I.1.1: Durées d'observations pour les pluies et les niveaux marins

La chronique de données brutes de pluie pour la durée $D1 = 1\ h$ est issue de trois stations différentes (figure I.1.8). Une étape de validation croisée a été réalisée avec les observations des trois stations. Cette validation croisée consiste à vérifier,

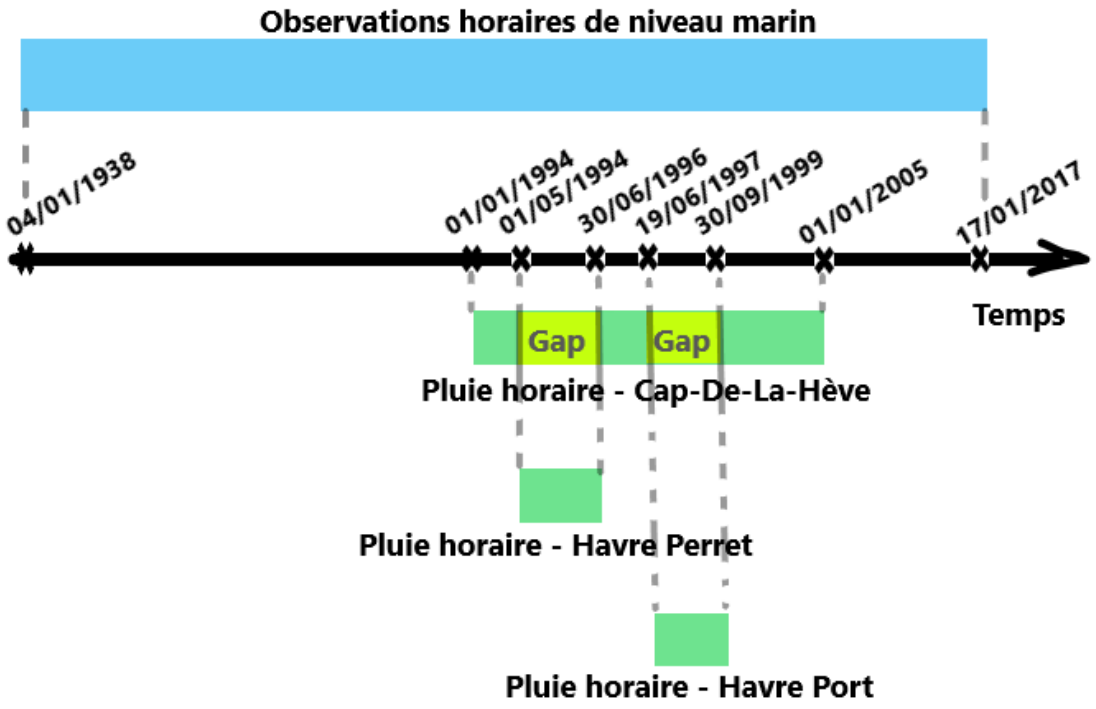


Figure I.1.8: Disponibilité des données de pluie et de niveau marin

à travers certains évènements en commun préalablement repérés, l'homogénéité et la cohérence de ces chroniques.

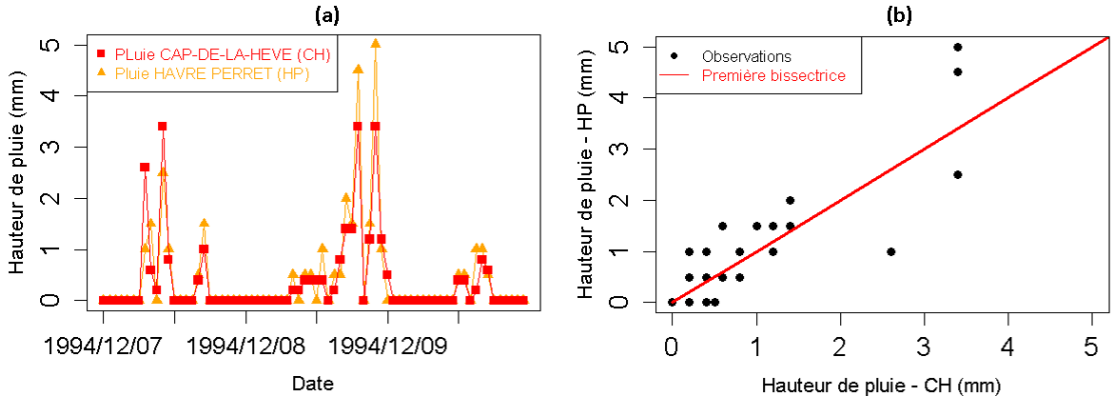


Figure I.1.9: Validation croisée des chroniques de pluies $D1 = 1\ h$ de Cap-De-La-Hève et de Havre-Perret pour la période du 07/12/1994 au 09/12/1994 ; vérification des évènements de hauteur de pluie en commun (a) et première bissectrice (b)

La figure I.1.9 (a) montre deux pics de hauteurs de pluie, un pic qui a eu lieu le 07/12/1994 et un autre le 09/12/1994. Ces deux pics correspondent bien à des évènements indépendants car ils sont séparés de 2 jours (avec l'hypothèse que ce délai de 2 jours soit suffisant pour séparer deux évènements indépendants). On remarque bien que les deux courbes (rouge et orange), correspondant aux stations Cap-De-La-Hève et Havre-Perret, enregistrent ces deux évènements le même jour, à une ou deux heures près de décalage. Ce décalage est lié à la distance géogra-

phique qui sépare les deux stations. Les données des deux stations semblent donc homogènes.
La figure I.1.9 (b) représente les observations de hauteurs de pluie pour les trois jours (du 07 au 09/12/1994). Le principe de ce graphe est simple, il faut que la première bissectrice (présentée en rouge) ajuste bien le nuage de points (points noirs) pour vérifier l'homogénéité des données. Le nuage de points, pour les trois jours représentés, est plus ou moins bien ajusté par la droite rouge. Cela confirme donc l'homogénéité des données des deux stations.
Une analyse identique a été menée entre les stations de Cap-De-La-Hève et Havre-port et a aussi révélé l'homogénéité des données de ces deux stations. Un échantillon mixte avec des données provenant des trois stations peut donc être construit en vue d'obtenir une série de données plus complète.

I.1.5.1 Statistiques descriptives des données

L'objectif de cette section est : (i) d'évaluer les caractéristiques de position (mode, valeurs minimale et maximale, moyenne, le premier quartile, la médiane et le troisième quartile). (ii) d'évaluer les caractéristiques de dispersion (étendue, intervalle interquartile et variance). Ce sont les grandeurs le plus souvent utilisées en statistique. L'étendue représente la différence entre les observations extrêmes. L'intervalle interquartile contient 50% de l'échantillon en éliminant 25% à chaque extrémité et il s'agit d'une caractéristique qui est nettement représentative que l'étendue. La variance est la caractéristique de dispersion la plus utilisée, elle indique de quelle manière notre variable aléatoire se disperse autour de sa moyenne ; (iii) de visualiser les séries temporelles.
Le tableau I.1.2 récapitule les statistiques descriptives des données de pluie et de niveau marin.

Statistiques		Pluie (mm/h)	Niveau marin (mCM (Côte Marine))
Position	Min	0.00	0.00
	1er Quantile	0.00	3.06
	Médiane	0.00	5.11
	Moyenne	0.10	4.91
	3ème Quantile	0.00	6.77
	Max	35.00	9.15
Dispersion	Étendue	35.00	9.15
	Intervalle interquarile	0.00	3.71
	Variance	0.32	4.16

Tableau I.1.2: Statistiques descriptives des séries de données de la pluie et du niveau marin

L'étendue est une mesure absolue de la dispersion des valeurs d'une VA (différence entre la valeur maximale et la valeur minimale). L'étendue des intensités de pluie est plus grande que celle du niveau marin. Ce critère ne permet pas de comparer la dispersion des intensités de pluie et celle du niveau marin car les distributions peuvent comporter des valeurs exceptionnelles (tel est le cas des intensités de pluie). À contrario, l'intervalle interquartille est un paramètre de dispersion absolue qui correspond à l'étendue de la distribution après avoir enlevé les 25% des valeurs les plus faibles et les 25% des valeurs les plus fortes, permettant ainsi de comparer deux distributions différentes. L'intervalle interquartile montre que 50% des intensités de pluies ont une valeur égale à 0 et que 50% des valeurs de niveau marin sont concentrées entre 3.06 mCM (1er quartile) et 6.77 mCM, c'est-à-dire sur un intervalle de 3.71 mCM. La distribution est donc plus forte pour le niveau marin que pour les intensités de pluie.

I.1.5.2 Corrélation des données

La corrélation statistique entre les pluies intenses et les surcotes extrêmes a fait l'objet de quelques travaux dans la littérature [111, 36, 74, 154] et plus récemment [94, 162, 165]. Svensson et Jones [140, 141] ont montré que la dépendance entre les pluies intenses et les surcotes est statistiquement significative plaidant ainsi la nécessité de sa prise en compte dans l'évaluation du risque d'inondation.

La corrélation globale entre les observations de niveau marin maximal et de pluies correspondante de durée $D1 = 1h$ sur la période commune (01/01/1994-01/01/2005) a d'abord été analysée. Il est important de signaler que la corrélation du niveau marin maximal et des pluie de durée $D1 = 6$ minutes pourrait être différente de celle évaluée avec la durée horaire de la pluie. Toutefois, dans l'application

de cette méthode, la durée de pluie $D1 = 1\ h$ est choisie car elle est très proche du temps de concentration du réseau d'assainissement du site du Havre (estimé entre 30 et 45 minutes à travers des simulations effectuées dans cette thèse). Des détails supplémentaires sont présentés plus loin dans ce chapitre.

La figure I.1.10 présente un "scatter plot" avec l'ensemble des couples $(V1, V2)$ observés. Il s'agit d'associer une probabilité empirique (portant sur les rangs des observations) aux couples de pluie et de niveau marin observés à la même date. Cette figure permet de déceler une certaine forme de dépendance dans le nuage des points (plus dense au niveau des valeurs extrêmes). Toutefois cette technique de "scatter plot" peut s'avérer subjective et ne permet pas de quantifier la dépendance. La technique du Chi-plot [58] a été utilisée pour détecter la dépendance entre les combinaisons de niveau marin et de pluie ($D1 = 1\ h$). Cette technique permet d'observer la variation de la dépendance (mesurée par la statistique de χ) en fonction de la distance (λ) des couples par rapport à leur médiane.

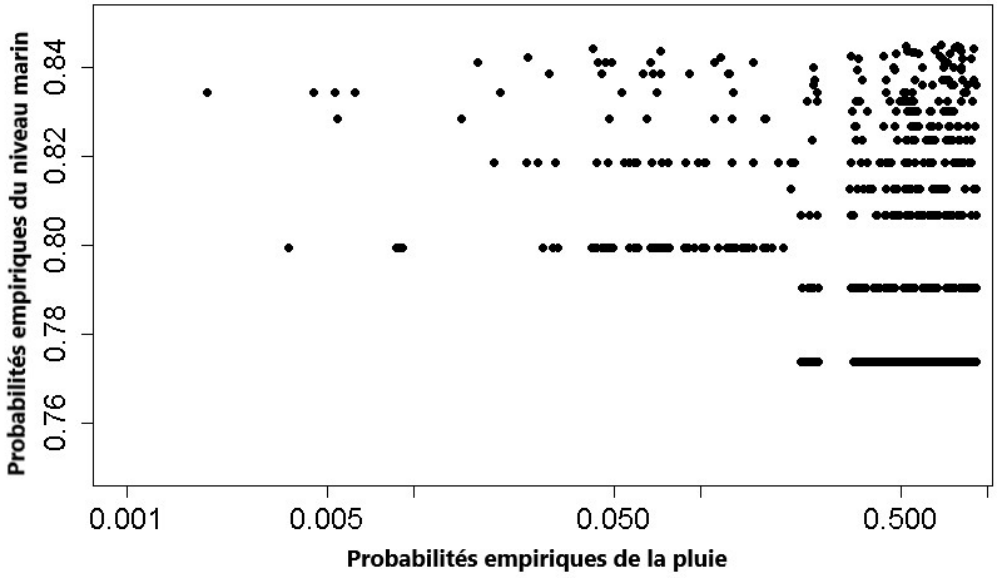


Figure I.1.10: Scatter plot des observations de pluie ($D1 = 1\ h$) et de niveau marin maximal (probabilités empiriques calculées à partir des rangs des observations de pluie et de niveau marin à la même date)

On note que la corrélation entre le niveau marin maximal et la pluie correspondante est quasi-nulle, elle est plus marquée entre la surcote et la pluie (figure I.1.11). Ce constat semble être cohérent avec les travaux de la littérature [141, 122, 166] qui avancent que la corrélation est mieux observable avec la surcote qu'avec le niveau marin. En effet, le signal de niveau marin contient la composante marée qui est complètement déterministe et indépendant de tout effet météorologique, et c'est à cause de cette composante que la corrélation est atténuée.

I.1.6 Conclusion

Le site du Havre a été d'abord présenté dans ce chapitre. Le choix de ce site est justifié par la configuration topographique de la ville, son exposition à des

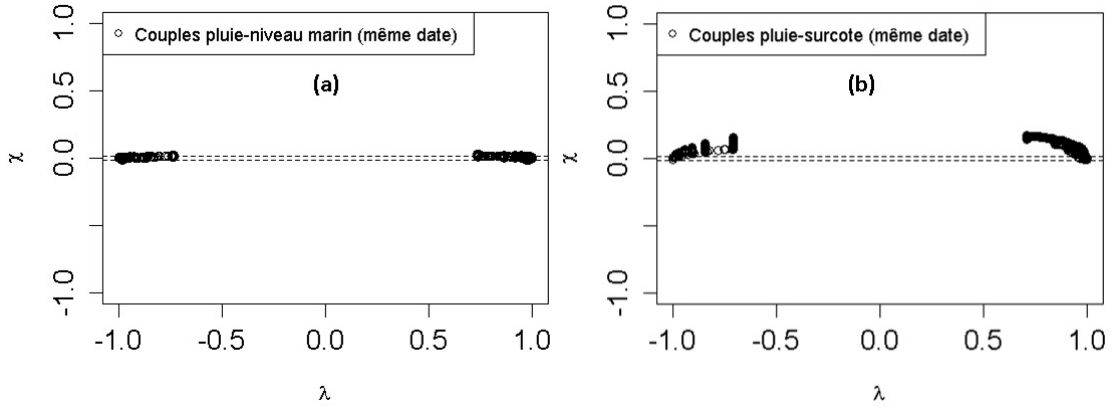


Figure I.1.11: Chi-plot d'un échantillon d'observations de pluie-niveau marin (à la même date) (a) et de pluie-surcote (à la même date) (b) ; les points représentent les mesures de dépendance obtenues avec la statistique de Chi (χ) en fonction de la distance (λ) des couples (pluie-niveau marin ou pluie-surcote) par rapport à leur médiane

inondations historiques liées aux évènements de tempêtes touchant la côte de la Manche ainsi que la disponibilité des données et d'un modèle hydraulique. La pluie et le niveau marin étant deux phénomènes initiés par ces tempêtes, ont été considérés et les VAs associées ont été présentées en tenant en compte des déphasages temporels possibles entre elles. L'étude des statistiques descriptives des données a montré que la distribution des données de niveau marin est plus forte que celle des intensités de pluie. La corrélation entre la pluie et le niveau marin est très faible. Cette corrélation est mieux visible entre la composante surcote du niveau marin et la pluie. Il serait intéressant pour les sites avec une marée dominante de considérer une convolution de la surcote et de la marée. Toutefois, l'échantillonnage sera réalisé directement sur les observations de niveau marin maximal pour ne pas rajouter un degré de complexité à l'application.

Deuxième partie

Échantillonnage - analyses univariées et multivariées

Introduction

Les analyses fréquentielles univariées et multivariées sont des méthodes d'échantillonnage permettant d'estimer des probabilités (ou des périodes de retour, faisant référence au temps durant lequel un évènement se produit en moyenne une fois) et de les associer à des niveaux de retour (intensité de pluie, niveau marin, etc.). Cependant, ces approches sont très souvent déployées dans des contextes déterministes par les concepteurs et les décideurs, généralement en ne prenant en compte qu'un seul scénario et un seul phénomène (e.g. le dimensionnement d'un réseau d'assainissement avec une pluie de période de retour 10 ans).
Les travaux menés dans cette thèse permettront d'exploiter les acquis scientifiques en analyse fréquentielle univariée et multivariée, et de les compléter en les intégrant dans un cadre théorique plus général et plus complexe qui est celui d'une approche probabiliste.

Cette partie du mémoire est structurée en deux chapitres dédiés respectivement aux méthodes d'échantillonnage sous les hypothèses de non-coïncidence et de coïncidence.

Chapitre II.1

Échantillonnage sous l'hypothèse de non-coïncidence

Introduction

L'estimation de la fréquence d'un évènement extrême d'inondation repose sur l'élaboration d'un modèle fréquentiel permettant de définir le comportement statistique du phénomène d'inondation initiateur de cet évènement extrême.
Il existe plusieurs modèles fréquentiels univariés qui reposent sur différentes méthodes d'estimation, on cite par exemple la méthode des Maxima par bloc ([104]) et la méthode dite de renouvellement ou en anglais Peaks Over Threshold (POT) ([100]). La première approche est basée sur la distribution des valeurs extrêmes généralisée (GEV) des maxima par blocs journaliers, mensuels ou annuels, qui consiste à prendre uniquement une valeur maximale de chaque bloc. Tandis que, l'idée de la deuxième approche repose sur le fait qu'il est possible de modéliser plus qu'une valeur maximale par bloc. En effet, toutes les observations supérieures à un seuil sont prises en compte. Dans le cadre de la théorie des valeurs extrêmes, le modèle fréquentiel POT est basé sur la distribution Pareto Généralisée (GP). Pour les deux approches, les paramètres de la fonction de distribution utilisée doivent être estimés avec, par exemple, la méthode du maximum de vraisemblance ([35]). La méthode des r-Largest Order Statistics ou r-LOS ([138, 142]) permet de retenir les r valeurs extrêmes de chaque bloc. C'est un modèle généralisé du modèle fréquentiel des Maxima par bloc (r=1). Pour plus de détails sur ces méthodes, le lecteur pourra se référer aux travaux suivants [35, 70].

L'analyse fréquentielle est alors une approche statistique de prédiction basée sur des calculs de probabilités utilisant l'historique des évènements pour prédire des fréquences d'apparitions futures. D'une manière générale, l'analyse fréquentielle univariée s'effectue en plusieurs étapes : (i) collecte et vérification des données brutes (ii) choix d'un modèle fréquentiel et extraction des échantillons des valeurs extrêmes correspondants ; (iii) vérification des hypothèses de base (indépendance, stationnarité et homogénéité). Si par exemple l'application du test de la stationnarité montre que les données sont non-stationnaires, le modèle fréquentiel doit être choisi en conséquence, (iv) choix de la fonction de distribution théorique, ajustement et vérification de l'adéquation. Si le ou les tests d'adéquation utilisés montrent que la fonction de distribution choisie n'est pas adéquate, cette étape doit être reprise en sélectionnant une autre distribution et (v) calcul des intervalles de confiance associés. Si ces intervalles sont trop larges, il faut choisir un modèle fréquentiel qui génère un échantillon d'une taille plus grande. L'organigramme de la figure II.1.1 résume ces étapes.

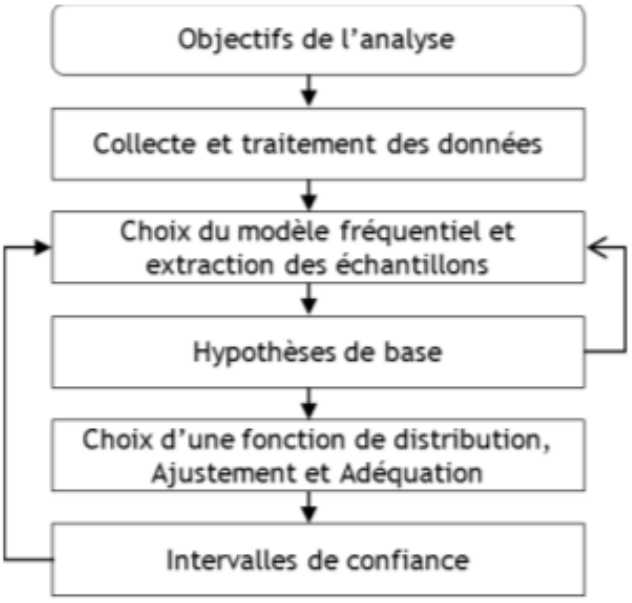


Figure II.1.1: Principales étapes d'une analyse fréquentielle univariée

L'analyse fréquentielle univariée est souvent utilisée dans des études déterministes de phénomènes.

La théorie des valeurs extrêmes est utilisée depuis plusieurs décennies dans le cadre de l'analyse fréquentielle univariée des évènements extrêmes [35]. Dans le cadre de cette thèse, un seul modèle sera utilisé pour ajuster une fonction de distribution théorique aux échantillons d'évènements d'intensité de pluie de durée 1 h et de niveau marin maximal. Cet ajustement nous permettra d'associer une probabilité (ou une période de retour) aux évènements extrêmes non observés.

Dans ce chapitre, nous verrons en premier lieu des aspects théoriques liés aux fonctions de distribution des méthodes d'échantillonnage MA et POT. Ensuite, nous utiliserons la méthode POT pour échantillonner les intensités de pluie et les niveaux marins maximums. Les distributions univariées issues de cette étape d'échantillonage seront utilisées plus loin dans ce rapport pour construire des scénarios.

II.1.1 Méthodes d'échantillonnage univariées

II.1.1.1 La méthode des Maxima Annuels (MA)

Cette méthode repose sur la loi limite du maximum d'un échantillon de variables aléatoires $X_1, X_2, .., X_n$ de fonction de distribution F. Cette loi limite du maximum $max(X_i)_{i\in 1,..,n}$ d'un échantillon, tel que $Pr\{max(Xi) < x\} = [F_X(x)]^n$, suppose l'existence de deux suites normalisantes A_n et B_n telles que :

$$\lim_{x\to\infty} Pr\{\frac{max(Xi) - B_n}{A_n} \leq x\} = \{F_X(A_n x + B_n)\}^n \to G_X(x) \qquad \text{(II.1.1)}$$

G_X est la fonction de distribution des valeurs extrêmes généralisées (distribution GEV).

$$G_X = \begin{cases} e^{-(1+\xi\frac{x-\mu}{\sigma})^{-1/\xi}} & \text{si } \xi \neq 0 \\ e^{-e^{-\frac{x-\mu}{\sigma}}} & \text{si } \xi = 0 \end{cases} \tag{II.1.2}$$

où μ, σ et ξ sont les paramètres de position, d'échelle et de forme, respectivement. De ce résultat asymptotique, il découle immédiatement que le comportement de la queue de distribution de la loi GEV est complètement caractérisé par un unique paramètre ξ qu'on appelle aussi l'indice des valeurs extrêmes. Le signe de ce paramètre est un indicateur essentiel de la forme de la queue de distribution et du domaine d'attraction (i.e., valeurs extrêmes).

II.1.1.2 La méthode POT

L'idée de la méthode POT est de partir d'un échantillon $x_1, x_2, .., x_n$, et de considérer les N_u observations dépassant un seuil u bien choisi. On note y_i, $i \in 1, .., N_u$, de nouvelles variables définissant les excès au-delà du seuil u, définies comme la différence entre l'observation et u. La fonction de répartition des excès au-delà de u est alors donnée par :

$$F_u(y) = Pr\{Y \leq y | X > u\} = \frac{F(u+y) - F(u)}{1 - F(u)} \tag{II.1.3}$$

D'après Pickands (1975), il existe une fonction σ et un réel ξ tels que la loi des excès F_u peut être uniformément approchée par une distribution de Pareto généralisée (GPD) notée $G_{\sigma,\xi}$:

$$G_{\sigma,\xi}(x) = \begin{cases} 1 - (1 + \frac{\xi \times (x-\mu)}{\sigma})^{-1/\xi} & \text{si } \xi \neq 0 \\ 1 - e^{-\frac{x-\mu}{\sigma}} & \text{si } \xi = 0 \end{cases} \tag{II.1.4}$$

pour $x \geq \mu$ quand $\xi \geq 0$, et $\mu \leq x \leq \mu - \sigma/\xi$ quand $\xi < 0$. Où σ et ξ sont les paramètres d'échelle et de forme, respectivement. La sélection du seuil u est souvent problématique et doit être choisi de façon objective ; en effet, un seuil relativement bas pourrait introduire un biais dans le modèle POT en prenant des évènements qui ne sont pas vraiment extrêmes. A l'inverse, en prenant un seuil trop élevé, nous pourrons perdre des évènements extrêmes et/ou ne pas avoir une taille d'échantillon suffisante pour mener une analyse fréquentielle (une variance élevée).

Il a été identifié dans la littérature que les principaux avantages de la méthode POT sont l'amélioration de l'échantillon (en éliminant les valeurs non extrêmes) et l'augmentation de sa taille ([35, 70]). Dans la suite de ce mémoire de thèse, seule la méthode POT est utilisée. Dans le contexte de la théorie des valeurs extrêmes, le modèle de fréquence POT est basé sur la distribution de Pareto généralisée (GPD). Les paramètres du CDF seront estimés à l'aide de la méthode du maximum de vraisemblance ([35]).

Le choix du seuil POT est basé sur plusieurs critères ([35]). La stabilité des quantiles élevés et du paramètre de forme et la linéarité du paramètre d'échelle de la

fonction de distribution ainsi qu'un nombre suffisant d'observations au-dessus du seuil sont les principaux critères permettant de faire ce choix.

II.1.1.3 Hypothèses de base

On cherche dans cette section à vérifier le critère aléatoire des variables, en s'assurant qu'elles soient indépendantes, stationnaires et homogènes. Cela se fait à l'aide de tests statistiques, à savoir le test d'indépendance de Wald-Wolfowitz, le test de stationnarité de Kendall et le test d'homogénéité de Wilcoxon. Par ailleurs, ces trois tests ont l'avantage d'être non-paramétriques, ils reposent sur les statistiques descriptives de la série telles que la moyenne et les rangs et ne font donc aucune hypothèse sur la fonction de distribution utilisée [70].
On cherche à vérifier en premier lieu l'indépendance statistique et mutuelle des observations d'un échantillon. L'hypothèse nulle H_0 du test de Wald-Wolfowitz consiste à vérifier l'absence d'auto-corrélation entre les observations.
Concernant l'homogénéité, le principe du test de Wilcoxon est de diviser l'échantillon en deux sous-échantillons et de vérifier si leurs moyennes sont significativement différentes. On se fixe alors comme hypothèse nulle H_0 l'homogénéité des observations.
Pour le test de Mann-Kendall, on s'intéresse à la tendance de la série. Il faut vérifier que les observations de la série n'ont tendance ni à augmenter ni à diminuer au fil du temps. L'hypothèse nulle H_0 dans ce cas est la stabilité des caractéristiques statistiques dans le temps.

Le résultat de chaque test est généralement sous la forme d'une «p-value». Un seuil observé ou p-value représente la probabilité au dépassement de la valeur calculée de la statistique du test. A un niveau de signification, généralement pris égal à 5%, un test est jugé comme réussi ou échoué, ce qui nous conduira respectivement à accepter ou refuser l'hypothèse nulle.

II.1.1.4 Adéquation des fonctions de distributions

Le test de Kolmogorov-Smirnov ([153]) est un test d'ajustement à une loi continue, qui prend en compte l'ensemble des quantiles. Pour un échantillon $(x_1, ..., x_n)$ de loi inconnue, l'hypothèse nulle est :
H_0 : la loi de l'échantillon a pour fonction de répartition F.
où F est la fonction de répartition d'une loi continue donnée (la GPD dans notre cas).
L'idée est la suivante : si l'hypothèse nulle H_0 est validée, alors la fonction de répartition empirique $\hat{F}$ de l'échantillon doit être proche de F. L'adéquation consiste à mesurer la distance de Kolmogorov-Smirnov, qui est la différence entre $\hat{F}$ et F aux points x_i. Le résultat de ce test d'adéquation est une probabilité (p-value) qui est comparée à un niveau de signification généralement pris égal à 5%.

II.1.2 Échantillonnage de l'intensité de pluie et du niveau marin maximal

La totalité des observations des séries de niveau marin (04/01/1938-17/01/2017) et de pluie (01/01/1994-01/01/2005) est utilisée pour construire les échantillons POT (figure II.1.2).

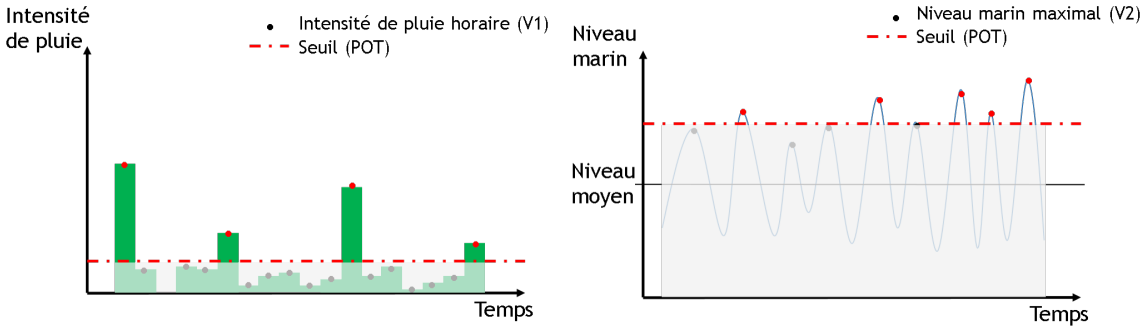


Figure II.1.2: Construction des échantillons d'intensité de pluie $V1$ (gauche) et de niveau marin maximal $V2$ (droite)

La figure II.1.3 montre la variation du paramètre de forme ξ, du quantile centennal et du nombre d'évènements pour les intensités de pluie $V1$ et le niveau marin maximal $V2$.
Une stabilité du quantile de niveau marin centennal est observée à partir d'un seuil POT égal à 8.5 mCM correspondant à une stabilité sur la courbe du paramètre de forme et un nombre d'évènements suffisamment élevé (69 évènements). Un seuil POT pour le niveau marin égal à 8.5 mCM semble alors justifiable.
De la même façon, un seuil POT égal à 8 mm/h est choisi pour l'intensité de pluie de durée 1 h.

Les seuils POT pour le niveau marin $V2$ et l'intensité de pluie $V1$ ainsi que les paramètres de la fonction de distribution GPD sont donnés dans le tableau II.1.1.

Echantillon	Seuil POT	σ	ξ	λ (evt/an)	Durée effective (années)
Niveau marin maximal	8.50 mCM	0.14	-0.03	1.47	46.80
Intensité de pluie $D = 1\ h$	8 mm/h	3.49	0.27	5.25	10.54

Tableau II.1.1: Paramètres des fonctions de distribution univariées pour la pluie et le niveau marin dans le cas de non-coïncidence

Le paramètre de forme ξ de l'ajustement d'intensité de pluie ($D1 = 1\ h$) est positif et relativement élevé donnant lieu nécessairement à une queue de distribution lourde, alors qu'il est légèrement négatif mais très proche de 0 donnant lieu à un ajustement exponentiel pour le niveau marin. Étant donné que ce paramètre gouverne le comportement de la queue de la distribution, cela signifie que la queue droite de la distribution est plus lourde pour le niveau de la mer que pour la pluie (figure II.1.4).
Les bornes de l'intervalle de confiance 70% de l'ajustement de l'intensité de pluie

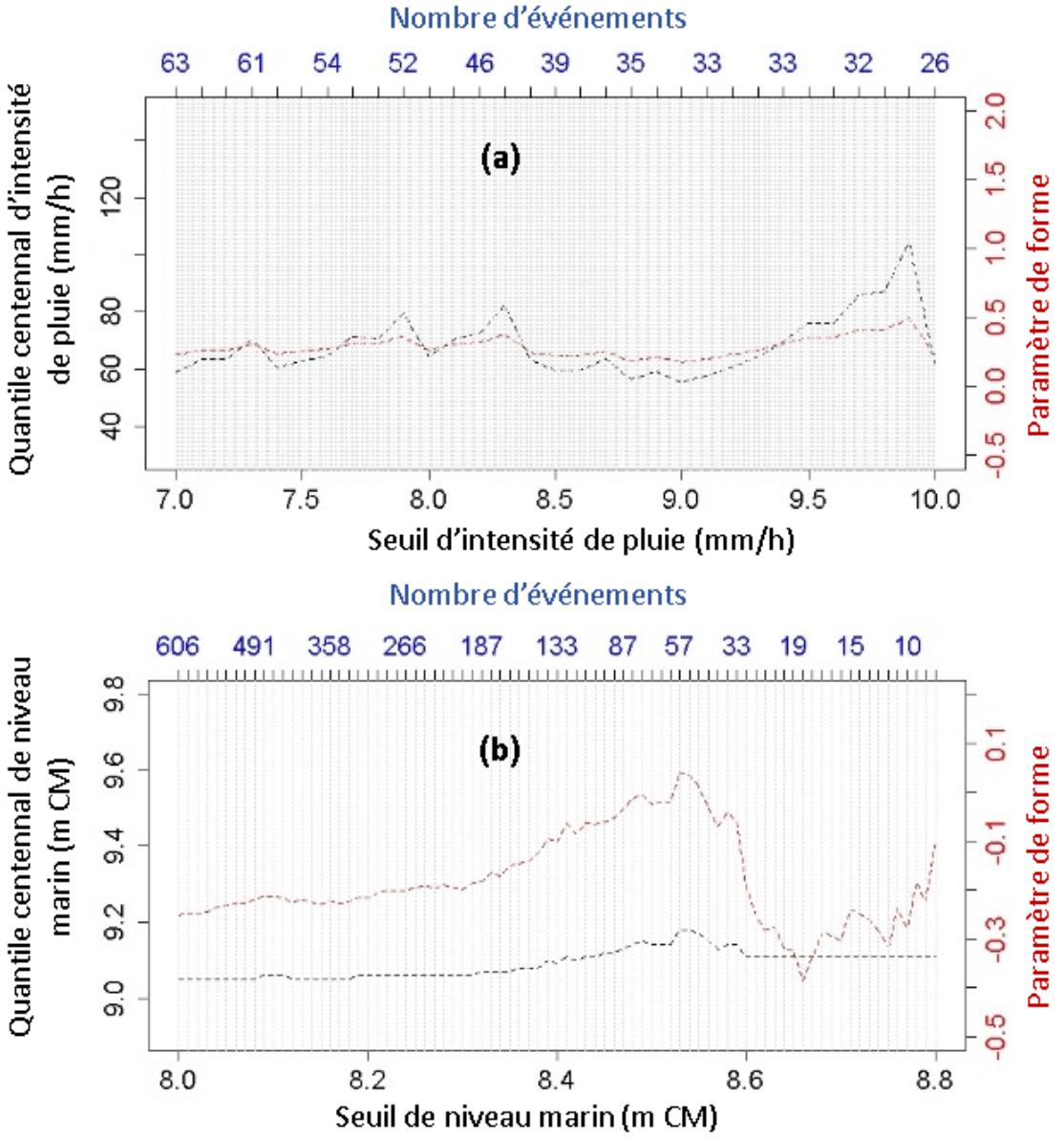


Figure II.1.3: Graphes de variation du quantile centennal (courbes noires), du paramètre de forme ξ (courbes rouges) et du nombre d'évènements (chiffres en bleu). (a) pour l'intensité de pluie 1h ; (b) pour le niveau marin maximal

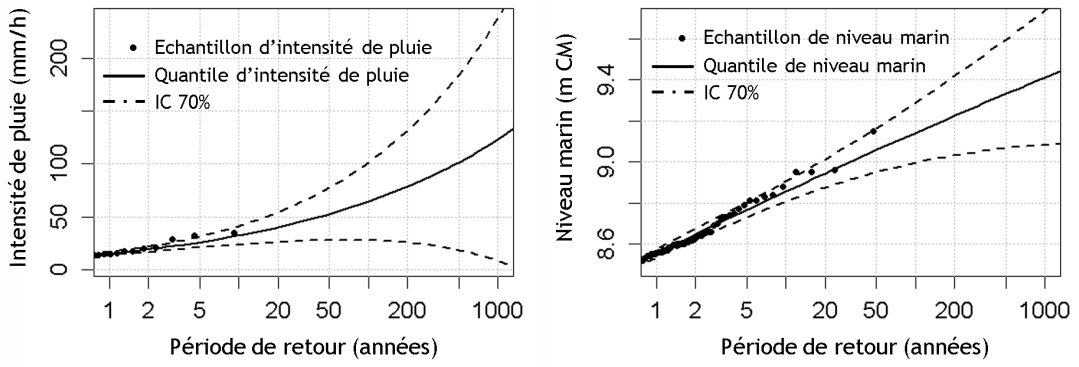


Figure II.1.4: Ajustement de la GPD à l'intensité de pluie 1 h (a) et au niveau marin maximal (b)

sont relativement larges aux valeurs extrêmes (e.g. période de retour 1000 *ans*) par rapport à ceux de l'ajustement du niveau marin (figure II.1.4). Cela signifie qu'il y a plus d'incertitude aléatoire (intrinsèque) liée aux données de pluie qu'au niveau marin. Notamment, la durée effective de l'échantillon est plus grande pour le niveau marin (47 *ans*) que pour la pluie (9 *ans*). La présence de données manquantes parmi les observations peut aussi être une source d'incertitude.

Les résultats du test d'adéquation de Kolmogorov-Smirnov sont donnés dans le tableau II.1.2.

Échantillon	p-value
Niveau marin maximal	0.80
Intensité de pluie $D = 1\ h$	0.87

Tableau II.1.2: Résultats du test d'adéquation de Kolmogorov-Smirnov

Les résultats du test de Kolmogorov-Smirnov présentés dans le tableau II.1.2 montrent que les ajustements de niveau marin et de pluie pour la durée $D1 = 1\ h$ sont adéquats (p-value $> 5\%$).

Les hypothèses de base sont ensuite vérifiées et le tableau II.1.3 présente les p-values des différents tests d'hypothèses.

Échantillon	Wald-Wolfowitz	Wilcoxon	Mann-Kendall
Niveau marin maximal $V2$	0.74	0.06	0.12
Intensité de pluie $V1$ ($D1 = 1h$)	0.57	0.87	0.53

Tableau II.1.3: Résultats des tests des hypothèses de base

Les p-values présentées dans le tableau II.1.3 sont toutes supérieures au niveau de signification 5%. Cela permet de valider les hypothèses nulles de ces tests statistiques (indépendance, homogénéité et stationnarité).

II.1.3 Conclusion

Ce chapitre dresse dans un premier lieu l'état de l'art sur ces approches d'échantillonnage univariées (sous l'hypothèse de non-coïncidence). Deux approches (GEV et GPD) faisant partie de la théorie générale des valeurs extrêmes ont été présentées. L'approche POT utilisant la fonction de distribution GPD présente un intérêt particulier car elle permet d'améliorer la taille des échantillons et d'éliminer les valeurs faibles, elle a été donc retenue. Cette méthode POT a été utilisée par la suite pour échantillonner les intensités de pluie et les niveaux marins maximums à partir des séries de données tout en vérifiant l'adéquation de ces ajustements obtenus à l'aide du test statistique de Kolmogorov-Smirnov. Ces ajustements serviront à construire des scénarios dans les prochains chapitres.

Chapitre II.2

Échantillonnage sous l'hypothèse de coïncidence

Introduction

La combinaison des phénomènes d'inondation est un sujet clé dans cette thèse. En effet, dans la nature, les phénomènes d'inondation sont souvent coïncidants (surviennent en même temps ou avec un certain déphasage). Les probabilités conjointes associées à ces phénomènes coïncidants sont différentes selon que ces derniers soient indépendants ou pas. Dans ce chapitre, on présentera d'abord les lois marginales de l'intensité de pluie et du niveau marin maximal puis on présentera les techniques utilisées dans l'estimation des probabilités conjointes.

II.2.1 Échantillonnage de la combinaison intensité de pluie-niveau marin maximal

L'échantillonnage d'une combinaison de variables aléatoires (e.g., intensité de pluie-niveau marin maximal) repose sur deux étapes : i) la construction des marginales univariées, dans cette étape les données de la période d'observation commune sont échantillonnées en utilisant la méthode POT, ii) le calcul des probabilités conjointes, s'ffectuant différemment dans les cas d'indépendance et de dépendance de l'intensité de pluie et du niveau marin maximal.

II.2.1.1 Construction des marginales univariées

Seulement la période d'observation commune des séries de niveau marin et de pluie (01/01/1994-01/01/2005) est utilisée pour construire les échantillons (figure II.2.1).

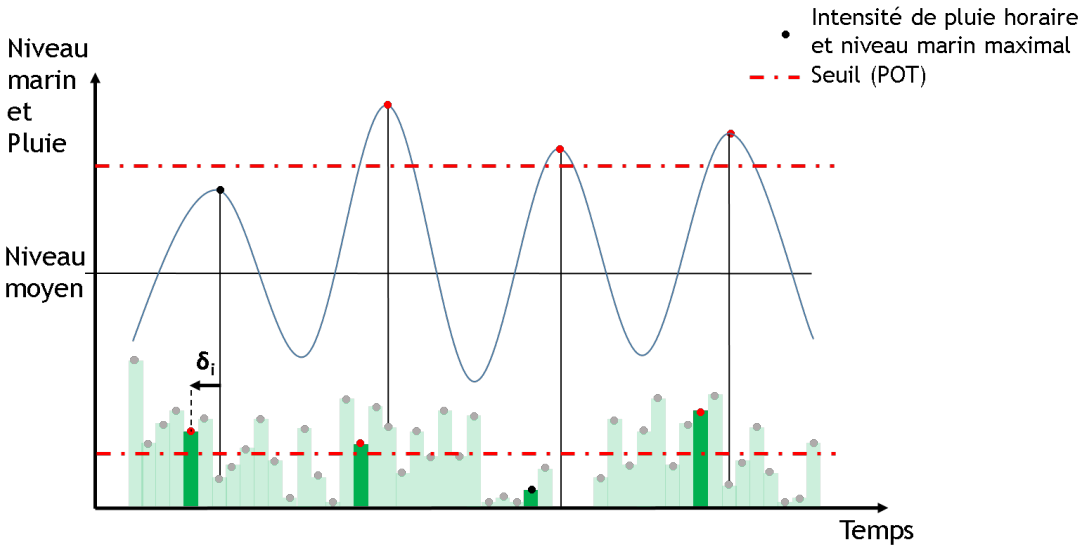


Figure II.2.1: Extraction des couples d'intensité de pluie $V1$ et de niveau marin maximal $V2$ afin de construire les marginales univariées (POT)

En cas de coïncidence des phénomènes d'inondation, des fonctions de distribution marginales sont utilisées. Ces distributions sont ajustées sur la période

d'observation commune et avec la méthode d'échantillonnage POT de la même façon que les distributions univariées de la pluie et du niveau marin présentées en section II.1.2.

Plusieurs travaux dans la littérature [97, 98, 86] révèlent que la pratique courante consiste à utiliser des distributions univariées mixtes. Selon Mazas et Hamm [97], la modélisation bivariée à partir des marginales (Inference From Marginals ou IFM) doit prendre en compte non seulement les valeurs extrêmes mais aussi les valeurs fréquentes des variables marginales. Les distributions mixtes permettent de couvrir à la fois le domaine fréquent (corps de la distribution) et le domaine des extrêmes (queue de la distribution). Le corps de la distribution est modélisé par une distribution empirique (échantillon) et la queue de distribution est modélisée par la théorie des valeurs extrêmes. Cette méthode de construction des distributions mixtes permet de rendre justifiable l'utilisation des quantiles fréquents dans la suite de la démarche. Cela n'est pas appliqué dans ce mémoire car nous utiliserons la méthode du Pseudo Maximum de Vraisemblance pour estimer les paramètres des copules (présentée plus loin dans ce chapitre).

La fonction de distribution GPD est ajustée à ces échantillons indépendants. Les paramètres de ces distributions marginales univariées sont présentés dans le tableau II.2.1.

Échantillon		Seuil POT	σ	ξ	λ (évt/an)	Durée effective (années)	Taille initiale (années)
Niveau marin maximal		8.45 mCM	0.11	0.03	2.62	10.67	10.67
Intensité de pluie	$\delta = -6\ h$	2.90 mm/h	2.02	-0.04	3.51		
	$\delta = -5\ h$	3.00 mm/h	2.25	0	3.79		
	$\delta = -4\ h$	3.00 mm/h	1.98	0.03	3.32		
	$\delta = -3\ h$	2.60 mm/h	1.59	0.48	4.17		
	$\delta = -2\ h$	3.00 mm/h	2.37	0.02	3.03		
	$\delta = -1\ h$	3.00 mm/h	2.12	-0.02	2.47		
	$\delta = 0\ h$	3.00 mm/h	2.00	0.015	2.37	10.54	10.67
	$\delta = +1\ h$	2.60 mm/h	1.68	0.30	3.60		
	$\delta = +2\ h$	3.00 mm/h	3.47	-0.02	2.84		
	$\delta = +3\ h$	3.00 mm/h	2.04	0.26	3.98		
	$\delta = +4\ h$	3.00 mm/h	2.78	0.03	2.84		
	$\delta = +5\ h$	2.90 mm/h	1.80	-0.01	3.51		
	$\delta = +6\ h$	3.00 mm/h	2.29	0.03	2.37		

Tableau II.2.1: Paramètres de la GPD pour la pluie et le niveau marin maximal dans le cas de coïncidence

Les taux d'évènements λ (tableau II.2.1) sont inférieurs à ceux présentés dans le tableau II.1.1, car les séries de données utilisées dans le cas de coïncidence sont plus courtes. De plus, le paramètre de forme ξ est presque nul pour la majorité des échantillons. Une distribution exponentielle peut être mieux adaptée à ces échantillons.
Les résultats du test d'adéquation de Kolmogorov-Smirnov sont donnés dans le tableau II.2.2.

Échantillon		p-value
Niveau marin maximal		0.59
Intensité de pluie	$\delta = -6\ h$	0.23
	$\delta = -5\ h$	0.64
	$\delta = -4\ h$	0.35
	$\delta = -3\ h$	0.55
	$\delta = -2\ h$	0.63
	$\delta = -1\ h$	0.62
	$\delta = 0\ h$	0.37
	$\delta = +1\ h$	0.50
	$\delta = +2\ h$	0.93
	$\delta = +3\ h$	0.75
	$\delta = +4\ h$	0.83
	$\delta = +5\ h$	0.47
	$\delta = +6\ h$	0.55

Tableau II.2.2: Résultats du test d'adéquation de Kolmogorov-Smirnov

Les résultats du test de Kolmogorov-Smirnov présentés dans le tableau II.2.2 montrent que les ajustements des marginales univariées de niveau marin et de pluie pour la durée $D1 = 1\ h$ sont adéquats (p-value $> 5\%$).

Les hypothèses de bases (i.e. indépendance, homogénéité, stationnarité) des échantillons de pluie et niveau marin de la période d'observation commune sont vérifiées. Les résultats des tests d'hypothèses sont présentés dans le tableau II.2.3.

Échantillon		Wald-Wolfowitz	Wilcoxon	Mann-Kendall
Niveau marin maximal		0.86	0.05	0.08
Intensité de pluie	$\delta = -6\ h$	0.27	0.14	0.08
	$\delta = -5\ h$	0.77	0.55	0.13
	$\delta = -4\ h$	0.77	0.71	0.90
	$\delta = -3\ h$	0.52	0.53	0.36
	$\delta = -2\ h$	0.22	0.50	0.87
	$\delta = -1\ h$	0.37	0.34	0.36
	$\delta = 0\ h$	0.29	0.36	0.11
	$\delta = +1\ h$	0.56	0.09	0.20
	$\delta = +2\ h$	0.29	0.98	0.66
	$\delta = +3\ h$	0.77	0.90	0.85
	$\delta = +4\ h$	0.42	0.74	0.87
	$\delta = +5\ h$	0.67	0.63	0.96
	$\delta = +6\ h$	0.47	0.85	0.42

Tableau II.2.3: Résultats des tests des hypothèses de base

Les p-values présentées dans le tableau II.2.3 sont toutes supérieures au niveau de signification 5%. Cela permet de valider les hypothèses nulles de ces tests statistiques (indépendance, homogénéité et stationnarité).

II.2.1.2 Calculs des probabilités conjointes

Lorsque les phénomènes sont coïncidants et indépendants, leurs probabilités conjointes sont données par le produit de leurs probabilités individuelles. Ces probabilités individuelles proviennent des distributions marginales univariées qui sont ajustées sur une période d'observation commune à $V1$ et à $V2$ (section II.2.1.1). Toutefois, les VAs environnementales sont souvent dépendantes. L'analyse fréquentielle univariée s'avère insuffisante pour analyser le risque associé aux VAs corrélées. On se trouve alors astreint à prendre en compte la structure de dépendance entre ces VAs en ayant recours à une analyse fréquentielle multivariée. Cette approche, qui réside dans l'utilisation de la théorie de copules [129, 102], a été intensivement employée durant la dernière décennie ([119, 39, 22]).
Dans les deux cas d'indépendance et de dépendance, les probabilités conjointes sont multipliées par la probabilité de coïncidence $Pr_c = \frac{1}{2\times365}$ afin d'obtenir les probabilités que les évènements d'intensité de pluie et de niveau marin maximal se produisent pendant le même cycle de marée.
On récapitule ainsi les probabilités conjointes dans le tableau suivant :

	Probabilités conjointes $Pr\{V1 > v1, V2 > v2\}$
Cas d'indépendance	$(1 - F_{V1}) \times Pr_c \times (1 - F_{V2}) \times Pr_c$
Cas de dépendance	$1 + C(V1, V2) - F_{V1} \times Pr_c - F_{V2} \times Pr_c$

Tableau II.2.4: Équations de calcul des probabilités conjointes dans les cas d'indépendance et de dépendance (variables coïncidantes)

Où C est une fonction de distribution bivariée (copule) modélisant la dépendance entre $V1$ et $V2$, F_{V1} et F_{V2} sont les marginales univariées et $Pr_c = \frac{1}{2\times365}$.

Analyse fréquentielle multivariée

L'analyse fréquentielle multivariée consiste à étudier la structure de dépendance de deux ou plusieurs variables à travers une fonction qui dépend de leurs fonctions de distribution individuelles (univariées). Les applications de l'analyse fréquentielle multivariée dans le domaine de l'hydrologie se sont multipliées durant ces dernières années. Les modèles multivariés ont été utilisés dans l'analyse et la modélisation des pluies intenses [38, 63, 163], l'étude de la sécheresse [116, 90, 162], la modélisation des tempêtes [40], la régionalisation [30, 31, 32], les crues fluviales [52, 33, 109], l'analyse fréquentielle des inondations [56, 83], l'analyse des précipitations et du ruissellement [164], la conception et la vérification des ouvrages hydrauliques [39, 114, 147]. Il a été constaté dans la littérature que plusieurs cas d'inondation comme ceux dus aux franchissements d'ouvrages de protection côtiers, sont provoqués par les actions combinées de deux ou plusieurs phénomènes tels que le niveau marin, les vagues, les vents extrêmes. Les combinaisons les plus courantes sont le niveau marin extrême avec des vagues extrêmes et les surcotes extrêmes avec une marée haute [97]. Dans ces cas, l'analyse fréquentielle multivariée permet de prendre en compte la dépendance.

Une analyse fréquentielle multivariée est une approche basée sur des fonctions de distribution à deux variables ou plus. Ces fonctions sont des copules ayant comme variables d'entrée les valeurs prises par les lois marginales univariées dont on connaît déjà les paramètres. D'une façon générale, la modélisation multivariée s'effectue suivant les étapes du logigramme de la figure II.2.2. Comme dans l'analyse univariée, l'estimation du paramètre de la copule se fait par la méthode du maximum de vraisemblance (ou des L-Moments). L'adéquation de la copule est vérifiée à travers des tests d'adéquation bivariés ou des critères AIC/BIC. Contrairement aux tests d'adéquation classiques, ces critères (AIC/BIC) devront permettre de choisir la meilleure copule parmi plusieurs. L'organigramme de la figure II.2.2 résume ces étapes.

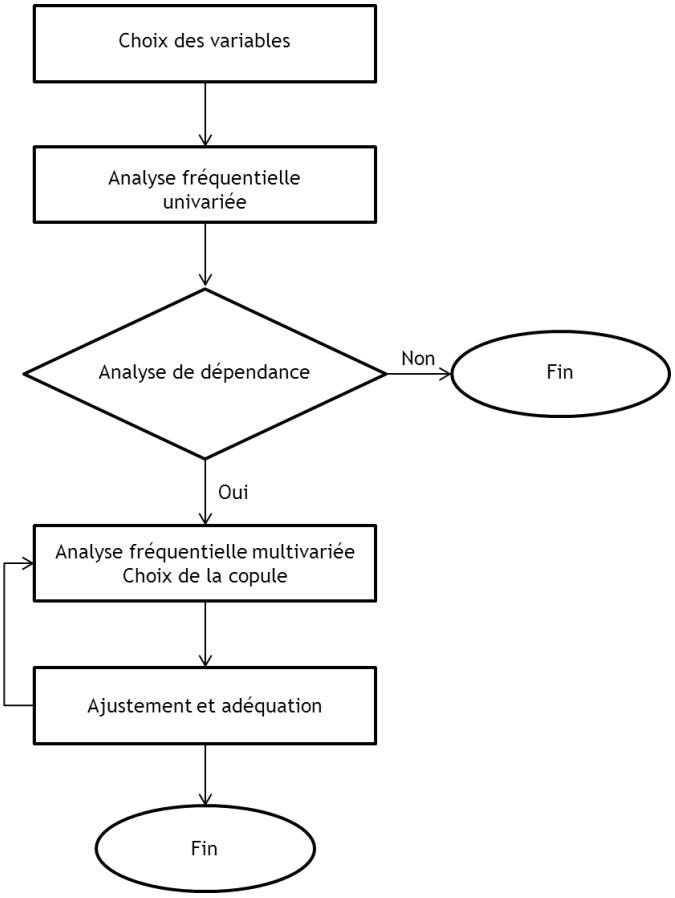


Figure II.2.2: Principales étapes d'une analyse fréquentielle multivariée

Le terme « copule » a été introduit pour la première fois par Sklar [129] en l'associant à un concept mathématique permettant de relier les distributions des variables en fonction de leur degré de dépendance. Une brève présentation de la théorie des copules est proposée dans cette section. Plus de détails peuvent être trouvés dans [82, 103, 119]. Cette théorie porte sur le théorème de Sklar [129] qui stipule que si $F_{X,Y}(x, y)$ est une fonction de distribution (CDF) ayant des fonctions de distributions marginales $F_X(x)$ et $F_Y(y)$ alors il existe une copule C telle que :

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y))$$

La construction du modèle fréquentiel bivarié repose sur la modélisation des lois marginales (univariées), le choix des copules candidates, l'estimation des paramètres ainsi que l'adéquation des copules. Ces différents sujets seront présentés dans les sections suivantes.

Mesures de dépendance

Il existe différentes mesures pour quantifier la relation de la dépendance entre deux VAs $V1$ et $V2$. Dans cette section, on présente brièvement des mesures de corrélation classiques (coefficient de corrélation de Pearson, tau de Kendall et le rho de Spearman) avant de détailler deux mesures de dépendance de queue.

La relation de dépendance linéaire entre deux VAs $V1$ et $V2$ peut être quantifiée à l'aide du coefficient de corrélation de Pearson. Si $V2 = a+bV1$ et $b > 0$, alors :

$$\rho_P(V1, V2) = \frac{Cov(V1, V2)}{\sqrt{Var(V1)Var(V2)}} = \frac{bVar(V1)}{\sqrt{Var(V1)b^2Var(V1)}} = 1 \qquad \text{(II.2.1)}$$

Inversement, si $V2 = a + bV1$, avec $b = -c$ où $c > 0$, on obtient :

$$\rho_P(V1, V2) = \frac{Cov(V1, V2)}{\sqrt{Var(V1)Var(V2)}} = \frac{-cVar(V1)}{\sqrt{Var(V1)c^2Var(V1)}} = -1 \qquad \text{(II.2.2)}$$

Contrairement au coefficient de corrélation de Pearson, le tau de Kendall et le rho de Spearman sont des mesures de corrélation portant sur les rangs des observations.
Pour le couple de VAs $(V1, V2)$, le tau de Kendall est défini par :

$$\tau(V1, V2) = Pr((V1 - V1')(V2 - V2') > 0) - Pr((V1 - V1')(V2 - V2') < 0) \qquad \text{(II.2.3)}$$

où $(V1', V2')$ est un couple (extrait de l'échantillon initial de (V1,V2)) ayant les mêmes marginales que $(V1, V2)$ tout en étant indépendant de $(V1, V2)$.
L'équation II.2.3 permet de quantifier le degré de concordance des VAs $V1$ et $V2$:

$$\tau(V1, V2) = Pr(\text{concordance}) - Pr(\text{discordance}) \qquad \text{(II.2.4)}$$

La valeur du tau de Kendall $\tau(V1, V2) = 0$ est associée à des VAs $V1$ et $V2$ indépendantes. La valeur du tau de Kendall ne dépend pas des marginales du couple $(V1, V2)$. Selon l'équation II.2.4, plus la relation de dépendance entre $(V1, V2)$ est positive (négative) plus la valeur s'approchera de 1 (-1).

Le rho de Spearman est donné par l'équation suivante :

$$\rho_S(V1, V2) = 3Pr((V1 - V1")(V2 - V2") > 0) - 3Pr((V1 - V1")(V2 - V2") < 0) \qquad \text{(II.2.5)}$$

où $(V1", V2")$ est un couple (extraits de l'échantillon initial de (V1,V2)) ayant les mêmes marginales que $(V1, V2)$ tout en étant indépendant de $(V1, V2)$.

La valeur du rho de Spearman $\rho_S(V1, V2) = 0$ est obtenue dans le cas où les VAs $V1$ et $V2$ sont indépendantes. Le signe du rho de Spearman (positif ou négatif) est associé à la concordance des VAs $V1$ et $V2$.
Il est possible d'établir une relation entre le rho de Spearman et le coefficient de corrélation de Pearson. Il existe également une relation entre le tau de Kendall et le rho de Spearman. Ces relations ne sont pas détaillées dans ce mémoire de thèse. Il est possible d'exprimer le tau de Kendall et le rho de Speaman d'un couple de VAs $(V1, V2)$ en fonction de la copule associée à leur distribution conjointe. Par ailleurs, il n'est pas possible d'appliquer directement le tau de Kendall et le rho de Spearman pour mesurer la dépendance entre deux VAs discrètes [55].

La dépendance de queue d'un couple de VAs continues $(V1, V2)$ est une notion importante dans l'analyse multivariée des valeurs extrêmes. Les coefficients de dépendance de queue (tail dependence en anglais) permettent de quantifier la dépendance au niveau du quadrant haut droit (bas gauche) d'une distribution bivariée [55, 82]. Selon Tosunoglu et Kisi [143] si la dépendance de queue n'est pas prise en compte par la copule choisie, cela pourra provoquer de fortes incertitudes dans l'estimation des quantiles extrêmes. La dépendance de queue est importante dans l'adéquation de la copule choisie. Cette dépendance de queue peut être quantifiée par un simple Chi-plot [3] ou par des estimateurs non-paramétriques de la dépendance de queue haute (upper tail) [126]. Le coefficient proposé par [26, 59] est couramment utilisé dans l'évaluation non-paramétrique de la dépendance de queue haute [110]. La comparaison de ce coefficient calculé pour une copule donnée avec celui de la copule empirique permet de vérifier l'adéquation et le choix de la copule. Au sein de la famille des copules Archimédiennes, les copules d'Ali-Mikhail-Haq et Frank ont une dépendance de queue nulle, contrairement aux copules de Gumbel et de Clayton qui possèdent une forte dépendance de queue (coefficient de dépendance de queue proche de 1 ou de -1 en fonction des données) [143].

Selon Joe ([82]), une distinction peut être faite entre les coefficients de dépendance de queue supérieure λ_U (upper tail dependence) et de dépendance de queue inférieure λ_L (lower tail dependence). Ces coefficients sont donnés par les relations suivantes :

$$\lambda_U(V1, V2) = \lim_{k \to 1^-} Pr\{V1 > F_{V1}^{-1}(k) | V2 > F_{V2}^{-1}(k)\} \tag{II.2.6}$$

si cette limite existe.

$$\lambda_L(V1, V2) = \lim_{k \to 0^+} Pr\{V1 \leq F_{V1}^{-1}(k) | V2 \leq F_{V2}^{-1}(k)\} \tag{II.2.7}$$

si cette limite existe.

Contrairement au tau de Kendall et au rho de Spearman qui mesurent une dépendance globale sur toute la distribution bivariée de la paire $(V1, V2)$, les coefficients de dépendance de queue sont des mesures locales. Ces coefficients permettent de donner une idée sur l'occurrence simultanée de valeurs élevées ou faibles de $(V1, V2)$. Ces coefficients de dépendance de queue dépendent uniquement de la distribution bivariée associée à des VAs $(V1, V2)$ continues.

Des valeurs $\lambda_U(V1, V2) \in (0, 1]$ ($\lambda_L(V1, V2) \in (0, 1]$) sont associées à une dépendance asymptotique des VAs $V1$ et $V2$ au niveau supérieur (inférieur) de la queue de distribution. Une valeur $\lambda_U(V1, V2) = 0$ ($\lambda_L(V1, V2) = 0$) est liée à une indépendance asymptotique des VAs $V1$ et $V2$ au niveau supérieur (inférieur) de la queue de distribution.
Dans le domaine des inondations, le concept de dépendance de queue supérieure se révèle important. Si $\lambda_U(V1, V2) \in (0, 1]$ et qu'un évènement extrême de pluie (ayant une intensité $V1$) a eu lieu, alors la probabilité qu'un évènement extrême de niveau marin (ayant une amplitude $V2$) survienne est non nulle.

Copules bivariées

État de l'art et différentes utilisations des copules

Il existe une grande variété de familles de copules avec des propriétés différentes [82, 103, 52]. Le premier critère de choix est la capacité de la copule à modéliser la structure de dépendance globale. Les copules Archimédiennes (Ali-Mikhail-Haq, Clayton, Frank et Gumbel) font partie des copules les plus utilisées en sciences environnementales en général et en hydrologie en particulier [60, 40, 144]. Ces copules Archimédiennes ont l'avantage d'être simples à construire et peuvent détecter la dépendance globale [27, 92, 116]. Selon Lian et al. [94], les distributions de probabilités bivariées telles que la distribution normale, de Gumbel-mixte [160], exponentielle [10], gamma [128], méta-gaussienne [85] ne permettent de modéliser qu'une partie limitée de la dépendance car elles prennent en compte une seule fonction de distribution pour les deux VAs [4]. Les copules permettent de mieux modéliser la structure de dépendance des variables indépendamment des propriétés des lois marginales [121].

Plus récemment, les copules ont été utilisées dans l'analyse de l'occurrence conjointe des phénomènes maritimes. De Waal et van Gelder [148] ont proposé une copule « Burr-Pareto-Logistic » pour modéliser la distribution conjointe des amplitudes et périodes de vagues extrêmes pendant les tempêtes. De Michele et al. [40] ont présenté un modèle multivarié basé sur une copule pour l'étude des variables des vagues lors des tempêtes (amplitude, période, direction, décalage de temps). Wahl et al. [149, 150] ont utilisé la copule de Gumbel pour modéliser la structure de dépendance du niveau marin maximal et de l'intensité de pluies. Ces derniers ont également fait l'objet d'une analyse bivariée sur la côte Allemande de la mer Baltique [117]. Yang and Zhang [159] ont étudié la distribution de probabilité conjointe des vents extrêmes et des vagues simulées à travers la copule de Gumbel.
La combinaison des crues fluviales et pluviales est rarement étudiée dans la littérature [5, 24]. Apel et al. [5] considèrent les crues fluviales et pluviales comme totalement indépendantes. La combinaison de ces deux aléas a été prise en compte directement dans un modèle hydraulique.

Mazas et Hamm [97] ont proposé une démarche pour étudier la combinaison des vagues et du niveau marin, dont les principales étapes sont : i) échantillonnage, ii) ajustement des lois marginales des vagues et des surcotes et modélisation

de la marée par une distribution empirique, iii) modélisation de la structure de dépendance des vagues et surcotes. En parallèle, la fonction de distribution du niveau marin est déterminée par la convolution de la marée astronomique et des surcotes en utilisant la méthode POT-JPM proposée par Mazas et al. [98], iv) la convolution de la distribution bivariée des vagues-surcotes avec la marée permet d'obtenir la distribution conjointe des vagues et du niveau marin.

Les copules de Gumbel, de Galambos et de Hüsler-Reiss ont été utilisées pour les vagues et les surcotes. Ces copules font partie de la famille des copules des valeurs extrêmes dont l'avantage par rapport aux copules Archimédiennes est qu'elles ne sont pas symétriques. Plus de détails sur la justification de l'utilisation des copules des valeurs extrêmes dans l'analyse fréquentielle multivariée des extrêmes sont présentés dans [64]. Les résultats ont été comparés avec une application faite avec le logiciel JOIN-SEA (HR-Wallingford) dans lequel seules les copules normale et "normale-mixte" sont implémentées. Cette comparaison a démontré l'intérêt d'utiliser des copules de valeurs extrêmes pour la modélisation des variables environnementales [97].

Masina et al. [95] ont étudié la combinaison de la houle et du niveau marin en prenant en compte les variables hauteur de vagues, période de pic des vagues, direction de vagues et le niveau marin ainsi que la variabilité saisonnière. La saison d'occurrence et la direction de propagation des vagues sont considérées comme des variables discrètes. La distribution de probabilité conjointe du niveau marin maximal et de la hauteur significative des vagues est estimée par les copules en considérant l'échantillon complet (toutes les observations), des échantillons par saisons et par directions. La distribution de probabilité conjointe de la hauteur significative des vagues et de la période de pic des vagues est déterminée à partir des observations de vagues. Des lois marginales mixtes ont été utilisées pour modéliser les distributions univariées du niveau marin maximal et de la hauteur significative des vagues.

Zheng et al. [166] ont étudié la structure de dépendance entre les surcotes et les pluies. La surcote a été extraite du niveau marin observé à travers une analyse harmonique décrite dans [111]. Le modèle logistique bivarié avec la méthode du seuil (Bivariate Logistic Threshold-Excess ou BLTE) [115] est utilisé pour modéliser la dépendance entre les pluies et les surcotes. Ce modèle est adapté pour les cas de dépendance asymptotique. La première étape consiste à choisir des seuils univariés pour les pluies et les surcotes afin d'ajuster la loi GPD. Les graphes de variation de la moyenne résiduelle et des paramètres de la GPD en fonction du seuil ont été utilisés pour effectuer ce choix de seuils. Zheng et al. [57] ont ensuite étudié la variabilité spatiale et saisonnale de la dépendance ainsi que l'impact des différentes durées de pluies et du décalage entre les pluies extrêmes et les surcotes sur la structure de dépendance. En effet, l'analyse a montré que plus la durée de pluies est grande plus la dépendance avec les surcotes est grande pour le site d'étude. Le degré de dépendance varie avec la durée des tempêtes. Zheng et al. [166] considèrent la dépendance entre les pluies et les surcotes comme typiquement faible. Selon ces auteurs la dépendance est asymptotique, elle peut rester constante comme elle peut croître aux faibles probabilités de dépassement. Plusieurs méthodes permettent la modélisation de cette dépendance extrémale i.e.

i) maxima par bloc [142], ii) méthode du seuil [115], iii) processus ponctuel [37], iv) méthode conditionnelle [76]. Ces méthodes ont été récemment utilisées dans [37, 84, 166].
Dans certains cas, la pratique consistant à modéliser la dépendance entre le niveau marin et les pluies intenses peut être trompeuse. Samuels and Burt [122] ont modélisé la dépendance entre les débits d'une rivière et le niveau marin (marée observée) et ont conclu à l'absence de corrélation. En revanche, l'analyse de dépendance par Svensson and Jones [141] pour le même cas d'étude a révélé une dépendance significative entre les débits journaliers et les surcotes. Selon Zheng et al. [166] cette contradiction est due à l'utilisation des données de niveau marin maixmal (qui est généralement dominé par la marée astronomique). En effet, la marée observée est composée d'une surcote et d'une marée astronomique. Seule la composante surcote est intéressante pour étudier la dépendance avec les pluies car elle est générée par les effets de pression atmosphérique et du vent et par conséquent elle peut être physiquement associée aux pluies extrêmes. Cela prouve que la sélection des données est un sujet critique dans l'étude de dépendance, d'autres détails peuvent être trouvés dans [72].

Plusieurs méthodes permettent d'estimer le(s) paramètre(s) d'une copule. Parmi ces méthodes, une méthode directe basée sur les mesures de corrélation comme le tau de Kendall, le rho de Spearman, la méthode du pseudo maximum de vraisemblance (PML) ou le maximum de vraisemblance canonique (CML), la méthode de l'inférence à partir des marginales (inference from margins ou IFM), la méthode du maximum de vraisemblance exact (EML) et la méthode des moments (MOM). Pour plus de détails, se référer à [60, 34, 82].

Fondements théoriques des copules

On présente une brève introduction aux copules avant de présenter des résultats de différentes copules et de mesures de dépendance.

Théorème de Sklar :

Soit $F \in \Gamma(F_1, ..., F_n)$ ayant des fonctions de répartition marginales $F_1, ..., F_n$. Alors, il existe une copule C telle que pour tout vecteur $\hat{x} \in \Re^n$:

$$F(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)) \tag{II.2.8}$$

Si $F_1, ..., F_n$ sont continues, alors C est unique. Inversement, si C est une copule et $F_1, ..., F_n$ sont des fonctions de répartitions alors la fonction définie par II.2.8 est une fonction de répartition multivariée.

Une copule bivariée $C(V1, V2)$ est une application $[0, 1]^n \rightarrow [0, 1]$ ayant les propriétés suivantes :

- 1. $C(0, 0) = 0$;
- 2. $C(1, 1) = 1$;
- 3. si une des composantes de C est v l'autre est 1, alors la copule prend la valeur v ;

— 4. si au moins une des composantes de C est 0, alors la copule prend la valeur 0 ;
— 5. la probabilité sur $(a_1, b_1]x(a_2, b_2]$ est positive i.e. :

$$\Delta_{a_1,b_1}\Delta_{a_2,b_2}C(V1, V2) \geq 0 \tag{II.2.9}$$

pour tout $a_1 \leq b_1$ et $a_2 \leq b_2$, où :

$$\Delta_{a_1,b_1}\Delta_{a_2,b_2}C(V1, V2) = \sum_{i_1=1}^{2}\sum_{i_2=1}^{2}(-1)^{i_1+i_2}C(V1_{i_1}, V2_{i_2}) \geq 0 \tag{II.2.10}$$

Dans la suite, on présente l'expression de certaines copules bivariées. Soient F_{V1} et F_{V2} les marginales des VAs $V1$ et $V2$, dont la distribution conjointe est définie par la copule C. $U1 = F_{V1}(V1)$ et $U2 = F_{V2}(V2)$.

Copule d'indépendance :

La copule d'indépendance est définie par :

$$C^I(U1, U2) = U1 \times U2 \tag{II.2.11}$$

où $Ui \in [0, 1], i \in 1, 2$.

Copule de Clayton :

L'expression de la copule de Clayton est :

$$C_\alpha(U1, U2) = \left(U1^{-\alpha} + U2^{-\alpha} - 2\right)^{-\frac{1}{\alpha}} \tag{II.2.12}$$

où $Ui \in [0, 1], i \in 1, 2$ et $\alpha > 0$.

Copule de Frank :

La copule de Frank est exprimée par :

$$C_\alpha(U1, U2) = -\frac{1}{\alpha}\ln\left(1 + \frac{(e^{-\alpha U1} - 1)(e^{-\alpha U2} - 1)}{e^{-\alpha} - 1}\right) \tag{II.2.13}$$

où $Ui \in [0, 1], i \in 1, 2$ et $\alpha > 0$.

Copule de Gumbel :

La copule de Gumbel est définie par :

$$C_\alpha(U1, U2) = e^{-((-\ln(U1))^\alpha + (-\ln(U2))^\alpha)} \tag{II.2.14}$$

où $Ui \in [0,1], i \in 1,2$ et $\alpha \geq 1$.

Copule normale (Gaussienne) :

L'expression de la copule normale est :

$$C_{\underline{\alpha}}(U1, U2) = \underline{\Phi}_{\underline{\alpha}}\left(\Phi^{-1}(U1), \Phi^{-1}(U2)\right) \tag{II.2.15}$$

où $Ui \in [0,1], i \in 1,2$ et avec :

$$\underline{\alpha} = s\begin{pmatrix} 1 & \alpha_{12} \\ \alpha_{12} & 1 \end{pmatrix} \tag{II.2.16}$$

où $\alpha_{12} \in [0,1]$, Φ est une loi normale univariée et $\underline{\Phi}$ est une loi normale bivariée standard avec une matrice de coefficients de corrélation $\underline{\alpha}$.

Copule de Student :

L'expression de la copule de Student est :

$$C_{\nu,\underline{\alpha}}(U1, U2) = t_{\nu,\underline{\alpha}}\left(t_\nu^{-1}(U1), t_\nu^{-1}(U2)\right) \tag{II.2.17}$$

où $Ui \in [0,1], i \in 1,2$ et avec :

$$\underline{\alpha} = s\begin{pmatrix} 1 & \alpha_{12} \\ \alpha_{12} & 1 \end{pmatrix} \tag{II.2.18}$$

où $\alpha_{12} \in [0,1]$, t est la loi de Student bivariée standard avec une matrice de coefficients de corrélation $\underline{\alpha}$.
Si $\underline{\alpha}$ est égale à la matrice identité, cela n'implique pas l'indépendance comme c'est le cas pour la copule normale.

Copule de Joe :

Pour être concis, le notation de (Joe, 1997) sera utilisiée : $\hat{U} = 1 - U$
L'expression de la copule de Joe est :

$$C_\alpha(U1, U2) = 1 - \left(\hat{U1}^{-\alpha} + \hat{U2}^{-\alpha} - \hat{U1}^{-\alpha}\hat{U2}^{-\alpha}\right)^{\frac{1}{\alpha}} \tag{II.2.19}$$

où $Ui \in [0,1], i \in 1,2$ et $\alpha \geq 1$.

Pour les copules de Frank, de Clayton et de Gumbel, on remarque que le paramètre de dépendance α est le même pour les deux VAs au sein de la copule. Pour le cas bivarié, les copules normale et de Student permettent aussi d'avoir le même paramètre de dépendance α. En revanche, pour un nombre de VAs supérieur à 2, ces deux copules (normale et de Student) permettent d'avoir des paramètres de dépendance différents selon les paires de VAs de la copule.

Ajustement des copules

Soient les VAs de niveau marin maximal $V2$ et d'intensité de pluie $V1_{/\delta i}$ (de durée $D1 = 1\ h$) aux différents déphasages.

Plusieurs copules ont été ajustées en utilisant les échantillons du tableau II.2.1.2 (sans application de seuils, comme l'indique la figure II.2.3). La méthode du Pseudo-Maximum de Vraisemblance (PML) a été utilisée pour estimer les paramètres de ces copules car elle a l'avantage de détecter la dépendance globale et d'avoir un grand nombres de couples (sans application de seuils).

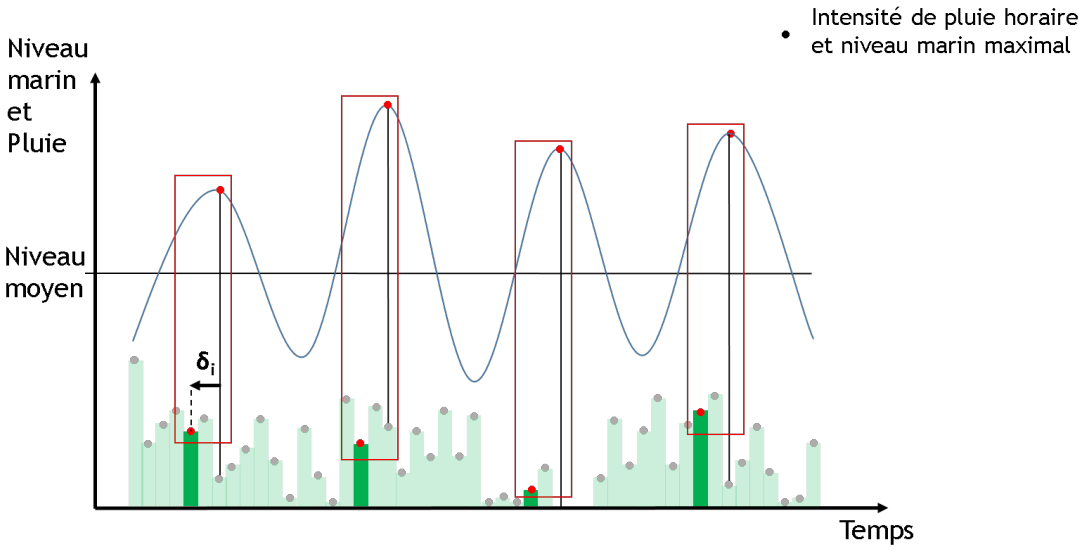


Figure II.2.3: Extraction des couples d'intensité de pluie $V1$ et de niveau marin maximal $V2$ afin de construire les marginales univariées (POT)

Échantillon		Durée effective (années)	Taille initiale (années)	Nombre de couples
Niveau marin maximal		10.67	10.67	Pluie - niveau marin
Intensité de pluie	$\delta = -6\ h$	10.54	10.67	6206
	$\delta = -5\ h$			6204
	$\delta = -4\ h$			6201
	$\delta = -3\ h$			6204
	$\delta = -2\ h$			6196
	$\delta = -1\ h$			6203
	$\delta = 0\ h$			6206
	$\delta = +1\ h$			6195
	$\delta = +2\ h$			6197
	$\delta = +3\ h$			6200
	$\delta = +4\ h$			6194
	$\delta = +5\ h$			6196
	$\delta = +6\ h$			6204

Tableau II.2.5: Échantillons ayant servi aux ajustements des copules

Le tableau II.2.6 présente, pour un déphasage $\delta = 0\ h$, les paramètres de dépendance pour les différentes copules testées ainsi que le tau de Kendall et les coefficients de dépendance de queue.
Les résultats en lien avec les autres déphasages sont mentionnés en annexe A.

Copule	par_1 (α)	par_2	τ	λ_L	λ_U
Gaussienne (normale)	0.09	0	0.05	0	0
Student	0.08	30	0.05	$1.39 \quad 10^{-5}$	$1.39 \quad 10^{-5}$
Clayton	0.07	0	0.03	$4.52 \quad 10^{-5}$	0
Survie de Clayton	0.06	0	0.03	0	$6.45 \quad 10^{-6}$
Gumbel	1.04	0	0.04	0	0.05
Survie de Gumbel	1.03	0	0.03	0.04	0
Frank	0.46	0	0.05	0	0
Joe	1.02	0	0.01	0	0.02

Tableau II.2.6: Paramètres des différentes copules pour $\delta = 0\ h$

Le tableau II.2.6 montre une faible dépendance sur l'ensemble de la distribution conjointe pour toutes les copules (τ de Kendall non nul). Cela signifie la présence de valeurs extrêmes (ou faibles) de $(V1, V2)$.
On remarque que les copules de Joe et de Gumbel ont des coefficients de dépendance de queue supérieure non nuls, cela montre que $V1$ et $V2$ sont asymptotiquement dépendantes au niveau supérieur de la queue de distribution. Toutefois, ce coefficient de dépendance de queue supérieure est nul ou quasi-nul pour le reste des copules. Ce résultat signifie aussi que les copules de Joe et de Gumbel sont des copules des valeurs extrêmes et permettent de mieux ajuster la queue de distribution supérieure que les autres copules. Toutefois, les coefficients de dépendance de queue inférieure sont nuls pour toutes les copules sauf celles de survie de Gumbel et de Clayton, ce qui signifie que ces deux copules sont adaptées aux VAs présentant une dépendance de queue inférieure.
La présence d'une dépendance de queue supérieure était déjà constatée dans les observations de $V1$ et $V2$. Le nuage de points des couples $(V1, V2)$ présenté dans

la figure I.1.10, montre bien cette dépendance de queue supérieure (densité des point et forme conique du nuage).

Adéquation et choix de copule

L'analyse de l'adéquation des copules s'effectue en deux étapes. Une première étape consiste au diagnostic visuel en comparant la copule ajustée avec la copule empirique [139] ou en simulant un large échantillon par la copule et ensuite le comparer avec les observations [97]. La deuxième étape consiste à utiliser des tests statistiques i.e. les tests de Kolmogorov-Smirnov (KS), d'Anderson-Darling (AD) d'Anderson-Darling Intégré (IAD). Selon Tosunoglu et Kisi [143], le test de Cramér-von Mises est puissant pour mesurer la performance des copules. Plus de détails sur la statistique de ce test (de Cramér-von Mises) peuvent être trouvés dans [62], la p-value associée peut être calculée à partir des bootstaps paramétriques [61] ou par une approche de "multiplicateur" [88]. Lorsque plusieurs copules sont utilisées, de nombreuses alternatives permettent de sélectionner une copule (la plus adéquate), i.e. le critère de l'information d'Akaike (AIC), le critère de l'information Bayésien (BIC). Dans la suite, les critères AIC et BIC seront utilisés.

Lian et al. [94] ont étudié la probabilité conjointe de la marée observée et des pluies extrêmes pour une ville côtière en utilisant les copules. La rivière a été modélisée par un modèle hydrodynamique développé avec HECRAS où 1713 sections transversales ont été utilisées. La fonction de distribution Pearson-3 a été utilisée comme loi marginale pour les pluies ainsi que pour la hauteur de marée. 5 copules différentes i.e. les copules Gaussienne, t-copule (Student), de Clayton, de Frank et de Gumbel ont été comparées. Leurs paramètres ont été estimés par l'approche du maximum de vraisemblance [152]. La copule de Gumbel a été choisie en se basant sur le critère des moindres carrés ordinaire (OLS) et le test de Kolmogorov-Smirnov (K-S).

Le tableau II.2.7 présente les résultats des critères AIC et BIC pour les différentes copules présentées précédemment (pour $\delta = 0h$).

Copule	par_1 (α)	par_2	AIC	BIC
Gaussienne	0.09	0	-14.02	-7.28
Student	0.08	30	-14.98	-1.51
Clayton	*0.07*	*0*	*-20.56*	*-13.83*
Survie de Clayton	0.06	0	-0.78	5.93
Gumbel	1.04	0	-4.93	1.80
Survie de Gumbel	1.03	0	-18.65	-11.92
Frank	0.46	0	-7.47	-0.74
Joe	1.02	0	1.30	8.03

Tableau II.2.7: Choix d'une copule sur la base des critères AIC et BIC - $\delta = 0\ h$

Pour le déphasage $\delta = 0\ h$, la copule de Clayton de paramètre $\alpha = 0.07$ est la copule la plus adéquate car elle minimise les critères AIC et BIC (-20.56 et

-13.83 respectivement). Toutefois, la copule de survie de Gumbel peut être aussi considérée comme un bon choix (classée deuxième selon les critères AIC/BIC). Selon [107], ne pas considérer la copule de Clayton (ou Gumbel de survie) peut conduire à sous-dimensionner fortement l'ouvrage. Le choix entre la copule de Gumbel et celle de Clayton est lié principalement à la dépendance de queue.
Les résultats pour les autres déphasages présentés en annexe A montrent que la copule de Clayton figure toujours parmi les 3 meilleures copules (selon les critères AIC et BIC). Pour des raisons d'homogénéisation, la copule de Clayton, dont les paramètres sont récapitulés dans le tableau II.2.8, a été choisie pour tous les déphasages.

Déphasage		par_1 (α)
Copule de Clayton	$\delta = -6\ h$	0.09
	$\delta = -5\ h$	0.11
	$\delta = -4\ h$	0.12
	$\delta = -3\ h$	0.10
	$\delta = -2\ h$	0.07
	$\delta = -1\ h$	0.07
	$\delta = 0\ h$	0.07
	$\delta = +1\ h$	0.07
	$\delta = +2\ h$	0.05
	$\delta = +3\ h$	0.06
	$\delta = +4\ h$	0.07
	$\delta = +5\ h$	0.06
	$\delta = +6\ h$	0.06

Tableau II.2.8: Paramètres de la copule de Clayton pour les différents déphasages entre la pluie et le niveau marin

II.2.2 Conclusion

Ce chapitre a dressé dans un premier temps l'état de l'art sur les approches multivariées permettant d'échantillonner les VAs coïncidantes. Les résultats obtenus ont montré une faible dépendance globale des VAs $V1$ et $V2$ (τ de Kendall faible mais non nul). Les coefficients de dépendance de queue ont été utiles pour quantifier la dépendance des VAs $V1$ et $V2$ au niveau de la queue de distribution (supérieure et inférieure). Ces coefficients ont également montré la capacité de certaines copules d'ajuster des données présentant une dépendance de queue. Le choix de la meilleure copule a porté sur les critères Bayésiens (AIC et BIC) et a permis de sélectionner la copule de Clayton.

Le tableau II.2.9 récapitule les méthodes d'échantillonnages de l'intensité de pluie et du niveau marin maximal sous les hypothèses de non-coïncidence et de coïncidence.

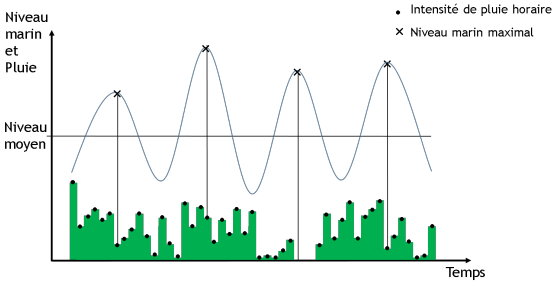
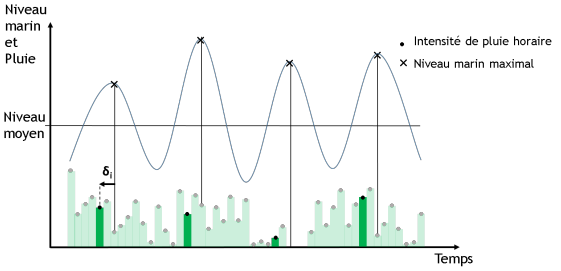
	Non-coïncidence	Coïncidence
Données	niveau marin : 01/1938-01/2017 ; pluie : 01/1994-01/2005	01/1994-01/2005
VAs		
Méthode d'échantillonnage	POT	POT (marginales) et PML (copules)
Nombre d'échantillons	pluie : 1 et niveau marin : 1	pluie : 13 et niveau marin : 1
Calculs des probabilités	pluie : $1 - F_{V1}$ et niveau marin : $1 - F_{V2}$	indépendance : $(1 - F_{V1}) \times (1 - F_{V2})$ et dépendance : $1 + C(V1, V2) - F_{V1} - F_{V2}$

Tableau II.2.9: Tableau récapitulatif de l'échantillonnage de la pluie et du niveau marin sous les hypothèses de non-coïncidence et de coïncidence

Troisième partie

Approche probabiliste

Introduction

Les approches probabilistes viennent compléter les approches déterministes. L'approche probabiliste présentée dans ce rapport est un concept composé de trois principales étapes : i) probabilisation de phénomènes d'inondation, ii) propagation de ces phénomènes de la source aux points d'intérêt sur le site, iii) agrégation des probabilités et construction de courbes d'aléa.
En effet, l'étape de probabilisation repose sur tous les outils déterministes de calcul des probabilités (approches univariée et multivariée). Dans cette étape, les analyses fréquentielles univariées, classiquement utilisées dans les approches déterministes, sont complétées par des approches multivariées qui prennent en compte la dépendance et la coïncidence des phénomènes d'inondation.
Tous ces savoir-faire utilisés dans les approches déterministes forment notamment une base pour mettre en place l'approche probabiliste proposée dans cette thèse. L'approche par scénarios ainsi que l'agrégation des courbes d'aléa font partie des principaux apports de cette thèse, sans oublier le sujet de la coïncidence qui a été introduit dans la combinaison des phénomènes d'inondation, lui aussi est un élément apporté par les travaux de cette thèse.

Dans cette troisième partie du mémoire, nous présenterons ces sujets clés abordés dans cette thèse. L'utilisation des distributions statistiques pour construire des scénarios et la propagation de ces derniers sera abordée dans un premier chapitre de cette partie. L'agrégation des sorties du modèle hydraulique et la construction de courbes d'aléa seront présentées dans un second chapitre.

Chapitre III.1

Scénarios d'inondation

III.1.1 Construction des scénarios

Le présent chapitre est dédié aux différentes méthodes de construction des scénarios de pluie et de niveau marin.

III.1.1.1 Définition d'un scénario

Un scénario est un signal temporel qui représente au mieux la physique observée dans les données de pluie et de niveau marin. Ce signal est d'abord défini par une forme (e.g. triangulaire, sinusoïdale, etc.). L'amplitude maximale du scénario (ou son intensité) est extraite des ajustements statistiques des VAs $V1$ et $V2$, sa durée est donnée par les VAs discrètes $D1$ et $D2$ et sa probabilité est fournie par la fonction de distribution des VAs $V1$ et $V2$ notées F_{V1} et F_{V2} respectivement (tableau II.2.4). Un scénario est donc un processus temporel dont la forme du signal dépend du phénomène étudié.
Cependant, une distinction est faite entre les scénarios simples ne concernant qu'un seul phénomène d'inondation à la fois et les scénarios composés de combinaisons de couples $(V1, V2)$. Dans le cas d'un scénario combiné, nous considérons la combinaison de deux signaux, l'un de $V1$ et l'autre de $V2$. Le déphasage temporel δ présenté dans la section I.1.3 du chapitre II.2 est défini entre les pics des deux signaux. Pour rappel, la probabilité conjointe de dépassement associée à un scénario composé est donnée par le tableau II.2.4.

Dans l'application de cette méthode, une forme triangulaire simple est utilisée pour les scénarios de pluie et une forme sinusoïdale est retenue pour les scénarios de niveau marin. Ces formes représentent plus ou moins bien les évènements observés de pluie et de niveau marin pour une mer à marée en régime macrotidal (e.g., le cas de la côte de la Manche).
Parmi toutes les fonctions développées dans le cadre de cette thèse, deux fonctions permettent, en leur donnant les valeurs de l'intensité $V1 = i$ de pluie (de durée $D1 = 1\ h$) et de l'amplitude du niveau marin $V2 = NM$ (de durée $D2 = 12\ h$), de construire des processus temporels triangulaire et sinusoïdale des scénarios de la pluie et du niveau marin respectivement (figure III.1.1). Le principe consiste à résoudre des équations du premier ordre :

- une fonction affine pour la pluie $f(t) = at + b$, f étant l'intensité de pluie à un instant t, a et b sont des paramètres déterminés en connaissant les valeurs de f aux instants $t = 0$ (0 mm/h), $t = D1/2$ ($I = 2i$ sur la figure III.1.1 (a)) et $t = D1 = 1\ h$ (0 mm/h) ;
- une fonction cosinus pour le niveau marin $g(t) = A \times cos(2\Pi \times t/T)$, g étant le niveau marin à un instant t, A est un paramètre déterminé en connaissant les valeurs de g aux instants $t = 0$ (NM) et $t = D2 = 12\ h$ (NM, figure III.1.1 (b)). T étant la période principale de la marée avec un régime semi-diurne qui est approximée à 12 h.

III.1.1.2 Cas de non-coïncidence

Des tirages dans les distributions de probabilités permettent de construire plusieurs scénarios simples de pluie et niveaux marins. Le tableau III.1.1 récapitule tous les niveaux marins maximums (de durée $D2 = 12\ h$) et les intensités de pluie (de durée $D1 = 1\ h$) de ces scénarios ainsi que leurs probabilités de dépassement issues des fonctions de distribution $1 - F_{V2}$ et $1 - F_{V1}$, respectivement. En appli-

Pluie		Niveau marin	
$V1 = i$ (mm/h)	$Pr\{V1 > v1\}$	$V2 = NM$ (mCM)	$Pr\{V2 > v2\}$
15.27	$1.91\ 10^{-1}$	8.55	$6.78\ 10^{-1}$
26.17	$3.81\ 10^{-2}$	8.77	$1.36\ 10^{-1}$
32.38	$1.93\ 10^{-2}$	8.86	$6.78\ 10^{-2}$
39.47	$1.00\ 10^{-2}$	8.94	$3.39\ 10^{-2}$
50.71	$4.32\ 10^{-3}$	9.06	$1.36\ 10^{-2}$
60.88	$2.30\ 10^{-3}$	9.14	$6.78\ 10^{-3}$
72.76	$1.24\ 10^{-3}$	9.22	$3.39\ 10^{-3}$
80.61	$8.62\ 10^{-4}$	9.27	$2.26\ 10^{-3}$
86.63	$6.68\ 10^{-4}$	9.31	$1.70\ 10^{-3}$
91.58	$5.48\ 10^{-4}$	9.33	$1.36\ 10^{-3}$
95.81	$4.67\ 10^{-4}$	9.35	$1.13\ 10^{-3}$
99.52	$4.08\ 10^{-4}$	9.37	$9.69\ 10^{-4}$
102.84	$3.62\ 10^{-4}$	9.39	$8.48\ 10^{-4}$
105.85	$3.27\ 10^{-4}$	9.40	$7.54\ 10^{-4}$
108.61	$2.98\ 10^{-4}$	9.41	$6.78\ 10^{-4}$
119.87	$2.09\ 10^{-4}$	9.46	$4.52\ 10^{-4}$
128.51	$1.63\ 10^{-4}$	9.49	$3.39\ 10^{-4}$
141.66	$1.14\ 10^{-4}$	9.53	$2.26\ 10^{-4}$
151.74	$8.90\ 10^{-5}$	9.57	$1.70\ 10^{-4}$
160.02	$7.34\ 10^{-5}$	9.59	$1.36\ 10^{-4}$
188.55	$4.03\ 10^{-5}$	9.67	$6.78\ 10^{-5}$

Tableau III.1.1: Intensité des scénarios de pluie (de durée 1 h et de forme triangulaire simple) et de niveau marin (de durée 12 h et de forme sinusoïdale) dans le cas de non-coïncidence

quant la méthode de construction de scénarios, on obtient les scénarios présentés en figure III.1.1. Ces scénarios seront par la suite injectés un par un dans les

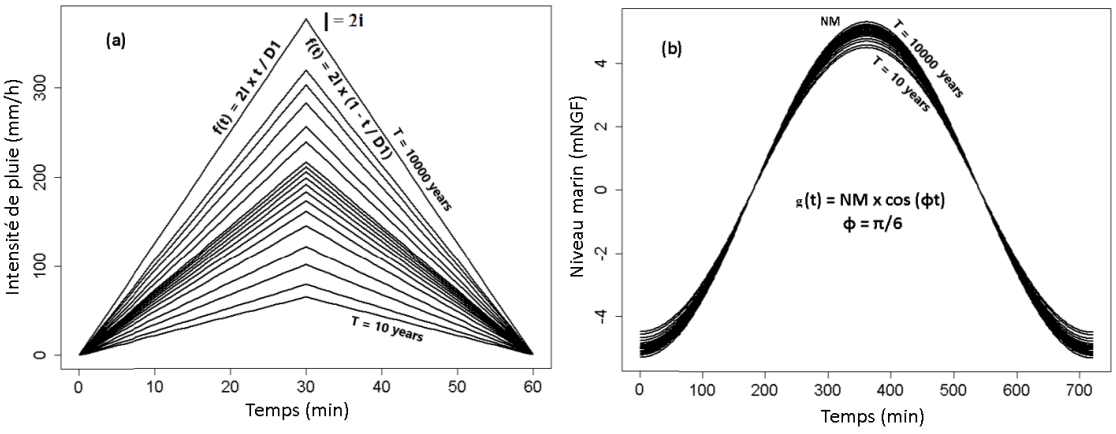


Figure III.1.1: Scénarios simples de pluie (a) et de niveau marin (b)

conditions aux limites du modèle hydraulique. Cette étape fera l'objet du chapitre suivant lié à la propagation hydraulique des scénarios.

III.1.1.3 Cas de coïncidence

Des combinaisons de 13 niveaux marins (de forme sinusoïdale) avec 14 pluies (de forme triangulaire) ont donné lieu à $13 \times 14 = 182$ scénarios. En outre, 13 déphasages équiprobables $\delta = \{-6\ h, -5\ h, -4\ h, -3\ h, -2\ h, -1\ h, 0\ h, +1\ h, +2\ h, +3\ h, +4\ h, +5\ h, +6\ h\ \}$ sont considérés entre le pic du signal du niveau marin ($V2$) et celui du signal de la pluie ($V1$). Nous avons donc 182 scénarios par déphasage et un total de 2366 scénarios composés.

Le tableau III.1.2 montre certains scénarios composés ainsi que leurs probabilités conjointes de dépassement en cas d'indépendance et de dépendance calculées selon le tableau II.2.4.

Pluie		Niveau marin		Prob. conjointes	
$V1 = i$ (mm/h)	$Pr\{V1 > v1\}$	$V2 = NM$ (mCM)	$Pr\{V2 > v2\}$	Pr_{indep}	Pr_{dep}
$\delta = -6\ h$					
22.29	$6.26\ \ 10^{-2}$	0	1	$6.26\ \ 10^{-2}$	$6.26\ \ 10^{-2}$
0	1	9.14	$6.65\ \ 10^{-3}$	$6.65\ \ 10^{-3}$	$6.65\ \ 10^{-3}$
22.29	$3.5\ \ 10^{-6}$	9.14	$5.22\ \ 10^{-6}$	$1.85\ \ 10^{-11}$	$8.79\ \ 10^{-13}$

Tableau III.1.2: Scénarios composés de pluie (de durée 1 h et de forme triangulaire simple) et de niveau marin (de durée 12 h et de forme sinusoïdale) pour différents déphasages (liste non-exhaustive, $\delta = -6\ h$)

Il est à noter que le cas de coïncidence comprend implicitement des scénarios avec une pluie nulle et d'autres avec un niveau marin nul (e.g., première et seconde lignes du tableau III.1.2). Pour ces scénarios, les probabilités sont issues des fonctions de distribution issues des échantillons sous l'hypothèse de non-coïncidence (sur une période d'observation complète). L'hypothèse de coïncidence intègre donc implicitement l'hypothèse de non-coïncidence.
Les probabilités sous l'hypothèse de coïncidence sont multipliées par $Pr_c = \frac{1}{2\times365}$ [1], sauf les scénarios de pluie seule et de niveau marin seul (non-coïncidence).

III.1.2 Propagation de la source aux points d'intérêt sur le site - modèle hydraulique

La propagation hydraulique est un élément important dans les travaux de cette thèse. C'est une étape qui permet de joindre les connaissances déterministes au concept probabiliste. Elle permet de propager l'ensemble des scénarios construits à partir des outils statistiques de la source (station de mesure) aux points d'intérêt dans le site.

1. nombre de cycles de marées par an

Le modèle hydraulique utilisé dans cette thèse est d'abord présenté dans une première section. Les caractéristiques de ce modèle et ses conditions initiales font l'objet des deuxième et troisième sections. Un point d'intérêt est choisi en section 4 afin de pouvoir présenter les résultats de la propagation des scénarios de pluie et de niveau marin. La propagation de ces scénarios est présentée en section 5 avant de montrer les courbes d'aléa dans la dernière section.

III.1.2.1 Modèle hydraulique

Il existe une variété de modèles hydrauliques allant des modèles simplifiés à casiers avec des relations de débits liant les casiers, aux modèles 2D, plus complexes, composés de mailles et de nœuds et fondés sur la résolution numérique d'équations mathématiques (e.g., équations de Navier Stokes, Barré de Saint-Venant).

Le modèle du site du Havre a été fourni par la COmmunauté de l'Agglomération Havraise (CODAH). Il s'agit d'un modèle couplé 1D-2D, construit avec le logiciel Infoworks-ICM et concerne uniquement la partie basse de la ville du Havre.
Les caractéristiques du modèle seront présentées en un premier temps. Ensuite, le choix d'un point d'intérêt sera détaillé. Par la suite, les conditions initiales et aux limites seront présentées avant de montrer la méthode de propagation des scénarios ainsi que les sorties du modèle.

Partie 1D

La partie 1D simule les écoulements dans le réseau d'assainissement. Elle est connectée à la partie 2D à travers des "noeuds 2D" permettant le transfert des flux d'un modèle à l'autre. Le tableau III.1.3 récapitule les composantes du modèle 1D.

Nombre total de nœuds modélisés	996	Déversoirs	55
Nœuds (regards)	975	Vannes	33
Nœuds Exutoires	17	Pompes	13
Nœuds Bassins	4	Clapets anti-retour	6
Nœuds de connexion au modèle 2D (nœuds 2D)	558	Longueur totale des réseaux modélisés	129km

Tableau III.1.3: Composantes du modèle 1D du Havre

Partie 2D

La partie 2D est composée d'un maillage issu d'un modèle numérique de terrain (MNT), de bâtiments, de murets de protection et simule les écoulements surfaciques. L'étendue de la zone 2D modélisée est donnée par la figure III.1.2. L'aire totale de la zone modélisée est de 1621.17 ha et les limites de cette zone modélisée sont considérées dans le modèle comme des murs verticaux. La pluie étant prise en compte dans la partie 1D du modèle hydraulique en les injectant dans des sous-bassins versants définis autour des regards.

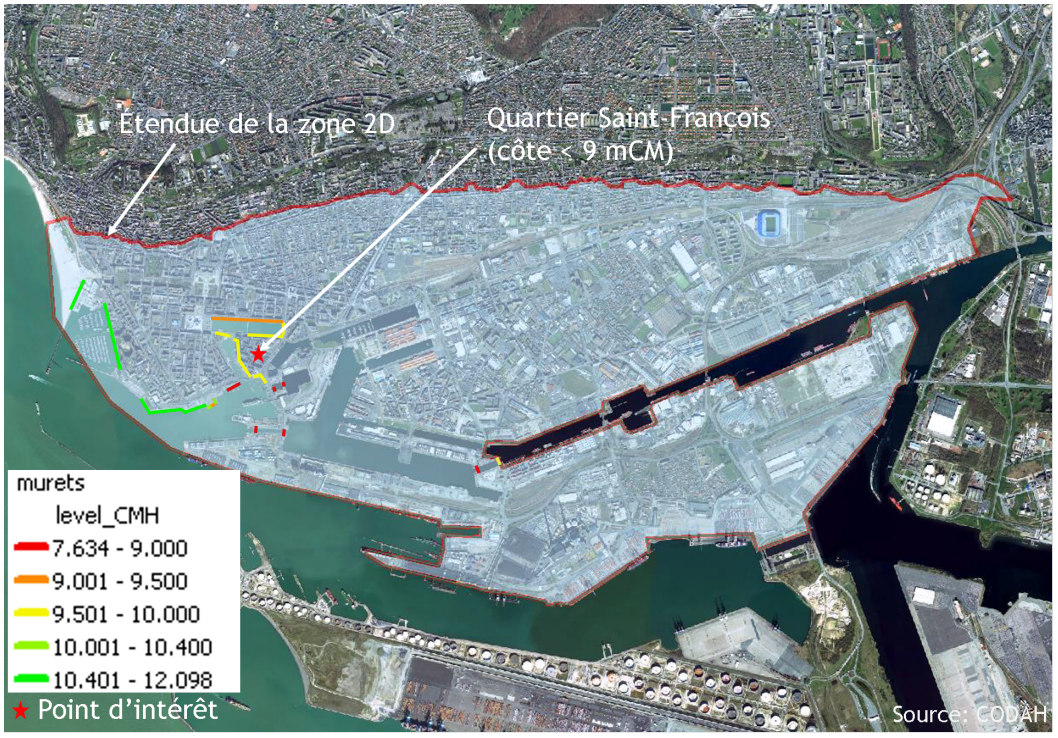


Figure III.1.2: Étendue de la zone 2D du modèle du Havre

Le MNT est d'une résolution $1m \times 1m$ et le maillage utilisé est de type triangulaire irrégulier. Ce maillage a donné lieu à 76590 mailles et 54133 noeuds de maillage.

Conditions initiales

Les conditions initiales permettent de définir l'état initial des composantes du modèle hydraulique. Un profil des eaux usées avec un débit journalier de $130l/jour$ est utilisé. Une élévation initiale de 1.22 mNGF est définie pour le niveau d'eau dans les différents bassins de la zone modélisée.

Point d'intérêt

Des tests sur le modèle ont permis de déterminer les zones inondables de la partie basse de la ville du Havre. Il a été constaté que le quartier Saint-François est parmi les premiers endroits à être inondés. Ce quartier s'apparente à une presque île car il est entouré par l'eau de mer. Des murets de protection ont été érigés en 1984 autour du quartier Saint-François afin de se prémunir du risque de débordement des bassins (figure III.1.3). Pour toutes ces raisons, un point d'intérêt a été choisi dans le quartier Saint-François (étoile rouge sur la figure III.1.3). Ce point d'intérêt est parmi les points les plus bas du quartier (altitude = 3.87 m NGF).

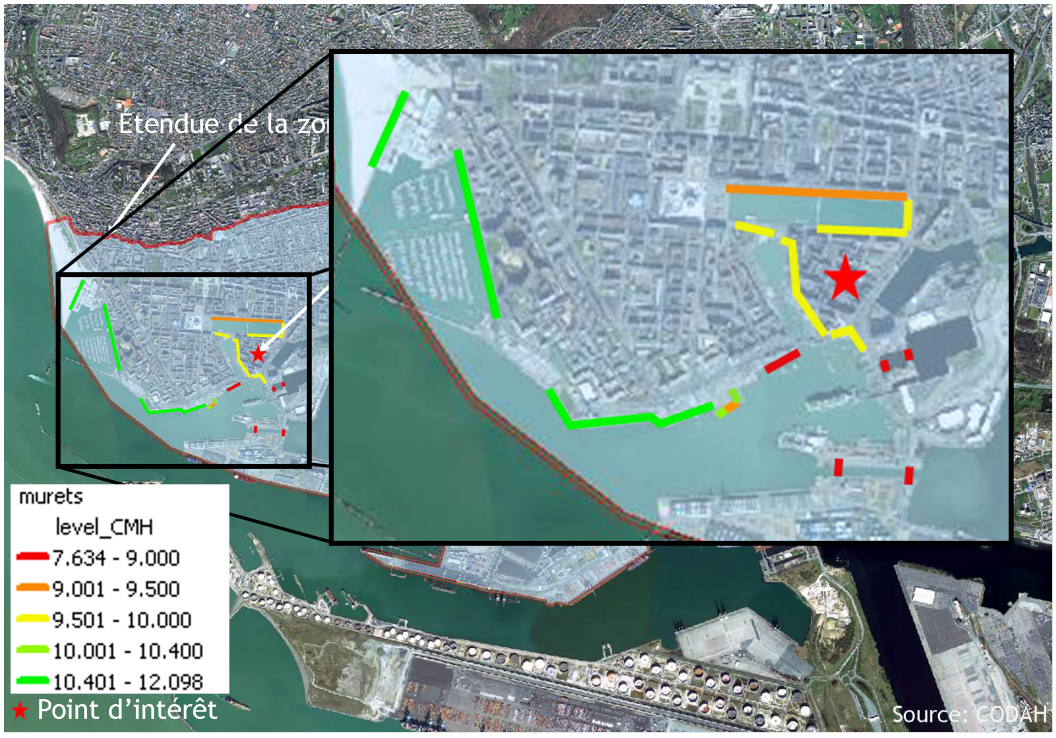


Figure III.1.3: Murets de protection autour du quartier Saint-François et point d'intérêt (étoile rouge)

III.1.2.2 Propagation des scénarios

Le modèle hydraulique du site est utilisé pour propager l'eau depuis les sources des phénomènes ϕ_1 et ϕ_2 jusqu'au point d'intérêt choisi dans le quartier Saint-François. Un scénario est injecté dans les conditions aux limites du modèle et une variation de la hauteur d'eau est obtenue en sortie, comme présenté dans l'exemple de la figure III.1.4.

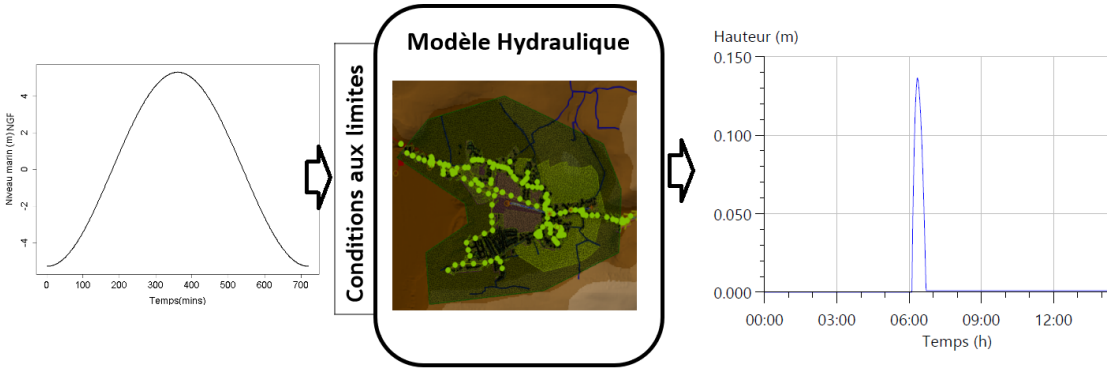


Figure III.1.4: Exemple de propagation d'un scénario (ici un scénario simple de niveau marin est présenté)

Un des points forts de la méthode proposée dans cette thèse est qu'elle prend en compte tout un ensemble de scénarios simples et composés. Comme vu dans les sections précédentes, les scénarios sont construits à partir de pluies et de niveaux marins extrêmes mais aussi fréquents.

Cas de non-coïncidence

Dans ce premier cas de figure, les scénarios de pluie et de niveau marin présentés dans le tableau III.1.1 (figure III.1.1 du chapitre précédent) sont propagés. La figure III.1.5 montre les sorties de modèle obtenues en propageant tous ces scénarios.

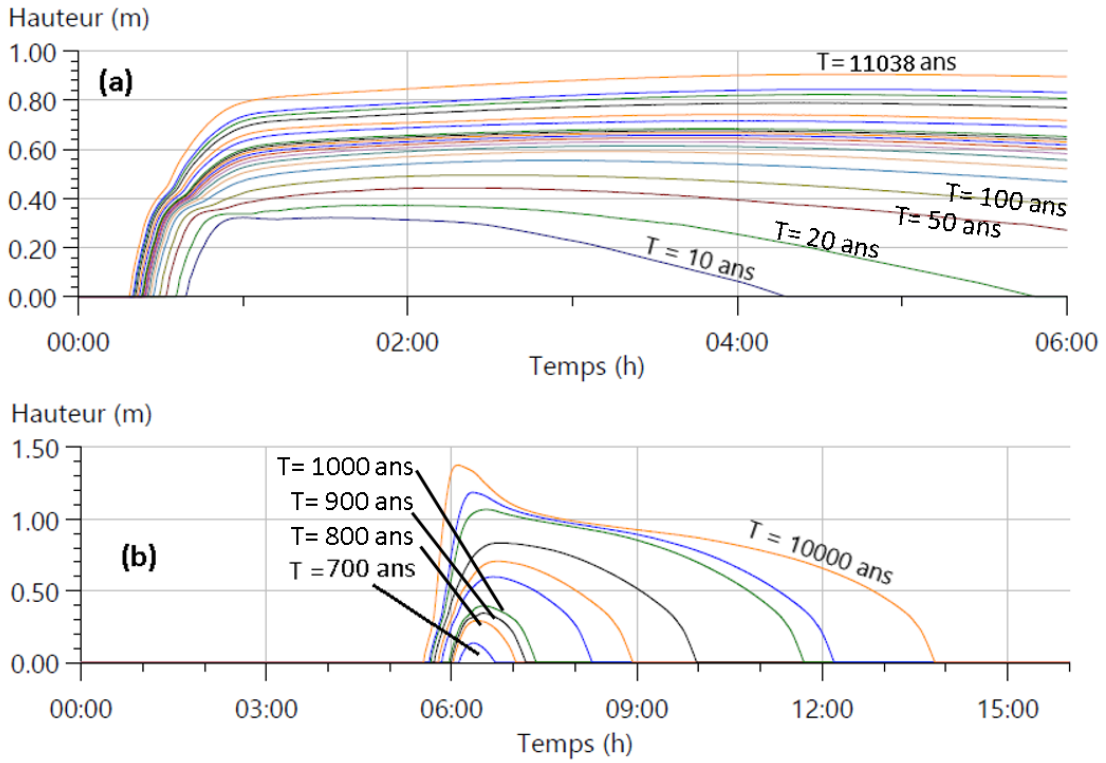


Figure III.1.5: Hauteurs d'eau issues de la propagation des scénarios simples de pluie (a) et de niveau marin (b)

Les courbes de la figure III.1.5 (a) montrent la réponse du modèle hydraulique suite à l'injection des scénarios de pluie de forme triangulaire simple et de durée $D1 = 1\ h$.
L'eau commence à s'accumuler au point d'intérêt au bout de 30 minutes, c'est-à-dire aux alentours des pics d'intensité des scénarios. Cela signifie que le temps de concentration du réseau (qui est le temps le plus long que met une goutte d'eau sur le sol pour atteindre l'exutoire du réseau si on suppose que la pluie est uniformément répartie) est d'environ 30 minutes. Il est important de noter que la durée de pluie doit être supérieure ou égale à ce temps de concentration pour qu'il y ait des débordements du réseau. Or, une pluie décennale de durée 1 h produit une réponse non nulle. La durée de pluie d'une heure est alors adéquate pour ce réseau d'évacuation d'eaux pluviales.
Il peut être constaté que pour des scénarios de plus en plus rares (figure III.1.5 (a)), le réseau d'assainissement met beaucoup plus de temps à évacuer l'eau de pluie. Une pluie décennale (courbe bleue tout en bas) met environ 3h pour s'évacuer. Cela confirme que ce réseau d'assainissement est correctement dimensionné. Il est toutefois nécessaire de rappeler que la forme de la pluie de projet compte beaucoup dans la réponse du réseau, une pluie double-triangle (classiquement utilisée dans

le dimensionnement) aurait peut-être mis moins de temps à s'évacuer.
A partir du scénario de pluie cinquantennale (troisième courbe du bas vers le haut sur la figure III.1.5 (a)), l'eau s'accumule pendant plus de 6 h au point d'intérêt (palier), le réseau d'assainissement est alors quasi-dysfonctionnel.

Les courbes de la figure III.1.5 (b) montrent la réponse du modèle hydraulique suite à l'injection des scénarios de niveau marin.
Le premier scénario à inonder le point d'intérêt est de période de retour 700 ans. Cela est dû à la présence de murets entourant le quartier Saint-François (cf. figure III.1.6), dont la côte moyenne est d'environ 5.2 m NGF.
On peut constater sur la figure III.1.6 que la partie du niveau marin qui dépasse la côte moyenne des murets du quartier Saint-François ne dure qu'une heure trente minutes environ pour le scénario de période de retour 10000 ans et à peu près 40 minutes seulement pour le scénario de période de retour 700 ans. Il en résulte que les volumes d'eau provenant des niveaux marins de petites périodes de retour sont faibles et donc relativement vite évacués par le réseau d'assainissement. Les scénarios correspondant aux périodes de retour élevées (à partir de 2000 ans voire 3000 ans) causent une saturation du réseau, l'eau met donc beaucoup plus de temps à s'évacuer (ce qui explique la présence de paliers).
Les hauteurs d'eau maximales sont atteintes quasiment en même temps que les pics des scénarios de niveaux marin (à 6 h). Cela est dû au fait que le point d'intérêt se trouve à proximité de la mer. Ces pics de hauteurs d'eau sont de plus en plus proches de 6 h pour les scénarios de périodes de retour élevées (10000 ans par exemple). En effet, les volumes de déversements sont élevés pour ces scénarios rares, cela se traduit notamment par des débits importants qui font que le réseau d'évacuation est de plus en plus vite saturé.

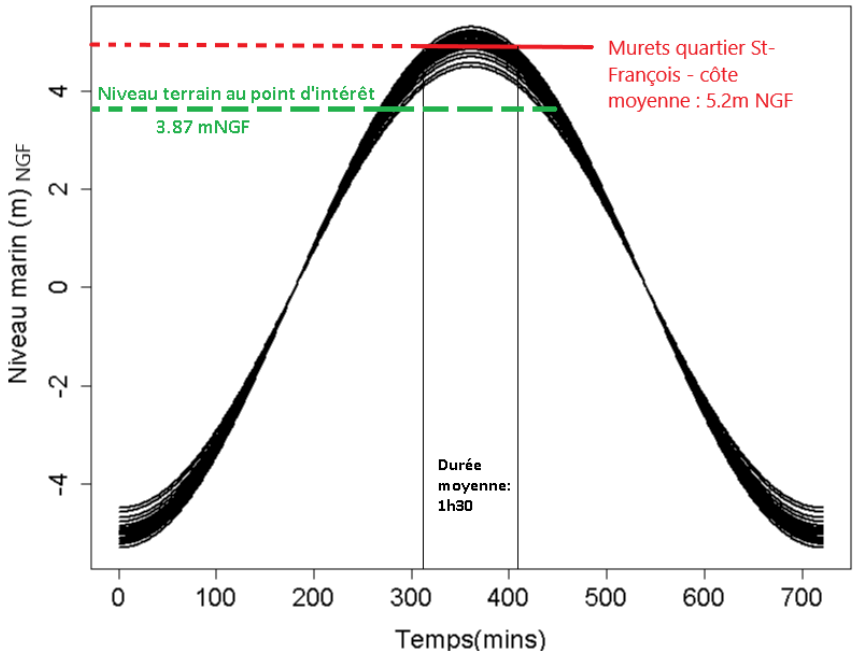


Figure III.1.6: Ligne de côte du quartier Saint-François - scénarios de niveau marin

Cas de coïncidence

Dans le cas de coïncidence, on combine 14 scénarios de pluie avec 13 scénarios de niveau marin, au total 2366 scénarios composés de couples pluie et niveau marin pour chaque valeur de déphasage $\delta = \{-6\ h, -5\ h, -4\ h, -3\ h, -2\ h, -1\ h, 0\ h, +1\ h, +2\ h, +3\ h, +4\ h, +5\ h, +6\ h\}$ sont propagés. La figure III.1.7 montre les sorties de modèle obtenues en propageant les scénarios avec seulement le déphasage $\delta = -6\ h$.

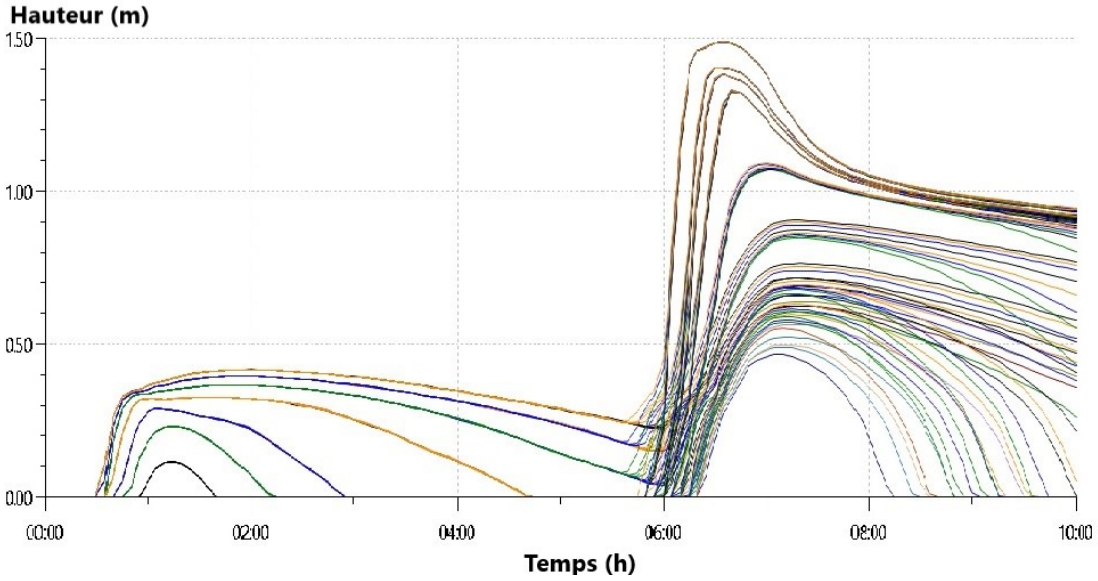


Figure III.1.7: Hauteurs d'eau issues de la propagation des scénarios composés de pluie et de niveau marin déphasés de $\delta = -6\ h$

Les courbes de la figure III.1.7 montrent la réponse du modèle hydraulique suite à l'injection de 182 scénarios composés de pluies (de forme triangulaire) et de niveau marin (de forme sinusoïdale) déphasés de $\delta = -6\ h$.
On peut noter la présence de deux ensembles de pics, le premier ensemble représente les pics de hauteurs d'eau provoqués par la pluie et le second représente les pics provoqués par le niveau marin. En effet, le réseau d'évacuation des eaux pluviales a eu suffisamment de temps (le temps séparant deux pics de pluie et de niveau marin étant de 6 h) pour évacuer l'eau provenant des pluies.
On peut compter seulement 7 pics dans le premier ensemble (de la pluie), alors que le nombre de scénarios est de 182. Cela montre que plusieurs scénarios donnent lieu à une même réponse du réseau. En effet, lorsque la pluie est combinée avec le niveau marin, le comportement du réseau n'est pas le même que lorsqu'on a de la pluie seulement. Bien qu'il n'y ait pas de volumes de niveau marin franchissant la côte du quartier Saint-François (pour le premier ensemble de pics), la combinaison de niveaux marins (extrêmes ici) impacte fortement la réponse du réseau de drainage. Cela confirme encore une fois l'intérêt de considérer la coïncidence de la pluie et du niveau marin.

III.1.3 Conclusion

La première section de ce chapitre a été dédiée aux scénarios d'inondation sous les deux hypothèses de non-coïncidence et de coïncidence. La seconde section a d'abord décrit le modèle 1D-2D du Havre qui a permis de propager les scénarios simples et combinés.
Les résultats de la propagation des scénarios ont montré dans un premier temps que le temps de concentration du réseau est d'environ 30 minutes. Il a été vérifié que le réseau d'évacuation du Havre est correctement dimensionné avec une pluie décennale. Toutefois, la forme de triangulaire simple utilisée pour les scénarios de pluie doit être prise en compte dans l'interprétation des résultats. Il a été également mis en évidence l'existence d'intéractions entre la mer et le réseau 1D (condition aval).
Ces résultats seront traduits dans le prochain chapitre par des courbes d'aléa qui seront par la suite agrégées.

Chapitre III.2

Agrégation - courbe d'aléa finale

Introduction

L'agrégation permet d'apporter aux décideurs une information plus complète sur les périodes de dépassement des hauteurs d'eau (ou d'autres variables d'intérêt) en un point d'intérêt dans le site. C'est une étape clé et constitue notamment un des apports de cette thèse. Cette étape permet d'obtenir une probabilité statistique unique et agrégée (totale) à partir de plusieurs probabilités et de plusieurs phénomènes pouvant être dépendants ou pas, coïncidants ou pas.

Une première section présente une méthode d'agrégation, assez générale mais précise, portant sur un cadre mathématique et logique. Une déclinaison de cette méthode sous les hypothèses de non-coïncidence et de coïncidence fait l'objet d'une deuxième section.

III.2.1 Méthode d'agrégation

Il est rappelé qu'une variable d'intérêt est choisie en fonction des besoins et des intérêts scientifiques de la suite des travaux. Dans cette thèse, la variable d'intérêt est la hauteur d'eau maximale au point d'intérêt.

- $X1$ est une VA décrivant les hauteurs d'eau maximales issues des scénarios de pluie ;
- $X2$ est une VA décrivant les hauteurs d'eau maximales issues des scénarios de niveau marin ;
- Z_{δ_i} est une VA décrivant les hauteurs d'eau maximales issues d'un ensemble de scénarios composés de couples de pluie et de niveau marin déphasés de δ_i ;

III.2.1.1 Cas de non-coïncidence

Soit h une hauteur d'eau quelconque.
On suppose que $Pr\{X1 > h\} = 1 - F_{V1}(h)$ (relation bi-univoque entre le cumul de pluie et la hauteur maximale observée si une seule variable est considérée). De même, $Pr\{X2 > h\} = 1 - F_{V2}(h)$.
Les courbes $h = f(T)$ sont appelées des courbes d'aléa (figure III.2.1). Où $T(h) = \frac{1}{1-F(h)}$, $F(h) = Pr\{X \leq h\}$.
Il peut être noté que la courbe d'aléa liée à la pluie (figure III.2.1 (a)) possède une allure proche de l'allure de la fonction de distribution GPD qui a servi à l'extraction des scénarios de pluie (figure II.1.4 (a)). En effet, cette conservation de la forme de la distribution est potentiellement due au fait que les scénarios de pluie déversent des volumes relativement faibles dans le quartier Saint-François. Ces volumes d'eau sont assimilés par le réseau d'assainissement qui évacue l'eau vers la partie nord du quartier comme indiqué sur la figure III.2.2. Cela signifie donc que le réseau réseau de drainage (1D) est en bon état de fonctionnement. Cela traduit aussi que l'hypothèse faite sur les probabilités de pluie est vérifiée. En effet, une pluie de période de retour T donne lieu à une hauteur d'eau maximale de même période de retour.

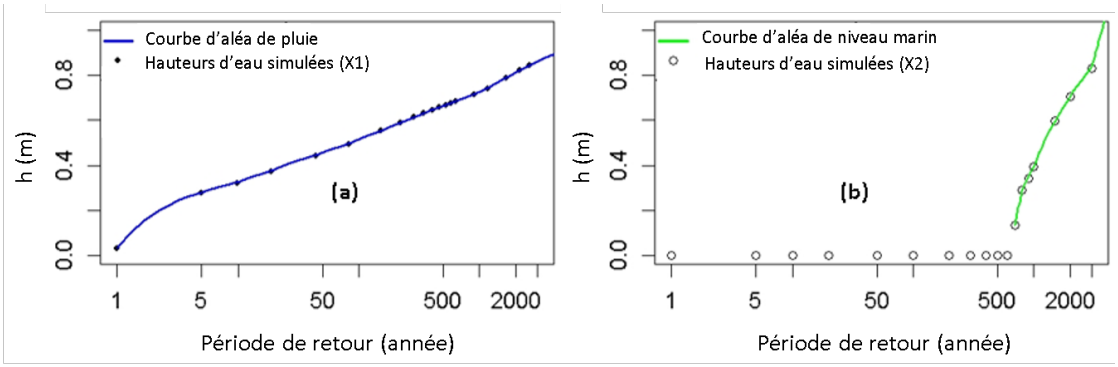


Figure III.2.1: Courbes d'aléa $h = f(T)$ liées à la pluie (a) et au niveau marin (b)

Figure III.2.2: Sens d'écoulement de l'eau dans le réseau 1D du quartier Saint-François

Contrairement à la pluie, la courbe d'aléa du niveau marin (figure III.2.1 (b)) a une allure plus ou moins différente de celle de la fonction de distribution GPD (figure II.1.4 (b)). Cela est dû d'une part, à la présence d'ouvrages de protection (murets) qui font que les scénarios fréquents ne donnent pas lieu à des déversements. D'autre part, cela peut être expliqué par le dysfonctionnement du réseau d'évacuation des eaux pluviales qui n'est pas en mesure de bien assimiler des scénarios rares de niveau marin. Ce réseau est en dépassement de capacité à cause des volumes d'eau importants déversés dans le réseau 1D (figure III.1.5 (b)). Dans ces conditions, uniquement la partie 2D du modèle hydraulique fonctionne, induisant ainsi cet effet de non-linéarité. L'hypothèse de conservation de la distribution pour les scénarios de niveau marin est alors plus ou moins critiquable.
En revanche, les scénarios de pluie même de faible période de retour, donnent tous lieu à une inondation sur le site. Cela pourrait être dû à la forme triangulaire simple des scénarios de pluie, qui est différente de la pluie double triangle (modèle de Desbordes) classiquement utilisé dans la modélisation des pluies de projets.

On cherche par la suite à calculer une probabilité totale de dépasser une hauteur d'inondation h quelconque. Il y aura une partie de cette probabilité totale qui viendra de chacune des 2 courbes d'aléa, parce que h peut être dépassée par la pluie seule OU par le niveau marin seul.

Théoriquement, il faudrait calculer la probabilité totale suivante :

$$\begin{aligned}
P_{totale}(h) &= Pr\{X > h\} = Pr\{(X1 > h) ou (X2 > h)\} \\
&= Pr\{(X1 > h) \cup (X2 > h)\} \\
&= Pr\{X1 > h\} + Pr\{X2 > h\} - Pr\{(X1 > h) \cap (X2 > h)\} \\
&= Pr\{X1 > h\} + Pr\{X2 > h\} - Pr_c \times Pr\{X1 > h\} \times Pr\{X2 > h\}
\end{aligned} \tag{III.2.1}$$

Le troisième terme de l'équation III.2.1 est négligeable. Les termes de probabilités de l'équation III.2.1 sont multipliés par leurs taux annuels d'évènements λ, en vue de les ramener à l'année. L'équation des probabilités totales de dépasser h devient alors :

$$P_{totale}(h) = \lambda_1 \times Pr\{X1 > h\} + \lambda_2 \times Pr\{X2 > h\} - Pr_c \times \lambda_1 \times \lambda_2 \times Pr\{X1 > h\} \times Pr\{X2 > h\} \tag{III.2.2}$$

[1]

où :

- λ_1 : est le taux annuel d'évènements d'intensité de pluie $V1$ (de durée $D1 = 1$)
- λ_2 : est le taux annuel d'évènements de niveau marin maximal $V2$
- $Pr_c = \frac{1}{2 \times 365}$, c'est la pobabilité que deux évènements se produisant la même année, se produisent pendant le même cycle de marée.

Les périodes de retour de la courbe d'aléa agrégée sont calculées par l'équation suivante :

$$\begin{aligned}
T_{totale} &= \frac{1}{P_{totale}} \\
&= \frac{1}{\lambda_1 \times Pr\{X1 > h\} + \lambda_2 \times Pr\{X2 > h\} - Pr_c \times \lambda_1 \times \lambda_2 \times Pr\{X1 > h\} \times Pr\{X2 > h\}}
\end{aligned} \tag{III.2.3}$$

Dans l'application de cette méthode, λ_1 et λ_2 sont issues des analyses fréquentielles univariées effectuées dans le chapitre II.1. On note donc : $\lambda_1 = 5.25$ évènements/an et $\lambda_2 = 1.47$ évènements/an.

Le tableau III.2.1 montre, pour quelques hauteurs d'eau h_i, les calculs des probabilités totales annuelles de dépassement et des périodes de retour totales de dépassement effectués avec les équations III.2.2 et III.2.3.

1. Approximation d'un processus Poissonien

	X1 (Pluie)			X2 (niveau marin)			Courbe d'aléa agrégée		
h_i (m)	$Pr\{X1 > h_i\}$		$T_{X1>h_i}$ (années)	$Pr\{X2 > h_i\}$		$T_{X2>h_i}$ (années)	Pa_{totale} (/an)		T_{totale} (années)
0.03	1.91	10^{-1}	1.58	5.29	10^{-1}	1.85	8.31	10^{-1}	1.20
0.16	1.08	10^{-1}	2.32	9.45	10^{-4}	716.80	4.32	10^{-1}	2.31
0.54	1.49	10^{-3}	128.30	5.15	10^{-4}	1316.29	8.55	10^{-3}	116.99
0.81	9.86	10^{-5}	1933.67	2.46	10^{-4}	2754.80	8.80	10^{-4}	1136.41

Tableau III.2.1: Calcul des probabilités totales annuelles de dépassement (équation III.2.2) et des périodes de retour totales (équation III.2.3) pour quelques hauteurs d'eau h_i - cas de non-coïncidence

Des calculs similaires à ceux du tableau III.2.1 mais plus complets avec un nombre de points plus grand, ont permis d'obtenir la courbe agrégée $h = f(T_{totale})$ présentée en noir sur la figure III.2.3.

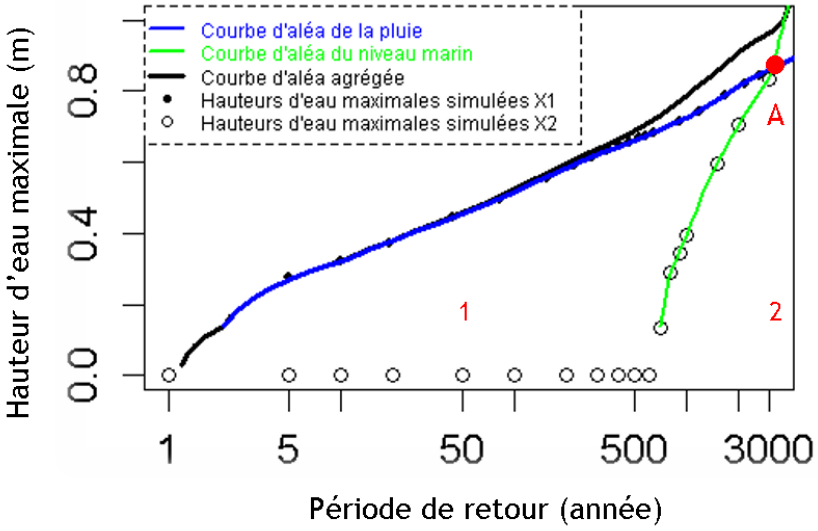


Figure III.2.3: Courbes d'aléa de pluie (bleue), de niveau marin (verte) et agrégée (noire) pour le cas de non-coïncidence

Il est rappelé que la figure III.2.3 ne tient compte que des évènements d'intensité de pluie et de niveau marin non-coïncidants (première hypothèse).
Pour les périodes de retour inférieures à 2300 ans, les hauteurs d'eau sont dépassées plus fréquemment du fait de la pluie (courbe bleue) que par le niveau marin (courbe verte). La courbe d'aléa agrégée (courbe noire) démarre superposée et demeure proche de la courbe des pluies (courbe bleue). En effet, une hauteur d'eau de 32 cm est dépassée plus fréquemment par la pluie (tous les 10 ans) que par le niveau marin (tous les 855 ans). La période de retour que cette hauteur d'eau (32 cm) soit dépassée par la pluie *ou* par le niveau marin est d'environ 10 ans.
La période de retour de dépassement de la hauteur d'eau de 75 cm est la même aussi bien pour le niveau marin que pour la pluie (T= 2300 ans, point de chevauchement des courbes bleue et verte). La période de retour totale est à peu près égale à la moitié de la période de retour de pluie (ou de niveau marin).

Au delà du point A (figure III.2.3, au delà de la période de retour T=2300 ans), la hauteur d'eau est dépassée bien plus fréquemment du fait du niveau marin que par la pluie, la courbe d'aléa du niveau marin passe au dessus de celle de la pluie. La période de retour totale de dépassement devient alors de plus en plus proche de celle du niveau marin.
Le point A est très important dans la compréhension du comportement du réseau. En effet, on a vu dans le chapitre III.1 que le réseau d'assainissement semble être en dépassement de capacité pour tous les scénarios de pluie de période de retour supérieure à 50 ans, et que ce réseau n'est pas impacté par les scénarios de niveau marin de période de retour inférieure à une valeur située entre 2000 ans et 3000 ans. Dans la zone 1 III.2.3, les hauteurs d'eau sont plus fréquemment dépassées par la pluie (la pluie domine) car le réseau d'évacuation est saturé. À la période de retour de 2300 ans (point A) le comportement du modèle hydraulique est identique aussi bien pour la pluie que pour le niveau marin, d'où la même hauteur d'eau. Dans la zone 2, le réseau 1D devient saturé pour les scénarios de niveau marin (cela est remarquable sur la figure III.1.5 du chapitre III.1), les hauteurs d'eau sont alors dépassées plus fréquemment par le niveau marin dans cette zone (le niveau marin domine).
A noter également que pour une même période de retour de dépassement, les niveaux d'eau donnés par la courbe d'aléa agrégée (courbe noire) sont toujours supérieurs à ceux donnés par la pluie seule ou le niveau marin seul (par exemple, pour une période de retour de 1000 ans, on constate une hauteur d'eau de 40 cm provenant du niveau marin, 70 cm provenant de la pluie, et 80 cm de la courbe d'aléa agrégée). Ne considérer qu'un seul phénomène sous-estime donc le niveau d'aléa. Le même raisonnement pourrait être fait sur les périodes de retour. En effet, pour une même hauteur d'eau, la période de retour totale de dépassement est inférieure aux périodes de retour données par la pluie seule ou le niveau marin seul.

III.2.1.2 Cas de coïncidence

Première étape d'agrégation

La situation est différente pour les scénarios combinés, car de nombreuses combinaisons avec des probabilités conjointes différentes peuvent donner lieu à une hauteur d'eau maximale Z_{δ_i} dépassant une même hauteur d'inondation h. Mais comment évaluer la probabilité totale de dépasser h ? Telle est la question qui sera traitée dans ce qui suit.

La figure III.2.4 représente tous les couples d'intensité de pluie-niveau marin. Ces couples ont été injectés dans le modèle hydraulique en considérant à chaque fois un déphasage différent et des hauteurs d'eau simulées Z sont obtenues.
La fonction "contourLinesR" du package $\{contourR\}$ développé sous l'environnement R, permet, en lui donnant les couples de la figure III.2.4 et les hauteurs d'eau simulées Z, de réaliser des interpolations 3D (intensité de pluie-niveau marin-hauteur d'eau) en utilisant un maillage de Delaunay afin de calculer les contours d'iso-hauteurs d'inondation de la figure III.2.5 (seulement 3 contours correspon-

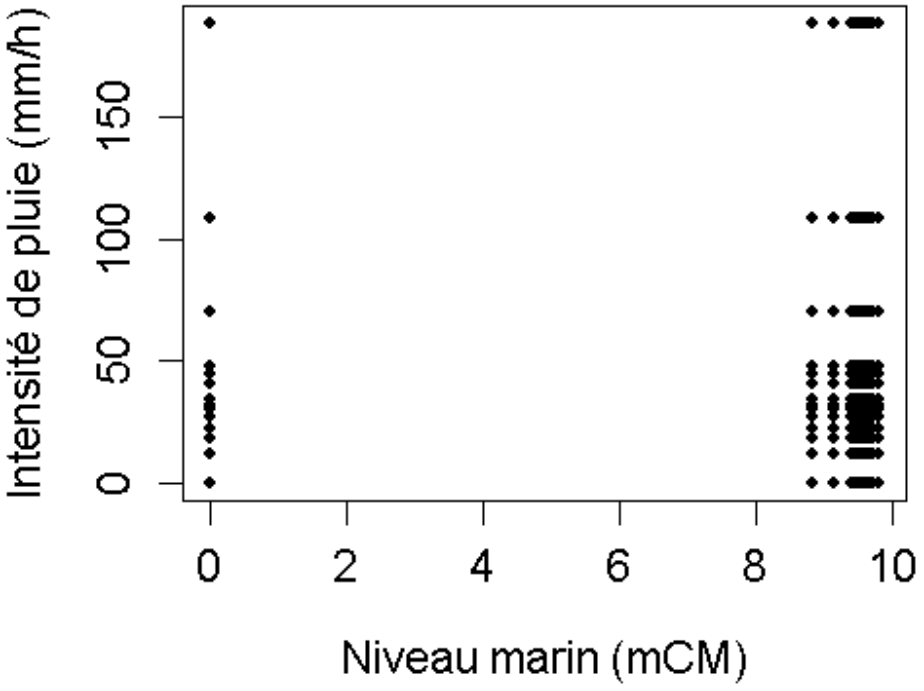


Figure III.2.4: Couples d'intensité de pluie-niveau marin

dant à trois hauteurs d'inondation $h = \{0.6, 0.5, 0.4\}$ m sont tracés).

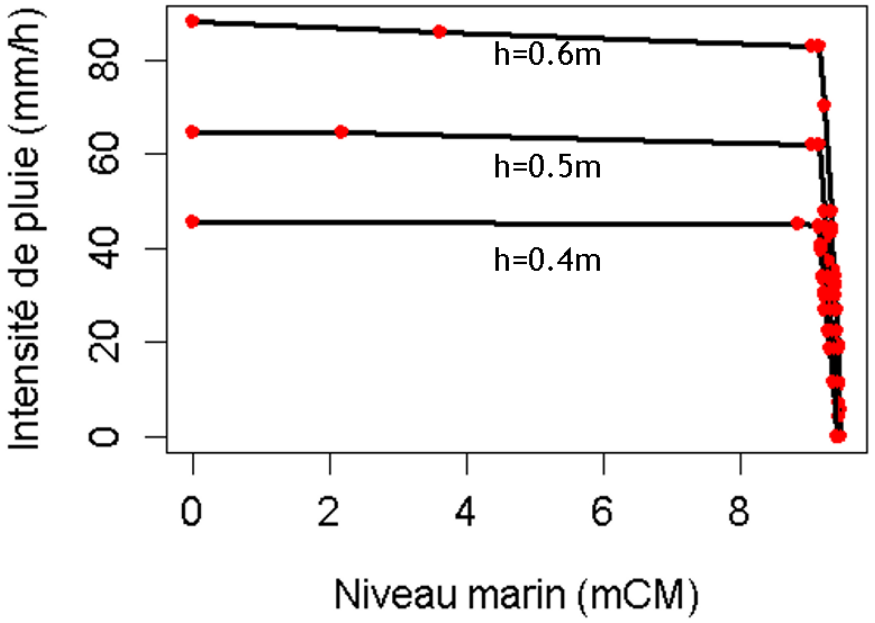


Figure III.2.5: Contours d'iso-hauteurs d'inondation h

On cherche dans la suite à agréger tous ces contours de la figure III.2.5 afin d'obtenir une courbe d'aléa. On considère dans un premier temps un seul contour comme illustré sur la figure III.2.6. La probabilité totale de dépasser la valeur h de ce contour est donnée par l'intégrale double sur Ω, qui est le domaine de l'espace $\{X, Y\}$ tel que $Z(x, y) > h$ (aire grisée). Cela se traduit par l'équation III.2.4.

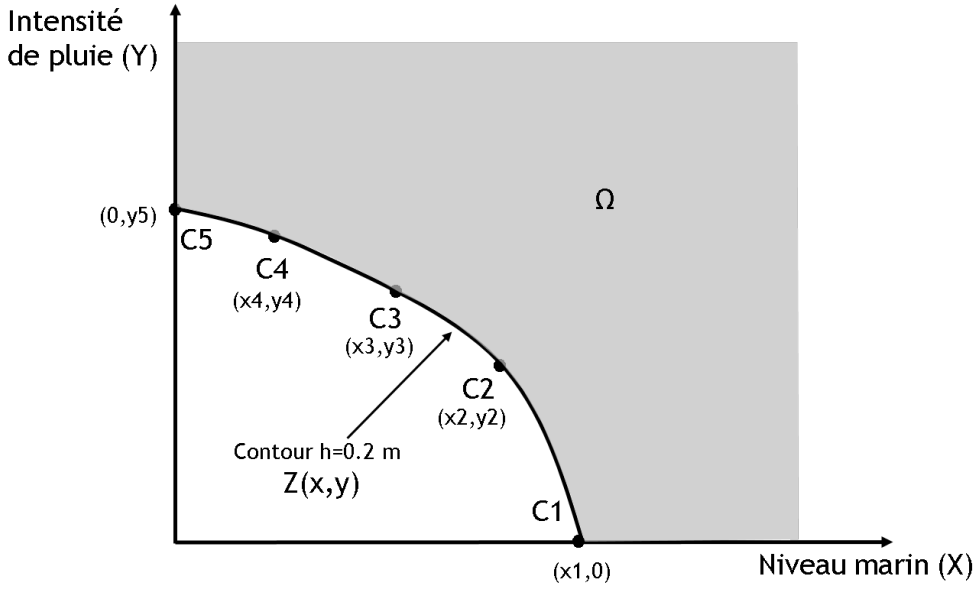


Figure III.2.6: Calcul de la probabilité totale de dépasser h avec la méthode de l'intégrale

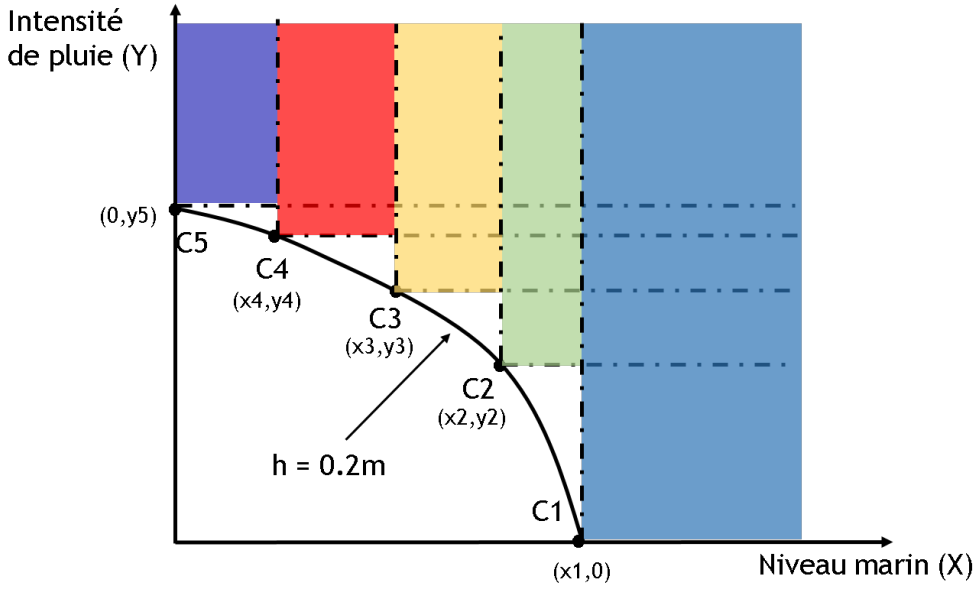


Figure III.2.7: Méthode graphique de construction d'une courbe d'aléa à partir des contours d'iso-hauteurs d'eau : les points C_j représentent les couples pluie-niveau marin simulés à travers le modèle hydraulique et donnant lieu à des hauteurs d'eau maximales Z

$$Pr\{Z > h\} = \int\int_{\Omega} Z(x, y)dxdy \tag{III.2.4}$$

où $Z(x, y)$ est la densité conjointe du niveau marin et de l'intenstié de pluie. Classiquement dans la littérature (e.g., [41, 73]), une méthode graphique est utilisée pour approcher le calcul de cette intégrale en sommant les probabilités associées aux rectangles colorés sur la figure III.2.7. Nous avons démontré dans ce travail de thèse que cette méthode graphique peut être formalisée en considérant l'équation suivante :

$$Pr\{Z > h\} = \sum_{k=1}^{n} Pr\{X > x_k \cap Y > y_k\} - \sum_{k=2}^{n} Pr\{X > x_{k-1} \cap Y > y_k\} \quad \text{(III.2.5)}$$

Les termes de probabilités conjointes de l'équation III.2.5 sont calculées suivant le tableau II.2.4 (étape d'échantillonnage), ce sont des probabilités rapportées aux cycles de marées. Dans cette première étape d'agrégation, ces probabilités conjointes sont remultipliées par $\frac{1}{Pr_c} = 2 \times 365$, qui est la probabilité que les deux évènements se produisant au cours de la même année, se produisent lors du même cycle de marée. L'équation III.2.5 devient alors :

$$Pr\{Z > h\} = \frac{1}{Pr_c} \times (\sum_{k=1}^{n} Pr\{X > x_k \cap Y > y_k\} - \sum_{k=2}^{n} Pr\{X > x_{k-1} \cap Y > y_k\}) \quad \text{(III.2.6)}$$

Il est important de noter que les probabilités associées à la pluie seule et au niveau marin seul ne sont pas concernées par cette multiplication par $\frac{1}{Pr_c}$. Ces probabilités liées à la non-coïncidence sont représentées par les aires des rectangles blue et violet sur la figure III.2.8. En appliquant l'équation III.2.6 pour les différents contours

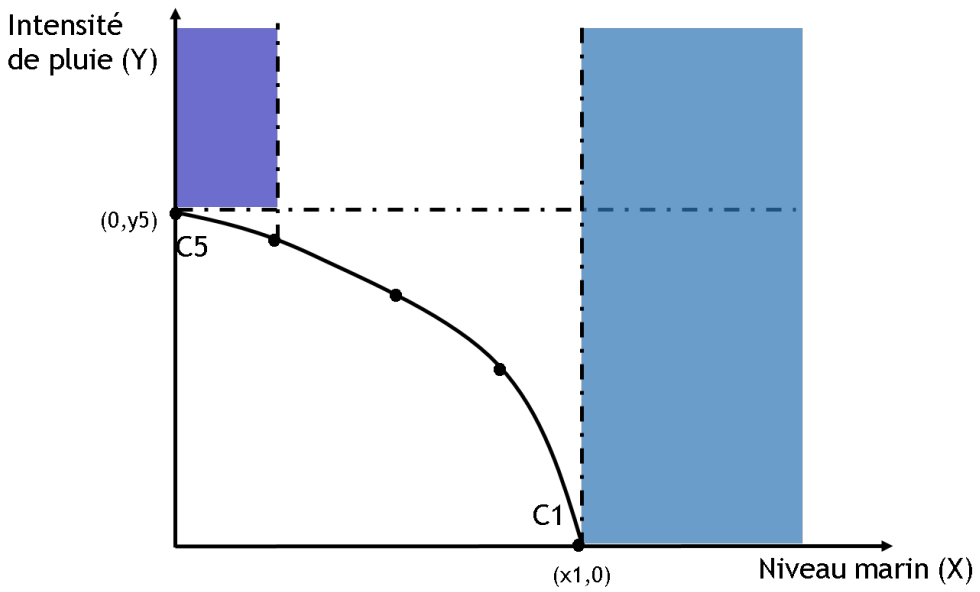


Figure III.2.8: Surfaces liées aux probabilités de non-coïncidence de pluie et de niveau marin

d'iso-hauteurs d'inondation (comme effectué avec le contour de la figure III.2.7). Une courbe d'aléa $h = f(T)$ est obtenue, où $T(h) = \frac{1}{Pr\{Z>h\}}$. La courbe d'aléa de la figure III.2.9 est obtenue pour $\delta = -6\ h$.

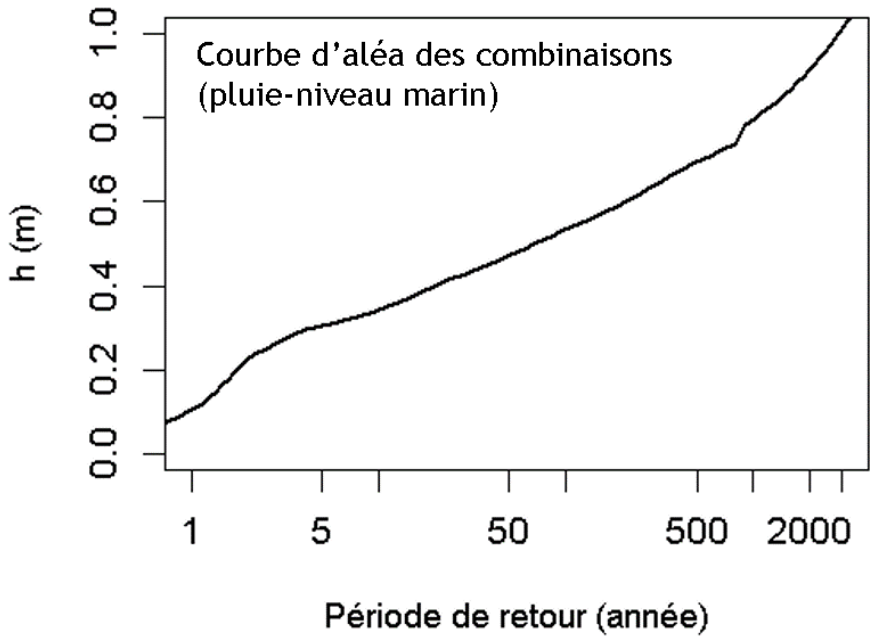


Figure III.2.9: Courbe d'aléa pour $\delta = -6\ h$ (même courbe pour le cas cas d'indépendance et de dépendance)

La même méthode est utilisée pour construire les courbes d'aléa pour le autres déphasages. Nous obtenons ainsi un total de 13 courbes d'aléa pour le cas d'indépendance (figure III.2.10) et 13 courbes d'aléa pour le cas de dépendance (du même type que celles de la figure III.2.10).

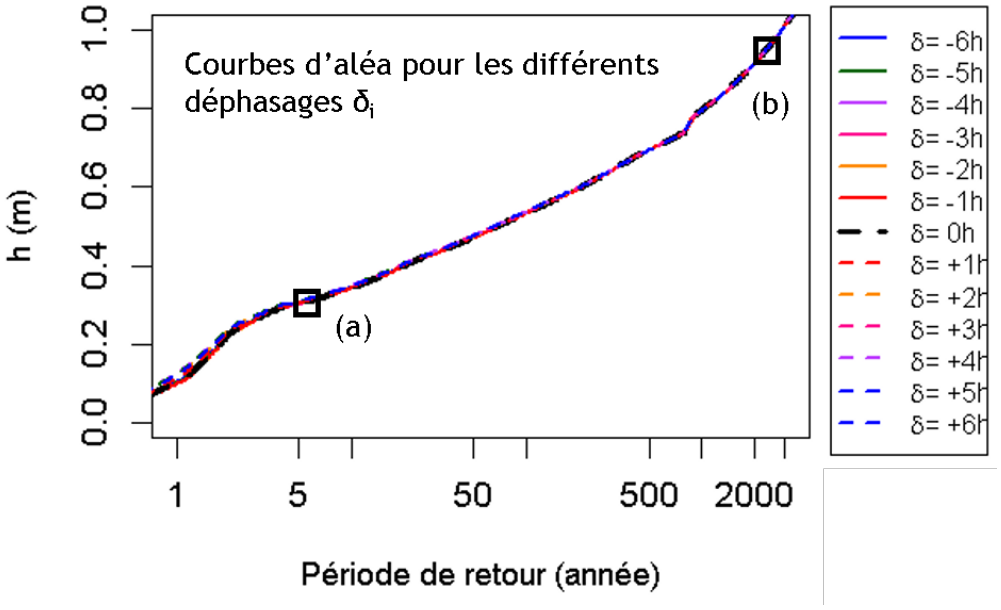


Figure III.2.10: Courbes d'aléa pour tous les déphasages (mêmes courbes pour les cas de dépendance et d'indépendance)

Il peut être noté sur la figure III.2.10 que l'écart entre les courbes d'aléa est très faible. Cela est assuré avec une bonne méthode d'interpolation 3D lors de la construction des contours d'iso-hauteurs d'eau.

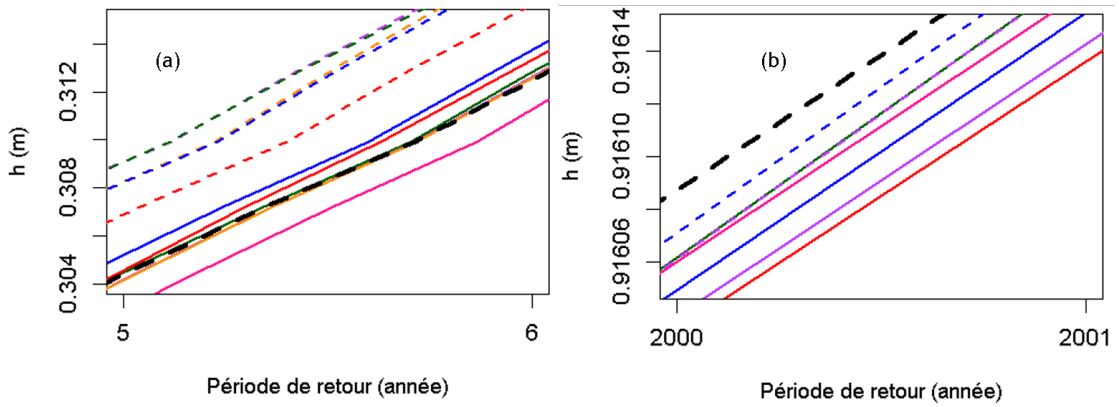


Figure III.2.11: Agrandissements des courbes d'aléa aux périodes de retour fréquentes (a) et rares (b)

On note un léger changement de la pente des courbes d'aléa pour les différents déphasages pour une hauteur d'eau de 0.70 m correspondant à une période de retour d'environ 700 ans. Cela pourrait être dû à la contribution du niveau marin qui devient considérable à partir de cette période de retour.
Les courbes de la figure III.2.10 peuvent être enveloppées par deux courbes d'aléa correspondant à $\delta = [-3h, +3h]$ aux faibles périodes de retour et à $\delta = [-1h, 0h]$ aux fortes périodes de retour. Cela signifie qu'aux faibles périodes de retour, la coïncidence induit des effets relativement importants uniquement si la pluie survient dans l'intervalle $[-3h, +3h]$ par rapport au moment du niveau marin maximal, et dans l'intervalle $[-1h, 0h]$ aux fortes périodes de retour.

Seconde étape d'agrégation

Dans le cas de la coïncidence, on a $n = 13$ courbes d'aléa, chacune correspondant à un déphasage δ_i différent entre les VAs $V1$ et $V2$. On cherche alors à calculer une probabilité totale de dépasser une hauteur d'inondation h quelconque, sachant les hauteurs d'eau Z issues des combinaisons de pluie et de niveau marin.

La variable de déphasage δ ne peut pas être égale à δ_i et δ_j $(i \neq j)$ en même temps, ces éléments sont alors disjoints. Les déphasages sont équiprobables $(Pr\{\delta = \delta_i\}_{i \in 1..n} = \frac{1}{n})$. Par conséquence, l'équation des probabilités totales peut être approximée comme suit, en éléminant les termes de probabilités négligeables et en multipliant les probabilités annuelles par leurs taux annuels d'évènements $\lambda_{m_{\delta_i}}$:

$$P_{totale}(h) = Pr\{Z > h\} = \frac{1}{n} \times \sum_{i=1}^{n} \lambda_{m_{\delta_i}} \times Pr\{Z_{\delta_i} > h\} \qquad \text{(III.2.7)}$$

où $\lambda_{m_{\delta_i}}$ est le taux annuel moyen d'évènements de $(V1, V2)$ dont le déphasage est δ_i. Il est important de noter que les probabilités associées à la pluie seule et au niveau marin seul ne sont pas multipliées par $\lambda_{m_{\delta_i}}$ mais par les taux annuels d'évènements issus des calculs sous l'hypothèse de non-coïncidence.

Les périodes de retour de la courbe d'aléa agrégée sont calculées avec l'équation suivante :

$$
\begin{aligned}
T_{totale} &= \frac{1}{P_{totale}} \\
&= \frac{1}{\frac{1}{n} \times \sum_{i=1}^{n} \lambda_{m_{\delta_i}} \times Pr\{Z_{\delta_i} > h\}}
\end{aligned}
\tag{III.2.8}
$$

Dans l'application de cette méthode, les taux annuels moyens d'évènements $\lambda_{m_{\delta_i}}$ sont calculés sur la base de l'analyse fréquentielle POT des marginales univariées sur une période d'observation commune (section II.2.1.1), et ce pour les différents déphasages. Il en résulte les valeurs présentées dans le tableau III.2.2. Ces valeurs représentent des moyennes des taux de pluie et ceux du niveau marin et sont donc différentes de celles présentées dans le tableau II.2.1.

Taux	Valeur (évènement/an)
$\lambda_{m_{\delta=-6\ h}}$	3.07
$\lambda_{m_{\delta=-5\ h}}$	3.21
$\lambda_{m_{\delta=-4\ h}}$	2.97
$\lambda_{m_{\delta=-3\ h}}$	3.40
$\lambda_{m_{\delta=-2\ h}}$	2.83
$\lambda_{m_{\delta=-1\ h}}$	2.54
$\lambda_{m_{\delta=0\ h}}$	2.50
$\lambda_{m_{\delta=+1\ h}}$	3.11
$\lambda_{m_{\delta=+2\ h}}$	2.73
$\lambda_{m_{\delta=+3\ h}}$	3.30
$\lambda_{m_{\delta=+4\ h}}$	2.73
$\lambda_{m_{\delta=+5\ h}}$	3.07
$\lambda_{m_{\delta=+6\ h}}$	2.50

Tableau III.2.2: Taux annuels moyens d'évènements pour les différents déphasages

Les calculs des probabilités totales annuelles de dépassement et des périodes de retour totales de dépassement ont permis d'obtenir les courbes d'aléa agrégées $h = f(T_{totale})$ pour le cas d'indépendance et de dépendance (figure III.2.12).
La courbe d'aléa finale (en noir dans la figure III.2.12) représente la moyenne des autres courbes d'aléa. Elle combine tous ces déphasages avec un même poids ($\frac{1}{13}$). Supposer que la variable δ est déterministe est une façon intuitive de l'intégrer dans les calculs. L'écart entre les courbes d'aléa est très faible dans le cas de cette application. En perspective de ces travaux, il serait intéressant d'associer des probabilités à cette variable δ, en fonction des décalages observés (a priori). Cela permettrait de donner des poids plus significatifs aux déphasages les plus impactant, notamment et comme constaté précédemment les intervalles $\delta = [-3h, +3h]$ (faibles périodes de retour) et $\delta = [-1h, 0h]$ (fortes périodes de retour) mériteraientt un peu plus de poids que les autres déphasages.

La courbe bleue de la figure III.2.13, traduisant l'hypothèse de coïncidence, est superposée avec celle issue de l'hypothèse de non-coïncidence. Cela montre que la coïncidence entre la pluie et le niveau marin est négligeable au Havre. Ce

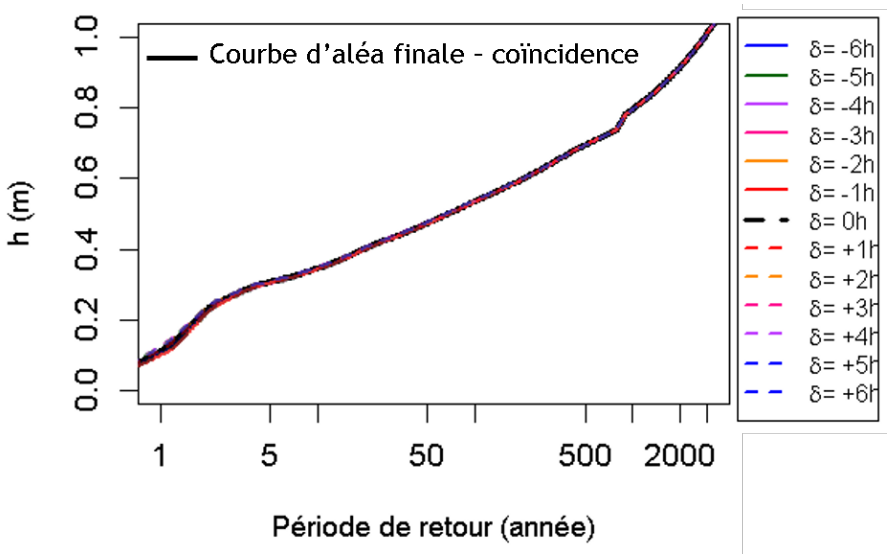


Figure III.2.12: Courbe d'aléa finale (noire continue) et courbes d'aléa d'aléa liées aux différents déphasage (superposées)

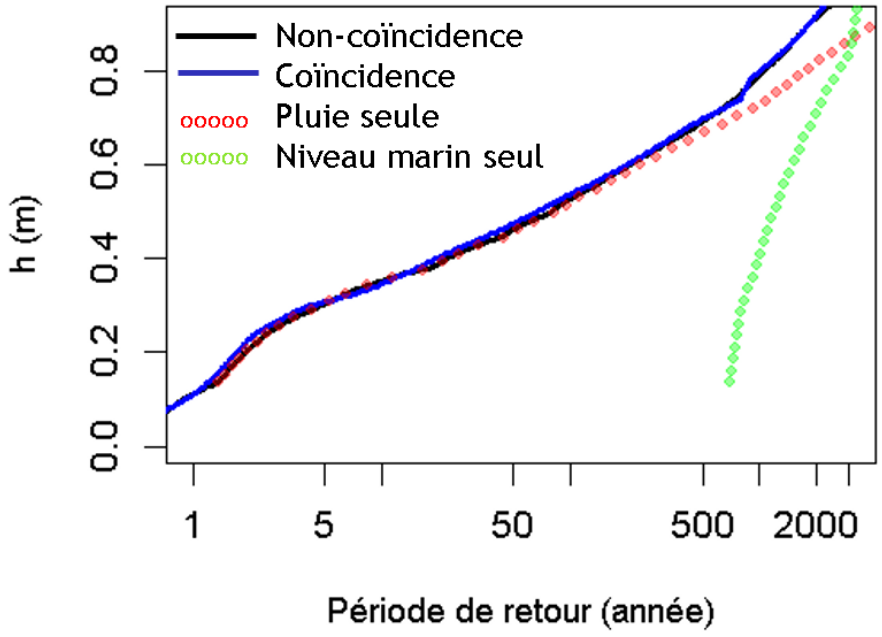


Figure III.2.13: Comparaison des courbes d'aléa finales pour les cas de non-coïncidence (en noir) et de coïncidence (en bleu) ; courbes superposées

résultat est lié à la faible fréquence des évènements provoquant des inondations au Havre, notamment des évènements liés au niveau de mer et donc à la probabilité encore bien plus faible de concomitance de deux événements inondants de pluie et de niveaux de mer lors du même cycle de marée. Cela peut être expliqué par la présence des murets de protection autour du quartier Saint-François empêchant des effets combinés de la pluie et du niveau marin.

Les résultats de la figure III.2.13 peuvent être biaisés pour les faibles périodes de retour (<100 ans). Cela est dû à l'utilisation d'une fonction de distribution de valeurs extrêmes (GPD) pour extraire des scénarios fréquents. On peut contourner cette limite en choisissant une nouvelle fonction de distribution mieux adaptée au besoin ou en construisant une fonction de distribution mixte comme fait dans [97, 98, 86] (composée du GPD pour les valeurs extrêmes et d'une fonction de distribution adaptée pour le corps de la distribution). Par ailleurs, l'extrapolation à des périodes de retour élevées est aussi une source d'incertitudes car la taille des séries de données n'est pas assez élevée.

III.2.2 Conclusion

Nous avons exploité des connaissances à la pointe de l'état de l'art pour mettre en place une démarche permettant de combiner les phénomènes d'inondations dans un cadre d'une méthode probabiliste générale (PFHA). Peu d'idées ont été proposées dans la littérature pour s'attaquer à la coïncidence des phénomènes d'inondation. Le présent travail propose un nouveau concept pour prendre en compte cette notion de coïncidence dans un contexte de dépendance et dans un cadre général de PFHA. Deux hypothèses ont été examinées. La première (non-coïncidence) a montré, de manière non combinatoire, une courbe d'aléa finale (agrégée) synthétisant les courbes d'aléa univariées (classiques). La seconde hypothèse a porté sur un concept plus complexe selon lequel les phénomènes d'inondation peuvent survenir séparément, simultanément ou avec un déphasage temporel.

Les résultats en termes de courbes d'aléa dans les deux cas de coïncidence et de non-coïncidence sont analysés. Globalement, l'application a montré que la pluie est un phénomène dominant pour de faibles périodes de retour, au Havre. Le niveau marin devient dominant aux périodes de retour élevées. De plus, comme indiqué tout au long des parties présentant les résultats (figure III.2.9, par exemple), la dépendance entre la pluie et le niveau marin est négligeable au Havre. La probabilité de coïncidence de la pluie et du niveau marin susceptibles de provoquer des inondations lors du même cycle de marée, est négligeable au Havre. Cela est expliqué par la présence d'ouvrages de protection empêchant les effets combinés des phénomènes.

Les incertitudes sont présentes dans toutes les étapes de ce travail. En effet, la taille des séries de données (indiquée dans les tableaux II.1.1 et II.2.1) n'est pas assez grande pour pouvoir extrapoler à de rares scénarios de pluie et de niveau marin. En outre, le réchauffement climatique a un effet direct sur les phénomènes d'inondation. Il induira globalement une forte élévation du niveau marin et une modification des caractéristiques des tempêtes (modification de l'intensité et de la

fréquence des tempêtes). Cet impact rendra probablement de plus en plus fréquents les évènements extrêmes liés au niveau marin et aux pluies. La coïncidence de ces deux phénomènes et leur dépendance sont également susceptibles d'évoluer sous l'impact du changement climatique. Il est donc suggéré que la question du changement climatique soit essentielle à prendre en compte dans les travaux que nous proposons. Les travaux de la dernière partie de cette thèse permettront d'apporter des détails sur l'évaluation et l'intégration des incertitudes et du changement climatique.

À titre récapitulatif le tableau III.2.3 décrit les méthodes d'agrégation des courbes d'aléa sous les hypothèses de non-coïncidence et de coïncidence.

Non-coïncidence		
Pr(h)	$\lambda_1 \times Pr\{X1 > h\} + \lambda_2 \times Pr\{X2 > h\}$	
Coïncidence		
Étape	Étape 1 (δ_i)	Étape 2 (tous déphasages)
Pr(h)	$Pr\{Z_{\delta_i} > h\} = \frac{1}{Pr_c} \times (\sum_{k=1}^{n} Pr\{X > x_k \cap Y > y_k\} - \sum_{k=2}^{n} Pr\{X > x_{k-1} \cap Y > y_k\})$	$\frac{1}{13} \times \sum_{i=1}^{13} \lambda_{m_{\delta_i}} \times Pr\{Z_{\delta_i} > h\}$

Tableau III.2.3: Tableau récapitulatif des méthodes d'agrégation sous les hypothèses de non-coïncidence et de coïncidence

Quatrième partie

Incertitudes et changement climatique

Chapitre IV.1

Evaluation et intégration des incertitudes et de l'impact du changement climatique

Introduction

Dans de nombreux secteurs applicatifs comme en hydrologie et en modélisation des inondations, il s'agit souvent de fournir des outils d'aide à la décision à partir d'informations observées. La représentation des incertitudes constitue un des principaux challenges auxquels est confronté la communauté scientifique.

Le manque d'informations est parmi les problèmes auxquels sont confrontés les scientifiques dans la modélisation des aléas naturels. La modélisation de ces informations incomplètes repose généralement sur la considération d'un a priori (souvent une loi Gaussienne, comme dans le calcul des intervalles de confiance). En effet, la non-prise en compte de toutes les informations disponibles d'une façon correcte peut engendrer des erreurs conséquentes dans les résultats. La modélisation des informations incomplètes est alors un challenge de taille.

La théorie des probabilités est un cadre général portant sur des expériences aléatoires dont les résultats ne sont pas certains. L'existence de cette théorie et son utilisation assez répandue a empêché la formalisation de la distinction entre les incertitudes aléatoires (irréductibles et dues à la variabilité des phénomènes aléatoires) et les incertitudes épistémiques (réductibles et dues au manque d'informations). Selon [118], il existe différentes lacunes dans la théorie classique des probabilités. En effet, cette théorie ne fait pas la distinction entre le cas d'ignorance totale et celui de l'équiprobabilité ; une distribution de probabilité uniforme est considérée pour représenter les deux cas de figure. D'autre part, une difficulté particulière réside dans le choix de la fonction de distribution des probabilités en cas de manque d'expérience. La théorie classique des probabilités est uniquement adaptée à la représentation des incertitudes aléatoires [9].

D'autres cadres théoriques plus généraux permettent de faire la distinction entre les incertitudes épistémiques et aléatoires [50]. Dans ces théories de l'incertain, les informations sont représentées par une familles de distributions au lieu d'une seule distribution. Cette famille étant d'autant plus grande que l'information est rare. En pratique, le cadre le plus général de représentation des informations incertaines est celui de la théorie des probabilités imprécises. Elle a été introduite par Walley [151] et reste pour le moment une théorie relativement difficile à manipuler. Les probabilités imprécises forment un encadrement d'une distribution de probabilité par deux distributions basse et haute. La théorie des sous-ensembles flous (Fuzzy Sets Theory) développée par Zadeh [161] permet une modélisation pertinente des connaissances imprécises. Cette théorie permet de gérer des informations issues d'un langage naturel. e.g., un objet x est un élément d'un sous-ensemble flou A avec un certain degré d'appartenance. La théorie des possibilités, développée notamment par Dubois et Prade [47], est théorie plus générale que celle des sous ensembles flous. La théorie des fonctions de croyance (connue aussi sous le nom de la théorie de Dempster-Shafer) a été introduite par Dempster [43] et présente un cadre théorique très séduisant souvent présenté comme un bon compromis entre la théorie des probabilités imprécises et la théorie des possibilités. Elle permet dans un même formalisme d'encoder l'imprécision liée aux connaissances imparfaites et l'information incertaine. Ph. Smets a continué à améliorer la théorie développée par Dempster [43] et Shafer [127] à travers le modèle des croyances transférables

[137, 134]. En outre, la théorie des fonctions de croyance permet de combiner plusieurs sources d'informations incertaines. Il existe d'autres concepts relativement moins importants tels que les P-box, les ensembles approximatifs, etc. La figure IV.1.1 récapitule les principales approches. En France, plusieurs chercheurs s'intéressent à ces cadres théoriques et leur applications, e.g., [50, 44].

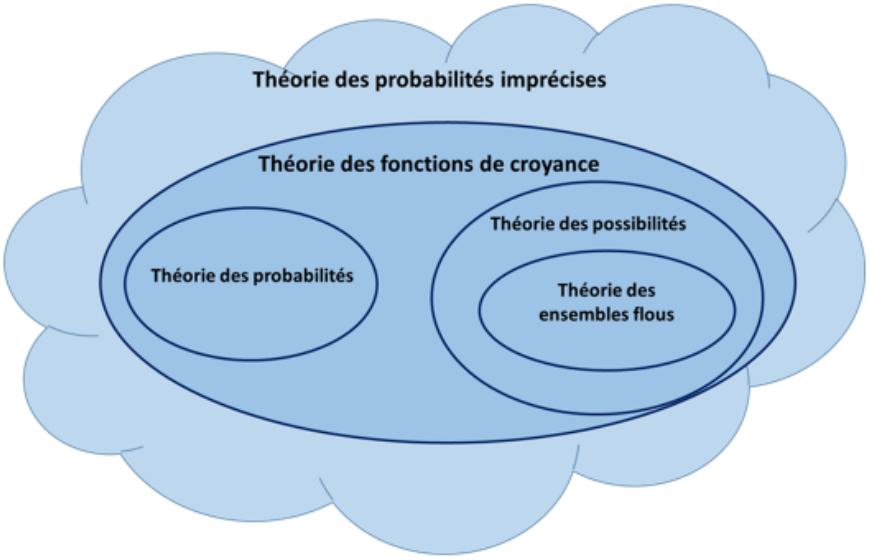


Figure IV.1.1: Principales théories de modélisation des incertitudes

Dans les travaux de cette thèse, les incertitudes sont omniprésentes dans presque toutes les étapes de l'approche proposée (PFHA). Pour cela, la question des incertitudes, leur identification, leur quantification et leur intégration dans les résultats constituent l'objet du présent chapitre. Ce dernier est dédié au formalisme des fonctions de croyance. Ce formalisme est particulièrement intéressant pour la simplicité relative de sa mise en œuvre et sa capacité de gérer à la fois imprécision et incertitudes (aléatoire et épistémique). Ce modèle est suffisamment souple pour être appliqué dans le cadre de cette thèse. Ce chapitre dresse d'abord une revue de la littérature sur les fonctions de croyance. Par la suite, une description du modèle des croyances et de ses fondements permettra au lecteur néophyte de se familiariser avec les fondements mathématiques liés au concept théorique et à son application à différents problèmes. Enfin, cette théorie sera appliquée pour la modélisation de quelques sources d'incertitudes aléatoires et épistémiques liées à l'approche développée dans cette thèse.

IV.1.1 Théorie des fonctions de croyance

La théorie des fonctions de croyance (connue aussi sous le nom de la théorie de Dempster-Shafer) a été introduite par les travaux de Dempster sur les bornes (inférieure et supérieure) d'une famille de fonctions de distributions de probabilités [43] (figure IV.1.2). Ces travaux ont servi à Shafer pour poser et améliorer les fondements de la théorie des fonctions de croyance [127] et montrer ensuite l'intérêt de ces dernières pour la modélisation de l'incertain. Les travaux de Smets [132, 133]

ont montré par la suite l'utilité de ces fonctions de croyance, comme une alternative à la théorie classique des probabilités.

Les applications de la théorie de Dempster sont nombreuses telles que dans l'analyse de données, le diagnostic, etc. Pour plus d'exemple pratiques, le lecteur pourra se référer au guide [131].

Nous présentons d'abord les fondement théoriques des fonctions de croyance. Puis, deux familles de modèles utilisées dans cette thèse seront présentées ; les modèles basés sur l'exploitation des fonctions de vraisemblance, les modèles basés sur les avis d'experts. Ensuite la combinaison des fonctions de croyance dans un contexte de fusion des informations incertaines sera présenté. Enfin, nous détaillerons les méthodes d'exploitation des résultats de modélisation des incertitudes avec les fonctions de croyance.

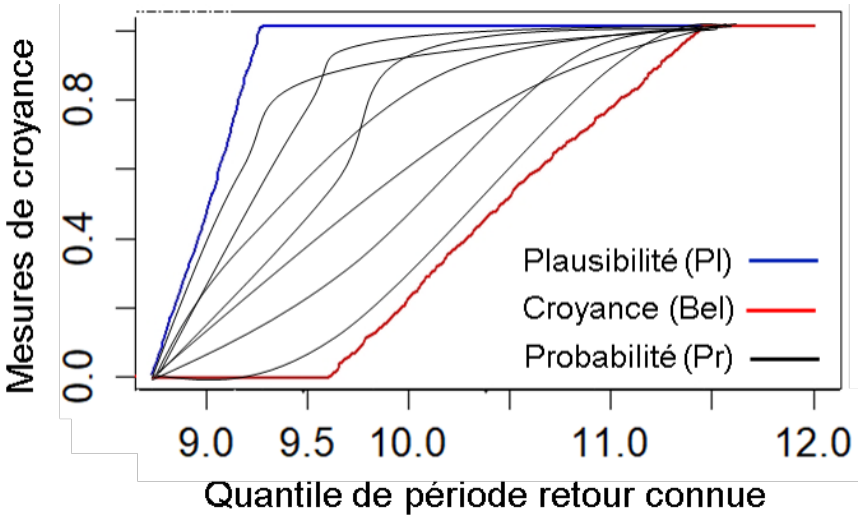


Figure IV.1.2: Bornes d'une famille de fonctions de distributions de probabilités

IV.1.1.1 Fondements théoriques des fonctions de croyance

Les fonctions de croyance représentent le degré avec lequel l'ensemble des informations disponibles plaide en faveur de l'hypothèse selon laquelle la valeur d'un élément est comprise dans un ensemble $\theta \in A$ (A est un sous-ensemble du cadre de discernement Ω), et aucune hypothèse plus spécifique. Des masses (équation IV.1.1) sont affectées à des sous ensembles du cadre de discernement Ω au lieu uniquement des éléments (singletons) habituellement considérés dans la théorie classique des probabilités. Les éléments focaux sont les sous-ensembles du cadre de discernement ayant des masses strictement positives. La modélisation fidèle des informations incertaines constitue un avantage majeur des fonctions de croyance. La fonction de masse est dite Bayésienne et les résultats de la théorie des fonctions de croyance sont ceux de la théorie des probabilités lorsque tous les éléments focaux sont des singletons. La fonction de masse qui a un unique élément focal (singleton) traduit la certitude totale. A l'inverse, l'ignorance totale est codée par une fonction de masse vide dont l'unique élément focal est le cadre de discernement

en entier. Cette modélisation n'introduit pas une information précise comme dans le cas de l'équi-probabilité en théorie des fonctions de probabilité.

$$m : 2^{\Omega} \rightarrow [0, 1] \tag{IV.1.1}$$

Dans la théorie probabiliste, une « masse de probabilité » est affectée à des valeurs ponctuelles (singletons). La théorie des fonctions de croyance affecte des masses de probabilité plutôt à des « éléments focaux ». La somme de toutes les masses doit être égale à 1.

La théorie des fonctions de croyance permet d'utiliser un cadre unique pour traiter des deux sources d'incertitude. Dans cette théorie, la mesure de vraisemblance d'une proposition est qualifiée à l'aide de deux indicateurs : la croyance (notée Bel, comme "Belief") et la plausibilité (notée Pl). La croyance (équation IV.1.2 est la somme des probabilités de tous les éléments focaux qui impliquent forcément la proposition recherchée. Tandis que la plausibilité (équation IV.1.3) est la somme des probabilités de tous les éléments focaux qui ne contredisent pas nécessairement la proposition recherchée. Le couple $\{Bel; Pl\}$ peut s'interpréter comme les mesures de probabilités, respectivement basse et haute, encadrant $Pr(A)$ (équation IV.1.5. Bel et Pl peuvent découler l'une de l'autre à travers l'équation IV.1.4.

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{IV.1.2}$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \tag{IV.1.3}$$

$$Pl(A) = 1 - Bel(A^c) \text{ pour tout } A \subseteq \Omega \tag{IV.1.4}$$

$$Bel(A) \leq Pr(A) \leq Pl(A) \text{ pour tout } A \subseteq \Omega \tag{IV.1.5}$$

où A^c est le complémentaire de A.
L'application concrète de la théorie des croyances requiert plusieurs étapes de modélisation. La définition du cadre de discernement qui dépend de l'application envisagée se fait en premier. Vient ensuite la modélisation des fonctions de masse. La définition de ces masses est un problème crucial qui n'obéit à aucune méthode universelle. Il faut que les éléments focaux soient pertinemment choisis. Le choix des fonctions de masse et surtout l'optimisation des paramètres du modèle se font en second lieu. Ensuite, l'étape de fusion permet de combiner toutes les fonctions de croyance. La dernière étape consiste à choisir un critère de décision (e.g., plausibilité, probabilité pignistique). Ces critères seront détaillés plus loin dans ce mémoire.

Chaque source d'information est donc codée sous forme d'une fonction de croyance et peut provenir d'origines différentes (statistiques, avis d'experts, etc.). Les modèles à mettre en œuvre peuvent varier en fonction du type d'incertitude. Les incertitudes statistiques sont modélisées par des fonctions de vraisemblance alors que les avis d'experts sont modélisés à l'aide d'une fonction de croyance dont la construction est basée sur les possibilités. Ces différents aspects sont présentés dans les sections suivantes.

IV.1.1.2 Modélisation des informations statistiques

L'approche de la vraisemblance est une alternative aux inférences classiques. Cette approche repose sur le principe de la vraisemblance [20] qui stipule que toutes les informations nécessaires à l'évaluation d'une VA d'intérêt sont contenues dans la vraisemblance. Les fonctions de plausibilité et de croyance découlant du principe de la vraisemblance (L) sont données par les équations suivantes :

$$Pl(A;x) = \underset{\theta \in A}{Sup}\ pl(\theta;x) = \frac{\underset{\theta \in A}{SupL(\theta;x)}}{\underset{\theta' \in \Omega}{SupL(\theta';x)}} \tag{IV.1.6}$$

$$Bel(A;x) = 1 - \underset{\theta \notin A}{Sup}\ pl(\theta;x) = 1 - \frac{\underset{\theta \notin A}{SupL(\theta;x)}}{\underset{\theta' \in \Omega}{SupL(\theta';x)}} \tag{IV.1.7}$$

où $\theta \in A$ et $A \subseteq \Omega$.

IV.1.1.3 Modélisation d'autres informations : avis d'experts

Une des principales difficultés dans la construction de fonctions de croyance consiste à modéliser la connaissance disponible en initialisant ces fonctions de croyance de manière adéquate et fidèle à la réalité. Les fonctions de croyance permettent de coder l'ignorance partielle ou totale sous la forme d'une masse [51]. Cela représente un avantage des fonctions de croyance de pouvoir modéliser les informations fidèlement sans introduire un a priori, contrairement au concept probabiliste qui associe souvent une loi Gaussienne ou une loi uniforme aux fonctions de distribution dans certains cas de manque d'informations. Guyonnet et al. (1999) ([68]) ont étudié ces avantages de l'utilisation des fonctions de croyance au lieu des formalismes classiques. En effet, la considération d'un a priori, sans justification avec les informations disponibles, peut sous-estimer les résultats [67]. Les travaux de thèse de Ben Abdallah, N. [13, 14] ont également montré l'importance des fonctions de croyance dans la modélisation et la combinaison des avis d'experts dans un contexte de changement climatique.

IV.1.1.4 Combinaison de fonctions de croyance

L'utilisation des fonctions de croyance présente un intérêt important pour la fusion d'informations [7, 135]. En effet, cette étape de combinaison d'informations permet de fusionner des informations, pouvant être incertaines, incomplètes ou redondantes, afin d'obtenir une information plus fiable et efficace pour les décideurs. Cela permet de diminuer l'incertitude sur une VA d'intérêt en utilisant la redondance et la complémentarité des informations utilisées dans chacune de ces fonctions de croyance.

Plusieurs règles de combinaison propres à la la théorie des fonctions de croyance existent dans la littérature. Ces règles permettant de prendre en compte les interactions entre les différentes sources d'information (fiabilité des sources, dépendance,

conflits, etc.). Le plus couramment utilisé lorsque toutes les sources sont fiables et indépendantes est un mode de combinaison conjonctif, basé sur l'intersection des éléments focaux de toutes les masses (règle de Dempster, équation IV.1.8). Un mode disjonctif permet une combinaison en considérant l'union des éléments focaux des masses élémentaires. Cela est adapté aux cas où au moins l'une des sources (inconnue) est fiable (version disjonctive de la règle de Dempster proposée par Dubois et Prade [46, 49]). La règle disjonctive est souvent très conservatrice et ne permet pas d'aboutir à des résultats informatifs. Une revue des techniques de combinaison existe dans [124, 136].

$$m_{1\cap 2}(A) = m_1 \cap m_2(A) = \sum_{B\cap C=A} m_1(B)m_2(C),\ \forall A \subseteq \Omega \tag{IV.1.8}$$

Dans le cas où l'hypothèse de l'indépendance n'est pas vérifiée, il existe d'autres méthodes (e.g., moyenne simple avec des poids égaux) qui permettront de combiner les sources d'évidence. Une revue de ces méthodes peut être trouvée dans [13].

IV.1.1.5 Propagation d'incertitudes

La propagation des incertitudes dans un modèle est une étape aussi importante que la quantification de ces incertitudes et leur combinaison.

Soit un modèle défini par la fonction suivante :

$$\begin{aligned} &f:\ U_1 \times U_2 \times .. \times U_n \rightarrow V \\ &(X_1, X_2, .., X_n) \rightarrow Y = f(X_1, X_2, .., X_n) \end{aligned} \tag{IV.1.9}$$

où f est la fonction décrivant le modèle, X_i sont les paramètres incertains prenant des valeurs dans U_i, Y est la réponse de ce modèle dont les valeurs sont dans V. Ce modèle peut représenter une fonction de distribution GPD, un modèle hydraulique (avec des équations connues) ou simplement une surface de réponse d'un modèle dont les équations sont inconnues.

On cherche à propager les incertitudes liés aux VAs d'entrée $(X_1, X_2, .., X_n)$ à travers le modèle décrit par la fonction f et à calculer sa réponse Y. Pour cela, on suppose que chacune de ces VAs X_i est codé par un corps d'évidence. Cela signifie que chacune de ces VAs a été modélisée avec la théorie des croyance et qu'on possède une masse m_i associé à chaque VA. Si les VAs X_i sont indépendantes, la masse conjointe m est définie selon Dubois [48] par un produit cartésien aléatoire (équation IV.1.10).

$$m(A) = m_1(A_1) \times m_2(A_2) \times .. \times m_n(A_n),\ \forall A = A_1 \times A_2 \times .. \times A_n \in \Omega \tag{IV.1.10}$$

Le principe d'extension [156] permet de propager le corps d'évidence résultant associé à la masse m dans le modèle f. La nouvelle masse M propagée est donnée par l'équation IV.1.11.

$$M(B) = \sum_{A\in\Omega | f(A)=B} m(A) \tag{IV.1.11}$$

Le degré de complexité des algorithmes permettant d'effectuer ces calculs dépend des propriétés de la fonction f (monotonicité, dimension, continuité, etc.). Si f

est monotone, la résolution peut s'effectuer facilement à l'aide de la méthode du "vertex" [45]. Cette méthode calcule les bornes inférieure et supérieure des éléments focaux conjoints A.

Une méthode de simulation de Monte Carlo pourrait remplacer le principe d'extension de Yager [157] lorsque les fonctions de masse des VAs d'entrée ne peuvent pas être construites à partir des informations disponibles. Les valeurs des VAs d'entrée sont d'abord générées à partir de leur fonctions de distribution (supposées connues) puis ces valeurs sont propagées dans le modèle. Cette opération est répétée jusqu'à avoir un nombre suffisant de réalisations de la réponse (Y) qui servira à la détermination de la fonction de distribution de Y.
Toutefois, l'indépendance des VAs d'entrée est une condition dans les deux méthodes citées (principe d'extension et méthode de Monte Carlo). Si l'indépendance n'est pas vérifiée, les calculs de propagation des éléments focaux deviennent compliqués. Le sujet de la propagation d'incertitudes en cas de dépendance des VAs d'entrée est un sujet qui a fait l'objet de certains travaux. Berleant [21] utilise une méthode de Détermination d'une Distribution Enveloppe (Distribution Envelope Determination ou DEnv). Une méthode de convolution des bornes de dépendance a été proposée dans [155] et une méthode de calcul d'un intervalle de probabilités a été présentée dans [151]. Regan [113] a montré que malgré que ces méthodes ont été construites différemment, elles convergent toutes vers les résultats de la methode de Yager [157] lorsque l'hypothèse d'indépendance est vérifiée. Une revue des techniques utilisées pour prendre en compte la dépendance dans la théorie des fonctions de croyance peut être trouvée dans le rapport de Ferson [57].

IV.1.1.6 Prise de décision dans l'incertain en utilisant les fonctions de croyance

Dans le cadre de la théorie des fonctions de croyance, la dernière étape concerne la prise de décision. Cette prise de décision portant sur le résultat des fonctions de croyance est un sujet assez complexe. En effet, prendre une décision d_i se fait en maximisant un critère. Plusieurs approches existent dans la littérature.

Les fonctions de croyance et de plausibilité construites à partir de la combinaison de toutes les information représentent respectivement la croyance minimale et la croyance maximale. Ainsi, une décision idéale d_i correspond au cas où sa croyance est toujours supérieure à la plausibilité de toute autre décision d_k (équation IV.1.12), ce qui en pratique est rarement vérifié.

$$\forall\ d_k \neq d_i, Pl(d_i) \geq Bel(d_i) \geq Pl(d_k) \geq Bel(d_k) \tag{IV.1.12}$$

La décision doit donc, dans tous les cas, porter sur un critère de choix d'un singleton d_i optimal à partir des fonctions de croyance et de plausibilité. Il serait plus pertinent de garder toute l'information, ou de la résumer de façon synthétique sans cette dernière étape.

Un premier critère consiste à sélectionner un élément d_i correspondant au maximum de plausibilité. Ce critère est optimal pour des fonctions de masse dérivées des probabilités [6]. En effet, quand les informations sont bayésiennes, ce maximum

de plausibilité rejoint une probabilité. Ce critère peut être trop optimiste. Le fait de sélectionner un élément di ayant le maximum de croyance est bien plus sélectif et pessimiste que le critère précédent. Un compromis entre le critère du maximum de plausibilité et celui du maximum de croyance a été proposé par Smets dans [130], où il introduit la probabilité pignistique. Ce nom provient de pignus en latin signifiant pari (bet en anglais). La probabilité pignistique est une mesure permettant d'approcher le couple (croyance, plausibilité) en répartissant les masses des éléments d'un singleton équitablement sur tous ces éléments. La probabilité pignistique est alors définie par :

$$bet(d_i) = \sum_{A \in 2^D, d_i \in A} \frac{m(A)}{|A|(1 - m(\emptyset))} \tag{IV.1.13}$$

où $|A|$ représente le cardinal de A. Le critère qu'on cherche étant la probabilité pignistique maximale qui est donnée par :

$$bet(d_i) = \underset{1 \leq k \leq n}{max}\, bet(d_k) \tag{IV.1.14}$$

Ce critère est donc un critère prudent, bien adapté aux situations ambiguës (manque d'informations, beaucoup d'incertitudes épistémiques, etc.). Certains auteurs, proposent d'abord de collecter d'autres informations afin de réduire les incertitudes épistémiques.
Un exemple d'application peut être trouvé dans [91]. Dans cet exemple, la théorie des fonctions de croyance a été appliquée aux analyses de risques (EPR). Les auteurs ont détaillé plusieurs alternatives de prise de décision, dont certaines ont été liées à un seuil de sûreté. D'autres travaux dans le même sens peuvent être trouvés dans [67, 11, 44].

IV.1.2 Quantification et intégration des incertitudes

Les sources d'incertitudes du PFHA sont multiples. Une partie des incertitudes aléatoires est déjà quantifiée à travers les fonctions de distribution. Toutefois, les incertitudes sur les paramètres de ces fonctions de distribution (GPD) ne le sont pas. Parmi les incertitudes épistémiques on retrouve le choix du modèle statistique, le choix du seuil POT, le choix du pré-seuil d'indépendance, etc. L'objectif principal de cette section est d'appliquer la théorie des fonctions de croyance dans le cadre de la démarche globale du PFHA définie dans les chapitres précédents. Seront donc considérées dans cette application les incertitudes sur les paramètres des fonctions de distribution univariées, les incertitudes sur le choix du seuil POT ainsi que les incertitudes sur l'impact du changement climatique.

IV.1.2.1 Incertitudes statistiques

Dans cette section, la théorie des fonctions de croyance est appliquée pour estimer les incertitudes liées aux paramètres de la fonction de distribution GPD (σ et

ξ). Le paramètre μ représente le seuil POT, il est supposé constant (valeur choisie). L'incertitude sur ce choix de seuil fera l'objet de la section suivante. Comme détaillé dans la section IV.1.1.2, l'approche porte sur le codage des informations contenues dans la vraisemblance de la GPD (équation IV.1.15).

$$L(\sigma,\xi,x)=\prod_{i=1}^{n} g(x_i,\mu,\sigma)\begin{cases}\frac{1}{\sigma}^{n}\times 1+\frac{\xi(x-\mu)}{\sigma}^{-\frac{n}{\xi}-n} \text{si } \xi\neq 0\\ 1-e^{-\frac{x-\mu}{\sigma}^{n}} \quad \text{si } \xi=0\end{cases} \tag{IV.1.15}$$

On cherche à étudier l'impact des incertitudes des paramètres sur un quantile de période de retour donnée x_T. L'expression de la vraisemblance reparamétrisée en fonction de x_T est alors donnée par l'équation IV.1.16.

$$L(x_T,\xi,x)=\prod_{i=1}^{n} g(x_i,\mu,x_T)\begin{cases}\frac{1}{\sigma_T}^{n}\times 1+\frac{\xi(x-\mu)}{\sigma_T}^{-\frac{n}{\xi}-n} \text{si } \xi\neq 0\\ 1-e^{-\frac{x-\mu}{\sigma_T}^{n}} \quad \text{si } \xi=0\end{cases} \tag{IV.1.16}$$

où $\sigma_T=\frac{\xi\times(x_T-\mu)}{\lambda\times T^{\xi}-1}$. La plausibilité étant la vraisemblance normalisée, s'exprimera ainsi :

$$pl(x_T,\xi,x)=\frac{L(x_T,\xi,x)}{\underset{x_T,\xi}{Sup}L(x_T,\xi,x)} \tag{IV.1.17}$$

La variable d'intérêt étant x_T. ξ est alors un paramètre de nuisance qui sera éliminé en considérant le profil de vraisemblance normalisé. On obtient alors l'expression suivante :

$$pl(x_T,x)=\underset{\xi}{Sup}\ pl(x_T,\xi,x) \tag{IV.1.18}$$

L'équation IV.1.18 encode toute l'évidence statistique sur x_T. La figure IV.1.3 représente les courbes de plausibilité obtenues avec l'équation IV.1.18 pour un quantile centennal de niveau marin égal à 9.14 mCM (a) et de pluie égal à 64.30 mm/h (b).

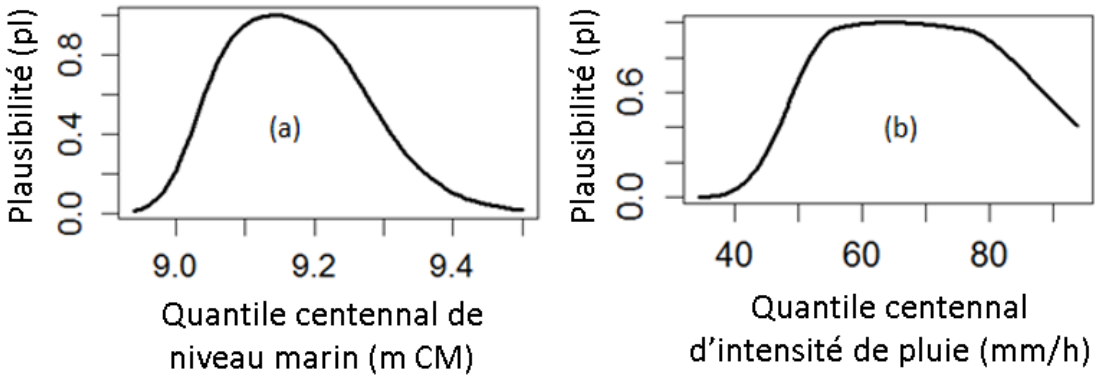


Figure IV.1.3: Courbes de plausibilité pour un quantile centennal de niveau marin (a) et de pluie (b)

IV.1.2.2 Incertitudes liées au choix du seuil POT

La fonction de distribution GPD n'est pas monotone en fonction du seuil (paramètre μ). Par conséquent, la construction d'une plausibilité comme décrit à la section IV.1.1.3 n'est pas possible. En effet, la monotonie de la GPD complique la propagation de la fonction de plausibilité associée au seuil dans le modèle POT en vue d'avoir une fonction de plausibilité exprimée en fonction de la variable d'intérêt qui est x_T. Nous proposons dans la suite une méthode de quantification des incertitudes épistémiques liées au choix de ce paramètre (seuil POT) et de leurs combinaison par la suite avec les autres sources d'incertitudes. L'approche illustrée par la figure IV.1.4 consiste à considérer l'algorithme suivant :

- 1) Calculer ${x_T}^*$ pour la valeur du seuil POT μ^* choisie en se basant sur les variations du quantile et du paramètre de forme de la GPD en fonction du seuil (stabilité des grandeurs) ;
- 2) Considérer une nouvelle valeur du seuil μ_i (autour de μ^* ou dans un intervalle choisi ou simplement une autre valeur qui constitue un doute) ;
- 3) Calculer le quantile ${x_T}^i$ donné par cette nouvelle valeur de seuil ;
- 4) Refaire les étapes 2) et 3) un grand nombre de fois (n = 1000 par exemple) ;
- 5) Calculer les erreurs $\Delta_{{x_T}^{min}} = {x_T}^{min} - {x_T}^*$ et $\Delta_{{x_T}^{max}} = {x_T}^{max} - {x_T}^*$;
- 6) Retenir $\Delta_{{x_T}}{}^{min}$ et $\Delta_{{x_T}}{}^{max}$, ces deux bornes représentent les incertitudes épistémiques sur le choix du seuil POT. Elles seront directement combinées avec les autres sources d'incertitudes, comme détaillé plus loin.

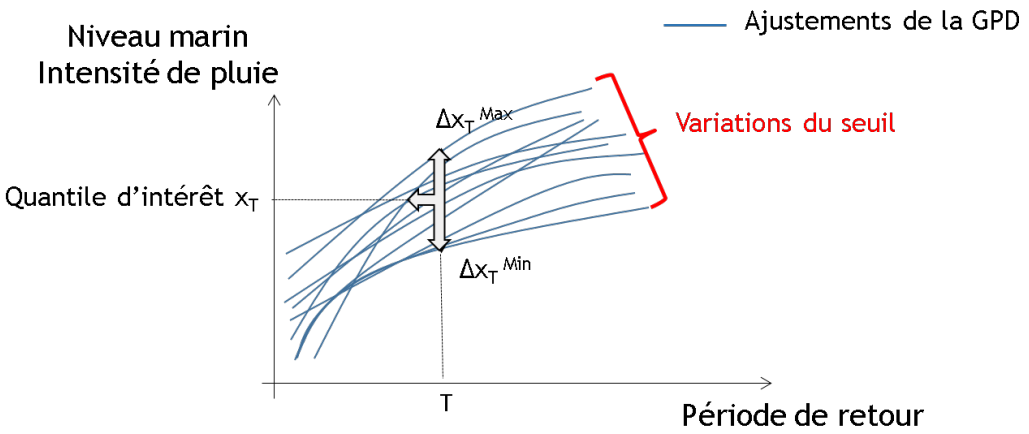


Figure IV.1.4: Illustration de la méthode de calcul de l'incertitudes liées au seuil POT sur un quantile d'intérêt

Cet algorithme a été exécuté avec $n = 1000$ pour estimer les incertitudes sur les quantiles centennaux de pluie et de niveau marin. Les valeurs du seuil ont été prises dans un intervalle autour de la valeur μ^* (choisie initialement). Les bornes obtenues ($\Delta_{{x_T}}{}^{min}; \Delta_{{x_T}}{}^{max}$) sont récapitulées dans le tableau IV.1.1

	Quantile d'intérêt ($x_{T=100\ ans}$)	Variations du seuil	Δx_T^{Min}	Δx_T^{Max}
Intensité de pluie (mm/h)	64.30	[7.76-8.24]	0.01	30.85
Niveau marin (m CM)	9.14	[7.30-8.75]	-0.09	0.15

Tableau IV.1.1: Incertitudes liées aux choix du seuil POT sur les quantiles centennaux d'intensité de pluie et de niveau marin maximal

IV.1.2.3 Incertitudes liées au changement climatique

Le réchauffement climatique a un effet direct sur les phénomènes d'inondation. Il induit avant tout une forte remontée du niveau marin et une évolution des caractéristiques des tempêtes (changement de l'intensité et de la fréquence des tempêtes). Il est fort probable que cet impact du réchauffement climatique fera que des évènements extrêmes de niveau marin et de pluie deviennent de plus en plus fréquents. La coïncidence de ces deux phénomènes et leur dépendance sont aussi susceptibles d'évoluer sous l'impact du changement climatique. Toutefois, les incertitudes sur la vitesse de la montée du niveau marin dans les futures décennies sont élevées.
Une présentation de la notion du changement climatique ainsi que les scénarios considérés par les experts pour le caractériser sont d'abord considérés. En second lieu nous présenterons les impacts de ce changement climatique sur le niveau marin et sur les pluies à travers une revue des avis d'experts, avant de modéliser ces avis avec la théorie des fonctions de croyance.

Scénarios de réchauffement climatique

L'évaluation des scénarios de changement climatique est un sujet d'actualité durant les dernières décennies. Le GIEC, un groupement intergouvernemental d'experts du climat, a montré à travers son rapport de 2018 [80] (édition spéciale) que la température du globe a augmenté de 1°C (par rapport aux niveaux pré-industriels (1860)) suite aux activités humaines. Certains effets de ce réchauffement climatique seront irréversibles (perte de certains écosystèmes, etc.) autour de l'an 2040 si nous continuons à suivre la même trajectoire. Pour éviter cela, l'humanité devra réduire ses émissions de gaz à effet de serre de 45% en 2030 par rapport à 2010 et les annuler d'ici 2050, ce qui constitue un défi mondial de taille en termes de changements industriels, économiques et sociétaux.

Le GIEC définit dans son 5ème rapport de 2014 ([79]) des scénarios de changement climatique représentant les trajectoires d'émissions et de concentrations de gaz à effet de serre, d'ozone et d'aérosols, ainsi que d'occupation des sols baptisés RCP (« Representative Concentration Pathways » ou « Profils représentatifs d'évolution de concentration »). Les 4 profils RCP (RCP2.6, RCP4.5, RCP6 et RCP8.5) remplaçent les 4 anciennes familles (A1, A2, B1 et B2) des anciens rapports. Chacun des 4 RCP est traduit par les différents effets sur l'environnement (niveau marin, température, etc.). Le profil RCP 8.5 est le plus extrême (pessimiste) et le profil RCP 2.6 est une nouveauté importante et sa réalisation implique

la limitation du réchauffement planétaire à 2°C (en 2100).

Plusieurs études alarmantes (e.g., [25]) montrent qu'en suivant la trajectoire actuelle en termes d'émissions de CO2 et d'autres gaz à effet de serre, le réchauffement climatique dépassera un seuil de 4°C d'ici à la fin de ce siècle. Cela suggère de considérer le scénario le plus pessimiste donné par le 5ème rapport d'évaluation du GIEC de 2014 [79]. Le scénario RCP 8,5 correspond à la prolongation des émissions de CO2 actuelles et projette de 2081 à 2100 une fourchette d'augmentation de température possible entre +2,6°C et +4,8°C par rapport aux niveaux pré-industriels (1860). Cependant, la trajectoire en termes d'émissions de CO2 et d'autres gaz à effet de serre dépend beaucoup des politiques actuelles et futures des Etats. Dans ce mémoire, le scénario pessimiste RCP 8.5 est considéré.

Impact du changement climatique sur le niveau marin

Plusieurs experts se sont basés sur les RCP définis dans le rapport du GIEC pour donner leur avis sur l'impact du changement climatique sur la montée du niveau marin. Le tableau suivant montre deux avis d'experts parmi les plus récents, ceux de la National Oceanic and Atmospheric Administration (NOAA) [105] et du GIEC [79].

Montée du niveau marin (SLR)	GIEC		NOAA	
	SLR [m]	Intervalle de confiance [%]	SLR [m]	Probabilités de dépassement [%]
Inf	0.52	5	0.30	100
Inf - intermédiaire	-	-	0.50	96
Intermédiaire	0.74	50	1.00	17
Intermédiaire - sup	-	-	1.50	1.3
Sup	0.98	95	2.00	0.3
Extrême	-	-	2.50	0.1

Tableau IV.1.2: Avis du GIEC et de la NOAA concernant l'impact du changement climatique sur la montée du niveau marin en 2100 (scénario RCP 8.5)

David et Hinkel [145] ont dressé l'état de l'art sur les différentes représentations des incertitudes liées à l'impact du changement climatique sur la montée du niveau marin. La théorie des probabilités imprécises est identifiée comme un bon moyen pour caractériser ces incertitudes. Dans la suite de ce travail, les avis d'experts liés à l'impact du changement climatique sur la montée du niveau marin et sur les prévisions de pluies seront modélisés avec la théorie des fonctions de croyance. L'application des outils de cette théorie permettra de combiner tous ces avis disponibles.

Comme décrit dans la section IV.1.1.3, la construction d'une plausibilité repose sur les masses des corps d'évidence. On cherche, dans un premier temps, à encoder l'avis du GIEC avec la théorie des fonctions de croyance. Le cadre de discernement est $\Omega = [0, +\infty[$. Les informations données par le GIEC peuvent être reformulées

comme suit : en donnant la valeur médiane de la montée du niveau marin (SLR), la croyance est distribuée équitablement entre $[0, SLR_{50\%}]$ et $[SLR_{50\%}, +\infty[$. En donnant la borne supérieure (inférieure) de l'intervalle de confiance 95%, le GIEC raffine sa croyance en chaque sous-ensemble, en affectant 5% de chance que la valeur de SLR dépasse (se trouve en dessous) la borne $SLR_{95\%}$ ($SLR_{5\%}$). Sans aucune hypothèse supplémentaire, ces informations sont exprimées dans le cadre de la théorie des fonctions de croyance de la façon suivante :

— $m([0, SLR_{5\%}]) = m([0\ m, 0.52\ m]) = 0.05$;
— $m(]SLR_{5\%}, SLR_{50\%}]) = m(]0.52m, 0.74\ m]) = 0.45$;
— $m(]SLR_{50\%}, SLR_{95\%}]) = m(]0.74, 0.98\ m]) = 0.45$;
— $m(]SLR_{95\%}, +\infty[) = m(]0.98, +\infty[) = 0.05$.

La courbe en bleu sur la figure IV.1.5 représente la plausibilité obtenue à partir de l'avis du GIEC.

Contrairement à l'avis du GIEC, la NOAA a fournis des probabilités de dépassements associées aux projections de montée du niveau marin. Ces probabilités ont été calculées à partir des hypothèses météorologiques de Kopp et al. [89]. L'utilisation des probabilités pour représenter ce type d'information est désormais discutable. En effet, il n'existe pas une unique fonction de distribution compatible avec l'évidence (les quantiles). Le choix d'un modèle probabiliste particulier est pour le moins arbitraire. D'autre part, le modèle de probabilités affecte un degré de croyance exact à chaque valeur de montée de niveau marin, et donc ne reflète pas le manque de précision dans les connaissances de l'expert. Les modèles alternatifs de modélisation des incertitudes (tels que la théorie des fonctions de croyance) décrivent mieux l'imprécision dans les avis d'experts, sans pour autant introduire des hypothèses subjectives (comme les probabilités). David et Hinkel [145] ont fait cette distinction à travers une revue complète sur les modèles des probabilités précises et les modèles des probabilités imprécises. Baudrit [11] a décrit plusieurs modèles de probabilités imprécises pour représenter les informations incertaines données par des avis d'experts. Toutefois, il existe des méthodes pour prendre en compte ces probabilités comme des informations faisant partie de l'évidence disponible. Dubois et al. ont montré dans [123] comment construire un modèle possibiliste à partir des estimations de probabilités.

La même démarche que celle utilisée pour l'avis du GIEC est appliquée pour l'avis de la NOAA, les informations encodées dans le cadre de la théorie des fonctions de croyance sont :

— $m([0\ m, 0.30\ m]) = 0$;
— $m(]0.30\ m, 0.50\ m]) = 0.04$;
— $m(]0.50\ m, 1.00\ m]) = 0.79$;
— $m(]1.00\ m, 1.50\ m]) = 0.157$;
— $m(]1.50\ m, 2.00\ m]) = 0.01$;
— $m(]2.00\ m, 2.50\ m]) = 0.002$;
— $m(]2.50\ m, +\infty[) = 0$.

La courbe de plausibilité de l'avis de la NOAA est représentée en vert dans la figure IV.1.5.

L'utilisation de la règle de Dempster présentée en section IV.1.1.4 pour combiner les incertitudes doit d'abord satisfaire la condition d'indépendance entre les

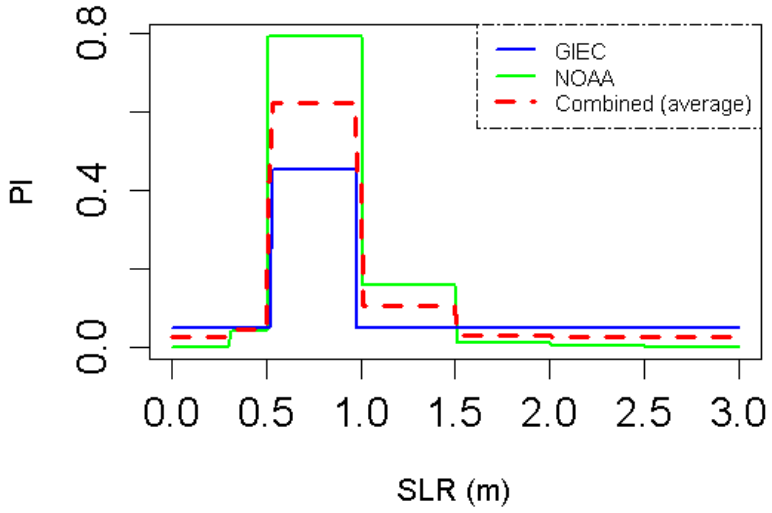


Figure IV.1.5: Plausibilité de la montée du niveau marin portant sur l'avis du GIEC (bleu) de la NOAA (vert) et leur combinaison (rouge)

sources d'évidence. Or dans ce cas, les experts ont tous les deux utilisé le même scénario de réchauffement climatique à savoir le RCP 8.5. La condition d'indépendance n'est donc pas satisfaite. En l'absence d'informations supplémentaires sur les experts, une règle de moyenne simple avec des poids égaux semble être le moyen le plus naturel et approprié pour réaliser cette combinaison. La courbe rouge de la figure IV.1.5 est alors obtenue.

Impact du changement climatique sur la pluie

Plusieurs évènements extrêmes ont eu lieu dans le monde causant d'importants dégâts. Les fortes pluies qui ont touché le département de l'Aude en Octobre 2018, causant la mort d'au moins 14 personnes et faisant plusieurs blessés graves. Les orages de mousson du mois d'août 2018 ont touché le sud-ouest de l'Inde provoquant la mort de plusieurs centaines de personnes et entraînant l'évacuation des populations. Sur les côtes méditerranéennes, des crues d'automne causées par des pluies diluviennes qui se sont abattues sur la péninsule du nord-est de la Tunisie (Cap Bon) en septembre 2018 donnant lieu à d'impressionnantes inondations qui ont fait au moins 4 morts et ont détruit routes, commerces et véhicules par dizaines.

Le lien entre le changement climatique et ces évènements extrêmes est loin d'être facile à expliquer. Le problème est complexe car le changement climatique peut affecter à la fois la fréquence et l'intensité des évènements.

Les ressources dans la littérature sont rares en termes de projections de pluies sous l'impact du changement climatique. Cela est en partie du à une difficulté majeure qui est l'aspect très localisé des précipitations. Ouzeau et al. [108] ont estimé des projections de pluies sous l'impact du changement climatique en France. Ces projections issus des simulations de modèles (Aladin-Climat et WRF) portent sur l'évolution temporelle des précipitations et températures estivales et hivernales. Les simulations s'appuient sur trois des quatre scénarios considérés dans le dernier rapport d'évaluation du GIEC (n°5, [79]). Suite au projet GICC DRIAS,

un portail (http ://www.drias-climat.fr) met à disposition des projections climatiques régionalisées sous différentes formes (observations, indices, représentations cartographiques). Ces dernières traduisent à l'échelle de la France les projections globales données par le rapport du GIEC de 2012.

Mise à part la rareté des informations, les résultats ne permettent pas de dégager une tendance claire à la fin du siècle sur l'évolution des précipitations estivales, les modèles WRF et Aladin-Climat étant globalement incertains sur la tendance du changement [108]. La considération des séries de données de pluies complètes (année entière) constitue une autre difficulté liée aux travaux de cette thèse. Pour toutes ces raisons, l'impact du changement climatique sur les évènements pluviométriques ne sera pas intégré dans les travaux de cette thèse.

IV.1.2.4 Combinaison des incertitudes

On cherche à calculer un nouveau quantile de période de retour T sachant toutes les incertitudes statistiques, liées au choix du seuil POT de la fonction de distribution GPD et à l'impact du changement climatique. En assumant l'hypothèse d'indépendance des sources d'évidence (i.e., les incertitudes ont des origines différentes), le nouveau quantile x'_T sera exprimé ainsi :

$$x'_T = x_T + \Delta_{x_T}{}^{seuil} + \Delta_{x_T}{}^{ChangementClim} \tag{IV.1.19}$$

où x_T, Δ_{x_T} et $\Delta_{x_T}{}^{ChangementClim}$ sont estimés dans les sections IV.1.2.1, IV.1.2.2 et IV.1.2.3, respectivement.

Les fonctions de croyance (Bel) et de plausibilité (Pl) cumulées sont difficiles à trouver analytiquement. Elles peuvent être facilement approximées en utilisant une approche Monte Carlo selon les étapes suivantes :

- Générer N éléments aléatoires $\omega_1, .., \omega_N$ de (Ω, P). Où : $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$ (faisant référence aux 3 sources d'incertitudes) et $P = P_1 \times P_2 \times P_3$ (faisant référence à leur masses) ;
- Pour chaque élément ω_i $(i = 1, .., N)$, générer $\Gamma(\omega_i) = [u_i, v_i]$, où $u_i = U1(\omega_i) + \Delta_{x_T}{}^{min} + U3(\omega_i)$ et $v_i = V1(\omega_i) + \Delta_{x_T}{}^{max} + V3(\omega_i)$, $(U1, U3)$ et $(V1, V3)$ sont les intervalles aléatoires liées aux incertitudes statistiques (x_T), et à $\Delta_{x_T}{}^{ChangementClim}$ respectivement ;
- Pour chaque intervalle I, les quantités $Bel(I)$ et $Pl(I)$ sont estimées par :

$$\hat{Bel}(I) = \frac{1}{N} card(\{1 \leq i \leq N | [u_i, v_i] \subseteq I\}) \tag{IV.1.20}$$

et

$$\hat{Pl}(I) = \frac{1}{N} card(\{1 \leq i \leq N | [u_i, v_i] \cap I \neq \emptyset\}) \tag{IV.1.21}$$

Les courbes de plausibilité et de croyance cumulées pour le quantile centennal du niveau marin sont présentées à la figure IV.1.6 (valeur estimée sans incertitudes=9.14 m CM).

L'écart entre la plausibilité (Pl) et la croyance (Bel) sur la figure IV.1.6 renseigne sur l'importance des incertitudes. Plus cet écart est grand, plus les incertitudes sont élevées. En fait, pour réduire cet écart, l'utilisateur pourra collecter de

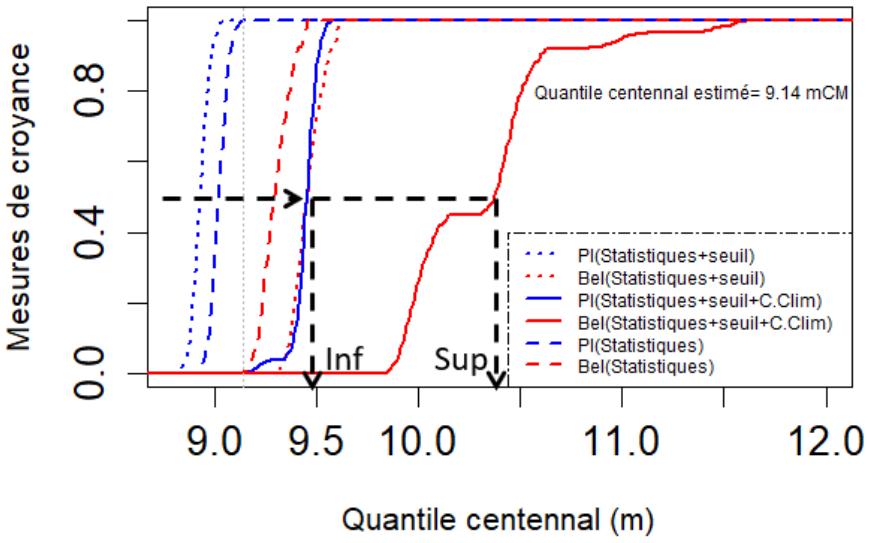


Figure IV.1.6: Plausibilité (bleu) et croyance (rouge) cumulées pour un niveau marin centennal (a) ; incertitudes statistiques uniquement (courbes discontinues) ; incertitudes statistiques combinées avec les incertitudes liées au seuil POT (courbes en pointillé) ; toutes incertitudes combinées (courbes continues)

nouvelles informations afin d'enrichir les corps d'évidence et réduire ainsi les incertitudes épistémiques. Cet espace se trouvant entre une plausibilité et une croyance correspond en effet à la famille des distributions possibles. Comme expliqué dans la partie théorique des fonctions de croyance, la plausibilité et la croyance définissent des bornes supérieure et inférieure, respectivement.

La figure IV.1.6 montre que les incertitudes sur le changement climatique dominent les deux autres sources d'incertitudes (choix du seuil POT et statistiques). En effet, l'écart entre les courbes continues est plus grand par rapport à celui des autres courbes.

L'étape suivante consiste à définir un critère de décision. Ce critère permettra de choisir de nouveaux quantiles qui intègrent les incertitudes. Le choix de ce critère dépend énormément de l'usage des résultats, il n'existe aucun consensus scientifique sur ce sujet. Dans les travaux de cette thèse, on cherche principalement à montrer le côté applicatif de cette méthode dans le cadre du PFHA. De ce fait, on se propose de choisir les quantiles correspondant à 50% de croyance et 50% de plausibilité, comme présenté sur la figure IV.1.6. Cette démarche a été appliquée pour tous les quantiles de niveau marin et de pluie, ce qui a donné lieu à deux bornes (inférieure et supérieure) pour chacune des courbes d'ajustements obtenues dans les chapitres précédents (figure IV.1.7).

La figure IV.1.7 montre que toutes les incertitudes sont négligeables aux faibles périodes de retour, à l'exception des incertitudes sur l'impact du changement climatique sur la montée du niveau marin (courbes noires continues de la figure IV.1.7 (a)).

Les incertitudes sur le choix du seuil POT sont négligeables pour le niveau marin (figure IV.1.7 (a)), les courbes noires pointillées sont très proches des courbes noires discontinues associées aux incertitudes statistiques.

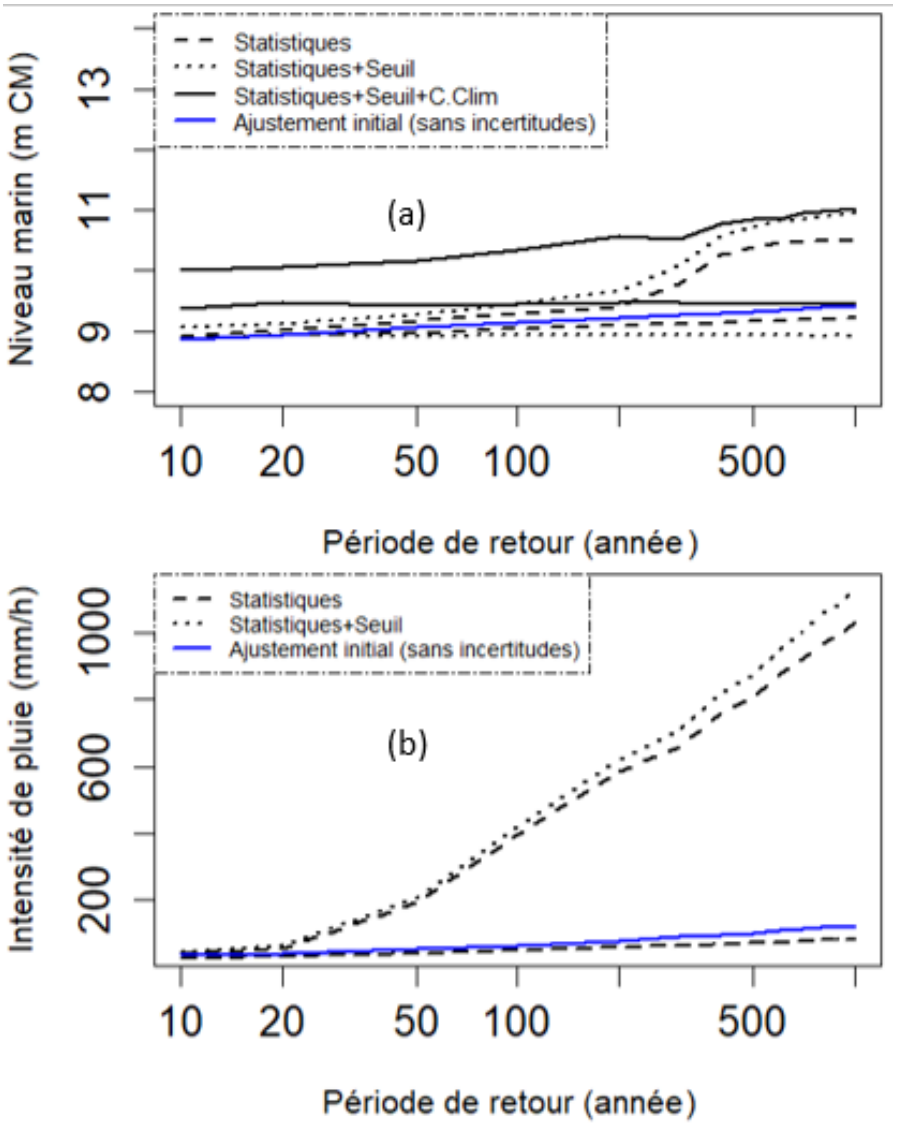


Figure IV.1.7: Ajustements de la fonction de distribution GPD aux données de niveau marin (a) et de pluie (b) ; ajustements initiaux sans incertitudes (en bleu) ; nouveaux ajustements en intégrant les incertitudes statistiques (discontinu) avec les incertitudes sur le choix du seuil POT (pointillé) et avec les incertitudes sur l'impact du changement climatique à l'horizon de 2100 (noire continue) dans le cadre de la théorie des fonctions de croyance (en choisissant les quantiles médians correspondant à 50% de croyance et 50% de plausibilité)

Les incertitudes liées au changement climatiques sont dominantes pour les niveaux marins de faibles périodes de retour (<700 ans). Les incertitudes statistiques et celles liées au choix du seuil POT sont dominantes aux fortes périodes de retour.

Il en va de même pour la pluie, les incertitudes sont principalement d'origine statistique (figure IV.1.7 (b)).

IV.1.2.5 Propagation des incertitudes et courbes d'aléa

La propagation des incertitudes dans un modèle nécessite la connaissance des équations de ce modèle ainsi que leur paramètres. Dans le cadre des travaux de cette thèse, le modèle utilisé est développé sous Infoworks. On ne dispose pas de l'accès au code du logiciel. Une autre difficulté réside dans le couplage du modèle développé sous Infoworks avec le code du PFHA développé avec le logiciel R. Le nombre élevé des scénarios de la démarche du PFHA présente à son une difficulté de taille. Dans le cadre de l'application de cette démarche de prise en compte des incertitudes à travers les fonctions de croyance, uniquement les courbes d'aléa associées aux cas de non-coïncidence seront réévaluées.

Les scénarios de non-coïncidence sont resimulés en intégrant les incertitudes statistiques, celles liées au choix du seuil POT et celles liées au changement climatique de la même façon telle que présentée dans la figure IV.1.7 (courbes noires continues), les résultats en termes de nouvelles courbes d'aléa sont présentés dans les sections suivantes.

Les courbes d'aléa de la pluie et du niveau marin sont présentées dans la figure IV.1.8 (courbes verte et bleue en trait continu). Les intervalles de confiances obtenus à partir de 50% de plausibilité et de 50% de croyance sont présentés en bleu transparent et en vert transparent pour la pluie et le niveau marin respectivement. Ces intervalles de confiance intègrent les incertitudes statistiques et celles liées au choix du seuil POT pour la pluie et les incertitudes statistiques, celles liées au choix du seuil POT et celles liées à l'impact du changement climatique à l'horizon de 2100 pour le niveau marin.

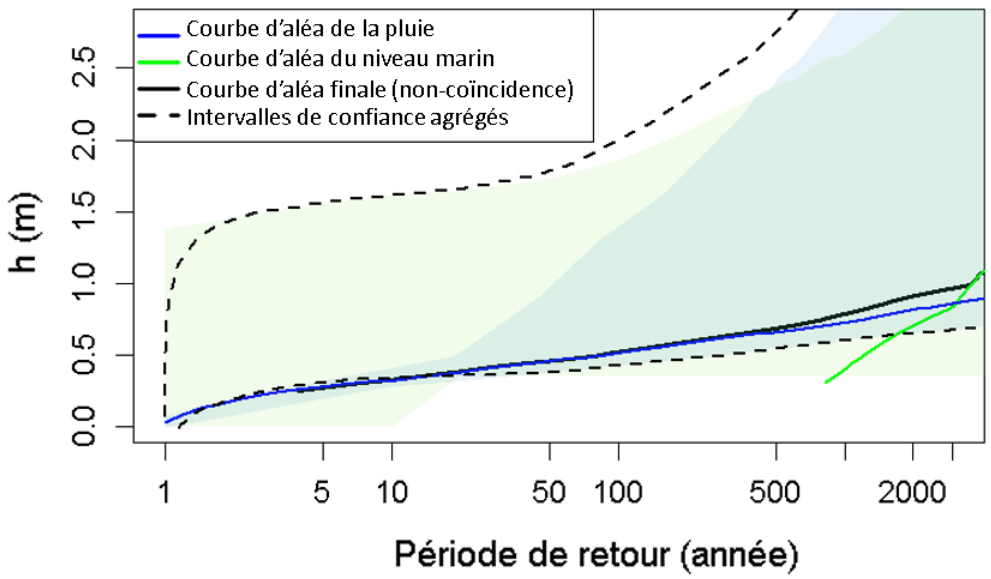


Figure IV.1.8: Courbes d'aléa du niveau marin (vert continue), de la pluie (bleue continue) ; les bornes inférieure et supérieure des intervalles de confiance obtenues à partir de 50% de plausibilité et de 50% de croyance (en bleu et vert transparents) ; la courbe d'aléa agrégée en trait noir continu et les intervalles de confiances agrégés en trait noir discontinu

L'intervalle de confiance associé aux hauteurs d'eaux maximales provenant de la pluie (en bleu transparent sur la figure IV.1.8) s'élargit aux fortes périodes de retour, cela est dû à la propagation des incertitudes statistiques dans le modèle d'analyse fréquentielle (POT) et en particulier à la faible taille de la série des données (8 ans). La largeur de l'intervalle de confiance du niveau marin (en vert transparent sur la figure IV.1.8) est due en grande partie aux incertitudes liées au changement climatique (aux faibles périodes de retour) et aux incertitudes statistiques aux fortes périodes de retour, comme constaté dans la figure IV.1.7 (a) avant même la resimulation des scénarios. L'intervalle de confiance de la courbe d'aléa agrégée (en noir discontinu sur la figure IV.1.8) est calculé avec la même méthode d'agrégation donnée par l'équation III.2.3. Sans incertitudes, une hauteur d'eau maximale de 0.5 m (courbe noire continue sur la figure IV.1.8) est dépassée en moyenne tous les 100 ans. En revanche, sous l'impact de toutes incertitudes combinées, cette hauteur d'eau qui est dépassée tous les 100 ans est d'une amplitude plus grande (1.5 m). La fréquence d'une même hauteur d'eau est aussi augmentée car cette même hauteur d'eau de 0.5 m sera dépassée tous les ans sous l'impact des incertitudes.

IV.1.3 Conclusion

Cette dernière partie du mémoire de thèse a été dédiée aux incertitudes. D'abord, l'état de l'art sur les différentes méthodes d'évaluation des incertitudes a été présenté. Nous avons ensuite mis en relief l'importance de la théorie des fonctions de croyance par rapport aux autres formalismes. En effet, la théorie des fonctions de croyance permet une modélisation fidèle des informations disponibles et

possède l'avantage de combiner plusieurs sources d'incertitudes (statistiques, avis d'experts, etc.). Les fondements théoriques de cette théorie ont été présentés et appliqués par la suite sur le cas de la démarche probabiliste proposée dans cette thèse, afin d'évaluer les incertitudes statistiques, celles liées au choix du seuil de la loi GPD et celles liées à l'impact du changement climatique. La démarche a été appliquée en considérant les incertitudes sous l'hypothèse de non-coïncidence. Les résultats obtenus montrent que les incertitudes statistiques sont dominantes pour la pluie et que ces dernières sont dominantes aussi pour les quantiles de niveau marin de fortes périodes de retour ($>$ 700 ans). L'impact du changement climatique est dominant seulement pour les niveaux marins de faibles périodes de retour ($<$ 700 ans). L'agrégation des intervalles de confiances dans le cas de non-coïncidence a montré que les incertitudes totales (combinées) peuvent atteindre 1 m pour une hauteur d'eau centennale. Cette incertitude assez élevée est due à la combinaison des incertitudes liées au changement climatique (dominantes aux faibles périodes de retour du niveau marin) et des incertitudes statistiques (dominantes pour la pluie).

Conclusion

Dans ces travaux de thèse, une démarche probabiliste a été mise en place pour l'aléa inondation dans un cadre général de PFHA. Peu d'idées ont été proposées dans la littérature pour prendre en compte la coïncidence des phénomènes d'inondation. Cette thèse présente un travail méthodologique et original de prise en compte des effets combinés de deux processus dans un contexte de dépendance et dans un cadre de PFHA. Deux hypothèses ont été examinées ; la première (non-coïncidence) a montré, de manière non combinatoire, une courbe d'aléa finale (agrégée) synthétisant les courbes d'aléa univariées classiques. La seconde a traité un concept plus complexe selon lequel les phénomènes d'inondation peuvent se produire séparément, simultanément ou avec un déphasage temporel. Une autre considération dans ces travaux est d'appliquer et d'illustrer l'approche avec l'exemple des pluies et du niveau marin au Havre. Il est désormais possible d'appliquer la méthodologie à tout site susceptible de subir une inondation multi-phénomènes.

Le chapitre I.1 de cette thèse était introductif, il a fait l'objet d'une présentation du site du Havre, des phénomènes d'inondation (pluie et niveau marin) qui sont susceptible d'inonder cette ville, des hypothèses de non-coïncidence et de coïncidence, des données et des VAs de la démarche proposée. L'étude de la corrélation entre la pluie et le niveau marin a montré que celle-ci est négligeable au Havre.

L'idée de cette thèse était de construire une démarche probabiliste complète nourrie des avancées scientifiques dans le domaine des inondations. Les chapitres II.1 et II.2 ont été consacrés aux analyses fréquentielles univariée et multivariée. Une revue de la littérature sur les approches d'analyse fréquentielle univariée a été faite dans le chapitre II.1. La méthode POT a été retenue pour ses nombreux avantages permettant entre autres d'améliorer la taille de l'échantillon pour les séries de données courtes (telle que celle des pluies au Havre). Des ajustements de la loi GPD on été faits en appliquant cette méthode POT à l'intensité de pluie et au niveau marin maximal. La vérification de l'adéquation de cette loi à l'aide du test de Kolmogorov-Smirnov a montré que l'ajustement des intensités de pluie pourrait être amélioré en considérant par exemple une autre loi de type exponentielle. L'état de l'art sur la théorie des copule a été fait dans le chapitre II.2. La modélisation de la corrélation des intensités de pluie avec le niveau marin a été faite et la copule de Clayton a été choisie.

La troisième partie de ce mémoire de thèse (chapitres III.1 et III.2) a traité les questions des scénarios d'inondation, des courbes d'aléa et leur agrégation. Des scénarios ont d'abord été construits à partir des ajustements puis propagés dans le modèle hydraulique du Havre sous l'hypothèse de non-coïncidence et de coïncidence. Les résultats sous ces deux hypothèses ont été examinés. Globalement, l'application a montré que la pluie est un phénomène dominant aux faibles périodes de retour, au Havre. Le niveau marin devient dominant aux périodes de retour élevées. De plus, la dépendance entre la pluie et le niveau marin est négligeable au Havre. Dans ce cas d'étude particulier de la ville du Havre, la probabilité de coïncidence de la pluie et du niveau marin susceptibles de provoquer des inondations lors du même cycle de marée, est négligeable. Ces résultats dépendent désormais des caractéristiques du site (ouvrages de protection, etc.), des données et des méthodes statistiques utilisées et sous réserve des incertitudes associées.

La dernière partie de cette thèse porte sur l'évaluation et l'intégration des in-

certitudes. Le cadre théorique des fonctions de croyance a été utilisé pour proposer une démarche permettant de combiner certaines sources d'incertitudes de la démarche probabiliste proposée dans cette thèse. Ce cadre théorique des fonctions de croyance est bien séduisant car il permet une représentation fidèle des informations incertaines, ainsi que la distinction entre les incertitudes aléatoires et épistémiques et la combinaison de toutes ces incertitudes. Malgré ces avantages, on a souvent reproché à ce cadre théorique d'être assez compliqué à mettre en œuvre et nécessite des calculs assez conséquents. En effet, certains auteurs évoquent la nécessité de mettre en place des simplifications (considérer des valeurs précises, limiter la taille du cadre de discernement, etc.). Un autre apport de cette thèse réside dans l'application de cette démarche d'évaluation et de combinaison des incertitudes aléatoires et épistémiques. Cette application a été faite en considérant l'hypothèse de non-coïncidence. Elle a montré que les incertitudes liées à la pluie sont principalement statistiques. L'impact du changement climatique sur le niveau marin est négligeable aux fortes périodes de retour car ce sont les incertitudes statistiques qui dominent pour les fortes périodes de retour.

Perspectives

Ce travail de thèse visait en premier lieu à mettre en place une démarche probabiliste pour la modélisation de l'aléa inondation avec prise en compte des incertitudes. La démarche proposée dans cette thèse a nécessité un énorme travail de revue de la littérature, sur la notion de coïncidence et la combinaison des phénomènes d'inondation, l'agrégation, les incertitudes, etc. Certains points méthodologiques peuvent être améliorés.

Les résultats de cette thèse ont nécessité le travail sur un champ pluridisciplinaire assez large et varié ; i.e., la modélisation du niveau marin, la modélisation des pluies, les méthodes de combinaison des aléas, la modélisation hydraulique des écoulements, l'agrégation des aléas, les incertitudes. Au-delà de mon côté curieux et des efforts de prise d'initiative, certaines hypothèses simplificatrices ont été considérées. Par exemple, le choix de la méthode POT, dans le chapitre II.1, pourrait poser problème lorsqu'on considère des scénarios de faible période de retour car la loi GPD fait partie des lois des valeurs extrêmes. Une solution pour contourner cette faiblesse consiste à construire une loi mixte (comme fait dans [97, 98, 86]), composée de la queue de distribution de la GPD pour les fortes périodes de retour et d'une autre loi (logarithmique, normale, etc.) pour le corps de la distribution. Un autre point d'amélioration consiste à considérer une convolution entre la marée et la surcote pour les sites à marée dominante. La forme des scénarios (de pluie et de niveau marin) et leurs durées sont simplifiées dans le chapitre III.1. Des travaux futurs pourront considérer d'autres formes de pluie (e.g., double-triangle, rectangle), d'autres durées de scénarios et peut-être aussi le rajout d'autres paramètres pour prendre en compte la dynamique des phénomènes d'inondation (e.g., considération des vagues et de leur interaction avec le niveau marin, du clapot). La quantification des incertitudes dans le chapitre IV.1 s'est limitée à l'hypothèse de non-coïncidence pour simplifier les calculs. Il serait désormais utile de refaire

les mêmes calculs sous l'hypothèse de coïncidence afin d'évaluer l'impact des incertitudes et du changement climatique sur les probabilités de coïncidence. La considération de trois sources d'incertitudes avait un simple objectif d'appliquer la démarche de prise en compte des incertitudes. D'autres sources d'incertitudes pourront être considérées (e.g., choix de la fonction de distribution, choix de la forme des scénarios, choix des durées).

Il serait notamment intéressant d'appliquer la démarche développée dans cette thèse sur d'autres sites présentant des risques d'inondation pour la tester et l'améliorer. Une application sur un site méditerranéen pourrait s'avérer particulièrement intéressante car le régime de marée est différent de celui de la côte de la Manche. Cela pourrait aussi concerner des applications avec prise en compte des franchissements au lieu des débordements (approche trivariée).

Bibliographie

[1] *U.S. Nuclear Regulatory Commission, Probabilistic Flood Hazard Assessment Research Plan, ADAMS Accession No. ML14296A442* (2014).

[2] *U.S. Nuclear Regulatory Commission, Workshop on Probabilistic Flood Hazard Assessment* (Accessed 24 May 2015).

[3] Abberger, K. A simple graphical method to explore tail-dependence in stock-return pairs. *Applied Financial Economics 15*, 1 (Jan. 2005), 43–51.

[4] AghaKouchak, A., Bardossy, A., and Habib, E. Conditional simulation of remotely sensed rainfall data using a non-Gaussian v-transformed copula. *Advances in Water Resources 33*, 6 (June 2010), 624–634. 00069.

[5] Apel, H., Martinez-Trepat, O., Hung, N. N., Chinh, D. T., Merz, B., and Dung, N. V. Combined fluvial and pluvial urban flood hazard analysis : concept development and application to Can Tho city, Mekong Delta, Vietnam. *Nat. Hazards Earth Syst. Sci. 16*, 4 (Apr. 2016), 941–961.

[6] Appriou, A. Probabilités et incertitudes en fusion de données multisenseurs. *Revue Scientique et Technique de la Défense* (1991).

[7] Appriou, A. *Aggregation and Fusion of Imperfect Information.* Physica-Verlag, 1998, ch. Uncertain Data Aggregation in Classification and Tracking Processes, pp. 231–260.

[8] ASN. Protection des installations nucleaires de base contre les inondations externes, guide no. 13, 40 pp. (in french). Tech. rep., ASN (Nuclear Safety Authority), 2013.

[9] Aven, T. On the need for restricting the probabilistic analysis in risk assessments to variability. *Risk analysis : an official publication of the Society for Risk Analysis 30* (11 2009), 354–60 ; author reply 381.

[10] Bacchi, B., Becciu, G., and Kottegoda, N. T. Bivariate exponential model applied to intensities and durations of extreme rainfall. *Journal of Hydrology 155*, 1 (Mar. 1994), 225–236. 00100.

[11] Baudrit, C., Guyonnet, D., and Dubois, D. Post-processing the hybrid method for addressing uncertainty in risk assessments. *Journal of Environmental Engineering* (2005).

[12] Beauval, C. *Analyse des incertitudes dans une estimation probabiliste de l'aléa sismique, exemple de la France.* PhD thesis, Université Joseph-Fourier - Grenoble I, 2003.

[13] Ben-Abdallah, N. *Modelling sea-level rise uncertainties for coastal defence adaptation using belief functions.* PhD thesis, University of Technology of Compiegne, 2014.

[14] Ben-Abdallah, N., Mouhous-Voyneau, N., and Denoeux, T. Combining statistical and expert evidence using belief functions : Application to centennial sea level estimation taking into account climate change. *International Journal of Approximate Reasoning, Volume 55, Issue 1, Part 3* (2014).

[15] Ben Daoued, A., Guimier, L., Hamdi, Y., Duluc, C.-M., and Rebour, V. Development of a probabilistic flood hazard assessment (pfha) for the nuclear safety. In *EGU General Assembly* (2016).

[16] Ben Daoued, A., Hamdi, Y., Mouhous-Voyneau, N., and Sergent, P. Modelling dependence and coincidence of marine flooding phenomena : methodology and simplified case study in le havre in france. In *EGU General Assembly* (2018).

[17] Ben Daoued, A., Mouhous-Voyneau, N., Hamdi, Y., Duluc, C.-M., and Sergent, P. Modeling coincidence and dependence of flood hazard phenomena in a probabilistic flood hazard assessment (pfha) framework : Case study in le havre. *Natural Hazards* (2019).

[18] Ben Daoued, A., Mouhous-Voyneau, N., Hamdi, Y., and Sergent, P. Development of a probabilistic multi-flood hazard approach considering uncertainties and climate change - application to the coastal flooding of the havre (france). *SPRINGER BOOK (EMCEI19)* (2019 [ACCEPTED]). Colloque international & publication.

[19] Bensi, M., and Kanney, J. Development of a framework for probabilistic storm surge hazard assessment for united states nuclear power plants. In *SMiRT-23* (2015).

[20] Berger, J., and Wolpert, R. The likelihood principle. *Institute of Mathematical Statistics, Hayward, CA* (1984).

[21] Berleant, D., and Goodman-Strauss, C. Bounding the results of arithmetic operations on random variables of unknown dependency using intervals. *Reliable Computing* (1998).

[22] Béliveau, J. Analyse fréquentielle multivariée de la pointe, du volume et de la durée de crue. Master's thesis, Univ. Laval, Quebec, 2006.

[23] Bozzoni, F., Corigliano, M., Lai, C., Salazar, W., Scandella, L., Zuccolo, E., Latchman, J., Lynch, L., and Robertson, R. Probabilistic seismic hazard assessment at the eastern caribbean islands. *Bulletin of the Seismological Society of America 101* (09 2011), 2499–2521.

[24] Breinl, K., Strasser, U., Bates, P., and Kienberger, S. A joint modelling framework for daily extremes of river discharge and precipitation in urban areas. *Journal of Flood Risk Management 10*, 1 (Mar. 2017), 97–114.

[25] Brown, P., and Caldeira, K. Greater future global warming inferred from earth's recent energy budget. *Nature* (2017).

[26] Capéraà, P., Fougères, A.-L., and Genest, C. A Stochastic Ordering Based on a Decomposition of Kendall's Tau. In *Distributions with given Marginals and Moment Problems*. Springer Netherlands, 1997, pp. 81–86. 00002 DOI : 10.1007/978-94-011-5532-8_9.

[27] Chang, J., Li, Y., Wang, Y., and Yuan, M. Copula-based drought risk assessment combined with an integrated index in the Wei River Basin, China. *Journal of Hydrology 540* (Sept. 2016), 824–834.

[28] Chauveau, E., Chadenas, C., Comentale, B., Pottier, P., Blanloeil, A., Feuillet, T., Mercier, D., Pourinet, L., Rollo, N., Tillier, I., and Trouillet, B. Xynthia : leçons d'une catastrophe. *Cybergo : European Journal of Geography [En ligne]* (2011).

[29] Chebana, F., and Ouarda, T. Multivariate quantiles in hydrological frequency analysis. *Environmetrics* (2011).

[30] Chebana, F., and Ouarda, T. B. Multivariate L-moment homogeneity test. *Water resources research 43*, 8 (2007).

[31] Chebana, F., and Ouarda, T. B. Index flood-based multivariate regional frequency analysis. *Water Resources Research 45*, 10 (2009).

[32] Chebana, F., Ouarda, T. B., Bruneau, P., Barbet, M., El Adlouni, S., and Latraverse, M. Multivariate homogeneity testing in a northern case study in the province of Quebec, Canada. *Hydrological processes 23*, 12 (2009), 1690–1700.

[33] Chen, A. S., Evans, B., Djordjevic, S., and Savic, D. A. Multi-layered coarse grid modelling in 2d urban flood simulations. *Journal of Hydrology 470-471* (Nov. 2012), 1–11. 00050.

[34] Cherubini, U., Luciano, E., and Vecchiato, W. *Copula methods in finance*. John Wiley & Sons, 2004.

[35] Coles, S. An introduction to statistical modeling of extreme values. *London : Springer Series in Statistics* (2001).

[36] Coles, S., Heffernan, J., and Tawn, J. Dependence Measures for Extreme Value Analyses. *Extremes 2*, 4 (Dec. 1999), 339–365.

[37] Coles, S., and Tawn, A. Statistical Methods for Multivariate Extremes : An Application to Structural Design. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 43*, 1 (1994), 1–48.

[38] De Michele, C., and Salvadori, G. A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas. *Journal of Geophysical Research : Atmospheres 108*, D2 (2003).

[39] De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., and Rosso, R. Bivariate statistical approach to check adequacy of dam spillway. *Journal of Hydrologic Engineering 10*, 1 (2005), 50–57.

[40] De Michele, C., Salvadori, G., Passoni, G., and Vezzoli, R. A multivariate model of sea storms using copulas. *Coastal Engineering 54*, 10 (2007), 734–751.

[41] DEFRA. Use of joint probability methodsin flood management - a guide to best practice. Tech. rep., DEFRA/Environment Agency, 2005.

[42] Deglaire, M. Le havre - etude sur le risque de submersion marine. Tech. rep., Université Paris 1 Panthéon-Sorbonne / DIRM - CODAH, 2010.

[43] Dempster, A. *Upper and lower probabilities induced by multivalued mapping.* Annals of Mathematical Statistics, 1967.

[44] Denoeux, T. Decision-making with belief functions : a review. *International Journal of Approximate Reasoning* (2019).

[45] Dong, W., and Shah, H. Vertex method for computing functions of fuzzy variables. *Fuzzy Sets and Systems* (1987).

[46] Dubois, D., and Prade, H. A set-theoretic view on belief functions : logical operations and approximations by fuzzy sets. *International Journal of General Systems* (1986).

[47] Dubois, D., and Prade, H. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence 4* (09 1988), 244–264.

[48] Dubois, D., and Prade, H. Random sets and fuzzy interval analysis. *Fuzzy Sets and Systems* (1991).

[49] Dubois, D., and Prade, H. On the combination of evidence in various mathematical frameworks. *Reliability Data Collection and Analysis* (1992).

[50] Dubois, D., and Prade, H. *Formal representations of uncertainty.* Decision-making Process : Concepts and Methods, 2010, ch. 3, pp. 85 – 156.

[51] Dubois, D., Prade, H., and Smets, P. Representing partial ignorance. *IEEE Systems, Machine and Cybernetics* (1996).

[52] Dung, N., Merz, B., Bardossy, A., and Apel, H. Handling uncertainty in bivariate quantile estimation - an application to flood hazard analysis in the mekong delta. *Journal of Hydrology* (2015).

[53] Elineau, S., Duperret, A., and Mallet, P. Coastal floods along the english channel : the study case of le havre town (nw france). *Caribbean Waves* (2013). (Poster).

[54] Elineau, S., Duperret, A., Mallet, P., and Caspar, R. Le havre : Une ville cotiere soumise aux submersions marines et aux instabilites de falaises littorales. In *Journées "Impacts du changement climatique sur les risques cotiers"* (2010).

[55] Etienne, M. *Modélisation et évaluation quantitative des risques en actuariat - Modèles sur une période.* Springer, 2013.

[56] Favre, A.-C., El Adlouni, S., Perreault, L., Thiémonge, N., and Bobée, B. Multivariate hydrological frequency analysis using copulas. *Water resources research 40*, 1 (2004).

[57] Ferson, S., Hajagos, J., Berleant, D., Zhang, J., Tucker, W., Ginzburg, L., and Oberkampf, W. Dependence in probabilistic modeling, dempster-shafer theory, and probability bounds analysis. Tech. rep., SAND-2004-3072. Albuquerque, NM, 2004.

[58] Fisher, N., and Switzer, P. Graphical assessment of dependence : Is a picture worth 100 tests ? *The American Statistician* (2001).

[59] Frahm, G., Junker, M., and Schmidt, R. Estimating the tail-dependence coefficient : Properties and pitfalls. *Insurance : Mathematics and Economics 37*, 1 (Aug. 2005), 80–100.

[60] Genest, C., and Favre, A.-c. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* (2007), 347–368.

[61] Genest, C., and Remillard, B. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'I.H.P. Probabilités et statistiques 44*, 6 (2008), 1096–1127. 00311.

[62] Genest, C., Remillard, B., and Beaudoin, D. Goodness-of-fit tests for copulas : A review and a power study. *Insurance : Mathematics and Economics 44*, 2 (Apr. 2009), 199–213.

[63] Grimald, S., and Serinaldi, F. Asymmetric copula in multivariate flood frequency analysis. *Advances in Water Resources 29*, 8 (2006), 1155–1167.

[64] Gudendorf, G., and Segers, J. *Extreme-Value Copulas*, vol. 198. 01 2010, pp. 127–145.

[65] Gupta, I. The state of the art in seismic hazard analysis. *ISET Journal of Earthquake Technology* (2002).

[66] Gupta, I. Probabilistic seismic hazard analysis method for mapping of spectral amplitudes and othe design-specific quantities to estimate the earthquake effects on man-made structures. *ISET Journal of Earthquake Technology* (2007).

[67] Guyonnet, D., Bourgine, B., Dubois, D., Fargier, H., Cme, B., and Chils, J. Hybrid approach for addressing uncertainty in risk assessments. *Journal of the Environmental Engineering Division, ASCE 129, 68-78* (2003).

[68] Guyonnet, D., Cme, B., Perrochet, P., and Parriaux, A. Comparing two methods for addressing uncertainty in risk assessments. *Journal of Environmental Engineering* (1999).

[69] Hallegatte, S., and Przyluski, V. *Gestion des risques naturels : Leçons de la tempête Xynthia.* "Matière à débattre et décider", 2012.

[70] Hamdi, Y., Bardet, L., Duluc, C.-M., and Rebour, V. Extreme storm surges : a comparative study of frequency analysis approaches. *Natural Hazards and Earth System Sciences 14* (08 2014).

[71] Hamdi, Y., Ben Daoued, A., Mouhous-Voyneau, N., and Sergent, P. Probabilistic multi-hazard risk assessment - development of an aggregation model based on the algebra of events. *Springer book (EMCEI2019)* (2019 [ACCEPTED]).

[72] Hawkes, P. J. Joint probability analysis for estimation of extremes. *Journal of Hydraulic Research 46*, sup2 (Dec. 2008), 246–256. 00000.

[73] Hawkes, P. J., Gouldby, B. P., Tawn, J. A., and Owen, M. W. The joint probability of waves and water levels in coastal engineering design. *Journal of Hydraulic Research 40*, 3 (May 2002), 241–251.

[74] Hawkes, P. J., and Svensson, C. Joint Probability : Dependence Mapping and Best Practice. *T02-06-16* (2006). 00000.

[75] Hebert, H., Abadie, S., Benoit, M., Creach, R., Duluc, C.-M., Gailler, A., Garziglia, S., Lemoine, A., Loevenbruck, A., Macary, O., Maspataud, A., Marcer, R., Morichon, D., Pedreros, R., Rebour, V., Ricchiuto, M., Silva Jacinto, R., Terrier, M., Toucanne, S., and Hayashi, Y. Le projet tandem (tsunamis en atlantique et manche : Definition des effets par modelisation) (2014-2017) : enjeux pour les vulnerabilites littorales aux tsunamis.

[76] Heffernan, J. E., and Tawn, J. A. A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society : Series B (Statistical Methodology) 66*, 3 (Aug. 2004), 497–546. 00388.

[77] Horrillo-Caraballo, J., Reeve, D., Simmonds, D., Pan, S., Fox, A., Thompson, R., Hoggart, S., S.H. Kwan, S., and Greaves, D. Application of a source-pathway-receptor-consequence (s-p-r-c) methodology to the teign estuary, uk. *Journal of Coastal Research SI65* (04 2013), 1939–1944.

[78] IAEA. Probabilistic safety assessment for seismic events. Tech. rep., International Atomic Energy Agency (IAEA), 1993.

[79] IPCC. Climate change : synthesis report. Tech. rep., IPCC, 2014.

[80] IPCC. Global warming of 1.5-c : Special report. Tech. rep., IPCC, 2018.

[81] IPSN. Rapport sur l'inondation du site du blayais survenue le 27 décembre 1999. Tech. rep., Institut de Protection et Sûreté Nucléaire (IPSN), 2000.

[82] Joe, H. *Multivariate models and multivariate dependence concepts*. CRC Press, 1997.

[83] Karmakar, S., and Simonovic, S. P. Bivariate flood frequency analysis. part 2 : a copula-based approach with mixed marginal distributions. *Journal of Flood Risk Management 2*, 1 (2009), 32–44.

[84] Keef, C., Papastathopoulos, I., and Tawn, J. A. Estimation of the conditional distribution of a multivariate variable given that one of its components is large : Additional constraints for the Heffernan and Tawn model. *Journal of Multivariate Analysis 115* (Mar. 2013), 396–404. 00035.

[85] Kelly, K. S., and Krzysztofowicz, R. A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrology and Hydraulics 11*, 1 (Feb. 1997), 17–31. 00169.

[86] Kergadallan, X. *Estimation des niveaux marins extremes avec et sans l'action des vagues le long du littoral métropolitain.* PhD thesis, Université Paris-Est, 2015. 00001.

[87] Klugel, J.-U. Probabilistic safety analysis of external floods – method and application. *Kerntechnik 78*, 2 (may 2013), 127–136.

[88] Kojadinovic, I., and Yan, J. A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. *Statistics and Computing 21*, 1 (Jan. 2011), 17–30. 00093.

[89] Kopp, R., Horton, R., Little, C., Mitrovica, J., Oppenheimer, M., Rasmussen, D., Stauss, B., and Tebaldi, C. Probabilistic 21st and 22nd century sea-level projections at a global network of tide-gauge sites. *Earth's Future* (2014).

[90] Kwak, J., Kim, S., Jung, J., Singh, V., Lee, D. R., and Kim, H. S. Assessment of meteorological drought in korea under climate change. *Advances in Meteorology 2016* (2016).

[91] Le Duy, T., Dieulle, L., Vasseur, D., Berenguer, C., and Couplet, M. An alternative comprehensive framework using belief functions for parameter and model uncertainty analysis in nuclear probabilistic risk assessment applications. *Proceedings of the Institution of Mechanical Engineers, Part O : Journal of Risk and Reliability 227* (10 2013), 471–490.

[92] Lee, T., Modarres, R., and Ouarda, T. B. M. J. Data-based analysis of bivariate copula tail dependence for drought duration and severity. *Hydrological Processes 27*, 10 (May 2013), 1454–1463. 00037.

[93] Letortu, P., Costa, S., and Cantat, O. Les submersions marines en manche orientale : Approche inductive et naturaliste pour la caractÉrisation des facteurs responsables des inondations par la mer. *Climatologie, vol. 9* (2012).

[94] Lian, J. J., Xu, K., and Ma, C. Joint impact of rainfall and tidal level on flood risk in a coastal city with a complex river network : a case study of Fuzhou City, China. *Hydrol. Earth Syst. Sci. 17*, 2 (Feb. 2013), 679–689.

[95] Masina, M., Lamberti, A., and Archetti, R. Coastal flooding : A copula based approach for estimating the joint probability of water levels and waves. *Coastal Engineering 97* (2015), 37–52.

[96] Maspataud, A., Elineau, S., Duperret, A., Ruz, M.-H., and Mallet, P. Impacts de niveaux d'eau extremes sur deux villes portuaires de la manche et mer du nord : Le havre et dunkerque. *Journées REFMAR* (2016).

[97] Mazas, F., and Hamm, L. An event-based approach for extreme joint probabilities of waves and sea levels. *Coastal Engineering 122* (2017), 44–59.

[98] Mazas, F., Kergadallan, X., Garat, P., and Hamm, L. Applying POT methods to the Revised Joint Probability Method for determining extreme sea levels. *Coastal Engineering 91* (Sept. 2014), 140–150. 00000.

[99] McGuire, R. K., and Arabasz, W. J. An introduction to probabilistic seismic hazard analysis. In *Geotechnical and Environmental Geophysics*. Society of Exploration Geophysicists, jan 1990, pp. 333–354.

[100] Miquel, J. *Guide pratique d'estimation des probabilités de crue*. Eyrolles, 1984.

[101] Narayan, S., Hanson, S., Nicholls, R., and Clarke, D. Use of the source - pathway - receptor - consequence model in coastal flood risk assessment.

[102] Nelsen, R. B. An introduction to copulas. *Springer (New York)* (1999).

[103] Nelsen, R. B. An introduction to copulas, 2nd. *New York : SpringerScience Business Media* (2006).

[104] NERC. Flood studies report. Tech. rep., London : Natural Environment Research Council, 1975.

[105] NOAA. Global and regional sea level rise scenarios for the united states. Tech. rep., NOAA, 2017.

[106] Norberto, C. N.-C., Jeffrey, A. M., Victor, M. G., and Andrew, T. C. Coastal storm hazards from virginia to maine. Tech. rep., US Army Corps of Engineers (USACE), 2015.

[107] Orcel, O., Sergent, P., and Ropert, F. Trivariate copula to design coastal structures. In *EVAN* (2019).

[108] Ouzeau, G., Deque, M., Jouini, M., Planton, S., Vautard, R., and Vrac, M. Scenarios regionalises - edition 2014 pour la metropole et les regions d'outre-mer. Tech. rep., Sous la direction de Jean Jouzel, Ministere de l'ecologie, du developpement durable et de l'energie, 2014.

[109] Pinya, M. A. S., Madsen, H., and Rosbjerg, D. Assessment of the risk of inland flooding in a tidal sluice regulated catchment using multi-variate statistical techniques. *Physics and Chemistry of the Earth, Parts A/B/C 34*, 10 (2009), 662–669.

[110] Poulin, A., Huard, D., Favre, A.-C., and Pugin, S. Importance of tail dependence in bivariate frequency analysis. *Journal of Hydrologic Engineering 12*, 4 (2007), 394–403.

[111] Pugh, D. Tides, surges and mean sea-level : a handbook for engineers and scientists, 472 pp, 1987.

[112] Rebour, V., Georgescu, G., Leteinturier, D., Raimond, E., La Rovere, S., Bernadara, P., Vasseur, D., Brinkman, H., Groudev, P., Ivanov, I., Turschmann, M., Sperbeck, S., Potempski, S., Hirata, K., and Kumar, M. Report 2 : Guidance document on practices to model and implement external flooding hazards in extended psa. Tech. rep., IRSN-PSN-RES-SAG-PSN-RES-SAG, 2016.

[113] Regan, H., Ferson, S., and Berleant, D. Equivalence of methods for uncertainty propagation of real-valued random variables. *International Journal of Approximate Reasoning* (2004).

[114] Requena, A. I., Mediero, L., and Garrote, L. A bivariate return period based on copulas for hydrologic dam design : accounting for reservoir routing in risk estimation. *Hydrology and Earth System Sciences 17*, 8 (2013), 3023.

[115] Resnick, S. I. Point Processes. In *Extreme Values, Regular Variation and Point Processes*, Springer Series in Operations Research and Financial Engineering. Springer New York, 1987, pp. 123–161. DOI : 10.1007/978-0-387-75953-1_4.

[116] Sadri, S., and Burn, D. Copula-based pooled frequency analysis of droughts in the canadian prairies. *Journal of Hydrologic Engineering* (2014).

[117] Salecker, D., Gruhn, A., Schlamkow, C., and Fröhle, P. Parameterisation of storm surge hydrographs using univariate bivariate statistical models. In *Proceedings of the 8th International Conference on Coastal and Port Engineer-ing in Developing Countries* (2012), p. 222.

[118] Sallak, M., Aguirre, F., and Schon, W. Incertitudes aléatoires et épistémiques, comment les distinguer et les manipuler dans les études de fiabilité ? *QUALITA2013 (Hal-00823114)* (2013).

[119] Salvadori, G., and De Michele, C. Frequency analysis via copulas : Theoretical aspects and applications to hydrological events. *Water Resour. Res.* (2004).

[120] Salvadori, G., de Michele, C., and Durante, F. On the return period and design in a multivariate framework. *Hydrology and Earth System Sciences 15* (Nov. 2011), 3293–3305.

[121] Salvatore, G., and Serinaldi, F. Design hyetograph analysis with 3-copula function. *Hydrological Sciences Journal 51*, 2 (2006), 223–238.

[122] Samuels, P. G., and Burt, N. A new joint probability appraisal of flood risk. *Proceedings of the Institution of Civil Engineers - Water and Maritime Engineering 154*, 2 (June 2002), 109–115.

[123] Sandri, S., Dubois, D., and Kalfsbeek, H. Elicitation, assessment and pooling of expert judgments using possibility theory. *IEEE Transactions on Fuzzy Systems* (1995).

[124] Sentz, K., and Ferson, S. Combination of evidence in dempster-shafer theory. Tech. rep., SAND2002-0835. USA, 2002.

[125] Sergent, P., Prevot, G., Mattarolo, G., Luck, M., Brossard, J., T Nguyen, D., Morel, G., Mar, N.-F., Benoit, M., Ropert, F., Guillou, N., Bouttes, F., Kergadallan, X., Trichet, J.-J., Delisle, J.-R., Menon, J.-M., Mallet, P., Voyneau, N., Lam, M., and Le Banner, G. Projet sao polo adaptation des structures cotieres au changement climatique.

[126] Serinaldi, F., Bardossy, A., and Kilsby, C. G. Upper tail dependence in rainfall extremes : would we know it if we saw it ? *Stochastic Environmental Research and Risk Assessment 29*, 4 (May 2015), 1211–1233. 00010.

[127] Shafer, G. *A Mathematical Theory of Evidence.* Princeton University Press, 1976.

[128] Skaugen, T. The spatial variability of snow water equivalent. *Hydrology and Earth System Sciences Discussions 4*, 3 (June 2007), 1465–1489.

[129] Sklar, M. *Fonctions de répartition à n dimensions et leurs marges.* Université Paris 8, 1959.

[130] Smets, P. Constructing the pignistic probability function in a context of uncertainty. *Uncertainty in Articial Intelligence* (1990).

[131] Smets, P. Practical uses of belief functions. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence* (1990).

[132] Smets, P. *International Joint Conference on Artificial Intelligence.* Morgan Kaufman, San Mateo, 1993, ch. An axiomatic justification for the use of belief function to quantify beliefs, pp. 598–603.

[133] Smets, P. *The normative representation of quantified beliefs by belief functions.* Artificial Intelligence, 1997.

[134] Smets, P. *Handbook of Defeasible Reasoning and Uncertainty Management Systems.* Kluwer, Doordrecht, The Netherlands, 1998, ch. The transferable belief model for quantified belief representation, pp. 267–301.

[135] Smets, P. Data fusion in the transferable belief model. In *Proceedings of FUSION-2000* (2000).

[136] Smets, P. Analyzing the combination of conflicting belief functions. *Information Fusion* (2007).

[137] Smets, P., and Kennes, R. *The Transferable Belief Model. Artificial Intelligence.* Artificial Intelligence, 1994.

[138] Smith, R. L. *Extreme value theory based on the r largest annual events.* J. Hydrol. 86, 1986.

[139] Sraj, M., Bezak, N., and Brilly, M. Bivariate flood frequency analysis using the copula function : a case study of the Litija station on the Sava River. *Hydrological Processes 29*, 2 (Jan. 2015), 225–238.

[140] Svensson, C., and Jones, D. A. Dependence between extreme sea surge, river flow and precipitation in eastern Britain. *International Journal of Climatology 22*, 10 (Aug. 2002), 1149–1168.

[141] Svensson, C., and Jones, D. A. Dependence between sea surge, river flow and precipitation in south and west Britain. *Hydrology and Earth System Sciences Discussions 8*, 5 (2004), 973–992.

[142] Tawn, J. A. Bivariate Extreme Value Theory : Models and Estimation. *Biometrika 75*, 3 (1988), 397–415.

[143] Tosunoglu, F., and Kisi, O. Joint modelling of annual maximum drought severity and corresponding duration. *Journal of Hydrology 543, Part B* (Dec. 2016), 406–422.

[144] van den Berg, M. J., Vandenberghe, S., de Baets, B., and Verhoest, N. E. C. Copula-based downscaling of spatial rainfall : a proof of concept. *Hydrology and Earth System Sciences 15* (May 2011), 1445–1457.

[145] van der Pol, T., and Hinkel, J. Uncertainty representations of mean sea-level change : a telephone game ? *Climatic Change* (01 2019).

[146] Volpi, E., and Fiori, A. Design event selection in bivariate hydrological frequency analysis. *Hydrological Sciences Journal 57*, 8 (Nov. 2012), 1506–1515. 00021.

[147] Volpi, E., and Fiori, A. Hydraulic structures subject to bivariate hydrological loads : Return period, design, and risk assessment. *Water Resources Research 50*, 2 (2014), 885–897.

[148] Waal, D. J. d., and Gelder, P. H. A. J. M. v. Modelling of extreme wave heights and periods through copulas. *Extremes 8*, 4 (Dec. 2005), 345–356. 00049.

[149] Wahl, T., Jensen, J., and Frank, T. On analysing sea level rise in the German Bight since 1844. *Natural Hazards and Earth System Science 10*, 2 (Feb. 2010), 171–179.

[150] Wahl, T., Jensen, J., and Mudersbach, C. A multivariate statistical model for advanced storm surge analyses in the north sea. *Coastal Engineering Proceedings 1*, 32 (2011), 19.

[151] Walley, P. *Statistical reasoning with imprecise probabilities.* New York : Chapman and Hall, 1991.

[152] Wang, C., Chang, N.-B., and Yeh, G.-T. Copula-based flood frequency (COFF) analysis at the confluences of river systems. *Hydrological Processes 23*, 10 (May 2009), 1471–1486.

[153] Wang, J., Wan Tsang, W., and Marsaglia, G. Evaluating kolmogorov's distribution. *Journal of Statistical Software 08* (12 2003).

[154] White, C. J. The use of joint probability analysis to predict flood frequency in estuaries and tidal rivers. *MyScienceWork* (Nov. 2007). 00000.

[155] Williamson, R., and Downs, T. Probabilistic arithmetic i : numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning* (1990).

[156] Yager, R. On the dempster-shafer framework and new combination rules. *Information Science* (1978).

[157] Yager, R. Arithmetic and other operations on dempster-shafer structures. *International Journal Man-Machine Studies* (1986).

[158] Yan, B., and Chen, L. Coincidence probability of precipitation for the middle route of South-to-North water transfer project in China. *Journal of Hydrology 499* (Aug. 2013), 19–26.

[159] Yang, X., and Zhang, Q. Joint probability distribution of winds and waves from wave simulation of 20 years (1989-2008) in Bohai Bay. *Water Science and Engineering 6*, 3 (July 2013), 296–307. 00000.

[160] Yue, S., Ouarda, T., Bobée, B., Legendre, P., and Bruneau, P. The gumbel mixed model for flood frequency analysis. *Journal of Hydrology* (1999).

[161] Zadeh, L. Fuzzy sets as a basis for theory of possibility. *Fuzzy Sets and Systems 1* (12 1999), 9–34.

[162] Zhang, D., Lin, Y., Zhao, P., Yu, X., Wang, S., Kang, H., and Ding, Y. The beijing extreme rainfall of 21 july 2012 : Right results but for wrong reasons. *Geophys. Res. Lett.* (2013).

[163] Zhang, L., and Singh, V. Trivariate flood frequency analysis using the gumbel-hougaard copula. *Journal of Hydrologic Engineering* (2007).

[164] Zhang, L., and Singh, V. P. Bivariate Rainfall and Runoff Analysis Using Entropy and Copula Theories. *Entropy 14* (Sept. 2012), 1784–1812.

[165] Zheng, F., Westra, S., Leonard, M., and Sisson, S. A. Modeling dependence between extreme rainfall and storm surge to estimate coastal flooding risk. *Water Resources Research 50*, 3 (Mar. 2014), 2050–2071. 00017.

[166] Zheng, F., Westra, S., and Sisson, S. A. Quantifying the dependence between extreme rainfall and storm surge in the coastal zone. *Journal of Hydrology 505* (Nov. 2013), 172–187.

Annexes

Annexe A

Ajustement et choix des copules pour différents déphasages

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.13	0	0.08	0	0	-35.27	-28.54
Student	0.12	30	0.08	$2.64 \quad 10^{-5}$	$2.64 \quad 10^{-5}$	-33.36	-19.89
Clayton	0.09	0	0.04	$6.82 \quad 10^{-4}$	0	-37.62	-30.88
Survie de Clayton	0.09	0	0.05	0	$6.98 \quad 10^{-4}$	-5.69	1.04
Gumbel	1.06	0	0.06	0	0.08	-15.38	-8.65
Survie de Gumbel	1.05	0	0.04	0.05	0	-5.69	1.04
Frank	0.81	0	0.09	0	0	-29.17	-22.44
Joe	1.03	0	0.02	0	0.04	-0.28	6.45

Tableau A.1: Paramètres des différentes copules pour $\delta = -6h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.15	0	0.10	0	0	-48.05	-41.32
Student	0.14	30	0.09	$3.50 \quad 10^{-5}$	$3.50 \quad 10^{-5}$	-44.31	-30.84
Clayton	0.11	0	0.05	$1.71 \quad 10^{-3}$	0	-48.09	-41.36
Survie de Clayton	0.11	0	0.05	0	$1.93 \quad 10^{-3}$	-8.50	-1.77
Gumbel	1.08	0	0.07	0	$9.69 \quad 10^{-2}$	-21.31	-14.58
Survie de Gumbel	1.05	0	0.05	$7.05 \quad 10^{-2}$	0	-41.09	-34.35
Frank	0.96	0	0.11	0	0	-41.64	-34.90
Joe	1.04	0	0.02	0	$5.23 \quad 10^{-2}$	-1.14	5.59

Tableau A.2: Paramètres des différentes copules pour $\delta = -5h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.15	0	0.10	0	0	-50.03	-43.30
Student	0.14	30.00	0.09	$3.44 \quad 10^{-5}$	$3.44 \quad 10^{-5}$	-49.41	-35.94
Clayton	0.12	0	0.06	$2.73 \quad 10^{-3}$	0	-58.94	-52.21
Survie de Clayton	0.10	0	0.05	0	$1.12 \quad 10^{-3}$	-6.98	-0.25
Gumbel	1.08	0	0.07	0	$9.59 \quad 10^{-2}$	-21.02	-14.29
Survie de Gumbel	1.06	0	0.06	$7.51 \quad 10^{-2}$	0	-53.26	-46.53
Frank	0.89	0	0.10	0	0	-35.09	-28.35
Joe	1.04	0	0.02	0	$4.72 \quad 10^{-2}$	-0.55	6.18

Tableau A.3: Paramètres des différentes copules pour $\delta = -4h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.15	0	0.10	0	0	-49.32	-42.59
Student	0.14	30.00	0.09	$3.39 \quad 10^{-5}$	$3.39 \quad 10^{-5}$	-41.46	-28.00
Clayton	0.10	0	0.05	$1.28 \quad 10^{-3}$	0	-45.45	-38.72
Survie de Clayton	0.12	0	0.06	0	$2.97 \quad 10^{-3}$	-10.87	-4.14
Gumbel	1.08	0	0.07	0	$9.52 \quad 10^{-2}$	-20.99	-14.25
Survie de Gumbel	1.05	0	0.05	$6.47 \quad 10^{-2}$	0	-35.20	-28.47
Frank	0.91	0	0.10	0	0	-39.66	-32.93
Joe	1.04	0	0.02	0	$5.45 \quad 10^{-2}$	-1.43	5.30

Tableau A.4: Paramètres des différentes copules pour $\delta = -3h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.11	0	0.07	0	0	-24.01	-17.28
Student	0.10	30.00	0.06	$1.94 \quad 10^{-5}$	$1.94 \quad 10^{-5}$	-17.75	-4.29
Clayton	0.07	0	0.04	$8.32 \quad 10^{-5}$	0	-22.46	-15.73
Survie de Clayton	0.09	0	0.04	0	$4.28 \quad 10^{-4}$	-5.01	1.72
Gumbel	1.05	0	0.05	0	$6.96 \quad 10^{-2}$	-9.66	-2.93
Survie de Gumbel	1.03	0	0.03	$4.41 \quad 10^{-2}$	0	-16.75	-10.02
Frank	0.67	0	0.07	0	0.00e+00	-19.46	-12.73
Joe	1.03	0	0.02	0	$3.93 \quad 10^{-2}$	0.24	6.97

Tableau A.5: Paramètres des différentes copules pour $\delta = -2h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.11	0	0.07	0	0	-21.41	-14.68
Student	0.09	30.00	0.06	$1.77 \quad 10^{-5}$	$1.77 \quad 10^{-5}$	-14.89	-1.42
Clayton	0.07	0	0.03	$6.05 \quad 10^{-5}$	0	-21.67	-14.93
Survie de Clayton	0.08	0	0.04	0	$2.52 \quad 10^{-4}$	-4.18	2.56
Gumbel	1.05	0	0.05	0	$6.22 \quad 10^{-2}$	-7.13	-0.40
Survie de Gumbel	1.03	0	0.03	$4.27 \quad 10^{-2}$	0	-21.98	-15.24
Frank	0.57	0	0.06	0	0	-12.98	-6.25
Joe	1.02	0	0.01	0	$3.35 \quad 10^{-2}$	0.72	7.45

Tableau A.6: Paramètres des différentes copules pour $\delta = -1h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.09	0	0.06	0	0	-16.39	-9.66
Student	0.09	30.00	0.05	$1.56 \quad 10^{-5}$	$1.56 \quad 10^{-5}$	-17.30	-3.84
Clayton	0.07	0	0.03	$5.80 \quad 10^{-5}$	0	-20.97	-14.24
Survie de Clayton	0.07	0	0.03	0	$4.24 \quad 10^{-5}$	-1.97	4.77
Gumbel	1.05	0	0.04	0	$6.01 \quad 10^{-2}$	-6.43	0.30
Survie de Gumbel	1.03	0	0.03	$.61 \quad 10^{-2}$	0	-18.98	-12.25
Frank	0.54	0	0.06	0	0	-11.31	-4.58
Joe	1.02	0	0.01	0	$2.99 \quad 10^{-2}$	0.99	7.72

Tableau A.7: Paramètres des différentes copules pour $\delta = +1h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.08	0	0.05	0	0	-11.13	-4.40
Student	0.08	30.00	0.05	$1.35 \quad 10^{-5}$	$1.35 \quad 10^{-5}$	-8.89	4.57
Clayton	0.05	0	0.03	$2.55 \quad 10^{-6}$	0	-10.54	-3.81
Survie de Clayton	0.07	0	0.03	0	$4.52 \quad 10^{-5}$	-2.00	4.73
Gumbel	1.04	0	0.04	0	$5.48 \quad 10^{-2}$	-4.92	1.81
Survie de Gumbel	1.03	0	0.03	$3.64 \quad 10^{-2}$	0	-8.97	-2.24
Frank	0.55	0	0.06	0	0	-12.04	-5.31
Joe	1.02	0	0.01	0	$3.11 \quad 10^{-2}$	0.90	7.64

Tableau A.8: Paramètres des différentes copules pour $\delta = +2h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.08	0	0.05	0	0	-12.60	-5.87
Student	0.08	30.00	0.05	$1.36 \quad 10^{-5}$	$1.36 \quad 10^{-5}$	-8.33	5.13
Clayton	0.06	0	0.03	$8.72 \quad 10^{-6}$	0	-13.78	-7.05
Survie de Clayton	0.06	0	0.03	0	$1.44 \quad 10^{-5}$	-1.21	5.52
Gumbel	1.04	0	0.04	0	$5.22 \quad 10^{-2}$	-4.24	2.49
Survie de Gumbel	1.03	0	0.03	$3.64 \quad 10^{-2}$	0	-10.22	-3.48
Frank	0.49	0	0.05	0	0	-8.58	-1.85
Joe	1.02	0	0.01	0	$2.61 \quad 10^{-2}$	1.23	7.96

Tableau A.9: Paramètres des différentes copules pour $\delta = +3h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.10	0	0.07	0	0	-20.37	-13.64
Student	0.09	30.00	0.06	$1.80 \quad 10^{-5}$	$1.80 \quad 10^{-5}$	-17.52	-4.06
Clayton	0.07	0	0.04	$7.68 \quad 10^{-5}$	0	-21.42	-14.69
Survie de Clayton	0.08	0	0.04	0	$1.55 \quad 10^{-4}$	-3.31	3.42
Gumbel	1.05	0	0.05	0	$6.66 \quad 10^{-2}$	-8.53	-1.80
Survie de Gumbel	1.04	0	0.04	$5.08 \quad 10^{-2}$	0	-22.20	-15.47
Frank	0.63	0	0.07	0	0	-16.67	-9.94
Joe	1.03	0	0.02	0	$3.55 \quad 10^{-2}$	0.57	7.30

Tableau A.10: Paramètres des différentes copules pour $\delta = +4h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.08	0	0.05	0	0	-12.22	-5.49
Student	0.08	30.00	0.05	$1.36 \quad 10^{-5}$	$1.36 \quad 10^{-5}$	-6.61	6.86
Clayton	0.06	0	0.03	$4.43 \quad 10^{-6}$	0	-11.97	-5.23
Survie de Clayton	0.06	0	0.03	0	$1.85 \quad 10^{-5}$	-1.31	5.42
Gumbel	1.04	0	0.04	0	$5.29 \quad 10^{-2}$	-4.40	2.33
Survie de Gumbel	1.03	0	0.03	$3.52 \quad 10^{-2}$	0	-9.25	-2.52
Frank	0.52	0	0.06	0	0	-10.07	-3.33
Joe	1.02	0	0.01	0	$2.81 \quad 10^{-2}$	1.11	7.84

Tableau A.11: Paramètres des différentes copules pour $\delta = +5h$

Copule	par_1 (α)	par_2	τ	λ_L	λ_U	AIC	BIC
Gaussienne (normale)	0.09	0	0.06	0	0	-15.46	-8.73
Student	0.09	30.00	0.05	$1.58 \quad 10^{-5}$	$1.58 \quad 10^{-5}$	-13.64	-0.17
Clayton	0.06	0	0.03	$2.07 \quad 10^{-5}$	0	-15.66	-8.93
Survie de Clayton	0.07	0	0.04	0	$8.02 \quad 10^{-5}$	-2.55	4.19
Gumbel	1.05	0	0.04	0	$6.09 \quad 10^{-2}$	-6.65	0.08
Survie de Gumbel	1.03	0	0.03	$4.54 \quad 10^{-2}$	0	-16.91	-10.17
Frank	0.61	0	0.07	0	0	-14.62	-7.89
Joe	1.02	0	0.01	0	$3.25 \quad 10^{-2}$	0.80	7.53

Tableau A.12: Paramètres des différentes copules pour $\delta = +6h$