VERS L'OPTIMISATION DE LA GESTION MUTUALISEE DE DEUX STATIONS D'EPURATION (STEP) : COORDINATION DES RESEAUX DE COLLECTE ET EXPLOITATION INTEGREE DES STEP D'EVRY-COURCOURONNES ET D'EXONA (FRANCE)

MEMOIRE POUR L'OBTENTION DU DIPLOME D'INGENIEUR 2IE AVEC GRADE DE **MASTER**

SPECIALITE : GENIE DE L'EAU, DE L'ASSAINISSEMENT ET DES AMENAGEMENTS HYDRO-AGRICOLES (GEAAH)

Présenté et soutenu publiquement le 22 juillet par

Djamilatou BAMOGO (2020 0387)

Directeur de mémoire : Pr. Harinaivo Anderson ANDRIANISA, Professeur Titulaire, 2iE

Encadrant 2iE : Dr. Ousmane Roland YONABA, Maître de Conférences (CAMES), 2iE

Encadrant externe : Mme Karine COUTURIER, Directrice Technique, SPL

Structure d'accueil du stage : SPL Confluence Seine Essonne Energie

Jury d'évaluation du stage :

Président : Pr. Yacouba KONATE

Membres et correcteurs : Dr Boukary SAWADOGO

M. Tontama SANOU

Promotion [2025/2026]

DEDICACE

Je dédie ce mémoire à mes parents, ***BAMOGO Somnongodo Rasmané*** et ***ZOROME Ramata***

Dont le soutien inconditionnel et les sacrifices consentis tout au long de mon parcours ont rendu ce travail possible.

A toute ma famille, qui a su m'entourer et m'encourager à chaque étape de ce chemin.

A Thierno Idrissa DIARRA pour son soutien indéfectible, sa présence constante et ses encouragements inlassables à chaque étape de ce travail.

REMERCIEMENTS

Ce mémoire marque l'aboutissement de plusieurs années de formation et n'aurait pas pu être réalisé sans le soutien de nombreuses personnes.

Je souhaite adresser mes sincères remerciements à l'Institut International d'Ingénierie de l'Eau et de l'Environnement (2iE) pour la qualité de l'enseignement reçu tout au long de mon cursus.

Je remercie l'École Nationale du Génie de l'Eau et de l'Environnement de Strasbourg (ENGEES) pour son accueil et la formation reçue dans le cadre de ma mobilité.

Je remercie mon Directeur de mémoire, Pr. Harinaivo Anderson ANDRIANISA, pour la confiance qu'il m'a témoignée tout au long de l'élaboration de ce mémoire.

Mes remerciements vont également au Dr Ousmane Roland YONABA pour sa disponibilité, ses conseils et ses orientations qui m'ont permis d'avancer sereinement dans la réalisation de ce mémoire.

Je remercie la SPL Confluence Seine Essonne Énergie de m'avoir accueillie pour ce stage et de m'avoir donné l'opportunité de travailler sur un sujet aussi enrichissant.

Je tiens à remercier tout particulièrement Mme Karine COUTURIER pour son encadrement, sa disponibilité et sa patience tout au long du stage. Ses conseils et son soutien ont été précieux pour mener à bien cette étude. Je remercie également M. Benoît VERNIÈRE pour son accueil et l'intérêt porté à mon travail.

Enfin, je remercie ma famille, mes proches et toutes les personnes qui m'ont soutenue et encouragée durant mes études et pendant la rédaction de ce mémoire.

A vous toutes et tous, merci.

RESUME

Ce mémoire porte sur l'optimisation de la gestion mutualisée des stations d'épuration d'Evry-Courcouronnes, dénommée Evry, et de Corbeil-Essonnes, dénommée Exona, dans un contexte marqué par une forte sensibilité aux épisodes pluvieux. La station d'Exona, limitée par une file d'eau unique et par une consigne actuelle de débit fixée à **1 000 m³/h**, connaît régulièrement des dépassements de capacité. Ces situations peuvent conduire à l'utilisation du by-pass vers la station d'Evry, dont le rôle est de sécuriser le fonctionnement du système, tout en évitant de reporter excessivement la contrainte hydraulique vers Evry. L'étude s'appuie sur l'analyse des données pluviométriques et hydrauliques de l'année **2024**. Les épisodes pluvieux ont été caractérisés selon leur cumul, leur durée et leur intensité. Les débits entrants à Exona ont ensuite été étudiés au pas de temps de **15 minutes** afin d'identifier les périodes de dépassement du seuil de **1 000 m³/h** et de mieux comprendre la réponse du réseau aux pluies. Les résultats indiquent que des dépassements se produisent régulièrement au cours de plusieurs mois de l'année, avec une intensité accrue durant la saison automnale. La distinction entre le débit normal et les eaux pluviales parasites souligne l'impact des précipitations sur les augmentations de débit. Toutefois, certains dépassements semblent également associés à un débit normal élevé ou à des effets hydrauliques qui se manifestent avec un certain retard. Pour prévoir le risque de by-pass dans **une heure** à venir, un arbre de décision a été élaboré. Le modèle montre que le débit actuel à Exona et le débit de l'heure précédente sont les variables les plus déterminantes. Il permet également d'identifier des seuils de vigilance autour de **920 m³/h** et **997 m³/h**. Ce travail propose ainsi une base d'aide à la décision pour une gestion plus anticipative du by-pass, en combinant le suivi du débit entrant à Exona et de la pluie récente.

Mots-clés :

1. Assainissement,
2. By-pass
3. Épisodes pluvieux
4. Gestion mutualisée
5. Station d'épuration

ABSTRACT

This thesis focuses on the optimization of the joint operation of the wastewater treatment plants of Evry-Courcouronnes, referred to as Evry, and Corbeil-Essonnes, referred to as Exona, in a context characterized by high sensitivity to rainfall events. The Exona plant, limited by a single treatment line and by a current flow setpoint of **1000 m³/h**, regularly experiences capacity exceedances. These situations may lead to the use of the bypass toward the Evry plant, whose role is to secure the overall system while avoiding excessive hydraulic pressure on Evry. The study is based on the analysis of rainfall and hydraulic data from **2024**. Rainfall events were characterized according to their cumulative rainfall, duration and intensity. The inflows to Exona were then analyzed at a **15 minute** time step in order to identify periods exceeding the 1,000 m³/h threshold and to better understand the network response to rainfall. The analysis shows that flow exceedances occur repeatedly during several periods of the year, with the highest frequency observed in autumn. Comparing normal flow conditions with parasitic stormwater contributions clearly illustrates the influence of rainfall on increases in flow rates. Nevertheless, not all exceedances can be attributed solely to precipitation. Some appear to be linked to elevated baseline flows, while others may result from hydraulic processes whose effects become noticeable only after a certain time lag. To better anticipate these events, a decision tree model was developed to estimate the likelihood of a bypass occurring within the following hour. The model shows that the current flow at Exona and the flow observed during the previous hour are the most decisive variables. It also identifies vigilance thresholds around **920 m³/h** and **997 m³/h**. This work thus provides a decision-support basis for a more anticipatory management of the by-pass, combining the monitoring of the incoming flow at Exona and recent rainfall.

Keywords:

1. By-pass
2. Rainfall events
3. Sanitation
4. Shared management
5. Wastewater treatment plant

LISTE DES ABREVIATIONS

2iE	:	Institut International d'Ingénierie de l'Eau et de l'Environnement
CA GPS	:	Communauté d'Agglomération Grand Paris Sud
CART	:	Classification And Regression Trees
CEE	:	Communauté Economique Européenne
EH	:	Equivalent Habitant
ENGEES	:	Ecole Nationale du Génie de l'Eau et de l'Environnement de Strasbourg
IETD	:	Inter-Event Time Definition
INSEE	:	Institut National de la Statistique et des Etudes Economiques
MBBR	:	Moving Bed Biofilm Reactor
PCB	:	Polychlorobiphényles
SAUR	:	Société d'Aménagement Urbain et Rural
SIARCE	:	Syndicat Intercommunal d'Aménagement, de Rivières et du Cycle de l'Eau
SPL	:	Société Publique Locale
STEP	:	Station de Traitement des Eaux Polluées
SYMETRIE	:	Syndicat Mixte d'Etudes et de Traitement Intersyndical des Eaux
US EPA	:	United States Environmental Protection Agency
WMO	:	World Meteorological Organization

SOMMAIRE

DEDICACE I

REMERCIEMENTS II

RESUME III

ABSTRACT IV

LISTE DES ABREVIATIONS V

SOMMAIRE VI

LISTE DES TABLEAUX VIII

LISTE DES FIGURES IX

INTRODUCTION 1

I. PRESENTATION DE LA STRUCTURE D'ACCUEIL ET DE LA ZONE D'ETUDE 4

I.1 Présentation de la structure d'accueil 4

I.2 Présentation de la zone d'étude 7

II. PRESENTATION DU PROJET 16

II.1 Contexte 16

II.2 Description du système épuratoire et des réseaux associés 18

II.3 Objectifs de l'étude 23

II.4 Résultats attendus 23

III. MATÉRIELS ET MÉTHODES 24

III.1 Données utilisées 24

III.2 Préparation et structuration des données 25

III.3 Identification et analyse des épisodes pluvieux 26

III.4 Analyse hydraulique de la station 29

III.5 Modélisation et aide à la décision 34

IV. RESULTATS D'ETUDE TECHNIQUE 39

IV.1 Caractérisation des épisodes pluvieux de l'année 2024 39

IV.2 Analyse mensuelle de la relation pluie-débit à Exona 43

IV.3 Analyse saisonnière des dépassements du seuil 49

IV.4 Identification et quantification des eaux parasites pluviales à la station Exona 52

IV.5 Modélisation prédictive du risque de bypass et définition de règles de gestion 60

V. LIMITES ET PERSPECTIVES 68

V.1 Coordination avec le réseau 68

V.2 Limites de l'étude 69

CONCLUSION 70

RECOMMANDATIONS 71

BIBLIOGRAPHIE 73

ANNEXES 76

LISTE DES TABLEAUX

Tableau 1: Charges nominales du système épuratoire Evry-Exona 22
Tableau 2: Variables utilisées dans le modèle de prédiction 36
Tableau 3. Exemple de matrice de confusion 37
Tableau 4: Synthèse statistique des épisodes pluvieux identifiés par saison 39
Tableau 5: Synthèse saisonnière des épisodes de dépassement du seuil de bypass 49
Tableau 6: Synthèse des règles opérationnelles de gestion du by-pass 67

LISTE DES FIGURES

Figure 1:Organigramme de la SPL Confluence Seine Essonne Energie 5
Figure 2:Carte de la localisation de la zone d'étude : A : Localisation de la région de l'Île de France ; B : Localisation du département de l'Essonne ; C : Localisation des communes d'Evry-Courcouronnes et de Corbeil-Essonnes 8
Figure 3:Vue aérienne de la localisation de la zone d'étude 9
Figure 4:Diagramme ombrothermique d'Evry-Courcouronnes et Corbeil-Essonnes, 2016-2025 (Source : Météo-France, Station Paris-Orly, WMO 07149 -prevision-meteo.ch) 12
Figure 5:Diagramme ombrothermique mensuel d'Evry-Courcouronnes et Corbeil-Essonnes, année 2024 (Source : Météo-France, Station Paris-Orly, WMO 07149 - prevision-meteo.ch) 13
Figure 6:Vue aérienne du milieu récepteur 14
Figure 7: Synoptique du système épuratoire Evry-Exona 21
Figure 8:Diagramme en boite de moustaches des épisodes pluvieux par saison de l'année 2024 ; a) Pluie cumulée par épisode ; b) Durée de épisodes ; c) Intensité moyenne ; d) Intensité maximale 40
Figure 9:Graphique de l'évaluation temporelle de la relation pluie-débit entrant à la station d'Exona en février 44
Figure 10:Graphique de l'évaluation temporelle de la relation pluie-débit entrant à la station Exona en avril 45
Figure 11:Graphique de l'évaluation temporelle de la relation pluie-débit entrant à la station Exona en juillet 46
Figure 12:Graphique de l'évaluation temporelle de la relation pluie-débit entrant à la station Exona en octobre 48
Figure 13:Diagramme en boite à moustaches des caractéristiques saisonnières des épisodes de dépassement du seuil de bypass ; a) Durée des dépassements ; b) Débit maximum ; c) Volume bypassé 50
Figure 14:Graphique de l'évaluation temporelle de la séparation du débit de base (Qbase) et du débit des eaux parasites pluviales (Qepp) à la station d'Exona en février 53
Figure 15:Graphique de l'évaluation temporelle de la séparation du débit de base (Qbase) et du débit des eaux parasites pluviales (Qepp) à la station d'Exona en avril 54
Figure 16:Graphique de l'évaluation temporelle de la séparation du débit de base (Qbase) et du débit des eaux parasites pluviales (Qepp) à la station d'Exona en juillet 56
Figure 17:Graphique de l'évaluation temporelle de la séparation du débit de base (Qbase) et du débit des eaux parasites pluviales (Qepp) à la station d'Exona en octobre 58
Figure 18:Matrice de confusion de l'arbre de décision pour la prédiction du by-pass à la station d'Exona 61
Figure 19:Courbe précision-rappel de l'arbre de décision pour la prédiction du by-pass à la station d'Exona 63
Figure 20:Importance relative des variables explicatives dans l'arbre de décision pour la prédiction du by-pass à la station d'Exona 64
Figure 21:Arbre de décision pour la prédiction du risque de by-pass à la station d'Exona ... 65

INTRODUCTION

L'assainissement des eaux usées occupe une place essentielle dans le fonctionnement des territoires urbains. Il permet à la fois de protéger la santé publique, de préserver les milieux aquatiques et de limiter les impacts des activités humaines sur l'environnement (Eau France, 2026). Dans les villes, cette mission devient cependant de plus en plus complexe. Les réseaux de collecte et les stations d'épuration doivent faire face à des volumes variables (France Environnement, 2024), à des infrastructures parfois anciennes, à des exigences réglementaires plus strictes et à des épisodes pluvieux qui peuvent fortement perturber leur fonctionnement.

Durant les périodes sèches, les stations d'épuration opèrent souvent dans des conditions qui se rapprochent de leur fonctionnement optimal. Les flux d'eau entrants sont assez constants, ce qui permet de gérer les processus de traitement de façon stable. En revanche, pendant les épisodes de pluie, les volumes d'eau dirigés vers les installations peuvent augmenter rapidement en raison des eaux indésirables, des apports dus au ruissellement ou de charges ponctuelles sur les réseaux. Ces apports supplémentaires peuvent provoquer une surcharge hydraulique des stations, réduire les temps de séjour dans les ouvrages et dégrader les performances de traitement. Plusieurs travaux soulignent ainsi que la gestion des eaux urbaines ne peut plus être abordée uniquement à l'échelle d'un ouvrage isolé, mais doit être pensée à l'échelle globale du système, depuis le bassin versant urbain jusqu'au milieu récepteur (Chocat, 1997; Schuetze, Butler et Beck, 2002).

Cette problématique prend tout son sens dans les systèmes d'assainissement confrontés à des variations rapides de débit. Quand les réseaux ou les stations atteignent leurs limites hydrauliques, des déversements vers le milieu naturel peuvent se produire. Ces rejets représentent un enjeu environnemental de premier plan, puisqu'ils risquent d'amener des eaux mal traitées jusque dans les cours d'eau. En France, l'arrêté du 21 juillet 2015 relatif aux systèmes d'assainissement insiste d'ailleurs sur la nécessité d'éviter tout rejet direct d'eaux usées par temps sec, et de limiter au maximum les déversements vers le milieu récepteur, sauf situations exceptionnelles liées à de fortes pluies.

Dans ce contexte, les exploitants disposent de plusieurs leviers pour sécuriser le fonctionnement des ouvrages. Parmi eux, le recours au by-pass peut permettre de protéger une station ou certains équipements contre une surcharge hydraulique. Toutefois, ce dispositif doit rester maîtrisé, car

il peut aussi déplacer la contrainte vers un autre ouvrage ou augmenter le risque de rejet vers le milieu naturel. Les retours d'expérience sur les systèmes d'assainissement montrent ainsi que la gestion des surcharges nécessite une coordination fine entre les réseaux, les stations d'épuration et les ouvrages de stockage ou de délestage (Schuetze, Butler et Beck, 2002; US EPA, 2004).

C'est dans ce cadre que s'inscrit le présent mémoire, consacré à l'optimisation de la gestion mutualisée des stations d'épuration d'Evry-Courcouronnes dénommé Evry et de Corbeil-Essonnes dénommé Exona. Ces deux stations, situées sur le territoire de Grand Paris Sud, ont été historiquement construites et exploitées par deux maîtres d'ouvrage distincts. La station d'Evry-Courcouronnes relève de la Communauté d'Agglomération Grand Paris Sud (CA GPS), tandis que la station d'Exona appartient au Syndicat Intercommunal d'Aménagement, de Rivières et du Cycle de l'Eau (SIARCE). Afin de mieux coordonner l'exploitation du site, les deux collectivités ont opté pour une gestion intégrée de la situation, concrétisée par la création de la SPL Confluence Seine Essonne Énergie. Cette gestion mutualisée s'appuie notamment sur un by-pass permettant de transférer une partie des effluents bruts d'Exona vers Evry en cas de besoin

Cependant, l'année 2024 a mis en évidence les limites de ce fonctionnement. En raison d'un fonctionnement difficile de la station d'Exona, principalement lié à une faible fonctionnalité du bassin d'anoxie, le débit entrant admissible a été réduit de 1 200 m³/h à 1 000 m³/h. Avec cette limitation, le by-pass est utilisé beaucoup plus souvent, presque tout le temps. De plus, il a beaucoup plu en 2024, ce qui a mis les réseaux et les stations sous forte pression. Pendant les épisodes de pluie intense, le by-pass d'Exona peut entrer en concurrence avec le siphon qui passe sous la Seine et qui amène les eaux usées de la rive droite vers Evry. Cela augmente le risque de refoulements et de rejets dans la Seine. On voit bien qu'une gestion commune entre les deux stations ne peut pas se résumer à simplement transférer des volumes d'eau : il faut comprendre comment le système se comporte d'un point de vue hydraulique, repérer les moments à risque, et fixer des règles simples et claires que les exploitants peuvent réellement appliquer. Le problème n'est donc pas seulement technique, il est aussi organisationnel : les décisions doivent être mieux coordonnées entre les stations d'épuration, les réseaux qui transportent l'eau, et les équipes qui gèrent le site au quotidien.

L'objectif de ce mémoire est de mieux comprendre dans quelles conditions hydrauliques et météo le by-pass entre Exona et Evry-Courcouronnes est utilisé, pour donner aux exploitants des outils qui les aident à prendre les bonnes décisions. Concrètement, il s'agit de trois choses : caractériser les épisodes de pluie de 2024, analyser comment Exona a fonctionné pendant ces épisodes, et proposer des règles simples pour aider à décider quand utiliser le by-pass.

Pour cela, la méthode utilise les données hydrauliques et météo enregistrées à Exona en 2024. On commence par repérer les épisodes de pluie et regarder leur durée, leur quantité et leur intensité. Ensuite, on étudie les moments où le seuil fixé sur le site a été dépassé. On distingue aussi le débit normal (sans pluie) du débit lié aux eaux parasites pluviales, pour mieux voir quelle part de l'eau qui arrive vient vraiment de la pluie. Enfin, un arbre de décision est utilisé pour aider à anticiper, à court terme, le risque de devoir utiliser le by-pass, à partir des données hydrauliques et météo disponibles.

Ce mémoire s'organise en plusieurs parties, consacrées respectivement à la structure d'accueil et à la zone d'étude, au projet et à ses enjeux, à la méthodologie adoptée, aux résultats obtenus, et enfin à leur discussion ainsi qu'aux recommandations associées.

I. PRESENTATION DE LA STRUCTURE D'ACCUEIL ET DE LA ZONE D'ETUDE

I.1 Présentation de la structure d'accueil

I.1.1 Genèse de la SPL

La Société Publique Locale Confluence Seine Essonne Energie a été créée par la Communauté d'agglomération Grand Paris Sud Seine-Essonne Sénart (CA GPS) et le Syndicat Intercommunal d'Aménagement, de Rivières et du Cycle de l'Eau (SIARCE), tous deux compétents en matière d'assainissement.

La CA GPS gère la station d'épuration d'Evry-Courcouronnes, d'une capacité de 250 000 équivalents-habitants, tandis que le SIARCE est propriétaire de la station Exona, mitoyenne, d'une capacité de 96 000 équivalents-habitants.

La gestion distincte de ces deux installations a fait l'objet de recommandations de la Cour des comptes dans son rapport annuel de 2013, invitant à une rationalisation de leur mode de gestion. Des études conjointes techniques, juridiques, financières et de gouvernance ont alors été menées, concluant à l'intérêt d'une mutualisation fonctionnelle des ouvrages, tout en permettant à chaque entité de conserver la propriété de ses équipements et la maîtrise de ses tarifs.

Cette réflexion a finalement abouti à la création d'une société publique locale. La décision a été actée par délibérations en décembre 2019, et la SPL a vu officiellement le jour le 12 mars 2021.

I.1.2 Missions de la SPL

La SPL Confluence Seine Essonne Énergie pilote l'exploitation mutualisée des stations d'épuration d'Evry-Courcouronnes et d'Exona. Son rôle ne se limite pas à la supervision administrative des ouvrages : elle définit également les orientations techniques, suit la performance des installations et coordonne les différents acteurs impliqués dans l'exploitation.

L'exploitation opérationnelle des deux stations est confiée à la société SAUR dans le cadre d'un contrat de prestation de service. La SPL assure donc un rôle de maître d'ouvrage et de pilote technique. Elle veille à ce que l'exploitation réponde aux objectifs de continuité de service, de conformité réglementaire, de performance environnementale et d'optimisation des ressources disponibles.

Dans ce cadre, ses principales missions peuvent être résumées autour de quatre axes :

- Le suivi du fonctionnement des stations d'épuration ;
- La coordination de l'exploitation entre les deux sites ;
- L'analyse des performances hydrauliques et épuratoires ;
- La définition de règles de gestion permettant d'améliorer la robustesse du système.

I.1.3 Organisation de la structure

La SPL Confluence Seine Essonne Énergie est une structure de taille relativement réduite, ce qui favorise une proximité entre les fonctions de décision, de suivi technique et d'exploitation. Elle est organisée autour d'une direction générale, d'une direction technique et d'une fonction administrative.

L'organigramme de la structure est présenté à la **Figure 1**.

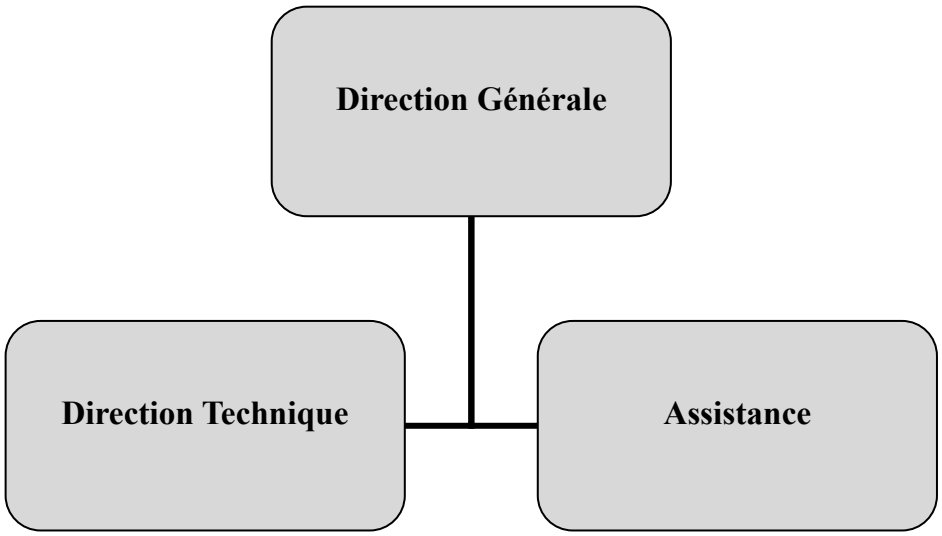


Figure 1:Organigramme de la SPL Confluence Seine Essonne Energie

I.1.4 Enjeux économiques de la mutualisation

Au-delà des aspects techniques et organisationnels, la mutualisation de la gestion des deux stations d'épuration répond avant tout à une logique économique. Dès 2000, une étude portée par le syndicat mixte d'études et de traitement intersyndical des eaux (SYMETRIE) avait déjà identifié un scénario d'équipements mutualisés permettant de réduire le coût global des investissements, option qui n'avait toutefois pas été retenue à l'époque par les collectivités concernées. L'absence de coordination qui a suivi a conduit chaque collectivité à engager séparément d'importants travaux de mise aux normes, une situation que la Cour des comptes a qualifiée, dans son rapport public annuel de 2013, de source de surcoût important pour les deux maîtres d'ouvrage. C'est cette expérience qui a motivé, une décennie plus tard, la création de la SPL Confluence Seine Essonne Énergie, avec l'objectif explicite d'éviter la répétition de ce type de surcoûts en confiant à une structure unique le pilotage de l'exploitation et des investissements futurs.

La gestion mutualisée est censée générer plusieurs types de bénéfices sur les coûts d'exploitation. Elle permet d'abord de regrouper, au sein d'un même contrat de prestation de service confié à la société SAUR, les moyens humains et techniques auparavant répartis entre deux structures distinctes, ce qui limite la duplication des fonctions de suivi et de maintenance. Elle facilite ensuite une meilleure allocation des investissements entre les deux stations, en priorisant les besoins les plus critiques plutôt qu'en imposant à chaque collectivité de financer isolément la totalité de ses propres équipements. Enfin, la mutualisation a permis à la SPL de développer une activité de valorisation énergétique des sous-produits de l'épuration, biogaz et biométhane, qui constitue un levier économique à part entière plutôt qu'un simple complément. Avant la mutualisation, le biogaz issu de la digestion des boues des deux stations était en grande partie brûlé en torchère, c'est-à-dire éliminé sans valorisation, ce qui en faisait une simple charge d'exploitation liée au traitement des boues. La mise en commun des moyens techniques et financiers des deux collectivités a permis d'investir dans une unité de méthanisation et d'épuration du biogaz commune aux deux sites, dimensionnée à une échelle qu'aucune des deux stations n'aurait pu justifier séparément au regard de sa taille. Le biométhane produit est injecté dans le réseau public de gaz dans le cadre d'un contrat d'achat à tarif réglementé, garanti pendant quinze ans, comme le prévoit la réglementation applicable à l'injection de biométhane. Ce tarif comprend un montant de référence, dégressif selon la capacité de l'installation, auquel

s'ajoute une prime spécifique aux intrants issus des boues de station d'épuration, ce qui rend la valorisation des boues d'Exona et d'Evry particulièrement incitative sur le plan financier. Contrairement aux redevances d'assainissement, dont le niveau dépend des volumes traités et reste contraint par l'acceptabilité tarifaire pour les usagers, cette rémunération constitue une recette additionnelle, prévisible sur quinze ans, qui vient directement alléger le coût net d'exploitation du système épuratoire mutualisé. Cette diversification s'inscrit par ailleurs dans une dynamique nationale favorable : la production française de biométhane injecté est passée d'environ 6 TWh en 2023 à un objectif de 44 TWh à l'horizon 2030 fixé par la programmation pluriannuelle de l'énergie, ce qui traduit un contexte réglementaire porteur pour ce type de projet. Pour Exona-Evry, une production de l'ordre de 13 à 14 GWhs de biométhane par an, soit l'équivalent de la consommation de chauffage de 2 000 à 3 300 logements neufs, illustre concrètement comment la mutualisation transforme un poste auparavant improductif en une ressource valorisée, sans se substituer pour autant à une gestion rigoureuse des coûts d'exploitation courants.

I.2 Présentation de la zone d'étude

I.2.1 Localisation géographique

La zone d'étude se situe dans le département de l'Essonne, au sud de la région Île-de-France. Elle concerne principalement le secteur d'Evry-Courcouronnes et de Corbeil-Essonnes (**Figure 2 et 3**), un territoire fortement urbanisé et marqué par la présence de la Seine, qui constitue le principal milieu récepteur des eaux traitées.

Les deux stations étudiées sont implantées à proximité immédiate l'une de l'autre. Cette proximité géographique a rendu possible la mise en place d'un by-pass hydraulique entre les deux ouvrages. Ce dispositif permet de transférer une partie des effluents bruts d'Exona vers Evry lorsque la situation l'exige, notamment en cas de maintenance ou de surcharge hydraulique.

Le territoire étudié est un secteur sensible en matière d'assainissement. On y trouve plusieurs communes raccordées aux réseaux de collecte, des ouvrages de transport conséquents, et surtout un milieu récepteur important. Gérer les flux d'eaux usées implique donc de prendre en compte plusieurs paramètres en même temps : la capacité des stations, les contraintes propres aux

réseaux, le risque de déversement, sans oublier les exigences réglementaires qui visent à protéger le milieu naturel.

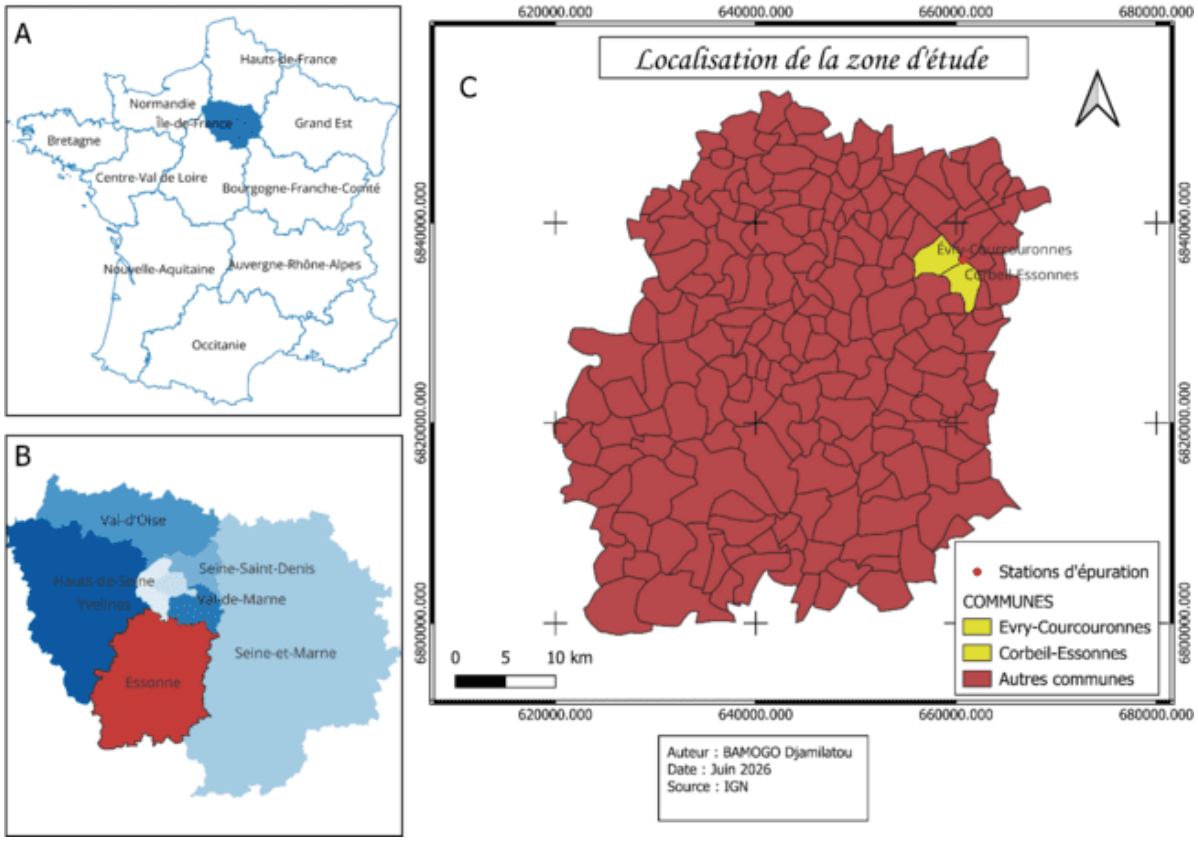


Figure 2:Carte de la localisation de la zone d'étude : A : Localisation de la région de l'Île de France ; B : Localisation du département de l'Essonne ; C : Localisation des communes d'Evry-Courcouronnes et de Corbeil-Essonnes

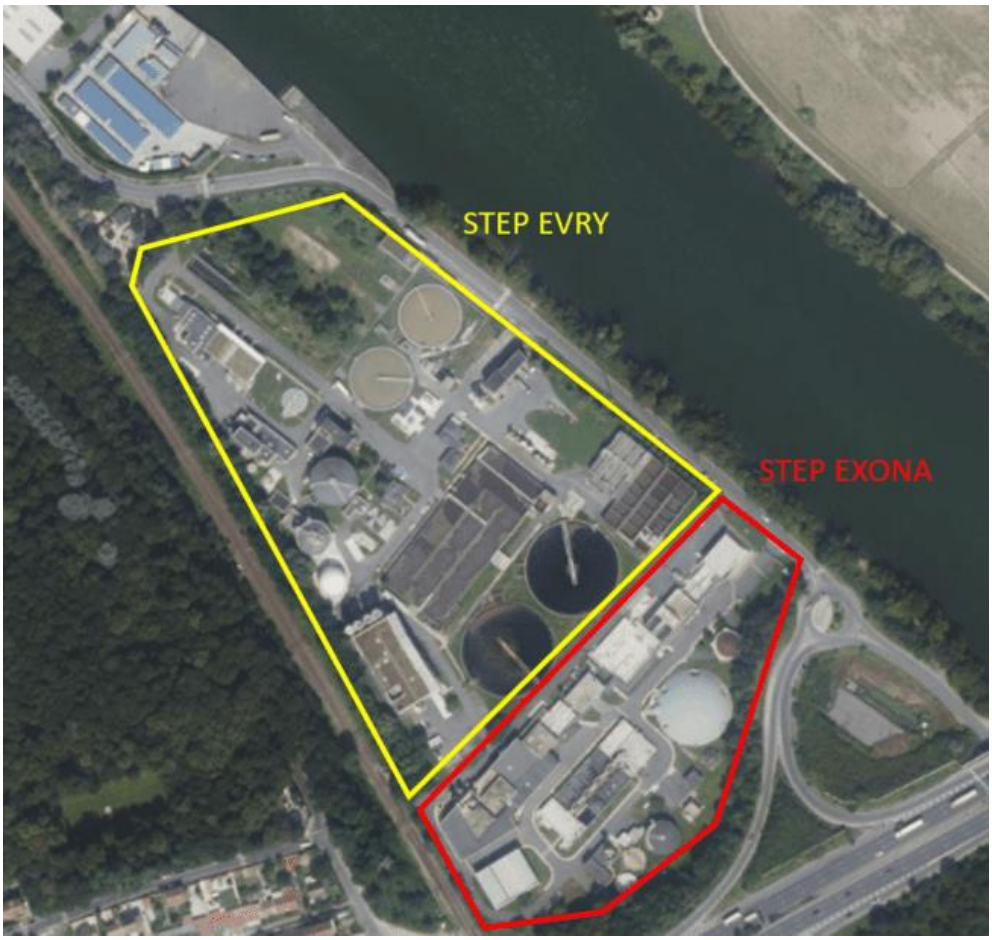


Figure 3:Vue aérienne de la localisation de la zone d'étude

I.2.2 Contexte démographique et occupation des sols

La zone d'étude regroupe une population dense et en croissance. Evry-Courcouronnes compte 66 517 habitants en 2024, pour une superficie d'environ 9 km², soit une densité d'environ 7 400 habitants par km² (INSEE, 2024). Corbeil-Essonnes compte, pour sa part, 54 471 habitants en 2023, pour une superficie de 11 km², soit une densité d'environ 4 950 habitants par km² (INSEE, 2023). Ces deux communes figurent parmi les plus densément peuplées du département de l'Essonne. Cette forte urbanisation se traduit par une occupation des sols largement dominée par le tissu urbain, les zones industrielles et commerciales et les réseaux de communication, les espaces verts et naturels occupant une part limitée du territoire. Cette artificialisation des sols entraîne une imperméabilisation importante des surfaces, qui accélère le ruissellement des eaux pluviales et accroît le risque de saturation des réseaux d'assainissement lors des épisodes

pluvieux intenses. Ce contexte territorial explique en grande partie la sensibilité du système Evry-Exona aux variations rapides de débit observées lors des épisodes de pluie.

I.2.3 Contexte urbain et hydraulique

Le système étudié se situe dans un environnement urbain dense, un contexte où les réseaux d'assainissement sont mis à rude épreuve, surtout lors des épisodes pluvieux. Même quand le réseau est en grande partie séparatif, on observe souvent des apports d'eaux parasites. Plusieurs causes peuvent expliquer ce phénomène : des infiltrations, de mauvais branchements, des remontées de nappe, ou encore des connexions accidentelles entre eaux pluviales et eaux usées.

Ces phénomènes pèsent particulièrement sur l'exploitation des stations d'épuration, car ils font varier rapidement les débits entrants. Une pluie intense ou prolongée peut entraîner une hausse significative du débit transporté par les réseaux, même si la charge polluante n'augmente pas dans les mêmes proportions. Cette situation provoque un effet de dilution, mais aussi une surcharge hydraulique des ouvrages. Les stations doivent alors traiter des volumes plus importants, parfois au détriment de leur stabilité hydraulique ou biologique (Inc *et al.*, 2003).

Dans le cas du système Evry-Exona, cette problématique est accentuée par l'interconnexion des deux stations. Lorsque le débit entrant à Exona dépasse la capacité admissible fixée à 1 000 m^3/h, une partie des effluents peut être transférée vers Evry. Cependant, la station d'Evry reçoit déjà d'autres apports, notamment ceux provenant de la rive gauche et de la rive droite via le siphon sous la Seine. En période pluvieuse, ces différents flux peuvent se cumuler et créer des situations de tension hydraulique.

I.2.4 Contexte climatique

La zone d'étude, englobant les communes d'Evry-Courcouronnes et de Corbeil-Essonnes, est caractérisée par un climat tempéré océanique dégradé, typique du Bassin parisien. Avec une température moyenne annuelle de 11,7 °C et une normale pluviométrique de 728 mm sur la période 1991-2020, ce climat ne connaît pas vraiment de saison sèche. Les précipitations sont réparties sur toute l'année, avec des cumuls mensuels qui varient entre 52 et 73 mm.

Sur la période 2016-2025, les cumuls annuels restent globalement proches de cette normale, même si on observe une variabilité d'une année à l'autre qui n'est pas négligeable. 2024 se démarque clairement : avec 798 mm, c'est l'année la plus pluvieuse de la période, soit 70 mm de plus que la normale, ce qui la place au premier rang des dix années de la série 2016-2025 (**Figure 4)**.

Ce choix méthodologique s'appuie sur le principe, reconnu en hydrologie urbaine, de l'année de référence observée. Contrairement à la construction d'une pluie de projet statistique, pertinente pour dimensionner un ouvrage neuf à partir d'une période de retour, cette approche consiste à retenir un épisode réellement survenu et localement représentatif afin de reconstituer et de diagnostiquer le fonctionnement effectif d'un système existant. Cette démarche est admise par la doctrine technique nationale en assainissement, qui reconnaît qu'une pluie historique faisant localement référence peut constituer une pluie de projet permettant de vérifier le fonctionnement d'un système. Dans le cas présent, l'année 2024 constitue le meilleur candidat pour cet exercice de diagnostic. Elle est à la fois l'année la plus contraignante sur le plan pluviométrique de la décennie récente et celle pour laquelle des données hydrauliques d'exploitation complètes ont pu être mobilisées, notamment les débits et les activations du by-pass. Cette combinaison permet une analyse couplée entre pluviométrie et réponse du système, représentative des conditions les plus sévères réellement rencontrées sur le site.

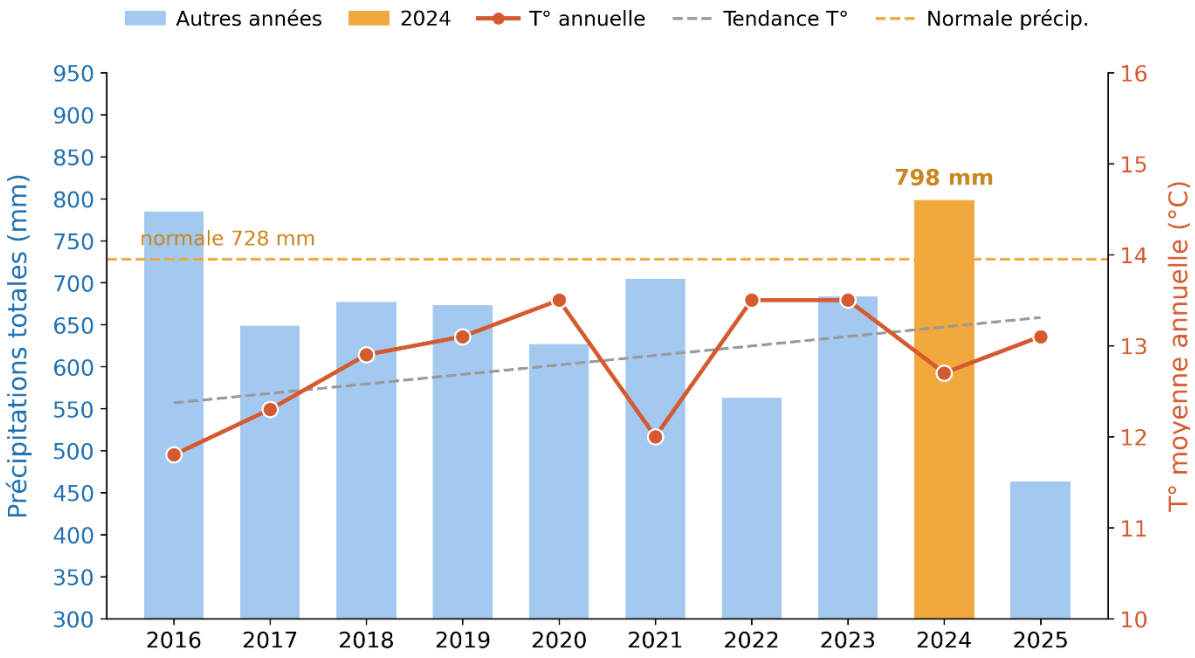


Figure 4:Diagramme ombrothermique d'Evry-Courcouronnes et Corbeil-Essonnes, 2016-2025 (Source : Météo-France, Station Paris-Orly, WMO 07149 -prevision-meteo.ch)

À l'échelle mensuelle, les précipitations de 2024 révèlent des contrastes saisonniers marqués. L'automne concentre les épisodes les plus importants, avec 116 mm en septembre et 98 mm en octobre, bien au-dessus des normales respectives de 52 mm et 65 mm. La **Figure 5** détaille cette répartition mensuelle.

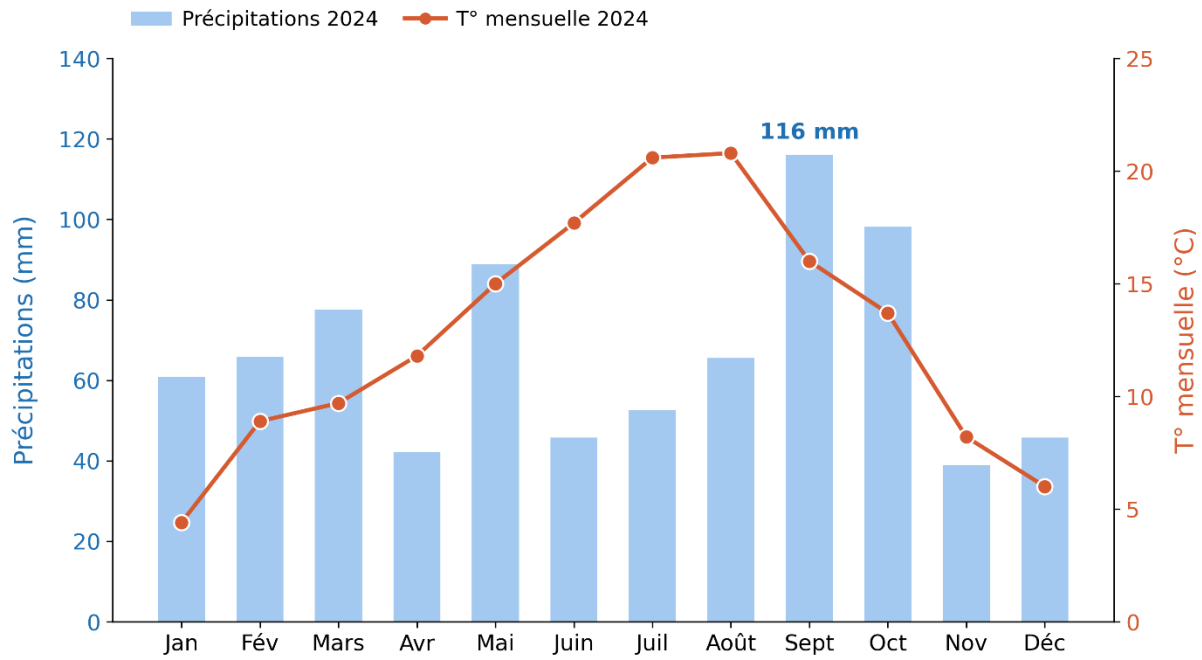


Figure 5:Diagramme ombrothermique mensuel d'Evry-Courcouronnes et Corbeil-Essonnes, année 2024 (Source : Météo-France, Station Paris-Orly, WMO 07149 - prevision-meteo.ch)

Pour faciliter l'analyse des épisodes pluvieux, l'année a été divisée selon les saisons météorologiques classiques :

- Hiver (décembre, janvier, février),
- Printemps (mars, avril, mai),
- Été (juin, juillet, août)
- Automne (septembre, octobre, novembre)

Cette approche permet de caractériser les périodes de surplus pluviométrique et d'en évaluer les effets sur le fonctionnement hydraulique du système, développés dans les sections suivantes.

I.2.5 Milieu récepteur : la Seine

La Seine dont un aperçu est donné à la **Figure 6** constitue le principal milieu récepteur des eaux traitées par les stations d'Evry-Courcouronnes et d'Exona. Elle joue donc un rôle central dans l'évaluation environnementale du système d'assainissement étudié. Les eaux traitées y sont rejetées après passage dans les ouvrages de traitement, dans le respect des prescriptions réglementaires applicables.

La qualité du milieu récepteur constitue un enjeu important. En effet, les déversements d'eaux insuffisamment traitées peuvent entraîner une dégradation temporaire de la qualité physico-chimique et biologique du cours d'eau. Ces impacts peuvent concerner l'oxygénation du milieu, les concentrations en matières organiques, les nutriments, les matières en suspension ou encore la qualité sanitaire de l'eau. La réglementation française encadre ainsi strictement le fonctionnement des systèmes d'assainissement et impose une limitation des rejets non traités ou insuffisamment traités vers le milieu naturel, en particulier en dehors des situations exceptionnelles de pluie intense (France, 2015).



Figure 6:Vue aérienne du milieu récepteur

I.2.6 Cadre réglementaire de la gestion des eaux usées et des déversements

La gestion des eaux usées et des déversements du système Evry-Exona s'inscrit dans un cadre réglementaire structuré aux échelles européennes, nationale et locale.

Au niveau européen, la directive 91/271/CEE du 21 mai 1991 relative au traitement des eaux urbaines résiduaires fixe les exigences de collecte et de traitement des eaux résiduaires urbaines ainsi que les valeurs limites de rejet au milieu récepteur. Cette directive sera progressivement remplacée par la directive (UE) 2024/3019 du 27 novembre 2024, applicable à compter du 1er août 2027, qui renforce notamment les exigences relatives aux déversements par temps de pluie et à la gestion des eaux de ruissellement urbain (Union européenne, 2024).

Au niveau national, le Code de l'environnement encadre le fonctionnement des systèmes d'assainissement collectif à travers son volet consacré à la police de l'eau. L'arrêté du 21 juillet 2015 relatif aux systèmes d'assainissement collectif précise les obligations opérationnelles applicables aux stations d'épuration et aux réseaux de collecte (France, 2015). Il impose notamment d'éviter tout rejet direct d'eaux usées en temps sec, hors situations exceptionnelles, et de limiter les déversements par temps de pluie aux seules situations où les capacités des ouvrages sont dépassées. Les déversoirs d'orage et les points de rejet vers le milieu récepteur doivent être localisés, décrits et surveillés dans le cadre de l'autosurveillance réglementaire, qui permet de statuer sur la conformité du système de collecte par temps de pluie (France, 2015).

A l'échelle du bassin, le Schéma Directeur d'Aménagement et de Gestion des Eaux Seine-Normandie fixe les objectifs de qualité et de quantité des masses d'eau du bassin, dont la Seine, et encadre les autorisations de rejet délivrées aux exploitants (Agence de l'eau Seine-Normandie, 2022). C'est dans ce cadre réglementaire que s'inscrit l'exploitation du système Evry-Exona, dont le fonctionnement par temps de pluie, et notamment le recours au by-pass entre les deux stations, doit rester compatible avec les exigences de conformité imposées par ces différents textes.

II. PRESENTATION DU PROJET

II.1 Contexte

En France, ce sont les collectivités territoriales qui gèrent les eaux usées, et elles peuvent se regrouper en syndicats pour mutualiser leurs moyens. C'est justement le cas dans l'Essonne, où les stations d'épuration d'Evry-Courcouronnes et d'Exona ne dépendent pas du même maître d'ouvrage. Exona relève du Syndicat Intercommunal d'Aménagement, de Rivières et du Cycle de l'Eau (SIARCE), tandis que Evry est rattachée à la Communauté d'Agglomération de Grand Paris Sud (CA GPS). Longtemps exploitées chacune de leur côté, ces deux stations ont été regroupées en 2020 au sein d'une Société Publique Locale (SPL), qui en assure aujourd'hui la gestion intégrée et qui a confié l'exploitation quotidienne à la société SAUR.

La station d'Exona présente une particularité technique notable, elle ne dispose que d'une seule file d'eau, sans équipement de traitement en parallèle. Cela veut dire que l'arrêt d'un seul équipement entraîne l'arrêt total de la station. Pour garantir la continuité du service, un canal de by-pass a été installé entre Exona et la station d'Evry. Ce canal permet de rediriger les eaux brutes d'Exona vers Evry pour qu'elles y soient traitées.

En 2024, une pluviométrie exceptionnellement élevée s'est ajoutée à une défaillance du bassin d'anoxie d'Exona, ce qui a poussé l'exploitant à réduire le débit admissible de la station, passant de 1 200 m³/h à 1 000 m³/h. Cette défaillance trouve son origine dans un encombrement progressif du bassin par des dépôts sédimentaires. Une inspection subaquatique réalisée en avril 2024 a mis en évidence une accumulation de 238 m³ de dépôts, un mélange de boues chargées en sable, de biomédias K1 (les supports bactériens du procédé MBBR) et de filasse concentrée autour de la crépine centrale, recouvrant 41,6 % du radier du bassin sur une épaisseur moyenne de 45,5 cm (cf. **Annexe 1**). Ces dépôts résultent d'un brassage insuffisant du bassin par les agitateurs en place, qui ne parviennent à mobiliser qu'une partie du volume. Or le bon fonctionnement des biomédias K1, dont l'efficacité dépend de leur maintien en suspension dans la colonne d'eau, est directement compromis par cet encombrement : privés d'un brassage suffisant, les K1 tendent à se stabiliser au fond du bassin, réduisant d'autant la capacité de dénitrification de l'ouvrage. Un suivi réalisé en novembre 2025 confirme la persistance, voire l'aggravation, de ce phénomène, avec un volume de dépôts porté à 381 m³, soit une hausse de

60 %, et un taux de recouvrement du radier passé à 50,7 % (cf. **Annexe 1**). Le recours au by-pass vers Evry est ainsi devenu quasi permanent.

Par ailleurs, lors des épisodes pluvieux importants, ce by-pass entre en compétition avec le siphon sous la Seine, qui achemine les eaux usées de la rive droite vers Evry, créant des risques de refoulements et de déversements non conformes dans la Seine.

II.2 Description du système épuratoire et des réseaux associés

II.2.1 Présentation du système épuratoire Evry-Exona

Le site épuratoire de la STEP d'Evry, qui appartient à la Communauté d'agglomération Grand Paris Sud (CA GPS), se trouve sur la commune d'Evry, tout près du site épuratoire de Corbeil-Essonnes appelé Exona, qui relève quant à lui du Syndicat Intercommunal d'Aménagement, de Rivières et du Cycle de l'Eau (SIARCE). Ensemble, ces deux stations constituent le site épuratoire Evry-Exona.

- **EVRY**

Mise en service en 1991 sur la rive gauche de la Seine, la station d'Evry a connu plusieurs phases de modernisation, en 2009 puis en 2013. Elle traite les eaux usées domestiques et assimilées de 16 communes collectées par un réseau majoritairement séparatif (99% de réseau séparatif et 1% de réseau unitaire), pour une capacité nominale de 250 000 EH, soit un débit journalier de référence de 48 000 m³/j et un débit de pointe de 3 600 m³/h.

Le traitement biologique repose sur deux files parallèles en boues activées. La file 1, plus ancienne, fonctionne en boues activées faible charge dans un bassin biologique de 24 000 m³, avec une élimination du phosphore par injection de chlorure ferrique et un apport de carbone par méthanol. Elle absorbe jusqu'à 60 % du débit de pointe nominal, soit 2 160 m³/h, ses vis de relevage pouvant monter jusqu'à 2 950 m³/h. La file 2, plus récente, associe un bassin biologique de 11 260 m³ à une filtration membranaire de type Biosep, sur 48 000 m² de surface membranaire installée, et traite les 40 % restants, soit 1 440 m³/h. Ce procédé, développé par Veolia, est un bioréacteur à membranes immergées, les membranes remplacent le clarificateur secondaire classique : plutôt que de laisser les boues décanter par gravité, l'eau traitée est directement aspirée à travers des membranes à fibres creuses plongées dans le bassin biologique, sous faible pression. Cette séparation par filtration permet de travailler avec une concentration de boues plus élevée que sur la file 1, ce qui réduit l'emprise au sol de l'ouvrage à capacité égale, et produit une eau de meilleure qualité, débarrassée des matières en suspension et d'une grande partie des germes, sans risque de perte de boues par foisonnement filamenteux. En sortie de file 1, un traitement tertiaire par tamis d'affinage complète l'épuration avant rejet en Seine. Les boues extraites des deux files sont épaissies puis stabilisées par digestion

anaérobie, avec valorisation du biogaz produit, avant d'être déshydratées par centrifugation, avec une siccité de 25 %, puis évacuées de la station. Elles ne sont plus destinées à l'épandage agricole direct depuis un épisode de pollution aux PCB survenu entre 2011 et 2014.
L'arrêté d'autorisation impose des rendements minimaux de 92 % pour la DBO5, 87 % pour la DCO, 95 % pour les MES, 85 % pour le NTK et 80 % pour l'azote global. Sur le plan de l'exploitation, la station présente une contrainte notable : le sécheur thermique des boues est à l'arrêt depuis 2020 pour des raisons techniques, économiques et de sécurité, ce qui explique le recours systématique à la déshydratation sans séchage préalable.

- **EXONA**

Également implantée sur la commune d'Evry-Courcouronnes, la station d'Exona traite les eaux usées de 12 communes, collectées par un réseau séparatif, pour une capacité nominale de 96 000 EH, même si elle absorbe en pratique des charges équivalentes à 100 000 EH. Le débit journalier de référence est de 15 620 m³/j pour un débit de pointe nominal de 1 800 m³/h.
Contrairement à Evry, Exona ne fonctionne pas en boues activées mais sur un procédé à cultures fixées. Le traitement s'organise en un bassin d'anoxie assurant la dénitrification sur bio-médias de type K1, un bassin carbone traitant la pollution carbonée sur bio-médias de type Chip, des flottateurs séparant les boues de l'eau traitée, puis un étage de biofiltres Biofor pour l'élimination de l'ammoniaque. Un post-traitement complète le dispositif, avec une étape de post-nitrification puis de post-dénitrification en réacteurs à lit mobile de type MBBR, cette dernière étape nécessitant un ajout de méthanol comme source de carbone. Entre 1 200 et 1 800 m³/h, le traitement fonctionne déjà en mode dégradé, avec une recirculation limitée vers le bassin d'anoxie, ce qui réduit d'autant la marge disponible avant d'atteindre la capacité nominale. Les boues, comme à Evry, sont épaissies puis digérées avant d'être envoyées, selon la période de l'année, vers deux filières distinctes : le compostage en casiers fermés, dont le produit normalisé est valorisé en agriculture, et l'épandage direct. La filière de déshydratation par filtre-presse, prévue pour compléter ce dispositif, n'est aujourd'hui pas fonctionnelle, ce qui constitue une contrainte d'exploitation supplémentaire en période de forte production de boues.
L'arrêté d'autorisation impose des rendements minimaux proches de ceux d'Evry, un peu plus stricts sur l'azote (93 % pour la DBO5, 89 % pour la DCO, 95 % pour les MES, et 75 à 80 % pour l'azote global selon la période). Plusieurs contraintes d'exploitation pèsent par ailleurs sur

la station. Elle fonctionne avec une seule file d'eau, sans ligne de traitement parallèle, ce qui la rend vulnérable à toute panne d'équipement, comme l'a montré la défaillance du bassin d'anoxie en 2024.

Le fonctionnement des deux stations de traitement, ainsi que les points de connexion entre les deux stations, sont représentés dans le synoptique de la **Figure 7**. Les schémas fonctionnels détaillés de chacun sont présentés en **Annexe 2**.

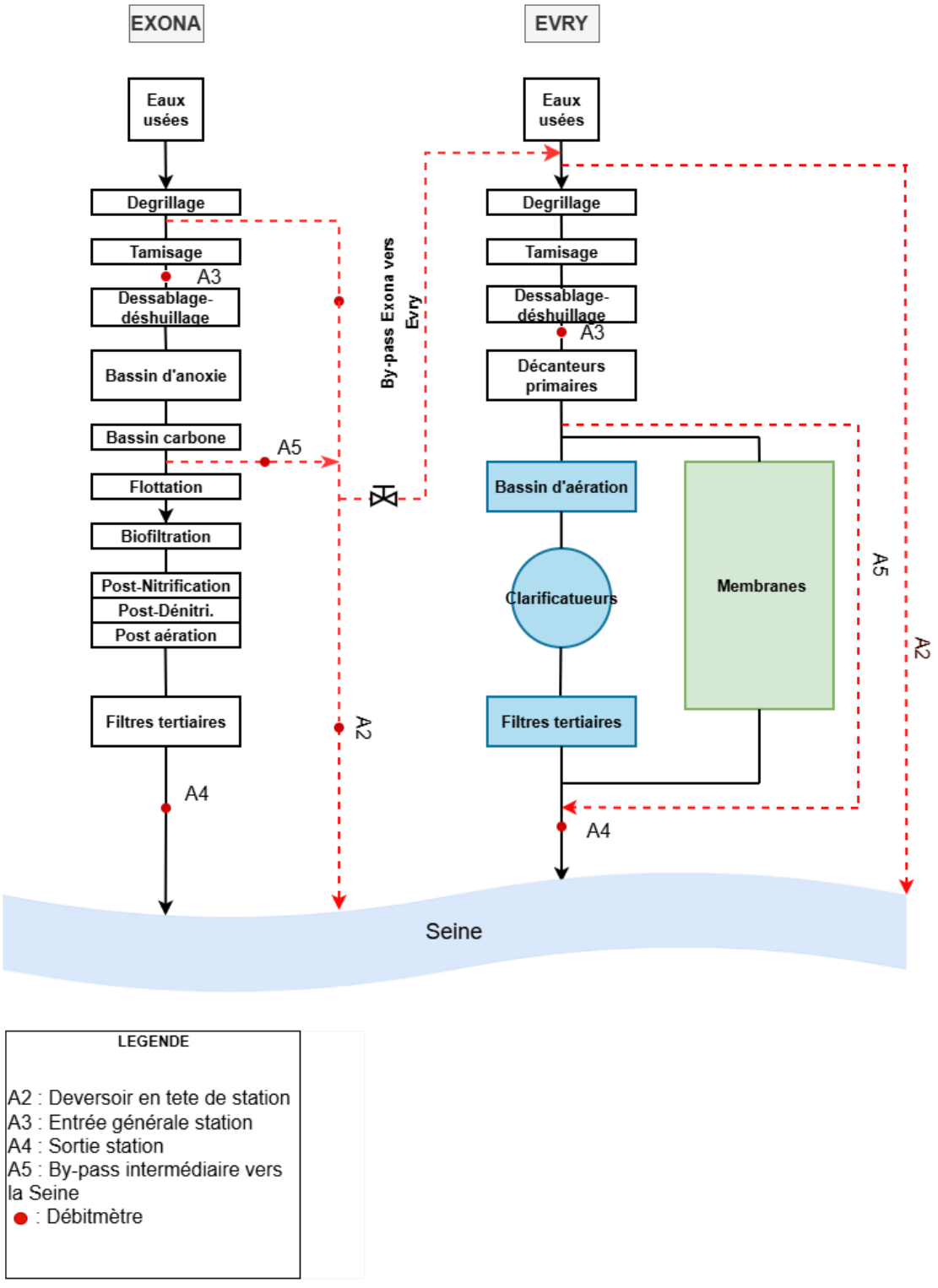


Figure 7: Synoptique du système épuratoire Evry-Exona

Les charges nominales des deux stations sont synthétisées dans le **Tableau 1**.

Tableau 1: Charges nominales du système épuratoire Evry-Exona

	Charges nominales	
Paramètres	**EVRY**	**EXONA**
Débit journalier	48 000 m^3/j	15 620 m^3/j
Débit de pointe	3 600 m^3/h	1 800 m^3/h
MES	18 300 kg/j	7 550 kg/j
DCO	32 100 kg/j	13 000 kg/j
DBO5	15 000 kg/j	5 800 kg/j
NTK	2 980 kg/j	1 300kg/j
Pt	510 kg/j	220 kg/j

II.2.2 Réseaux de collecte

Les réseaux de collecte tiennent une place centrale dans le fonctionnement du système Evry-Exona. C'est par eux que transitent les eaux usées, depuis les communes raccordées jusqu'aux stations d'épuration. Et c'est justement leur comportement hydraulique qui pèse sur les débits reçus par les stations, en particulier quand il pleut.

Le réseau qui alimente la station d'Evry est pour l'essentiel séparatif, avec seulement une petite portion en unitaire. Il collecte les eaux usées de plusieurs communes, aussi bien en rive gauche qu'en rive droite de la Seine. Les apports de la rive droite passent notamment par le siphon sous la Seine, un ouvrage qui occupe une place stratégique dans le système. On peut d'ailleurs le considérer comme un point sensible, dans la mesure où il permet aux effluents de franchir la Seine pour rejoindre la station d'Evry.

En période pluvieuse, les réseaux voient leur débit grimper rapidement. Et lorsque ces apports s'ajoutent aux volumes déjà transférés depuis Exona, on peut observer une forme de concurrence hydraulique entre les deux sources.

Le réseau d'Exona, lui, dessert plusieurs communes également et reste lui de configuration séparative. Pourtant, les données recueillies en 2024 montrent bien que les épisodes pluvieux pèsent de façon notable sur les débits entrants à la station. Ce n'est pas un détail. Cela montre

que même un réseau séparatif n'est pas à l'abri des eaux parasites pluviales. Et c'est justement cette sensibilité qui pousse à s'intéresser plus en profondeur aux épisodes de pluie, aux dépassements de débit, et à la part que prennent les eaux parasites dans le débit total.

II.3 Objectifs de l'étude

L'objectif général de ce mémoire est de développer un outil d'aide à la décision permettant d'optimiser la gestion hydraulique coordonnée des STEP d'Evry et d'Exona en contexte de forte variabilité.

Cet objectif général se décline en trois objectifs spécifiques :

- Identifier et caractériser les épisodes pluvieux survenus au cours de l'année 2024 ;
- Analyser le fonctionnement hydraulique de la station d'Exona pendant les épisodes pluvieux ;
- Proposer des règles opérationnelles d'aide à la décision pour la gestion du by-pass.

II.4 Résultats attendus

Les résultats attendus de cette étude sont les suivants :

- Les épisodes pluvieux survenus au cours de l'année 2024 ont été identifiés et caractérisés ;
- L'analyser du fonctionnement hydraulique de la station d'Exona pendant les épisodes pluvieux est effectuée ;
- Des règles opérationnelles d'aide à la décision pour la gestion du by-pass sont proposés.

III. MATÉRIELS ET MÉTHODES

III.1 Données utilisées

III.1.1 Nature et sources des données collectées

La réalisation de cette étude repose principalement sur l'exploitation de données hydrauliques, pluviométriques et opérationnelles relatives au fonctionnement de la station d'épuration d'Exona. Ces données proviennent principalement des outils de supervision et des documents techniques mis à disposition par la SPL Confluence Seine Essonne Énergie et par l'exploitant des installations. Elles couvrent principalement l'année 2024, qui constitue une année particulièrement intéressante pour l'étude en raison de sa pluviométrie élevée et de la sollicitation importante du système.

Les principales données utilisées sont les suivantes :

- Les débits entrants mesurés à la STEP d'Exona ;
- Les hauteurs de pluie enregistrées au pas de temps infra-journalier ;
- Les intensités pluviométriques ;
- Les volumes transférés via le by-pass Exona-Evry ;
- Les caractéristiques techniques de la station d'Exona ;

Le traitement de ces données a été réalisé à l'aide du logiciel RStudio, utilisé pour l'ensemble des traitements statistiques et graphiques. Pour cela, plusieurs scripts ont été développés (cf. **Annexe 3**). Ils permettent d'automatiser différentes étapes dont l'identification des épisodes pluvieux, la comparaison entre saisons, l'exploration des débits, l'analyse des dépassements, la séparation entre débit de base et eaux parasites pluviales, et enfin la construction d'un modèle prédictif simple.

III.2 Préparation et structuration des données

Avant toute analyse, les données brutes ont été préparées afin de garantir leur cohérence. Cette étape est essentielle, car les chroniques issues de l'exploitation peuvent contenir des valeurs manquantes, des formats de date hétérogènes, des pas de temps irréguliers ou des valeurs aberrantes.

Les données ont d'abord été organisées autour d'un pas de temps de 15 minutes, qui permet de suivre finement l'évolution de la pluie et des débits. Les variables principales utilisées sont la date et l'heure de mesure, la pluie en mm, l'intensité pluviométrique en mm/h, le débit entrant à Exona en m^3/h.

Pour les données d'identification des événements pluvieux, les dates sont au format jour/mois/année heure:minute. Lorsqu'une valeur de pluie ou d'intensité est manquante, elle est remplacée par zéro, en partant du principe qu'une absence de valeur signifie qu'il n'a pas plu. Cette hypothèse tient pour une analyse exploratoire, mais il faut la garder en tête au moment d'interpréter les résultats.

Un soin particulier a aussi été apporté à l'harmonisation des unités. Les débits sont exprimés en m^3/h, les pluies en mm et les intensités en mm/h. Cette homogénéité permet ensuite de calculer correctement les volumes, les cumuls et les dépassements.

III.3 Identification et analyse des épisodes pluvieux

III.3.1 Méthode d'identification (IETD)

L'identification des épisodes pluvieux constitue une étape centrale de la méthode mise en œuvre dans ce mémoire. Il ne suffit pas, en effet, de relever les jours pendant lesquels il pleut : il faut surtout parvenir à distinguer les événements qui sont réellement indépendants les uns des autres. Deux pluies rapprochées dans le temps peuvent ainsi appartenir à un même épisode si le réseau n'a pas eu le temps de revenir à un état sec entre les deux, alors que deux pluies séparées par une période sèche suffisamment longue seront, à l'inverse, considérées comme deux événements distincts.

Pour cela, la méthode retenue dans cette étude repose sur le principe de l'IETD (*Inter-Event Time Definition*), qui consiste à fixer une durée minimale sans pluie en deçà de laquelle deux épisodes rapprochés dans le temps sont regroupés en un seul événement. Plusieurs travaux (Thorndahl *et al.*, 2008; Balistrocchi et Bacchi, 2011) se sont penchés sur l'incidence de ce choix sur la caractérisation statistique des épisodes pluvieux et s'accordent sur le fait que des valeurs comprises entre 3 et 12 heures conviennent généralement aux contextes urbains, le choix précis dépendant à la fois de la taille et du degré d'imperméabilisation du bassin versant étudié. Dans cette étude, deux pluies sont ainsi considérées comme appartenant à des épisodes distincts lorsqu'elles sont séparées par 6 heures consécutives sans pluie (Balistrocchi et Bacchi, 2011) cette valeur correspondant par ailleurs à celle le plus couramment retenue pour les réseaux d'assainissement urbains de grande taille, à l'image de celui desservant la station d'Exona. Compte tenu de la résolution temporelle des données disponibles, cela revient concrètement à 24 pas de temps de 15 minutes.

Le choix d'un pas de temps de 15 minutes ne répond d'ailleurs pas uniquement aux besoins de l'identification des épisodes par la méthode IETD, il se justifie aussi par des considérations plus larges, propres au traitement des séries de débit. La séparation entre débit de base (Qbase) et eaux parasites pluviales (Qepp), qui permet d'évaluer la vulnérabilité d'un réseau aux intrusions d'eaux claires parasites, repose en effet sur des méthodes minimum glissant, filtres de séparation numérique de (Lyne, V. et Hollick, M., 1979), ou encore bilan volumique dont la robustesse dépend directement de la finesse du pas de temps d'analyse retenu. En comparant ces différentes approches sur des réseaux séparatifs en contexte urbain, (Thorndahl *et al.*, 2008) concluent que les méthodes fondées sur le filtrage numérique restent robustes face aux variations temporelles,

à condition que le pas de temps utilisé reste inférieur à 30 minutes. Ce constat conforte le choix d'un pas de 15 minutes pour l'ensemble des traitements menés dans ce mémoire.

Ce choix permet, au final, d'obtenir des épisodes cohérents du point de vue hydraulique : au-delà de plusieurs heures sans pluie, l'influence directe de l'événement précédent sur les débits commence en effet à s'atténuer, même si certains effets résiduels liés notamment au ressuyage progressif du bassin versant peuvent encore se faire sentir dans le réseau pendant un temps plus long.

III.3.2 Indicateurs calculés pour chaque épisode

Pour chaque épisode pluvieux identifié, plusieurs indicateurs sont calculés :

- La date et l'heure de début ;
- La date et l'heure de fin ;
- La durée totale de l'épisode ;
- Le cumul de pluie en mm ;
- L'intensité moyenne en mm/h ;
- L'intensité maximale en mm/h ;
- La saison météorologique ;

Les épisodes sans pluie réelle ou d'une durée inférieure ou égale à 15 minutes sont exclus afin d'éviter les artefacts. Cette étape permet de conserver uniquement les événements significatifs pour l'analyse hydraulique.

Les épisodes retenus permettent ensuite d'étudier la réponse du système aux pluies, en observant notamment le délai entre le début de la pluie et l'augmentation des débits à la station d'Exona.

III.3.3 Analyse saisonnière des épisodes pluvieux

Après l'identification des événements, une comparaison saisonnière est réalisée. Les épisodes sont regroupés selon les saisons météorologiques classiques :

- Hiver ;
- Printemps ;
- Été ;
- Automne.

Pour chaque saison, les caractéristiques des épisodes sont comparées à l'aide de boxplots ou boites à moustaches portant sur :

- La pluie cumulée ;
- La durée des épisodes ;
- L'intensité moyenne ;
- L'intensité maximale.

Cette analyse permet d'identifier les saisons les plus contributives en termes de fréquence, de cumul ou d'intensité des pluies. Elle permet aussi de repérer les épisodes atypiques susceptibles d'avoir fortement sollicité la station.

Le script de comparaison saisonnière produit une figure synthétique composée de quatre graphiques, ainsi qu'un tableau récapitulatif indiquant, pour chaque indicateur, la médiane et l'intervalle min-max par saison.

Cette étape est utile pour replacer les épisodes critiques dans leur contexte. Un événement peut être remarquable par son cumul, par son intensité ou par sa durée ; ces trois dimensions n'ont pas toujours les mêmes conséquences hydrauliques.

III.4 Analyse hydraulique de la station

III.4.1 Exploration des débits et de la relation pluie-débit

Une analyse exploratoire des chroniques de pluie et de débit est réalisée mois par mois. L'objectif est de visualiser la réponse hydraulique de la station d'Exona aux épisodes pluvieux et d'identifier les périodes où le débit dépasse le seuil de consigne.

Pour chaque mois, une figure en trois panneaux est produite :

- La pluie cumulée par pas de temps de 15 minutes ;
- L'intensité pluviométrique ;
- Le débit entrant à la STEP d'Exona.

Une ligne horizontale matérialise le seuil de 1 000 m³/h, qui correspond à la consigne actuelle de limitation du débit entrant à la station d'Exona.

Cette visualisation permet une lecture simple du comportement du système. Elle montre si les pics de débit suivent les épisodes de pluie, si les dépassements sont ponctuels ou prolongés, et si certains mois concentrent davantage de situations critiques.

III.4.2 Analyse des dépassements du seuil de 1 000 m³/h à Exona

III.4.2.1 Définition du dépassement

Le seuil de 1 000 m³/h constitue une valeur clé de l'étude. Il correspond au débit maximal actuellement admis à l'entrée de la station d'Exona en raison des limitations fonctionnelles observées sur la station. Lorsque le débit entrant dépasse ce seuil, le surplus représente un volume excédentaire susceptible d'être transféré vers Evry par le by-pass ou directement vers le milieu récepteur. Un dépassement est donc défini comme tout pas de temps satisfaisant le critère présenté à l'**Equation (1)** :

$$Q_{Exona} > 1000\ m^3/h \qquad (1)$$

Où Q_{Exona} représente le débit entrant mesuré à la station d'Exona.

III.4.2.2 Calcul du volume excédentaire

Pour chaque pas de temps de 15 minutes en dépassement, le volume excédentaire est calculé à partir de la différence entre le débit mesuré et le seuil de consigne. Comme le pas de temps est de 15 minutes, soit 0,25 heure, le volume excédentaire est calculé selon l'**Equation (2)** :

$$V_{exc} = \sum (Q_{Exona} - 1000) * 0{,}25 \qquad (2)$$

Avec :

- V_{exc} : volume excédentaire en m³ ;
- Q_{Exona} : débit entrant mesuré à Exona en m³/h ;
- 1000 : seuil de consigne en m³/h ;
- 0,25 : durée d'un pas de temps de 15 minute exprimé en heure.

Cette méthode permet d'estimer le volume qui dépasse la capacité admissible d'Exona. Elle donne donc une approximation du volume potentiellement concerné par le by-pass ou par une gestion spécifique.

III.4.3 Séparation du débit de base et des eaux parasites pluviales

III.4.3.1 Objectif de la séparation

Pour mieux comprendre l'influence de la pluie sur les débits entrants, le débit total mesuré à Exona est décomposé en deux composantes comme présente l'**Equation (3)** :

$$Q_{Exona} = Q_{base} + Q_{epp} \qquad (3)$$

Avec :

- Q_{base} : débit de base correspondant aux apports domestiques de référence ;
- Q_{epp} : contribution des eaux parasites pluviales.

Cette séparation permet de distinguer ce qui relève du fonctionnement normal du réseau (Q_{base}) et ce qui est directement lié aux apports de pluie (Q_{epp}). Elle est particulièrement utile pour identifier les épisodes où le dépassement du seuil de 1 000 m³/h est principalement dû aux eaux parasites pluviales.

III.4.3.2 Détermination de la référence temps sec (Q_{base})

Le débit de base (Q_{base}) est estimé à partir des périodes considérées comme réellement sèches. Un pas de temps est classé comme temps sec s'il n'y a eu aucune pluie dans les 48 heures précédentes et les 48 heures suivantes. Ce seuil, plus large que l'IETD retenu pour la seule séparation des épisodes pluvieux, répond à une exigence différente : il ne s'agit plus de distinguer deux événements l'un de l'autre, mais de s'assurer qu'une période servant de référence n'est affectée ni par la décrue d'un épisode précédent (la résorption des eaux parasites étant généralement plus lente que la réponse hydraulique directe) ni par l'approche d'un épisode à venir. Cette règle, qui consiste à exiger une marge de part et d'autre de la période considérée, s'inscrit dans la fourchette de 24 à 72 heures généralement retenue dans la littérature pour qualifier une période de « temps sec » de référence, suffisamment éloignée de toute influence pluviométrique (Rodríguez *et al.*, 2013).

Pour ces périodes sèches, une référence de débit est calculée en tenant compte :

- Du créneau horaire de la journée ;
- De la distinction entre jours ouvrés et week-ends ;
- D'une fenêtre glissante de ±21 jours autour de la date étudiée.

La distinction entre jours ouvrés et week-ends se justifie par le caractère structurellement différent des habitudes de consommation domestique selon le type de jour. Plusieurs études ont en effet montré que le débit observé le week-end diffère sensiblement de celui des jours ouvrés, notamment du fait de la réduction des flux pendulaires liés aux déplacements domicile-travail, ce qui se traduit par une baisse mesurable du débit moyen, en particulier le dimanche (Rodríguez *et al.*, 2013). Cette différence conduit la majorité des travaux portant sur la caractérisation du débit de temps sec à construire des profils diurnes distincts pour les jours ouvrés et pour le week-end, plutôt qu'un profil unique pour l'ensemble de la semaine. De la même manière, la prise en compte du créneau horaire répond au fait que les usages domestiques de l'eau ne se répartissent pas uniformément au cours de la journée : les habitudes de lever, les repas et les usages sanitaires génèrent des pics caractéristiques dont l'amplitude et la position varient au fil des heures, ce qui impose d'évaluer la référence de débit créneau par créneau plutôt qu'à l'échelle de la journée entière.

Quant à la fenêtre glissante de ±21 jours, son choix repose sur un compromis assez classique quand on cherche à estimer un débit de référence ; il faut une fenêtre assez longue pour lisser

la variabilité journalière aléatoire et obtenir une estimation statistiquement robuste, mais pas trop longue non plus, pour rester sensible aux variations de fond plus lentes qui touchent un réseau d'assainissement (évolution saisonnière de la nappe phréatique, variations démographiques, ou changements progressifs dans les habitudes de consommation). Ce principe de moyenne glissante sur une période de quelques semaines est une pratique reconnue dans la littérature consacrée à l'estimation du débit de temps sec (De Haas *et al.*, 2021).

Cette méthode permet d'obtenir un débit de base dynamique, capable de suivre les variations saisonnières et les habitudes journalières du réseau.

III.4.3.3 Calcul de la composante pluviale

A chaque instant, il faut déterminer si le réseau est sous l'influence d'une pluie ou non. Une période est considérée comme influencée par la pluie si au moins 0,5 mm de pluie est tombé au cours des 2 heures précédentes. Ce seuil relève davantage d'un choix pratique que d'une loi physique universelle. Le seuil de 0,5 mm sert avant tout à filtrer le bruit de mesure du pluviomètre : condensation, rosée, légère dérive du capteur. On reste volontairement au-dessus du seuil de détection habituel de 0,1 mm utilisé par l'Organisation météorologique mondiale pour qualifier un jour de pluie, afin d'éviter de classer à tort une pluie insignifiante comme un véritable épisode.

Le délai de 2 heures correspond grossièrement au temps que prend un réseau urbain pour réagir à une pluie : le débit ne bouge pas instantanément, il met parfois quelques dizaines de minutes, parfois plusieurs heures, à montrer une réaction visible. (Métadier et Bertrand-Krajewski, 2011) confirment ce principe : l'effet d'une pluie ne devient négligeable que lorsque le temps écoulé dépasse celui dont le bassin a besoin pour retrouver ses conditions de temps sec, ce délai étant propre à chaque réseau.

Une fois la période classée comme influencée ou non, le calcul devient direct. Pendant les périodes sous influence pluvieuse, la composante pluviale se calcule par différence entre le débit observé à la station d'Exona et le débit de base de référence comme présente l'**Equation (4)**.

$$Q_{epp} = \max(0\ ;\ Q_{Exona} - Q_{base}) \qquad (4)$$

En dehors des périodes sous influence pluvieuse, Q_{epp} est fixé à zéro. Le débit de base effectif est alors obtenu par l'équation 5 :

$$Q_{base_eff} = Q_{Exona} - Q_{epp} \qquad (5)$$

Ce mode de calcul garantit qu'à tout instant, les deux composantes recomposent exactement le débit total mesuré comme l'indique l'**Equation (6)** :

$$Q_{Exona} = Q_{base_eff} + Q_{epp} \qquad (6)$$

III.5 Modélisation et aide à la décision

III.5.1 Modélisation prédictive du risque de by-pass

III.5.1.1 Définition et objectif du modèle

Afin d'aller au-delà de l'analyse descriptive, un modèle simple d'aide à la décision a été développé pour prédire si un by-pass est susceptible de se produire à l'heure suivante, à partir de l'état actuel du système.

Le modèle utilisé dans cette étude est ce qu'on appelle un modèle oracle, ou arbre de décision. (CART, *Classification and Regression Trees*). Il fonctionne par une succession de questions simples sur les variables d'entrée (le débit dépasse-t-il un seuil, a-t-il plu plus de tant de millimètres dans la dernière heure). Chaque réponse oriente vers la question suivante jusqu'à une conclusion finale. (Breiman et al., 1984) ont posé les bases théoriques de cette méthode, fondée sur une segmentation récursive des variables explicatives produisant des règles simples et robustes.

Ce choix d'un arbre de décision est volontaire. L'objectif n'est pas seulement d'obtenir une bonne performance statistique, mais de produire des règles compréhensibles par les exploitants, vérifiables directement, contrairement à un modèle de type réseau de neurones qui reste une boîte noire. (Jalbert *et al.*, 2024) ont adopté la même logique pour prédire les débordements de réseaux unitaires à Montréal, en privilégiant un arbre interprétable plutôt qu'un modèle plus performant mais inexploitable au quotidien.

Le modèle est entraîné pour prédire un by-pass à un horizon d'une heure, à partir de l'état hydraulique et pluviométrique observé à l'instant présent. Il est qualifié de modèle oracle parce que cette évaluation est réalisée sur des données historiques pour lesquelles l'issue réelle, à savoir l'occurrence ou non d'un by-pass une heure plus tard, est déjà connue au moment de l'entraînement. L'objectif n'est donc pas encore de déployer un outil de prévision en temps réel, mais d'évaluer le potentiel prédictif maximal atteignable avec les variables retenues, avant d'investir dans une architecture réellement opérationnelle. Le choix de prédire un horizon futur plutôt que l'instant présent reste néanmoins intentionnel : il prépare directement le passage à une version opérationnelle, où l'état hydraulique actuel pourrait être combiné à une prévision météorologique plutôt qu'à la pluie observée, pour anticiper un by-pass avant qu'il ne survienne. L'intérêt de ce type de couplage entre prévision de pluie à courte échéance et modèle

hydraulique pour le contrôle en temps réel de réseaux d'assainissement est illustré par (Liguori *et al.*, 2012).

Dans l'étude les données sont d'abord agrégées au pas de temps horaire, en ne conservant que les heures complètes. Ce choix répond à deux exigences distinctes : d'une part, lisser le bruit de mesure inhérent aux données à pas de temps fin (15 minutes), qui peut introduire des variations erratiques peu représentatives de l'état réel du système ; d'autre part, s'aligner sur le temps de réaction réellement disponible pour les exploitants, pour qui une alerte émise à la minute près n'apporte pas de bénéfice opérationnel supplémentaire par rapport à une alerte horaire. Cette pratique est cohérente avec la littérature. (A. Bizer et Kirchhoff, 2022) ont construit leurs modèles de prédiction de débordement à partir de données de précipitation agrégées à l'heure, en identifiant des seuils d'intensité horaire critique au-delà desquels le débordement devient probable.

Le découpage des données entre apprentissage et test est effectué de manière chronologique, avec 80 % des observations pour l'apprentissage et 20 % pour le test, afin d'éviter toute fuite d'information temporelle. Ce ratio de 80/20 est une convention courante en apprentissage automatique : il réserve suffisamment de données à l'apprentissage pour que le modèle puisse dégager des règles robustes, tout en conservant un échantillon de test assez large pour évaluer sa performance de manière fiable.

III.5.1.2 Variables explicatives utilisées

Les variables utilisées pour prédire le risque de by-pass sont consignés dans le **Tableau 2** :

Tableau 2: Variables utilisées dans le modèle de prédiction

Symbole	Nom de la variable	Unité	Définition
Qexona_mean	QExona moyen de l'heure courante	m³/h	Débit moyen observé à la station Exona sur l'heure courante t
Qexona_lag1	QExona de l'heure précédente	m³/h	Débit moyen observé à la station Exona sur l'heure précédente t-1
dQ/dt	Variation du débit entre deux heures successives	m³/h	Différence entre le débit de l'heure courante et celui de l'heure précédente (Qexona,t − Qexona, t-1), traduisant la dynamique de montée ou de baisse du débit
Pluie_cumul_2 h	Cumul de pluie sur les deux dernières heures	mm	Somme des précipitations enregistrées sur les 2 heures précédant l'instant t
Pluie_cumul_6 h	Cumul de pluie sur les six dernières heures	mm	Somme des précipitations enregistrées sur les 6 heures précédant l'instant t
heure	Heure de la journée	heure	Heure courante de la journée permettant de capter le cycle journalier de consommation domestique
mois	Mois de l'année	mois	Mois de l'année permettant de capter les variations saisonnières du débit

Ces variables ont été choisies car elles traduisent à la fois l'état hydraulique immédiat du système, la dynamique récente du débit et l'influence de la pluie. Le débit courant renseigne sur la proximité du seuil de 1 000 m³/h. Le débit précédent et la variation du débit permettent de savoir si le système est stable ou en phase de montée. Les cumuls de pluie sur 2 h et 6 h donnent une information sur la sollicitation pluvieuse récente. Enfin, l'heure et le mois permettent de prendre en compte les rythmes journaliers et saisonniers du système.

III.5.1.3 Évaluation du modèle

Le modèle est évalué à l'aide de plusieurs indicateurs classiques, calculés à partir de la matrice de confusion. Une matrice de confusion est un tableau qui croise les prédictions du modèle avec les observations réelles, en distinguant les bonnes prédictions des erreurs. Le **Tableau 3** présente un exemple générique de matrice de confusion, utilisé ici pour expliquer le calcul des différents indicateurs.

Tableau 3. Exemple de matrice de confusion

	Classe prédite 1	**Classe prédite 2**	**Total observé**
Classe observée 1	a	b	$a+b$
Classe observée 2	c	d	$c+d$
Total prédit	$a+c$	$b+d$	$a+b+c+d$

Appliquée au cas du by-pass, cette matrice distingue quatre situations :

- a, le nombre de bypass prédit ET réellement survenu ;
- b, le nombre de bypass non prédit, mais réellement survenu (bypass manqué) ;
- c, le nombre de bypass prédit, mais non survenu (fausse alerte) ;
- d, le nombre de non-bypass prédit ET réellement non survenu

À partir de ces quatre valeurs, plusieurs indicateurs sont calculés :

- **L'exactitude globale (accuracy)**, définie par Exactitude = (a + d) / (a + b + c + d), mesure la proportion totale de prédictions correctes, toutes catégories confondues.
- **La précision**, définie par Précision = a / (a + c), mesure la proportion de bypass annoncés par le modèle qui se sont réellement produits.
- **Le rappel**, défini par Rappel = a / (a + b), mesure la proportion de bypass réellement survenus que le modèle a effectivement détectés.
- **Le F1-score**, défini par F1 = 2 × (Précision × Rappel) / (Précision + Rappel), combine la précision et le rappel en un seul indicateur, résumant le compromis entre les deux plutôt que de privilégier l'un au détriment de l'autre.

Ces indicateurs sont complétés par une courbe précision-rappel, construite sur une grille de seuils de classification allant de 0,05 à 0,95. Cette courbe permet de visualiser, pour chaque seuil testé, le compromis entre la précision et le rappel, et ainsi de choisir un seuil de décision adapté à l'arbitrage souhaité entre fausses alarmes et bypass manqués.

III.5.2 Définition des règles opérationnelles de gestion

À partir des résultats issus de la modélisation par arbre de décision, des règles de gestion opérationnelles du bypass ont été formalisées. L'idée derrière ces règles, c'est de donner aux exploitants une grille de lecture simple, qu'ils peuvent appliquer directement pour anticiper le risque de by-pass à court terme. Elles s'appuient sur deux types d'informations disponibles en temps réel : le débit instantané mesuré à la station d'Exona, et les conditions pluviométriques et temporelles récentes.

Ces règles ne sont pas à prendre comme des consignes figées, mais plutôt comme une base de travail pour aider à la décision. Elles doivent être discutées avec les exploitants et ajustées selon le retour d'expérience, les contraintes réglementaires et l'état réel des ouvrages.

IV. RESULTATS D'ETUDE TECHNIQUE

Cette section présente les résultats obtenus à partir du traitement des données pluviométriques et hydrauliques de l'année 2024. L'objectif principal est d'analyser le comportement de la station d'Exona en période pluvieuse, afin de mieux comprendre les conditions qui conduisent au dépassement du seuil de 1 000 m³/h et, par conséquent, à l'activation ou au maintien du by-pass vers la STEP d'Evry.

IV.1 Caractérisation des épisodes pluvieux de l'année 2024

Les épisodes ont été identifiés à partir des données de pluie au pas de temps de 15 minutes, en distinguant les événements indépendants selon une période minimale de 6h sans pluie (Balistrocchi et Bacchi, 2011). Cette approche a permis de ne pas se limiter à une analyse journalière, souvent trop grossière, et de mieux représenter la réalité hydraulique des réseaux.
L'application de la méthode IETD à la série pluviométrique 2024 a permis d'obtenir au total, 144 épisodes pluvieux, répartis sur l'ensemble des saisons (cf. **Annexe 4)**. Le printemps est la saison la plus active (54 épisodes), suivi de l'automne (34), de l'été (30) et de l'hiver (26). Le **Tableau 4** synthétise les principaux indicateurs statistiques par saison.

Tableau 4: Synthèse statistique des épisodes pluvieux identifiés par saison

Saison	Nb épisodes	Pluie médiane (mm)	Durée médiane (h)	Int. moy. médiane (mm/h)	Int. max. médiane (mm/h)
Printemps	54	2,9	2,8	0,91	7,6
Été	30	2,9	3,9	1,12	19,3
Automne	34	2,4	5,6	0,67	8,7
Hiver	26	4,0	5,0	0,42	4,8

Les résultats montrent des profils saisonniers bien différents les uns des autres. L'hiver se distingue avec la médiane de cumul la plus élevée (4,0 mm) et des épisodes plus longs, signe de pluies persistantes et diffuses. L'automne, lui, montre une forte variabilité, avec des intensités maximales qui peuvent grimper haut malgré un cumul médian plus faible (2,4 mm). Le printemps et l'été présentent des cumuls médians proches (2,9 mm), mais se distinguent par des

intensités ponctuelles plus marquées, en particulier en été, où l'intensité maximale médiane atteint 19,3 mm/h. La **Figure 8** illustre cette distribution saisonnière sous forme de boîtes à moustaches.

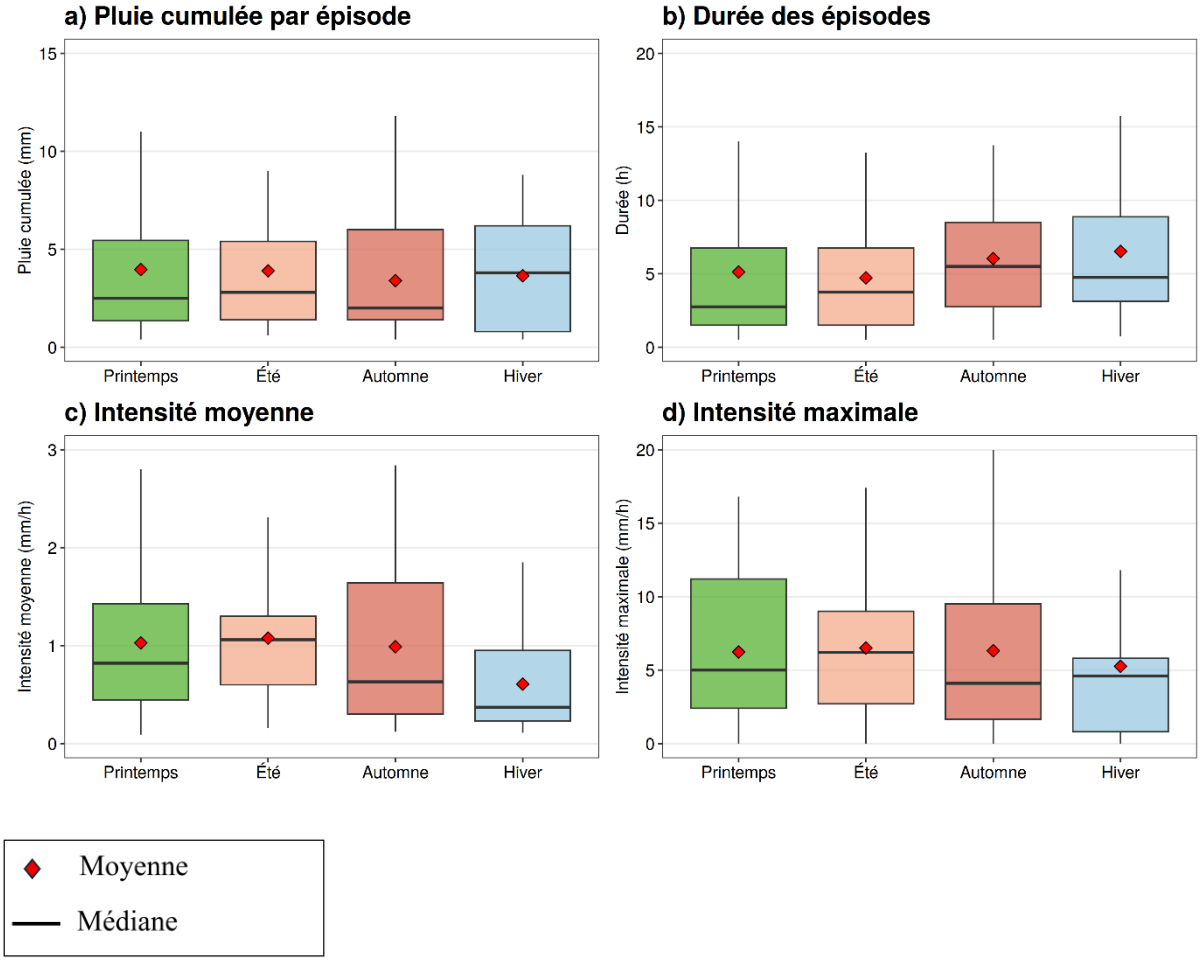


Figure 8:Diagramme en boite de moustaches des épisodes pluvieux par saison de l'année 2024 ; a) Pluie cumulée par épisode ; b) Durée de épisodes ; c) Intensité moyenne ; d) Intensité maximale.

IV.1.1 Pluie cumulée par épisode

Le boxplot de la pluie cumulée (**Figure 8a**) confirme que l'hiver présente les cumuls les plus élevés, avec une médiane de 4,0 mm. Toutefois, les quatre saisons restent relativement proches en termes de cumuls médians, ce qui suggère que c'est moins le volume total de pluie que sa dynamique intensité et durée qui différencie le comportement hydraulique du réseau selon les saisons.

La gravité d'un épisode ne se résume en effet pas au cumul total. Un cumul modéré mais très concentré dans le temps peut suffire à provoquer un pic hydraulique important, tandis qu'un cumul plus élevé mais étalé sur une longue durée donnera plutôt une montée progressive, laissant à l'exploitant davantage de temps pour réagir.

IV.1.2 Durée des épisodes

La durée des épisodes (**Figure 8b)** varie fortement d'une saison à l'autre. Les épisodes hivernaux et automnaux sont généralement les plus longs, certains pouvant dépasser une dizaine d'heures. L'hiver affiche d'ailleurs les durées les plus élevées, ce qui traduit des pluies plus persistantes. L'automne n'est pas en reste, avec lui aussi des épisodes prolongés, ce qui pourrait expliquer en partie la tension hydraulique observée sur Exona durant cette période.

Le printemps et l'été, en revanche, présentent des durées plus courtes et plus variables. Cette différence n'est pas anodine du point de vue hydraulique, un épisode long favorise une saturation progressive du réseau et le maintien de débits élevés sur plusieurs heures, alors qu'un épisode court mais intense peut provoquer un pic brutal, beaucoup plus difficile à anticiper (Nasrin, Sharma et Muttil, 2017).

IV.1.3 Intensité moyenne et intensité maximale

Au-delà des cumuls et des durées, c'est surtout l'intensité (**Figure 8c et 8d**) qui révèle le potentiel de réaction rapide du réseau. L'intensité moyenne médiane monte au printemps (0,91 mm/h) et en été (1,12 mm/h), ce qui confirme que ces deux saisons sont plutôt marquées par des pluies concentrées dans le temps. L'hiver, à l'inverse, reste sous 0,42 mm/h en moyenne médiane, ce qui confirme bien son caractère diffus et prolongé.

Pour l'intensité maximale médiane, elle grimpe particulièrement en été (19,3 mm/h) et en automne (8,7 mm/h). On voit donc apparaître des épisodes ponctuellement très intenses, capables de solliciter brutalement le réseau et ça, indépendamment même du cumul total de l'épisode.

Au final, chaque saison dessine un profil de risque hydraulique qui lui est propre. L'hiver et l'automne exposent surtout le système à une saturation progressive par accumulation, alors que le printemps et l'été l'exposent plutôt à des pics hydrauliques soudains, liés à des intensités ponctuelles élevées.

IV.2 Analyse mensuelle de la relation pluie-débit à Exona

L'exploration visuelle de la relation entre la pluie et le débit a été conduite mois par mois à partir de graphiques en séries temporelles triple panneaux, représentant au pas de 15 minutes les précipitations cumulées (a), l'intensité pluviométrique (b) et le débit entrant Q_{Exona} (c) avec le seuil de bypass à 1 000 m³/h matérialisé en ligne pointillée rouge. Quatre mois ont été sélectionnés comme représentatifs des quatre régimes saisonniers distincts identifiés sur l'année 2024 : février pour l'hiver (**Figure 9**), avril pour le printemps (**Figure 10**), juillet pour l'été (**Figure 11**) et octobre pour l'automne (**Figure 12**). Les graphiques des huit mois restants sont présentés en **Annexe 5**.

IV.2.1 Période hivernale : février

En février, les précipitations sont modérées en première moitié de mois, avec des épisodes réguliers mais d'intensité faible. A partir du 22 février, la situation bascule. Les cumuls augmentent, et le 23 février, une pointe d'intensité proche de 100 mm/h fait monter le débit jusqu'à 2 000 m³/h. Ce mois met en lumière une caractéristique importante du régime hivernal. Même sans pluie, le débit de base tourne déjà entre 800 et 1 000 m³/h, ce qui veut dire que le réseau reste en permanence tout proche du seuil de by-pass. Ce niveau de fond élevé résulte de la saturation progressive des sols et des nappes superficielles au cours de l'hiver, entretenue par une évapotranspiration très faible à cette période de l'année. Les sols n'ont pas le temps de s'assécher entre deux pluies, ce qui maintient un écoulement parasite quasi permanent vers le réseau. Dans ces conditions, il suffit d'un épisode pluvieux même modéré pour déclencher un dépassement. Le rythme circadien reste visible à travers les variations journalières du débit, mais le niveau de fond hivernal laisse très peu de marge avant la consigne.

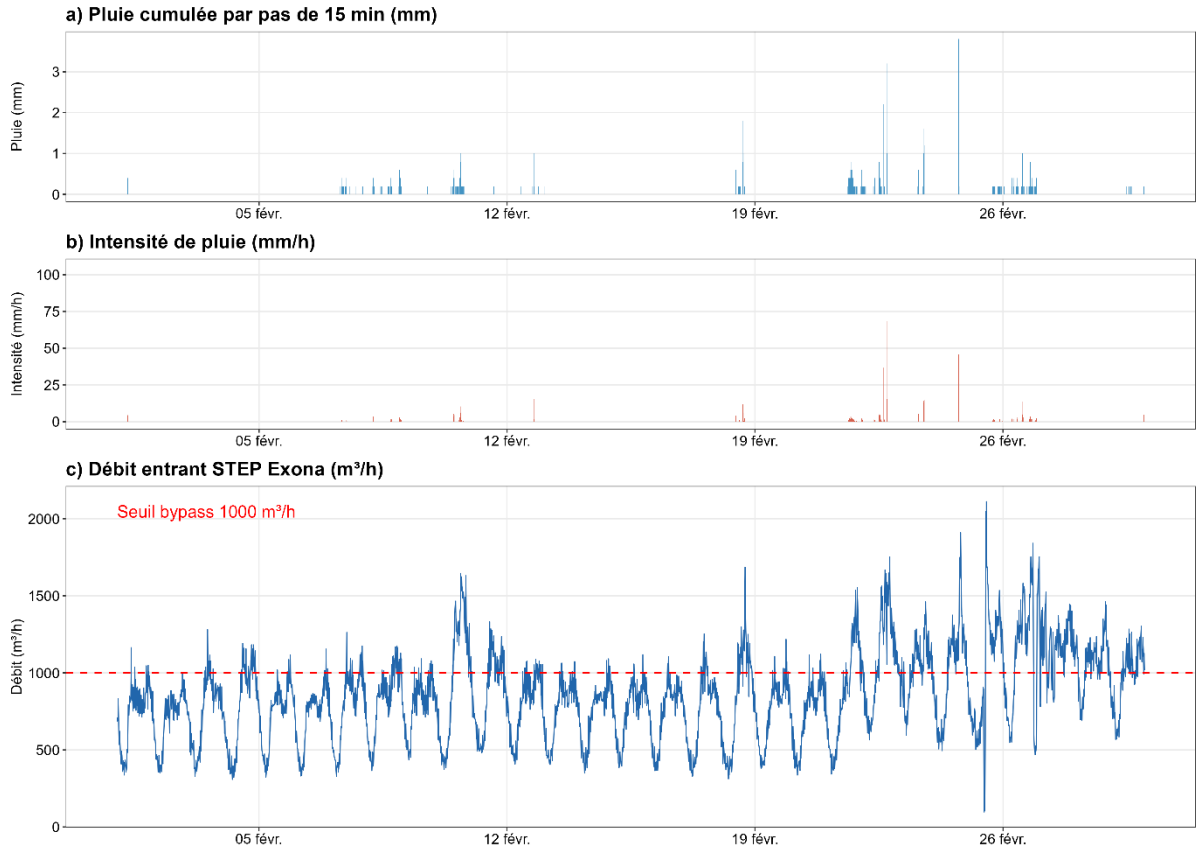


Figure 9:Graphique de l'évaluation temporelle de la relation pluie-débit entrant à la station d'Exona en février

IV.2.2 Période printanière : avril

Avril est représentatif du régime printanier. Les épisodes pluvieux sont nombreux et régulièrement répartis sur le mois, avec des intensités atteignant 30 mm/h autour des 8 et 15 avril. Le débit dépasse fréquemment les 1 000 m³/h avec des pointes à 1 500 m³/h, et bien qu'il redescende entre les épisodes, la succession rapide des pluies ne permet pas au réseau de se désaturer complètement. Ce comportement tient au redémarrage encore partiel de la végétation et à une évapotranspiration en hausse mais insuffisante pour compenser la fréquence des pluies. Chaque nouvel épisode survient alors que le sol est encore humidifié par le précédent, ce qui accroît la sensibilité du réseau et rapproche le débit du seuil plus vite qu'en été. C'est ce mécanisme qui produit le nombre élevé de dépassements au printemps, soit 824 sur l'année selon le **Tableau 5**, malgré des durées unitaires de dépassement relativement courtes.

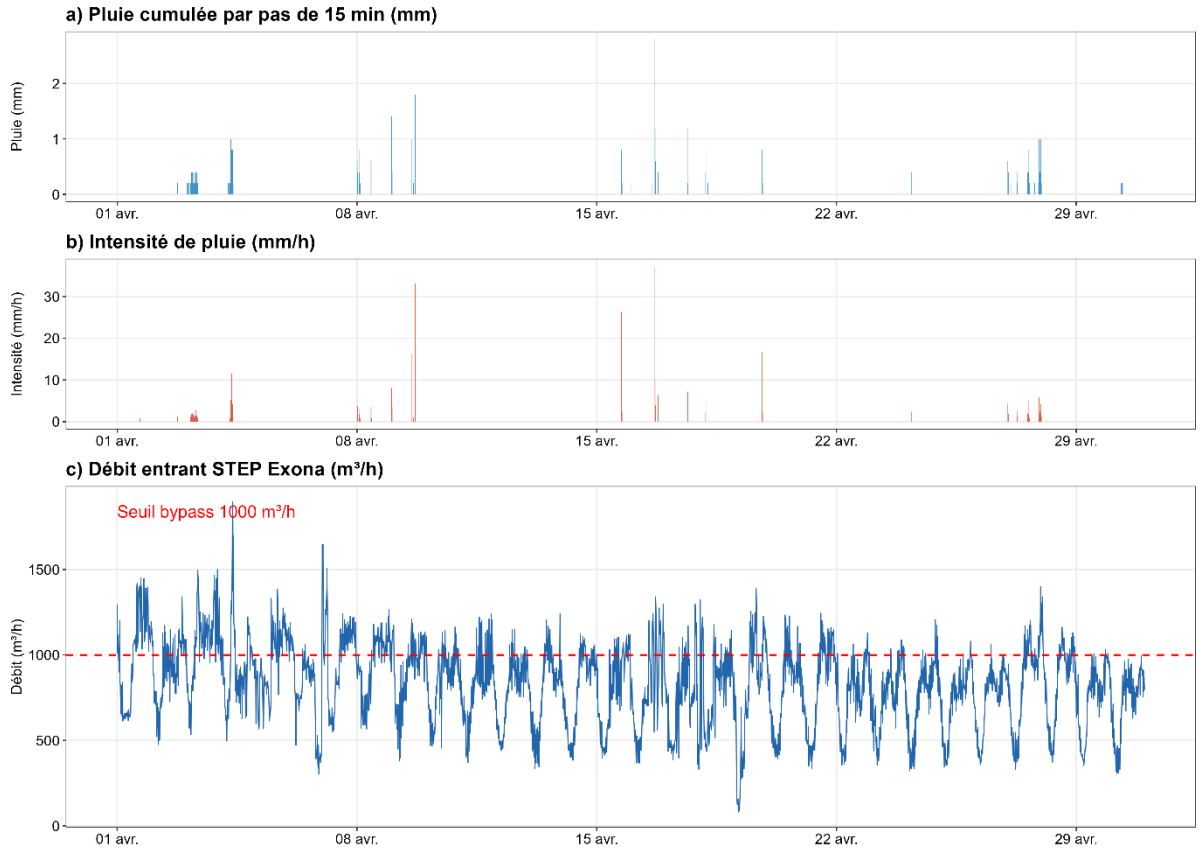


Figure 10:Graphique de l'évaluation temporelle de la relation pluie-débit entrant à la station Exona en avril

IV.2.3 Période estivale : juillet

Le mois de juillet se distingue nettement des mois précédents. Les épisodes pluvieux y sont rares, séparés par de longues périodes sans pluie. Cette rareté tient à une évapotranspiration estivale à son maximum, qui assèche les sols en profondeur et limite fortement les apports parasites en l'absence de pluie. Mais quand un épisode survient, comme autour du 13 juillet, l'intensité grimpe fort, avec une pointe proche de 50 mm/h, et le débit suit, montant rapidement jusqu'à 2 400 m³/h. Le retour sous le seuil ne tarde toutefois pas, il s'opère dans les heures qui suivent, car le sol sec, une fois l'intensité de la pluie retombée, réabsorbe rapidement l'humidité résiduelle et coupe court à l'apport parasite. Entre deux épisodes, le débit de base reste nettement sous les 1 000 m³/h, un contraste net avec le régime hivernal.

Ce profil explique d'ailleurs pourquoi l'été, malgré des pointes de débit élevées, finit par générer un volume bypassé annuel relativement modeste, 74 029 m³ selon le **Tableau 5**. La raison est simple, les épisodes sont trop brefs pour produire des dépassements qui durent.



Figure 11:Graphique de l'évaluation temporelle de la relation pluie-débit entrant à la station Exona en juillet

IV.2.4 Période automnale : octobre

Octobre est le mois le plus critique de l'année 2024 en termes de bypass. Entre le 7 et le 11 octobre, plusieurs épisodes pluvieux se succèdent sans laisser au réseau le temps de retrouver un fonctionnement normal. Le débit reste ainsi supérieur à 1 000 m³/h pendant plusieurs jours consécutifs, avec des pointes proches de 3 000 m³/h en fin de mois. Le 18 octobre, une intensité de 125 mm/h est enregistrée, correspondant à la valeur la plus élevée observée sur l'ensemble de l'année 2024. Même durant la seconde moitié du mois, alors que les pluies deviennent moins fréquentes, le débit peine à repasser durablement sous le seuil, témoignant d'une saturation persistante du réseau. Cette persistance vient de la transition saisonnière. Après l'été, les sols regagnent progressivement de l'humidité à mesure que l'évapotranspiration diminue, si bien que chaque nouvel épisode pluvieux d'octobre s'ajoute à un sol déjà moins capable d'absorber l'eau, contrairement à juillet où le sol repartait d'un état sec. Ce mois participe fortement au volume annuel bypassé et met en évidence le principal facteur de contrainte automnal pour l'exploitation de la station. Davantage que l'intensité des pluies, c'est leur répétition sur un réseau déjà chargé qui engendre les situations les plus critiques.

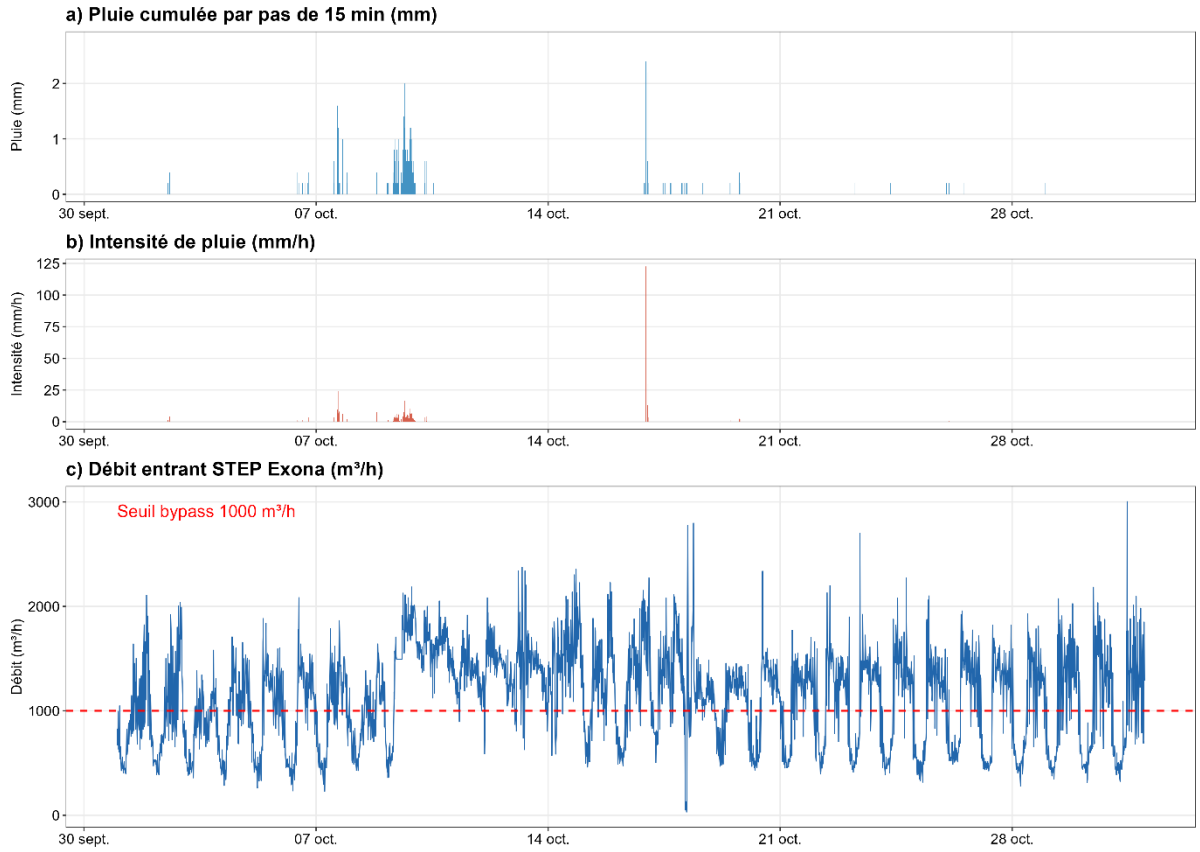


Figure 12:Graphique de l'évaluation temporelle de la relation pluie-débit entrant à la station Exona en octobre

Ces quatre mois mettent finalement en évidence un même phénomène, observé dans des contextes saisonniers différents. La réaction du réseau aux précipitations dépend moins de la pluie elle-même que de l'état d'humidité du bassin versant au moment où elle survient. Cet état résulte du bilan entre les précipitations accumulées et l'évapotranspiration des semaines précédentes. En été, ce bilan est généralement déficitaire : les sols restent secs, limitant les apports parasites hors épisode pluvieux et favorisant un retour rapide à la normale. En hiver, la situation est inverse : les sols saturés maintiennent un débit de fond élevé même en l'absence de pluie. Le printemps et l'automne constituent quant à eux des périodes de transition, où la fréquence des précipitations joue souvent un rôle plus déterminant que leur intensité.

Ainsi, un même débit observé à Exona ne traduit pas nécessairement la même situation selon la saison au cours de laquelle il est mesuré. En été, il signale presque toujours un épisode isolé et passager. En hiver ou en fin d'automne, en revanche, il peut traduire un état de saturation durable, bien plus difficile à résorber.

IV.3 Analyse saisonnière des dépassements du seuil

IV.3.1 Bilan des dépassements

Sur l'année 2024, 2 372 activations du bypass ont été recensés avec un débit dépassant le seuil de 1 000 m³/h, comme indiqué dans le **Tableau 5**. Ces dépassements représentent environ 593 heures de fonctionnement en régime de bypass et correspondent à un volume total bypassé de 707 126 m³.

Tableau 5: Synthèse saisonnière des épisodes de dépassement du seuil de bypass

Saison	Nombre d'activation du bypass	Durée médiane (h)	Débit max. médian (m³/h)	Volume bypassé total (m³)
Hiver	479	0,50	1 121	141 056
Printemps	824	0,25	1 119	129 923
Été	483	0,25	1 111	74 029
Automne	586	0,50	1 308	362 118

Les résultats montrent une répartition très inégale des volumes bypassés selon les saisons. A lui seul, l'automne totalise 362 118 m³, soit 51 % du volume annuel, et présente également le débit maximal médian le plus élevé, avec 1 308 m³/h. Le printemps enregistre le plus grand nombre de pas de temps en dépassement (824 occurrences), mais les épisodes restent courts avec une durée médiane de 0,25 heure. L'été présente les bypasses les moins sévères en volume (74 029 m³), ce qui s'explique par le caractère bref des orages convectifs qui génèrent des pointes de débit importantes sans provoquer de saturation durable du réseau.

IV.3.2 Distribution saisonnière des dépassements

Afin d'affiner ce bilan, la distribution saisonnière des dépassements est analysée à travers trois indicateurs complémentaires présenté à la **Figure 13** : la durée des épisodes de dépassement, le débit maximum atteint et le volume excédentaire associé. Ces trois indicateurs permettent de caractériser non seulement l'intensité des dépassements mais aussi leur persistance, deux dimensions essentielles pour la gestion opérationnelle du bypass.

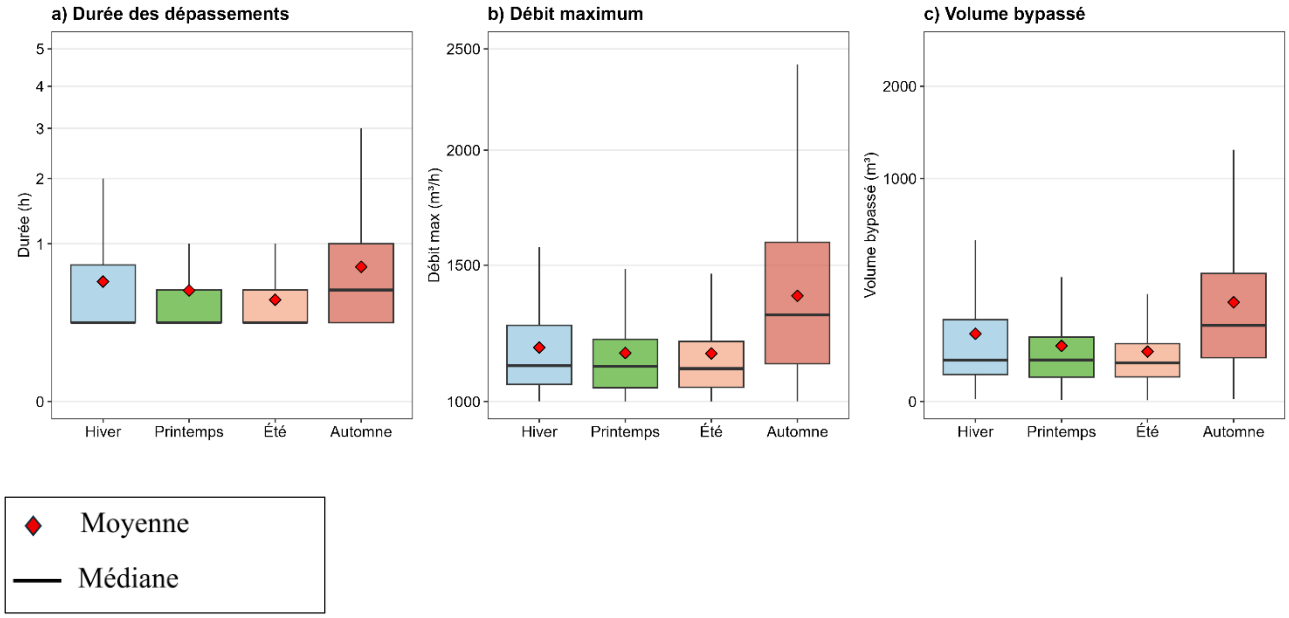


Figure 13:Diagramme en boite à moustaches des caractéristiques saisonnières des épisodes de dépassement du seuil de bypass ; a) Durée des dépassements ; b) Débit maximum ; c) Volume bypassé

IV.3.2.1 Durée des dépassements

Les dépassements sont généralement courts (**Figure 13a**), avec des durées médianes inférieures à une heure. L'automne se distingue toutefois par les durées de dépassement les plus importantes ainsi que par la plus forte variabilité. Certaines situations atteignent même près de 3 heures, traduisant des épisodes plus durables que dans les autres saisons.

L'hiver présente également des dépassements prolongés, mais avec une dispersion moins marquée que celle observée en automne. A l'inverse, au printemps et en été, les dépassements restent généralement de courte durée et sont le plus souvent associés à des pics ponctuels et isolés.

Cette distinction compte sur le plan opérationnel. Un dépassement bref se gère au cas par cas, sans grande difficulté. Un dépassement qui s'étire dans le temps, en revanche, oblige à mettre en place une coordination plus soutenue entre Exona, Evry et les réseaux de transport.

IV.3.2.2 Débit maximum pendant les dépassements

Le débit maximum (**Figure 13b**) atteint pendant les dépassements est nettement plus élevé en automne. La médiane automnale est supérieure à celles des autres saisons, et les valeurs maximales peuvent dépasser 2 400 m³/h. Cela confirme que l'automne reste la saison la plus critique pour Exona.

L'hiver affiche lui aussi des valeurs élevées, mais généralement plus modérées que celles observées en automne. Le printemps et l'été, en revanche, présentent des débits maximums plus faibles en médiane, même si certains épisodes isolés peuvent rester importants.

Cette observation montre bien que la saisonnalité joue un rôle réel dans le risque hydraulique. Les épisodes automnaux combinent probablement plusieurs facteurs, des pluies plus fréquentes, une humidité qui persiste plus longtemps, et des débits de base déjà élevés, ce qui réduit d'autant les marges disponibles à Exona.

IV.3.2.3 Volume excédentaire ou volume by-passé

Le volume excédentaire (**Figure 13c**) est également plus important en automne. La dispersion est forte, avec certains épisodes dépassant largement 1 000 m³. L'hiver arrive ensuite, avec des volumes également significatifs. Le printemps et l'été présentent des volumes généralement plus faibles, bien que certains épisodes puissent générer des volumes non négligeables.

Ces résultats sont cohérents avec les observations précédentes. L'automne cumule en quelque sorte deux facteurs aggravants, des dépassements plus longs et des débits maximums plus élevés, ce qui conduit logiquement à des volumes excédentaires plus importants.

Pour la gestion du by-pass, cette information compte énormément. Le risque ne se résume pas au seul débit maximum atteint, il dépend tout autant de la durée pendant laquelle ce dépassement se maintient. Un pic très élevé mais très bref peut au final générer un volume limité, alors qu'un débit seulement modérément supérieur au seuil, mais qui se prolonge sur plusieurs heures, peut produire un volume bien plus important à transférer ou à gérer.

IV.4 Identification et quantification des eaux parasites pluviales à la station Exona

La séparation entre débit de base et eaux parasites pluviales a aussi été analysée mois par mois à partir de graphiques à trois panneaux représentant, au pas de temps de 15 minutes, la pluie cumulée, le débit entrant Q_{Exona} et le débit de base Q_{base} par rapport au seuil de bypass de 1 000 m³/h, ainsi que le débit d'eaux parasites pluviales Q_{epp}. Les quatre mois représentatifs des régimes saisonniers identifiés sur l'année 2024 sont présentés ci-après par les **Figures 14 ; 15 ;16 et 17** ; les résultats des huit mois restants sont disponibles en **Annexe 6**.

IV.4.1 Période hivernale : février

Du 1er au 20 février, les courbes de débit de base Q_{base} et de débit entrant Q_{exona} se confondent quasiment, sans qu'on puisse y mesurer la moindre contribution des eaux parasites pluviales. Et pourtant, le débit dépasse régulièrement les 1 000 m³/h pendant cette période sèche. Ce comportement s'explique par la saturation des sols et des nappes superficielles en hiver. Les pluies tombées au cours des semaines précédentes, associées à une évapotranspiration très faible à cette période de l'année, maintiennent les sols et les drains dans un état d'humidité élevé. Le réseau reçoit alors un apport parasite quasiment permanent, qui n'est pas directement lié à une pluie récente. Il s'agit davantage d'un écoulement progressif provenant des nappes que d'un véritable phénomène de ruissellement.

A partir du 22 février, des pluies plus marquées (2 à 3 mm par épisode) viennent s'ajouter à ce fond déjà élevé. Le débit grimpe alors jusqu'à 2 100 m³/h, avec un léger décalage de quelques heures entre la pluie et la réaction du débit. Ce délai correspond simplement au temps que prennent les eaux pour transiter dans les collecteurs avant d'atteindre la station, un effet de propagation hydraulique tout à fait classique en réseau d'assainissement, qu'il faut bien distinguer du temps de réaction du sol lui-même.

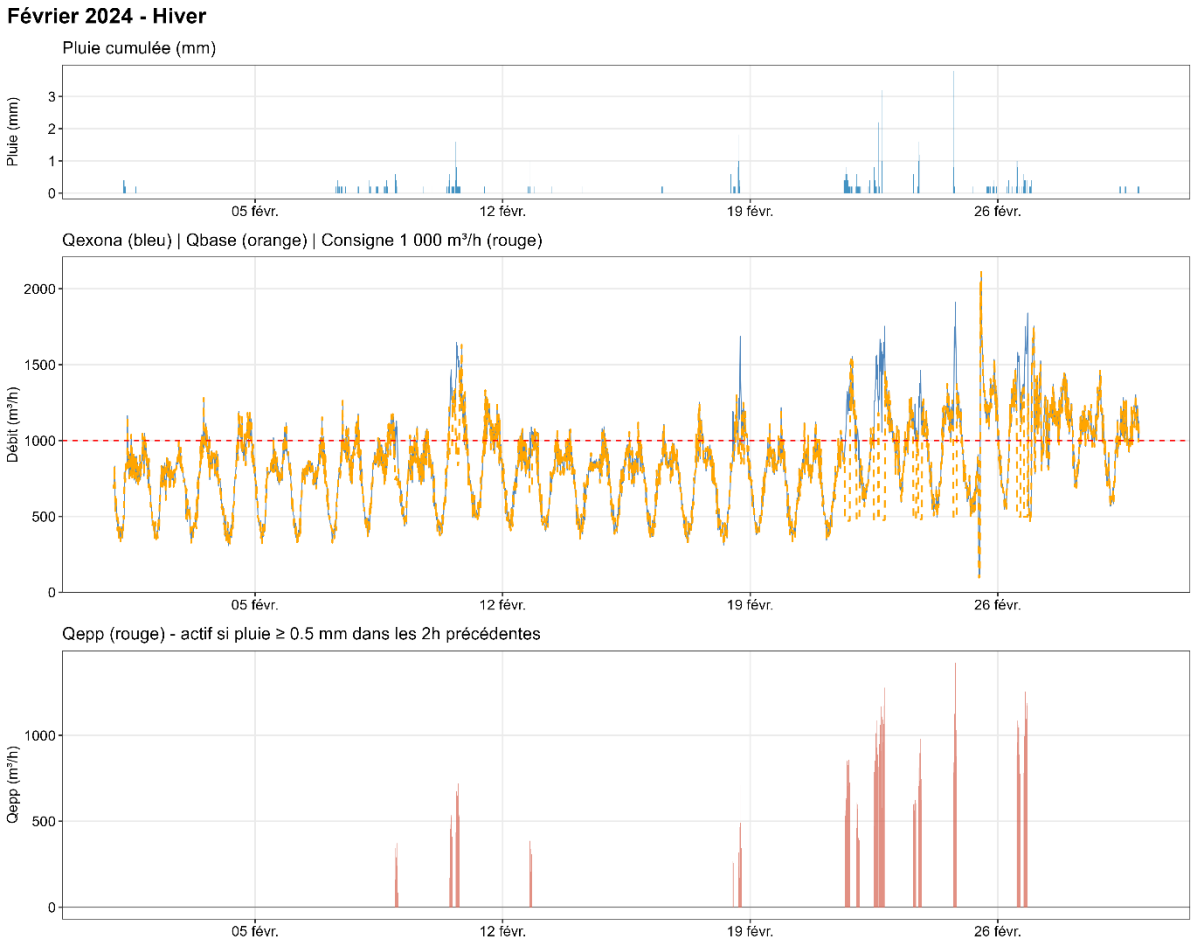


Figure 14:Graphique de l'évaluation temporelle de la séparation du débit de base (Qbase) et du débit des eaux parasites pluviales (Qepp) à la station d'Exona en février

IV.4.2 Période printanière : avril

En avril, le débit de base Q_{base} oscille entre 400 et 900 m³/h, nettement en dessous du niveau observé en février. Cette baisse s'explique par le redémarrage de la végétation et la hausse de l'évapotranspiration propre au printemps, qui assèchent peu à peu les sols superficiels et réduisent du même coup l'apport parasite de fond.

Mais cette marge plus confortable se trouve compensée par un autre phénomène, les épisodes pluvieux se succèdent à un rythme rapproché, sans vraiment laisser au sol le temps de se ressuyer entre deux pluies. Chaque épisode produit un pic d'environ 750 m³/h, ce qui suffit largement à dépasser le seuil dès qu'il survient pendant la phase haute du cycle journalier de débit. Le risque ne vient donc pas tant de l'intensité des pluies, plutôt modérée ce mois-ci, que

de leur fréquence. Le sol, déjà humidifié par la pluie précédente, réagit alors plus vite et plus fort à la pluie suivante. C'est ce qu'on appelle un effet d'état antécédent du bassin versant, un phénomène bien connu en hydrologie urbaine.



Figure 15:Graphique de l'évaluation temporelle de la séparation du débit de base (Qbase) et du débit des eaux parasites pluviales (Qepp) à la station d'Exona en avril

IV.4.3 Période estivale : juillet

Le débit de base de juillet (200 à 700 m³/h) est le plus faible de toute l'année. L'évapotranspiration estivale, à son maximum à cette période, dessèche les sols en profondeur et réduit fortement les infiltrations résiduelles. Du coup, en l'absence de pluie, le réseau ne reçoit presque aucun débit parasite.

Quand un orage survient, comme celui du 13 juillet, le sol sec absorbe d'abord une partie de l'eau avant que le ruissellement n'atteigne le réseau. Mais l'intensité de la pluie, proche de 50 mm/h, suffit malgré tout à générer un pic de Q_{epp} de 1 500 m³/h, portant le débit jusqu'à 2 200 m³/h en quelques heures seulement. Le retour sous le seuil est tout aussi rapide, ce qui s'explique par la capacité du sol sec à réabsorber l'humidité résiduelle une fois la pluie passée.

C'est justement ce contraste, entre un sol sec en profondeur et une intensité pluvieuse parfois très élevée, qui explique le paradoxe apparent de l'été. On y observe certains des pics de débit les plus forts de l'année, mais un volume bypassé annuel finalement limité (74 029 m³ selon le **Tableau 5**), faute de durée suffisante pour accumuler du volume.



Figure 16:Graphique de l'évaluation temporelle de la séparation du débit de base (Qbase) et du débit des eaux parasites pluviales (Qepp) à la station d'Exona en juillet

IV.4.4 Période automnale : octobre

Octobre opère une transition inverse de celle observée au printemps. Après l'été, les sols regagnent peu à peu de l'humidité à mesure que l'évapotranspiration diminue, ce qui explique que le débit de base Q_{base} soit déjà élevé en tout début de mois, avant même l'arrivée de pluies significatives.

Entre le 7 et le 11 octobre, plusieurs épisodes pluvieux rapprochés viennent s'ajouter à ce débit de fond déjà élevé. Le débit atteint alors 1 200 m³/h et l'effet combiné de ces deux contributions le maintient au-dessus de 1 000 m³/h pendant plusieurs jours, avec des pointes pouvant atteindre 2 500 m³/h. Le phénomène le plus marquant apparaît toutefois en fin de mois. Alors que les précipitations deviennent moins fréquentes, le débit reste durablement élevé. Cette situation montre que les sols, fortement saturés par les épisodes précédents, continuent d'alimenter le réseau longtemps après la fin des pluies. Le bassin versant conserve ainsi une forme de mémoire hydrologique qui prolonge les effets des pluies dans le temps. C'est justement cette incapacité du système à retrouver son état initial entre deux épisodes qui explique qu'octobre concentre, à lui seul, 51 % du volume bypassé sur l'année.



Figure 17:Graphique de l'évaluation temporelle de la séparation du débit de base (Qbase) et du débit des eaux parasites pluviales (Qepp) à la station d'Exona en octobre

En croisant les trois panneaux sur ces quatre mois, on voit bien que la part de la pluie dans le débit entrant à Exona est tout à fait réelle. Elle se lit directement dans l'écart entre les courbes bleue et orange, et se trouve quantifiée par le panneau inférieur. Mais son effet sur le risque de by-pass change beaucoup d'une saison à l'autre.

En été, les dépassements sont principalement provoqués par les épisodes de pluie, mais le réseau retrouve rapidement un fonctionnement normal et les volumes bypassés demeurent relativement limités. En hiver, la situation est différente, le débit de fond devient prépondérant et le réseau reste fortement sollicité, même en l'absence de précipitations. À l'automne, ces deux phénomènes se cumulent sur un réseau qui ne dispose plus du temps nécessaire pour se désaturer entre deux épisodes pluvieux, ce qui engendre les dépassements les plus longs et les

volumes bypassés les plus importants de l'année. Au printemps, enfin, c'est surtout la répétition des épisodes de pluie qui maintient le réseau dans un état de sollicitation quasi permanent.

Ces quatre profils saisonniers confirment, au final, que le by-pass à Exona résulte d'une interaction entre le niveau du débit de base et la contribution des eaux parasites pluviales, une interaction qui évolue au fil des saisons tout au long de l'année 2024.

IV.5 Modélisation prédictive du risque de bypass et définition de règles de gestion

IV.5.1 Performance du modèle

La matrice de confusion présentée à la **Figure 18** permet d'évaluer la qualité des prédictions du modèle en comparant les classes réelles aux classes prédites. Sur les 1 639 observations analysées, le modèle prédit correctement 855 situations de Bypass et 592 situations de Non, soit un total de 1 447 classifications correctes. Les erreurs se répartissent entre 109 cas de Bypass identifiés à tort comme Non (faux négatifs) et 83 cas de Non classés à tort comme Bypass (faux positifs). Au final, le modèle atteint une exactitude globale de 88,3 %. En apprentissage automatique appliqué, une exactitude comprise entre 70 % et 90 % est généralement considérée comme une bonne performance pour un modèle de classification (Hosp et al., 2021), un seuil de 80 % étant couramment retenu comme repère au-delà duquel un modèle est jugé fiable. Avec 88,3 %, le modèle se situe donc nettement au-dessus de ce seuil, ce qui montre une bonne capacité à reproduire les observations réelles.

L'analyse de la matrice de confusion montre également que les performances restent assez homogènes entre les deux classes. Le modèle identifie correctement 88,7 % des situations de Bypass et 87,7 % des situations de Non. Les taux d'erreur observés sont donc très proches d'une classe à l'autre, avec 11,3 % de faux négatifs et 12,3 % de faux positifs. Cette répartition indique que le modèle ne favorise pas particulièrement une catégorie au détriment de l'autre et qu'il conserve un comportement stable malgré un jeu de données légèrement déséquilibré, composé de 964 cas de Bypass et de 675 cas de Non.

Cette tendance se retrouve également dans les valeurs du F1-score, qui permet d'évaluer simultanément la précision et le rappel. Pour la classe Bypass, la précision atteint 91,1 % et le rappel 88,7 %, ce qui conduit à un F1-score de 89,9 %. Pour la classe Non, la précision est de 84,5 % et le rappel de 87,7 %, soit un F1-score de 86,0 %. La moyenne macro obtenue est de 88,0 %, ce qui confirme le bon équilibre général des performances du modèle. En apprentissage automatique, un F1-score supérieur à 0,7 est généralement considéré comme correct, supérieur à 0,8 comme fort, et supérieur à 0,9 comme excellent (Analytics Vidhya, 2025). Les F1-scores obtenus ici, tous deux supérieurs à 0,85, situent donc le modèle dans la tranche haute de cette échelle, entre performance forte et excellente.

Ces deux scores, proches l'un de l'autre et tous deux supérieurs à 85 %, viennent ainsi corroborer l'observation faite à partir de la matrice de confusion : le modèle maintient un compromis satisfaisant entre précision et rappel pour les deux classes, ce qui en fait un outil fiable aussi bien pour identifier les cas de Bypass que pour écarter correctement les cas Non.

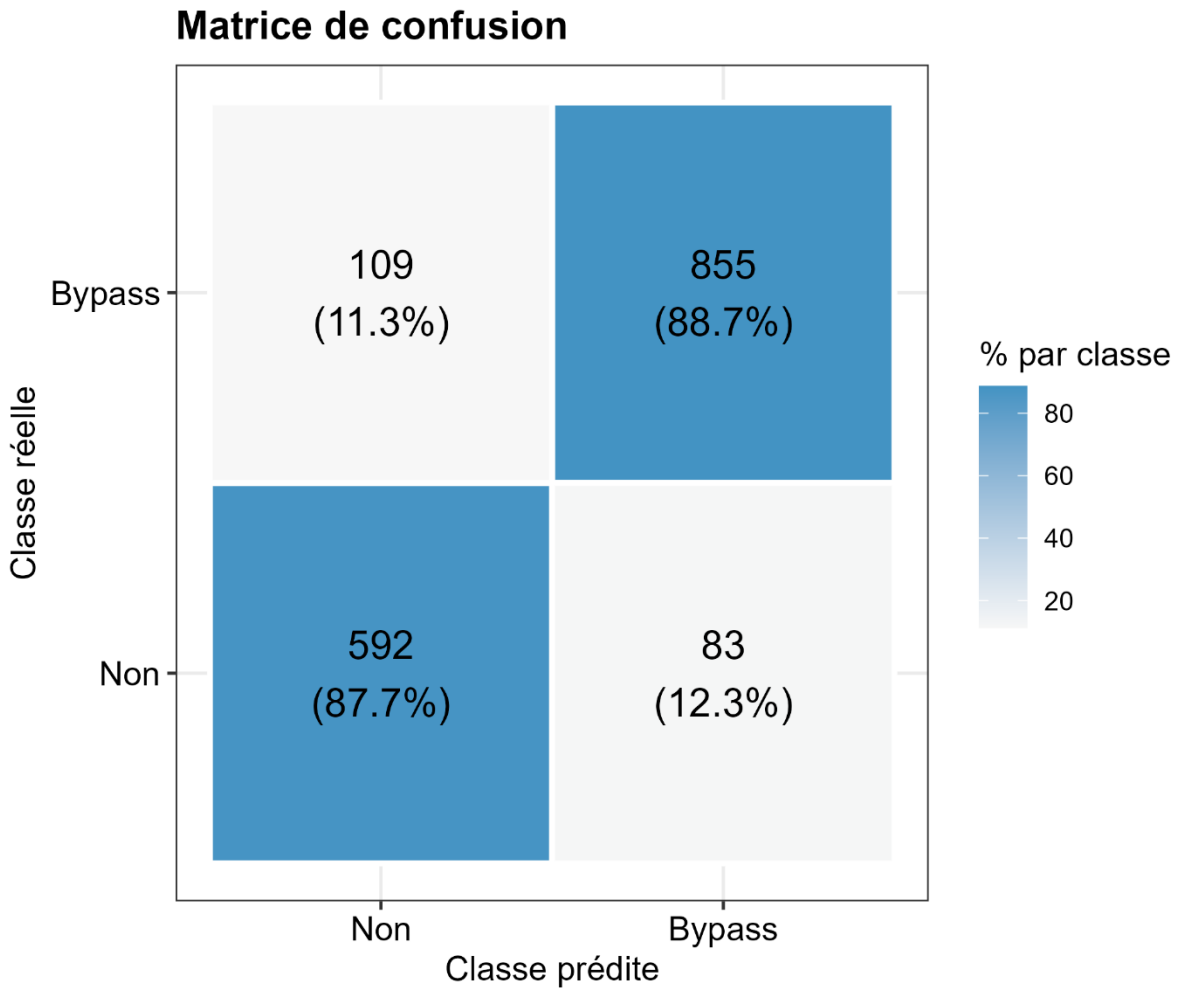


Figure 18:Matrice de confusion de l'arbre de décision pour la prédiction du by-pass à la station d'Exona

IV.5.2 Courbe précision-rappel de l'arbre de décision pour la prédiction du bypass

La courbe précision-rappel (**Figure 19**) illustre l'évolution du compromis entre ces deux métriques en fonction du seuil de décision appliqué par le modèle. Au seuil de 0,50 (cercle orange), le modèle présente un bon équilibre entre précision et rappel. Il atteint un rappel d'environ 88,7 % pour une précision de 91,2 %, des valeurs très proches de celles obtenues à partir de la matrice de confusion.

La courbe montre que la précision reste supérieure à 90 % tant que le rappel ne dépasse pas environ 87,5 %. Autrement dit, le modèle parvient à identifier une grande partie des cas positifs tout en conservant des prédictions fiables. En revanche, lorsque l'on cherche à augmenter davantage le rappel, la précision diminue plus rapidement. Pour des rappels proches de 95 %, elle passe sous les 86 %, ce qui traduit une augmentation du nombre de faux positifs. Les gains obtenus en détection deviennent alors moins intéressants au regard des erreurs supplémentaires générées.

Dans ce contexte, le seuil de 0,50 constitue un choix pertinent puisqu'il permet de conserver un bon niveau de détection sans dégrader excessivement la précision du modèle. Il pourrait néanmoins être ajusté à la baisse si l'enjeu prioritaire était de maximiser la détection des cas positifs (minimiser les faux négatifs), au prix d'une précision légèrement réduite, un choix qui devrait être guidé par les implications pratiques ou cliniques associées à chaque type d'erreur.

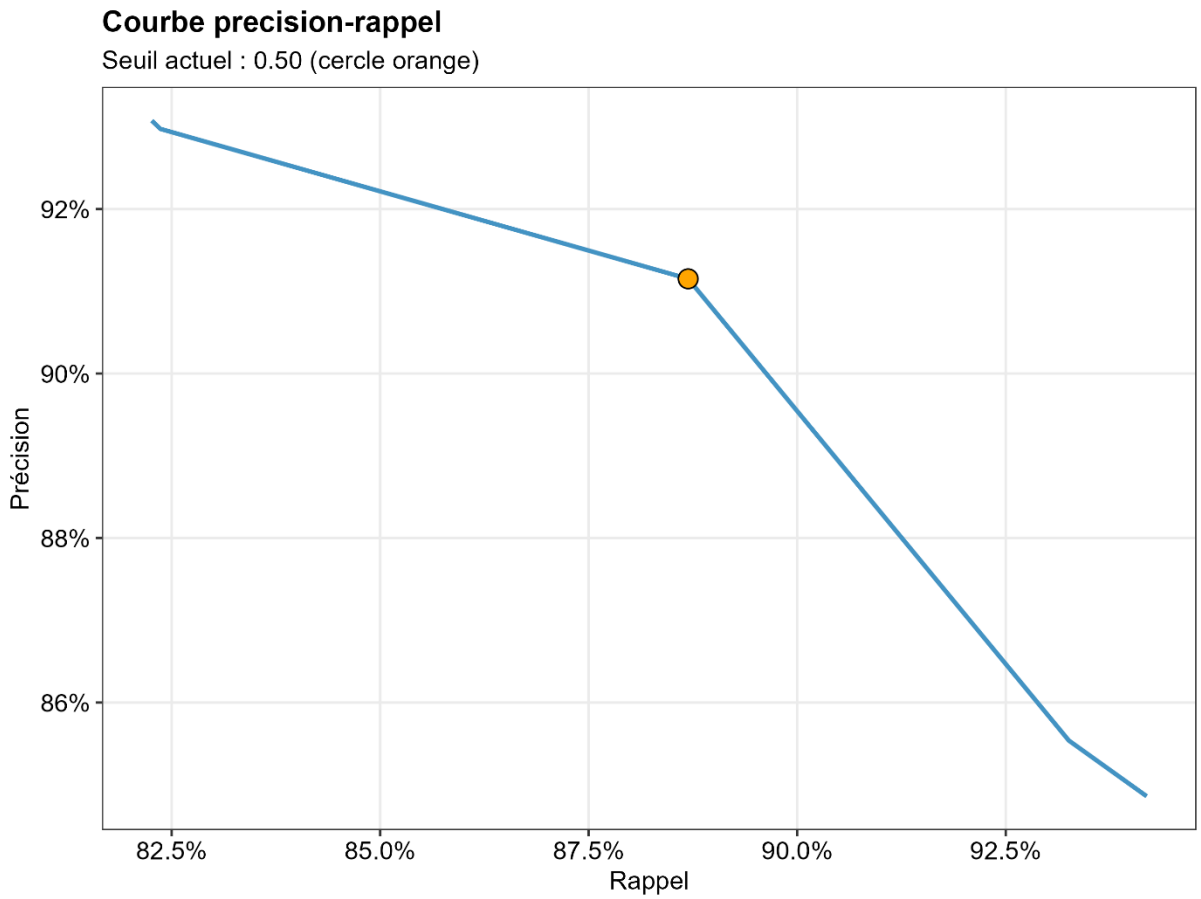


Figure 19:Courbe précision-rappel de l'arbre de décision pour la prédiction du by-pass à la station d'Exona

IV.5.3 Importance des variables

Les résultats de la **Figure 20** montrent que le débit moyen actuel Qexona_mean représente 50,3 % de l'importance totale du modèle. Le débit de l'heure précédente Qexona_lag1 représente 28,2 %.

Ces deux variables représentent donc près de 80 % de l'information utilisée par le modèle. Le risque de by-pass dépend ainsi d'abord de l'état hydraulique immédiat de la station et de son évolution récente.

Les variables pluviométriques ont un rôle secondaire mais utile. Le cumul de pluie sur 6 heures représente 6,5 %, tandis que le cumul sur 2 heures représente 4,7 %. Le mois contribue à hauteur de 5,6 %, ce qui confirme l'existence d'une saisonnalité. La variation du débit représente 3,5 %, et l'heure seulement 1,2 %.

Ces résultats montrent que la pluie influence bien le risque, mais de manière indirecte. La pluie est importante parce qu'elle modifie le débit. Mais pour l'exploitation, le signal le plus robuste reste le débit effectivement mesuré à Exona.

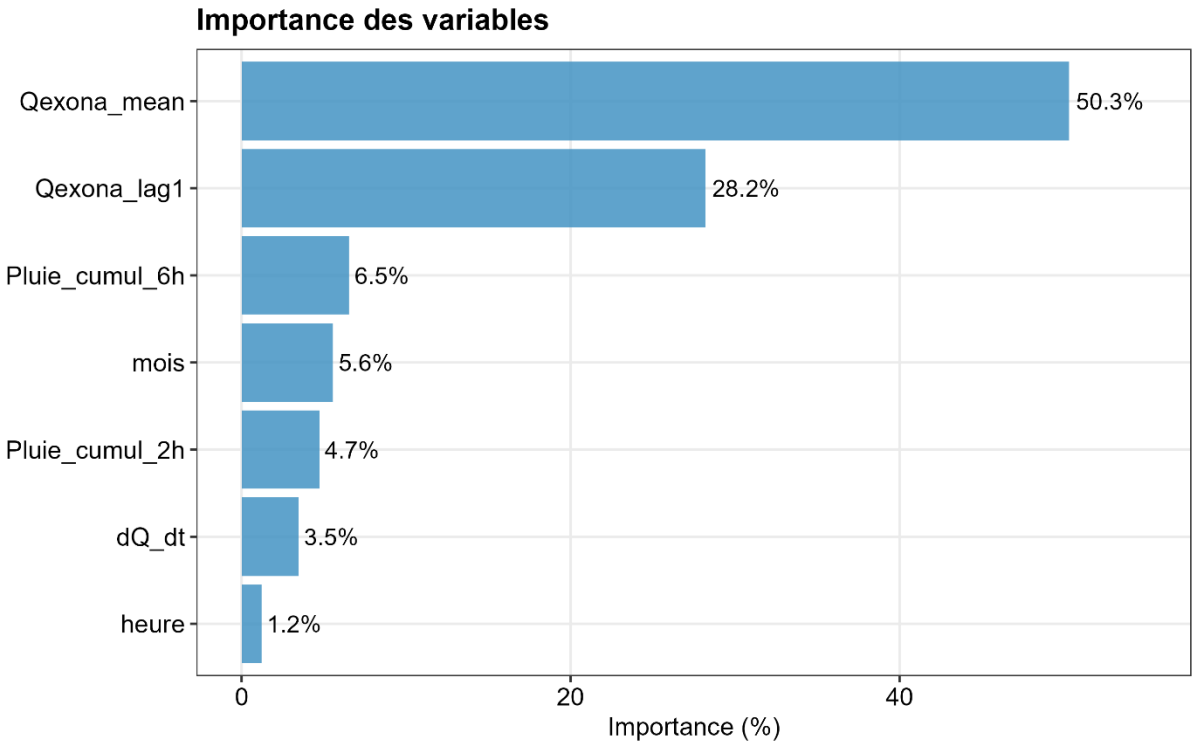


Figure 20:Importance relative des variables explicatives dans l'arbre de décision pour la prédiction du by-pass à la station d'Exona

IV.5.4 Règles de décision pour la gestion du by-pass

Afin d'identifier les conditions déclenchant un événement de bypass, un arbre de décision a été entraîné sur les données opérationnelles de la station d'Exona. Ce modèle repose sur trois variables explicatives : le débit moyen à l'exutoire (Qexona_mean), la pluie cumulée sur deux heures (Pluie_cumul_2h) et l'heure de la journée. La **Figure 21** présente la structure de l'arbre obtenu.

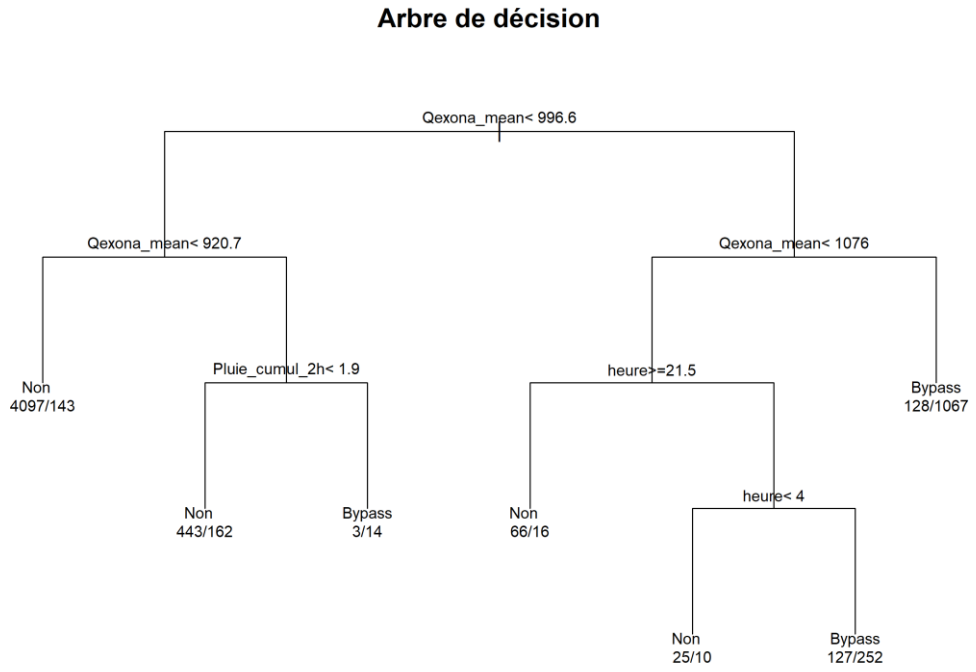


Figure 21:Arbre de décision pour la prédiction du risque de by-pass à la station d'Exona

L'arbre génère 6 règles de décision, dont l'interprétation est détaillée ci-après.

- **Règle 1 : Non-bypass certain**

Lorsque le débit Qexona_mean passe sous 921 m³/h, le modèle prédit systématiquement l'absence de bypass. Cette règle s'appuie sur 4 097 observations et ne compte que 143 erreurs de classification. Dans cette zone, le risque de bypass devient donc quasiment négligeable.

- **Règle 2 : Non-bypass en débit intermédiaire sans pluie**

Entre 921 et 997 m³/h, et en l'absence de précipitations significatives (pluie cumulée sur 2 heures inférieure à 1,9 mm), le modèle prédit également un non-bypass. Cette règle repose sur 443 observations, dont 162 erreurs, ce qui en fait une règle moins robuste que la précédente, même si le risque de bypass reste globalement limité.

- **Règle 3 : Bypass en débit intermédiaire avec pluie.**

Lorsque le débit se situe entre 921 et 997 m³/h et que la pluie cumulée sur les deux dernières heures dépasse 1,9 mm, le modèle prédit cette fois un bypass. Cette règle ne repose toutefois que sur 3 observations correctes pour 14 erreurs, ce qui invite à l'interpréter avec prudence, vu le faible effectif sur lequel elle s'appuie.

- **Règle 4 : Non-bypass en soirée pour débit élevé**

Pour des débits inférieurs à 997 m³/h, avant 22h00, le modèle prédit généralement l'absence de bypass à l'heure suivante. Cette règle est basée sur 66 observations et présente 16 erreurs.

- **Règle 5 : Non-bypass en nuit profonde pour débit élevé**

Pour des débits inférieurs à 997 m³/h, mais entre 22h00 et 4h00, le modèle prédit également l'absence de bypass. Cette règle repose sur 25 observations, avec 10 erreurs, ce qui traduit une fiabilité correcte, bien que limitée par le faible volume de données nocturnes.

- **Règle 6 : Bypass en journée pour débit élevé**

Entre 4h00 et 22h00, pour un débit inférieur à 997, le modèle prédit cette fois un bypass. Cette règle est validée par 252 classifications correctes contre 127 erreurs, ce qui montre que le bypass devient probable dans cette configuration, sans être pour autant certain.

Le **Tableau 6** synthétise l'ensemble de ces règles en y associant la probabilité de bypass observée P(bypass) et le nombre d'observations correspondant, permettant d'apprécier la fiabilité relative de chaque règle.

Tableau 6: Synthèse des règles opérationnelles de gestion du by-pass

Plage de débit	**Condition additionnelle**	**Décision**	**P(bypass)**	**nbre obs.**
< 921 m³/h	-	Pas de risque	3.40%	4240
921 – 997 m³/h	Pluie 2h < 1.9 mm	Surveillance légère	26.80%	605
921 – 997 m³/h	Pluie 2h ≥ 1.9 mm	**Alerte bypass**	82.40%	17
< 997 m³/h	heure < 22h	Pas de bypass immédiat	19.50%	82
< 997 m³/h	22h ≤ heure < 4h	Pas de bypass immédiat	28.60%	35
< 997 m³/h	4h ≤ heure < 22h	**Alerte bypass**	66.50%	379

V. LIMITES ET PERSPECTIVES

V.1 Coordination avec le réseau

Au-delà de l'analyse du débit à Exona, ces résultats amènent à s'interroger plus largement sur la coordination entre les réseaux de collecte et les deux stations, un enjeu qui figure d'ailleurs dès le titre de ce mémoire. Il convient toutefois de préciser ce que cette étude permet réellement d'affirmer sur ce point. Le fonctionnement du siphon sous la Seine et l'état des ouvrages de transport côté Evry n'ont pas fait l'objet de mesures ou de modélisation spécifiques dans ce travail, qui s'appuie uniquement sur les données de pluie et de débit disponibles à Exona.

Le contexte présenté en début de mémoire (cf. **II.1**) montre pourtant que le by-pass d'Exona et le siphon convergent vers le même point du réseau, en tête d'Evry. Le modèle proposé ici, fondé sur le seul débit d'Exona, ne tient donc pas compte de la situation hydraulique du reste du système au moment où la décision de by-pass est prise. Cette limite, déjà évoquée dans la partie « Limites de l'étude », a en réalité une portée bien plus large sur le plan opérationnel, une décision prise localement à Exona peut très bien reporter une contrainte vers Evry, sans que ce report soit anticipé au moment où il se produit.

Cela ne remet pas en cause la pertinence de l'arbre de décision proposé, qui répond bien à l'objectif initial, fournir un outil simple à partir des données disponibles à Exona. Cela ouvre plutôt une piste d'amélioration, faire évoluer le modèle pour qu'il intègre, à terme, des données sur l'état du siphon et la capacité disponible à Evry. Ce qui rejoint d'ailleurs la recommandation formulée plus loin sur la nécessité de formaliser un protocole de coordination entre les différents acteurs.

V.2 Limites de l'étude

Même si les résultats restent riches, plusieurs limites doivent être mentionnées.

La première concerne les données disponibles. Les analyses s'appuient sur les mesures de pluie et de débit issues des systèmes de supervision. Comme toute donnée d'exploitation, elles peuvent comporter des valeurs manquantes, des erreurs de capteur, ou encore des périodes perturbées.

La deuxième limite touche à l'estimation des eaux parasites pluviales. La séparation entre Q_{base} et Q_{epp} repose sur une méthode statistique qui permet une lecture utile, certes, mais qui ne remplace en rien une modélisation hydraulique complète du réseau.

La troisième limite concerne l'arbre de décision lui-même. Le modèle a volontairement été gardé simple, afin de rester interprétable. Il ne tient donc pas compte de toutes les variables qui pourraient pourtant améliorer la prédiction, comme le niveau dans les bâches, l'état des pompes, la capacité instantanée disponible à Evry, ou encore les conditions du siphon sous la Seine.

Enfin, les règles proposées doivent encore être validées avec les exploitants. Elles constituent surtout une base de discussion, pas une procédure définitive. Leur mise en œuvre demanderait une phase de test, de retour d'expérience, et d'ajustements progressifs.

CONCLUSION

Ce mémoire a porté sur l'optimisation de la gestion mutualisée des stations d'épuration d'Evry et d'Exona, dans un contexte de contraintes hydrauliques croissantes lors des épisodes pluvieux, à partir des données d'exploitation de 2024, année marquée par une forte fréquence de pluies et une sollicitation importante des ouvrages. La station d'Exona, limitée par une file d'eau unique et une capacité nominale de 1 000 m³/h, peut transférer une partie de ses effluents bruts vers Evry via un by-pass. Ce dispositif permet d'assurer la continuité du service, mais son usage doit rester maîtrisé, car il reporte une partie de la contrainte hydraulique sur Evry et ses réseaux.

L'analyse des épisodes pluvieux de 2024 met en évidence une grande diversité de situations. Les épisodes d'automne et d'hiver sont généralement plus longs et plus cumulés, tandis que ceux du printemps et de l'été sont plus courts mais souvent plus intenses, ce qui montre que le risque hydraulique dépend autant de la durée et de l'intensité des pluies que de l'état initial du réseau au moment de l'événement. L'étude des débits entrants à Exona confirme par ailleurs de nombreux dépassements du seuil de 1 000 m³/h, particulièrement marqués en automne. La séparation entre débit de base Qbase et eaux parasites pluviales Qepp montre que si les précipitations expliquent une grande partie des pics, certains dépassements sont liés soit à un débit de base déjà élevé, soit à des effets retardés dans le réseau. Dans ce contexte, la gestion du by-pass doit s'appuyer avant tout sur le débit réellement observé à Exona, plutôt que sur la seule intensité des pluies.

L'arbre de décision développé confirme que le débit moyen actuel constitue la variable la plus déterminante pour anticiper le risque de by-pass à une heure, suivi du débit de l'heure précédente, les variables pluviométriques restant complémentaires ; les seuils identifiés permettent d'esquisser une première grille de vigilance progressive, avec une alerte dès 920 m³/h. Ces résultats répondent aux trois objectifs fixés : caractériser les épisodes pluvieux de 2024, analyser le fonctionnement hydraulique d'Exona, et proposer des règles opérationnelles simples d'aide à la décision. La gestion mutualisée Evry-Exona doit ainsi être envisagée comme le pilotage dynamique d'un système global, dans lequel chaque décision prise à Exona a des répercussions directes sur le fonctionnement d'Evry, ce qui plaide pour une approche plus anticipative du by-pass et une coordination renforcée entre les différents acteurs du système.

RECOMMANDATIONS

A partir des résultats obtenus, plusieurs recommandations peuvent être formulées afin d'améliorer la gestion mutualisée des stations d'Evry-Courcouronnes et d'Exona. Chacune est adressée à l'acteur le plus pertinent pour sa mise en œuvre.

- **A l'exploitant : mettre en place une grille de vigilance progressive sur le débit d'Exona**

Les résultats montrent que le risque de by-pass augmente avant même que le seuil de 1 000 m³/h ne soit dépassé. Il est donc recommandé à l'exploitant de déclencher une vigilance opérationnelle dès l'approche de ce seuil, en s'appuyant sur la grille de seuils progressifs proposée (voir **Tableau 6**). Cette grille offrirait à l'exploitant une lecture rapide de la situation et lui permettrait d'adapter son niveau de vigilance en fonction de l'évolution du débit, quelle que soit la station depuis laquelle la surveillance est assurée.

- **A l'exploitant : croiser systématiquement le débit d'Exona avec la pluie récente**

La pluie récente ne suffit pas à expliquer tous les dépassements, mais elle reste un signal important. L'exploitant devrait suivre en continu le cumul de pluie sur 2 et 6 heures, l'intensité maximale récente et la tendance du débit entrant à Exona. Lorsque le débit se situe entre 920 et 1 000 m³/h avec un cumul de pluie significatif, la situation devrait être considérée comme une pré-alerte, déclenchant une vérification croisée de l'état des deux stations.

- **A l'exploitant, en lien avec la SPL : intégrer la capacité disponible à Evry dans la décision de by-pass**

Le by-pass permet de soulager Exona, mais peut déplacer la contrainte vers Evry. Avant toute décision de maintien ou d'augmentation du by-pass, il est nécessaire de vérifier le débit entrant à Evry, le niveau des ouvrages de tête, la sollicitation du siphon sous la Seine, la disponibilité des équipements et les risques de refoulement ou de déversement. Cette vérification devrait être formalisée dans une procédure validée par la SPL afin d'éviter qu'une décision favorable à Exona ne crée des difficultés en aval.

- **Au service patrimoine/réseau de la collectivité : poursuivre le diagnostic des eaux parasites pluviales**

L'analyse du débit de base Q_{base} montre que les eaux parasites pluviales Q_{epp} contribuent fortement à certains pics. Ce diagnostic relève du service réseau de la collectivité, qui devrait identifier les bassins versants les plus contributeurs, rechercher les branchements non

conformes, inspecter les ouvrages sensibles et prioriser les travaux de réhabilitation sur les secteurs les plus impactants.

- **A la SPL : faire développer un tableau de bord opérationnel**

Un tableau de bord simple pourrait être mis en place à destination de l'exploitant, soit via un prestataire externe, soit intégré aux outils de supervision existants. Il regrouperait les principaux indicateurs utiles à la décision : débit à Exona, cumul de pluie, niveau de vigilance, état du by-pass, débit à Evry et alertes de dépassement. L'objectif serait de disposer d'un outil lisible et directement exploitable pour améliorer la réactivité opérationnelle.

- **A la SPL : tester les règles proposées sur une période pilote**

Les seuils proposés dans ce mémoire restent une première base de travail. Avant toute intégration dans une procédure d'exploitation, la SPL devrait organiser, avec l'exploitant, une phase pilote permettant de vérifier la pertinence des seuils, le taux de fausses alertes, la capacité d'anticipation et l'acceptabilité de l'outil sur le terrain.

- **A la SPL et à la collectivité, à moyen terme : faire évoluer l'outil vers une aide à la décision plus complète**

Les résultats de ce mémoire pourraient servir de socle à un outil plus avancé, à porter conjointement par la SPL et la collectivité dans le cadre de leurs investissements futurs : intégration de prévisions météorologiques, de données temps réel sur les niveaux d'eau, des états des pompes et vannes, et d'une alerte automatique en cas de situation critique dans une logique de jumeau numérique ou de système d'aide à la décision en temps réel.

BIBLIOGRAPHIE

A. Bizer, M. et Kirchhoff, C.J. (2022) « Regression modeling of combined sewer overflows to assess system performance », *Water Science and Technology*, 86(11), p. 2848-2860. Disponible sur: https://doi.org/10.2166/wst.2022.362.

Agence de l'eau Seine-Normandie (2022) *SDAGE 2022-2027*, *Agence de l'Eau Seine-Normandie*. Disponible sur: https://www.eau-seine-normandie.fr/domaines-d-action/sdage (Consulté le: 28 juillet 2026).

Analytics Vidhya (2025) « What is F1 Score in Machine Learning? », *Analytics Vidhya*. Disponible sur: https://www.analyticsvidhya.com/blog/2025/12/what-is-f1-score-in-machine-learning/ (Consulté le: 28 juillet 2026).

Balistrocchi, M. et Bacchi, B. (2011) « Modelling the statistical dependence of rainfall event variables through copula functions », *Hydrology and Earth System Sciences*, 15(6), p. 1959-1977. Disponible sur: https://doi.org/10.5194/hess-15-1959-2011.

Breiman et al. (1984) *Arbres de classification et de régression*. Disponible sur: https://www.academia.edu/5867603/Classification_and_Regression_Trees (Consulté le: 20 juin 2026).

Chocat, B. (1997) *Encyclopédie de l'hydrologie urbaine et de l'assainissement*. Paris: Lavoisier. Disponible sur: https://www.oieau.fr/eaudoc/notice/ENCYCLOPEDIE-DE-LHYDROLOGIE-URBAINE-ET-DE-LASSAINISSEMENT (Consulté le: 1 juin 2026).

CORINE Land Cover (2012) *Occupation des sols sur Évry-Courcouronnes*, *Mairie d'Évry-Courcouronnes*. Disponible sur: https://www.annuaire-mairie.fr/occupation-des-sols-evry-courcouronnes.html (Consulté le: 28 juillet 2026).

De Haas, D. *et al.* (2021) « Estimating ADWF at Sewage Treatment Plants », *Water e-Journal*, 6(1), p. 1-22. Disponible sur: https://doi.org/10.21139/wej.2021.005.

Eau France (2026) *L'assainissement des eaux usées domestiques*, *Eaufrance*. Disponible sur: https://www.eaufrance.fr/lassainissement-des-eaux-usees-domestiques (Consulté le: 1 juin 2026).

France (2015) *Arrêté du 21 juillet 2015 relatif aux systèmes d'assainissement collectif et aux installations d'assainissement non collectif, à l'exception des installations d'assainissement non collectif recevant une charge brute de pollution organique inférieure ou égale à 1,2 kg/j de DBO5 - Légifrance*, *Légifrance*. Disponible sur: https://www.legifrance.gouv.fr/loda/id/JORFTEXT000031052756/2021-01-01/ (Consulté le: 1 juin 2026).

France Environnement (2024) *Gestion des eaux pluviales en milieu urbain : défis et solutions*, *FranceEnvironnement*. Disponible sur: https://www.franceenvironnement.com/article/gestion-eaux-pluviales-milieu-urbain-defis-solutions-400010866 (Consulté le: 1 juin 2026).

Gis Sol (s.d.) *Artificialisation et imperméabilisation des sols*, *Gis Sol*. Disponible sur: https://www.gissol.fr/thematiques/artificialisation-des-sols-46 (Consulté le: 28 juillet 2026).

Hosp, B.W. *et al.* (2021) « Eye Movement Feature Classification for Soccer Goalkeeper Expertise Identification in Virtual Reality », *arXiv*. Disponible sur: https://arxiv.org/pdf/2009.11676 (Consulté le: 28 juillet 2026).

Inc, M.& E. *et al.* (2003) *Wastewater Engineering: Treatment and Reuse*. Boston: McGraw Hill Higher Education.

INSEE (2023) *Dossier complet – Commune de Corbeil-Essonnes (91174)*, *Insee*. Disponible sur: https://www.insee.fr/fr/statistiques/2011101?geo=COM-91174 (Consulté le: 28 juillet 2026).

INSEE (2024) *Dossier complet – Commune d'Évry-Courcouronnes (91228)*, *Insee*. Disponible sur: https://www.insee.fr/fr/statistiques/2011101?geo=COM-91228 (Consulté le: 28 juillet 2026).

Jalbert, J. *et al.* (2024) « Categorization of precipitation for predicting combined sewer overflows. Application to the City of Montréal », *Journal of Hydrology*, 637, p. 131333. Disponible sur: https://doi.org/10.1016/j.jhydrol.2024.131333.

Liguori, S. *et al.* (2012) « Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments », *Atmospheric Research*, 103, p. 80-95. Disponible sur: https://doi.org/10.1016/j.atmosres.2011.05.004.

Lyne, V. et Hollick, M. (1979) *Stochastic time-variable rainfall-runoff modelling*. Perth, Australie: Institution of Engineers, Australia. Disponible sur: https://www.academia.edu/17986414/Stochastic_time_variable_rainfall_runoff_modelling (Consulté le: 20 juin 2026).

Métadier, M. et Bertrand-Krajewski, J.-L. (2011) « Assessing dry weather flow contribution in TSS and COD storm events loads in combined sewer systems », *Water Science & Technology*, 63, p. 2983-2991. Disponible sur: https://doi.org/10.2166/wst.2011.185.

Nasrin, T., Sharma, A.K. et Muttil, N. (2017) « Impact of Short Duration Intense Rainfall Events on Sanitary Sewer Network Performance », *Water*, 9(3), p. 225. Disponible sur: https://doi.org/10.3390/w9030225.

Rodríguez, J.P. *et al.* (2013) « Generating time-series of dry weather loads to sewers », *Environmental Modelling & Software*, 43, p. 133-143. Disponible sur: https://doi.org/10.1016/j.envsoft.2013.02.007.

Schuetze, M., Butler, D. et Beck, M. (2002) *Modelling, Simulation and Control of Urban Wastewater Systems*. Disponible sur: https://doi.org/10.1007/978-1-4471-0157-4.

Thorndahl, S. *et al.* (2008) « Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology », *Journal of Hydrology*, 357(3), p. 421-437. Disponible sur: https://doi.org/10.1016/j.jhydrol.2008.05.027.

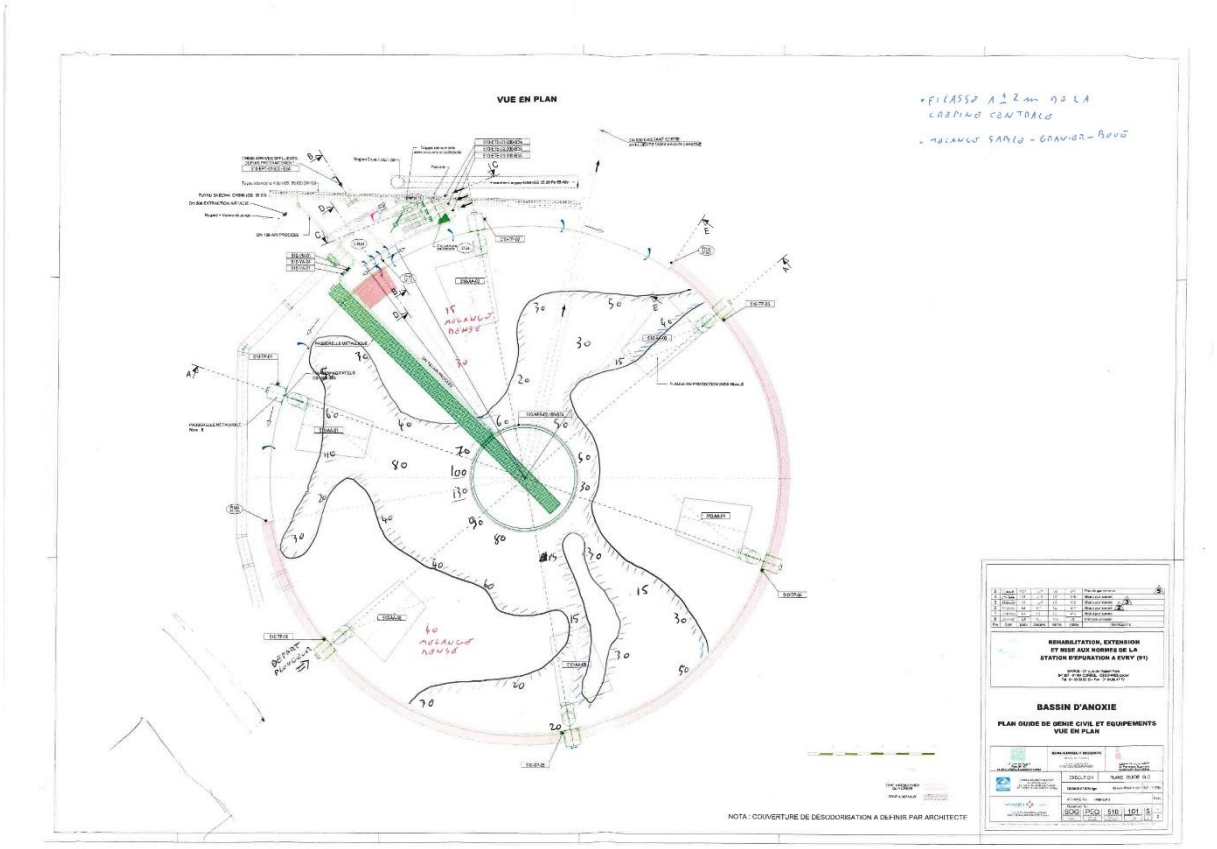
Union européenne (2024) « Directive (UE) 2024/3019 du Parlement européen et du Conseil du 27 novembre 2024 relative au traitement des eaux résiduaires urbaines », *EUR-Lex*. Disponible sur: https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A32024L3019 (Consulté le: 28 juillet 2026).

US EPA (2004) « Report to Congress on Impacts and Control of Combined Sewer Overflows and Sanitary Sewer Overflows », (EPA 833-R-04-001).

ANNEXES

Annexe 1 : Résultats de bathymétrie 2024 et 2025

Annexe 1.1 : Bathymétrie 2024



Annexe 1.2 : Bathymétrie 2025

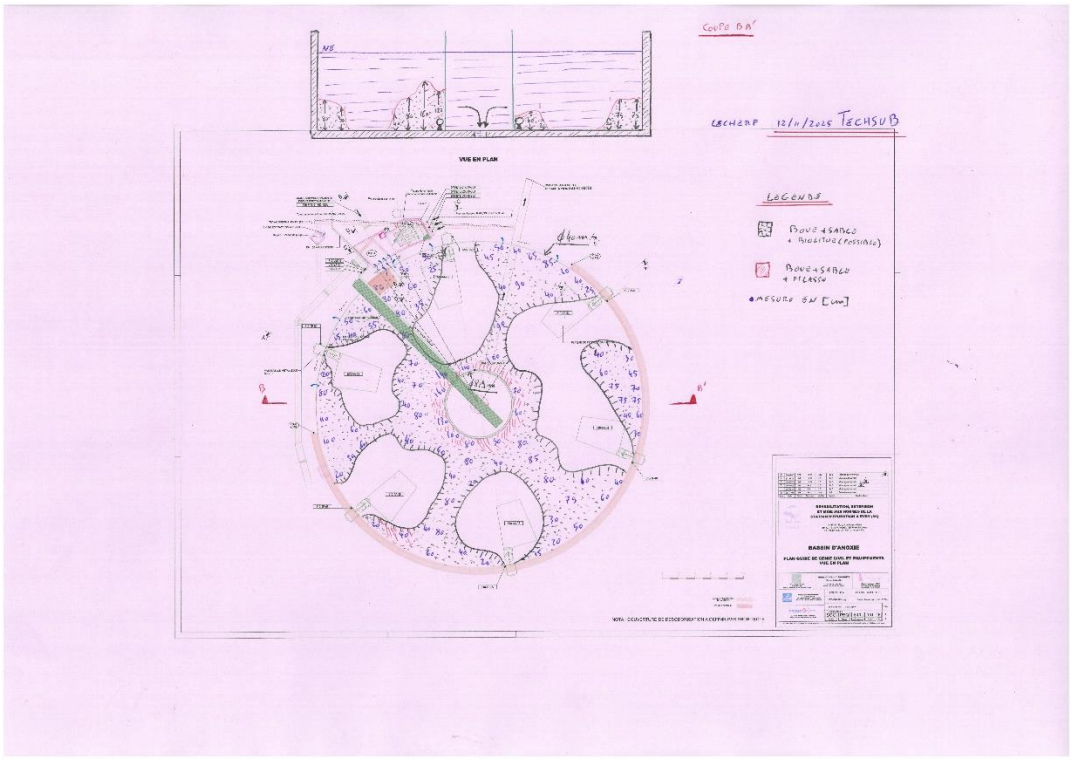
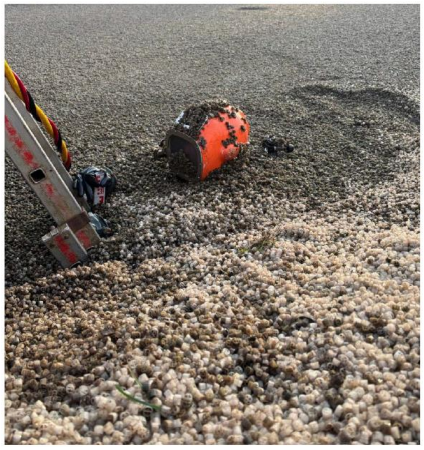


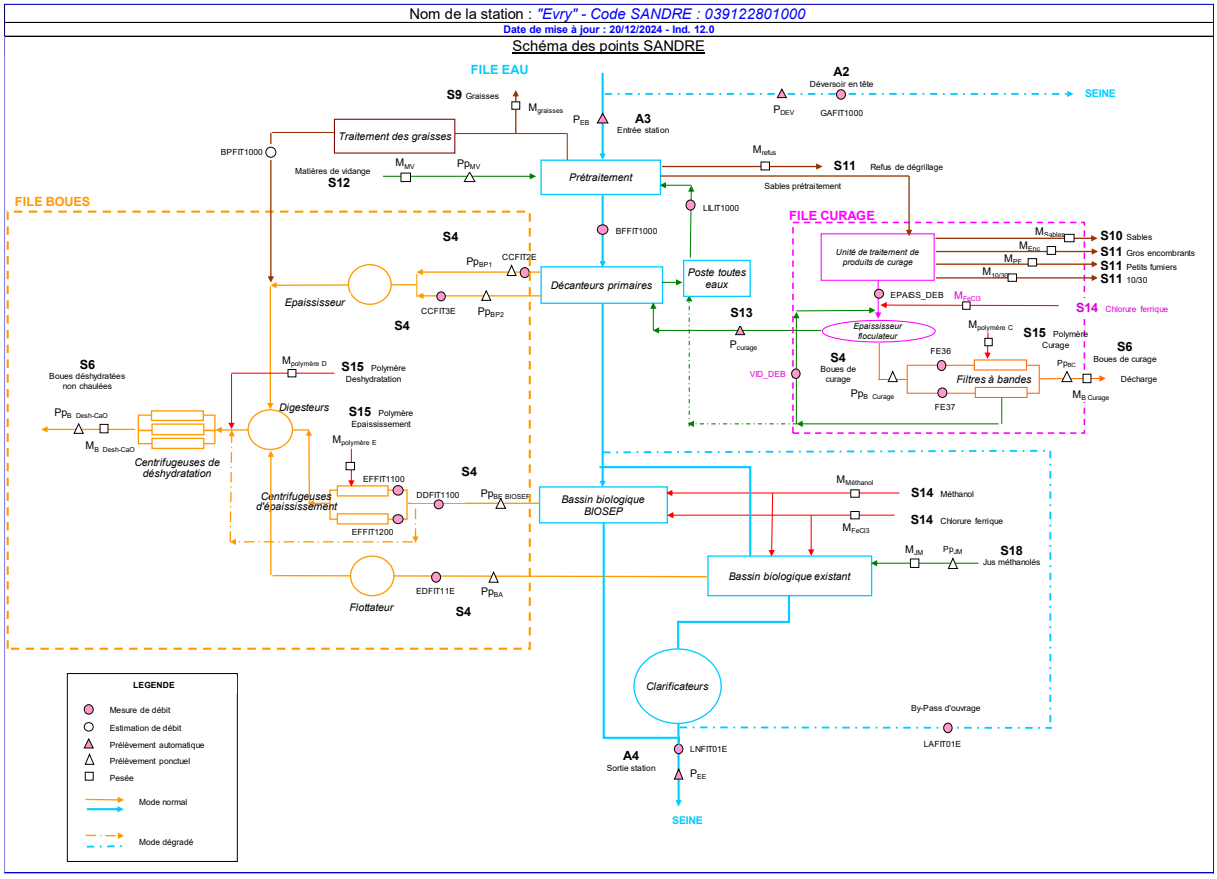
Tableau comparatif des 2 bathymétries.

	Avril /2024	*Novembre/2025*	*Delta Δ*
Surface recouverte	*523 m²*	*637 m²*	*+114m² soit + 22%*
% recouvert	*41.6 %*	*50.7 %*	*+9%*
Ep moyenne	*45.50 cm*	*59.88 cm*	*+14.38cm soit +32%*
Volume sédiment	*238 m³*	*381 m³*	*+143m³ soit +60%*

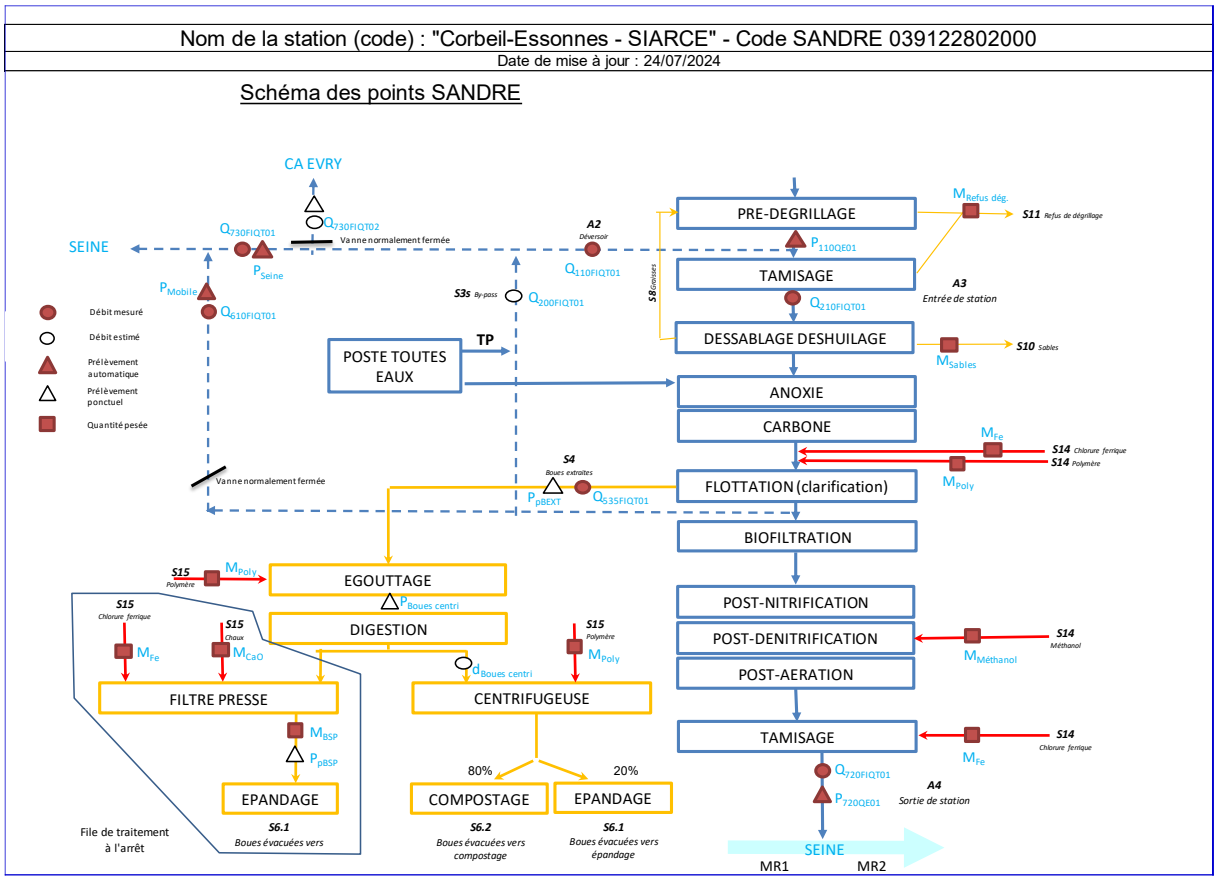


Annexe 2 : Synoptiques détaillés des stations d'Evry et d'Exona

Annexe 2.1 : Synoptique d'Evry



Annexe 2.2 : Synoptique d'Exona



Annexe 3 : Script R

Annexe 3.1 : Identification et caractérisation des épisodes pluvieux

Ce script identifie les épisodes pluvieux à partir des données au pas de 15 minutes, selon une règle de séparation inter-événements de 6 heures sans pluie. Il calcule pour chaque épisode sa durée, son cumul de pluie, ainsi que son intensité moyenne et maximale, et exporte les résultats sous forme de tableau.

```
root_folder <- dirname(rstudioapi::getActiveDocumentContext()$path)
setwd(root_folder)
Sys.setenv(TZ = "UTC")

# ============================================================================
# ANALYSE DES ÉPISODES DE PLUIE - STEP EXONA 2024
# ============================================================================
# Objectif : Identifier et caractériser les épisodes pluvieux par saison
# Règle de séparation : inter-événement de 6h minimum sans pluie (Pluie_mm = 0)
# Pas de temps : 15 minutes
# ============================================================================

library(dplyr)
library(tidyr)
library(lubridate)
library(readr)
library(zoo)

# ----------------------------------------------------------------------------
# 1. CHARGEMENT ET PRÉPARATION DES DONNÉES
# ----------------------------------------------------------------------------

df <- read_csv(
  "data.csv",
```

```
  col_types = cols(
    Date      = col_character(),
    Pluie_mm  = col_double(),
    Int_mmh   = col_double(),
    Qexona_m3h = col_double()
  )
)

# Parse des dates (format jj/mm/yyyy hh:mm)
df <- df %>%
  mutate(
    Date = dmy_hm(Date)
  ) %>%
  arrange(Date)

# Traitement des NA sur pluie : NA = pas de pluie (0)
df <- df %>%
  mutate(
    Pluie_mm = replace_na(Pluie_mm, 0),
    Int_mmh  = replace_na(Int_mmh,  0)
  )

# ------------------------------------------------------------------------
# 2. DÉFINITION DES SAISONS
# ------------------------------------------------------------------------

get_saison <- function(date) {
  m <- month(date)
  case_when(
    m %in% c(12, 1, 2)  ~ "Hiver",
    m %in% c(3, 4, 5)   ~ "Printemps",
```

```
    m %in% c(6, 7, 8)   ~ "Été",
    m %in% c(9, 10, 11) ~ "Automne"
  )
}

df <- df %>%
  mutate(saison = get_saison(Date))

# ------------------------------------------------------------------------
# 3. IDENTIFICATION DES ÉPISODES PLUVIEUX
# ------------------------------------------------------------------------
# Règle : un épisode débute dès que Pluie_mm > 0.
# Il se termine quand 6h consécutives (= 24 pas de 15 min) sont à Pluie_mm = 0.
# On utilise un algorithme de propagation : on marque chaque pas "actif"
# si une pluie a été observée dans les 24 derniers pas (fenêtre glissante arrière).

seuil_interev_pas <- 24   # 6h / 0.25h = 24 pas de 15 min

# Indicateur binaire de pluie
df <- df %>%
  mutate(pluie_active = Pluie_mm > 0)

# Propager l'activité : un pas reste "dans un épisode" si une pluie est tombée
# dans les 24 pas précédents (incluant le pas courant).
# Implémentation via rolling maximum sur une fenêtre de 24 pas.
df <- df %>%
  mutate(
    in_episode = rollapply(
      pluie_active,
      width  = seuil_interev_pas,
```

```
      FUN    = any,
      align  = "right",
      fill   = FALSE
    )
  )

# Numérotation des épisodes : changement de FALSE -> TRUE = nouveau départ
df <- df %>%
  mutate(
    episode_start = in_episode & !lag(in_episode, default = FALSE),
    episode_id    = cumsum(episode_start)
  )

# Conserver uniquement les pas appartenant à un épisode
episodes_pas <- df %>%
  filter(in_episode) %>%
  # Recalibrage : l'épisode commence réellement au 1er pas avec pluie,
  # pas au début de la fenêtre glissante. On ajuste episode_id en conséquence.
  group_by(episode_id)

# ------------------------------------------------------------------------------
# 4. CALCUL DES CARACTÉRISTIQUES PAR ÉPISODE
# ------------------------------------------------------------------------------

tableau_episodes <- episodes_pas %>%
  summarise(
    debut          = min(Date[Pluie_mm > 0]),           # 1er pas avec pluie
    fin            = max(Date[Pluie_mm > 0]),            # dernier pas avec pluie
    duree_h        = as.numeric(difftime(fin, debut, units = "hours")) + 0.25,
    pluie_cum_mm   = sum(Pluie_mm, na.rm = TRUE),
    int_moy_mmh    = round(pluie_cum_mm / duree_h, 2),
```

```
    int_max_mmh    = max(Int_mmh, na.rm = TRUE),
    saison         = first(get_saison(debut)),
    .groups = "drop"
  ) %>%
  # Filtrer les épisodes avec pluie réelle (exclure artéfacts)
  filter(pluie_cum_mm > 0) %>%
  filter(duree_h > 0.25) %>%
  # Réinitialiser la numérotation
  mutate(
    episode_id = row_number(),
    debut      = format(debut, "%d/%m/%Y %H:%M"),
    fin        = format(fin,   "%d/%m/%Y %H:%M"),
    duree_h    = round(duree_h, 2),
    pluie_cum_mm = round(pluie_cum_mm, 1),
    int_max_mmh  = round(int_max_mmh, 1)
  ) %>%
  select(
    episode_id, saison, debut, fin,
    duree_h, pluie_cum_mm, int_moy_mmh, int_max_mmh
  )

# Renommer les colonnes pour l'affichage
colnames(tableau_episodes) <- c(
  "N°", "Saison", "Début", "Fin",
  "Durée (h)", "Pluie cumulée (mm)", "Int. moy. (mm/h)", "Int. max. (mm/h)"
)

# ------------------------------------------------------------------------------
# 5. EXPORT
# ------------------------------------------------------------------------------
# Export CSV
```

```
write_excel_csv(tableau_episodes, "tableaux/1_episodes_pluie_2024.csv")

```

Annexe 3.2 : Comparaison des épisodes pluvieux par saison

Ce script construit des graphiques en boîtes à moustaches (boxplots) comparant, pour chaque saison, la pluie cumulée, la durée, l'intensité moyenne et l'intensité maximale des épisodes pluvieux identifiés dans le script précédent. Il produit également un tableau de synthèse statistique (médiane, min-max) par saison.

```
root_folder <- dirname(rstudioapi::getActiveDocumentContext()$path)
setwd(root_folder)
Sys.setenv(TZ = "UTC")

# ============================================================================
# VISUALISATION DES ÉPISODES DE PLUIE PAR SAISON - 2024
# ============================================================================

library(readr)
library(dplyr)
library(ggplot2)
library(patchwork)

# ----------------------------------------------------------------------------
# 1. CHARGEMENT DES DONNÉES
# ----------------------------------------------------------------------------
df <- read_csv("tableaux/1_episodes_pluie_2024.csv")

# Ordre logique des saisons
ordre_saisons <- c("Printemps", "Été", "Automne", "Hiver")

```

```
df <- df %>%
  mutate(Saison = factor(Saison, levels = ordre_saisons))

# ------------------------------------------------------------------------------
# 2. THÈME COMMUN
# ------------------------------------------------------------------------------

theme_episodes <- theme_bw(base_size = 12) +
  theme(
    panel.grid.minor  = element_blank(),
    panel.grid.major.x = element_blank(),
    axis.title.x      = element_blank(),
    axis.text         = element_text(size = 12, colour = "black"),
    plot.title        = element_text(face = "bold", size = 18, colour = "black"),
    legend.position   = "none"
  )

couleurs_saisons <- c(
  "Printemps" = "#4dac26",
  "Été"       = "#f4a582",
  "Automne"   = "#d6604d",
  "Hiver"     = "#92c5de"
)

# Fonction pour construire un panneau boxplot + moyenne (losange rouge)
make_panel <- function(data, var, ylab, titre, ylims) {
  ggplot(data, aes(x = Saison, y = .data[[var]], fill = Saison)) +
    geom_boxplot(alpha = 0.7, outlier.shape = NA) +
    stat_summary(fun = mean, geom = "point",
                 shape = 23, size = 3,
                 fill = "red", color = "black") +
```

```
    scale_fill_manual(values = couleurs_saisons) +
    labs(title = titre, y = ylab) +
    ylim(ylims) +
    theme_episodes
}

# ------------------------------------------------------------------------
# 3. CONSTRUCTION DES 4 PANNEAUX
# ------------------------------------------------------------------------

p1 <- make_panel(df,
                 var   = "Pluie cumulée (mm)",
                 ylab  = "Pluie cumulée (mm)",
                 titre = "a) Pluie cumulée par épisode", c(0,15))

p2 <- make_panel(df,
                 var   = "Durée (h)",
                 ylab  = "Durée (h)",
                 titre = "b) Durée des épisodes",c(0,20))

p3 <- make_panel(df,
                 var   = "Int. moy. (mm/h)",
                 ylab  = "Intensité moyenne (mm/h)",
                 titre = "c) Intensité moyenne", c(0,3))


p4 <- make_panel(df,
                 var   = "Int. max. (mm/h)",
                 ylab  = "Intensité maximale (mm/h)",
                 titre = "d) Intensité maximale", c(0,20))

```

```
# ------------------------------------------------------------------------------
# 4. ASSEMBLAGE ET EXPORT
# ------------------------------------------------------------------------------

figure <- (p1 | p2) / (p3 | p4)

ggsave(
  filename = "figures/1_episodes_pluie_boxplots.png",
  plot     = figure,
  width    = 12,
  height   = 8,
  dpi      = 350,
  bg       = "white"
)

# ------------------------------------------------------------------------------
# 5. TABLEAU SYNTHÈSE PAR SAISON
# ------------------------------------------------------------------------------

synthese <- df %>%
  group_by(Saison) %>%
  summarise(
    `Nb épisodes`                  = n(),

    `Pluie médiane (mm)`           = round(median(`Pluie cumulée (mm)`), 1),
    `Pluie min-max (mm)`           = paste0(round(min(`Pluie cumulée (mm)`), 1),
                                            " - ",
                                            round(max(`Pluie cumulée (mm)`),
1)),

    `Durée médiane (h)`            = round(median(`Durée (h)`), 1),
```

```
    `Durée min-max (h)`              = paste0(round(min(`Durée (h)`), 1),
                                              " - ",
                                              round(max(`Durée (h)`), 1)),

    `Int. moy. médiane (mm/h)`       = round(median(`Int. moy. (mm/h)`), 2),
    `Int. moy. min-max (mm/h)`       = paste0(round(min(`Int. moy. (mm/h)`), 2),
                                              " - ",
                                              round(max(`Int. moy. (mm/h)`), 2)),

    `Int. max. médiane (mm/h)`       = round(median(`Int. max. (mm/h)`), 1),
    `Int. max. min-max (mm/h)`       = paste0(round(min(`Int. max. (mm/h)`), 1),
                                              " - ",
                                              round(max(`Int. max. (mm/h)`), 1)),

    .groups = "drop"
  )

write_excel_csv(synthese, "tableaux/2_synthese_episodes_par_saison.csv")

```

Annexe 3.3 : Exploration mensuelle de la relation pluie-débit

Ce script génère, pour chacun des 12 mois de l'année 2024, un graphique en trois panneaux superposant la pluie cumulée, l'intensité pluviométrique et le débit entrant à la station d'Exona, avec le seuil de bypass de 1 000 m³/h matérialisé en ligne pointillée rouge. Les figures sont exportées automatiquement pour les 12 mois.

```
root_folder <- dirname(rstudioapi::getActiveDocumentContext()$path)
setwd(root_folder)
Sys.setenv(TZ = "UTC")

# ============================================================================
# VISUALISATION EXPLORATOIRE - TOUS LES MOIS 2024
# Pluie cumulée + Intensité + Débit Qexona + Seuil bypass
# ============================================================================

library(readr)
library(dplyr)
library(lubridate)
library(ggplot2)
library(patchwork)

# ----------------------------------------------------------------------------
# 1. CHARGEMENT COMPLET
# ----------------------------------------------------------------------------

df_all <- read_csv(
  "data.csv",
  col_types = cols(
    Date       = col_character(),
    Pluie_mm   = col_double(),
    Int_mmh    = col_double(),
    Qexona_m3h = col_double()
```

```
  )
) %>%
  mutate(
    Date     = dmy_hm(Date),
    Pluie_mm = replace_na(Pluie_mm, 0),
    Int_mmh  = replace_na(Int_mmh,  0)
  )

# ------------------------------------------------------------------------
# 2. THÈME COMMUN
# ------------------------------------------------------------------------

theme_visu <- theme_bw(base_size = 11) +
  theme(
    panel.grid.minor = element_blank(),
    axis.title.x     = element_blank(),
    plot.title       = element_text(face = "bold", size = 16, colour = "black"),
    axis.text        = element_text(size = 12, colour = "black"),
    axis.title       = element_text(size = 12, colour = "black")
  )

noms_mois <- c(
  "01_janvier", "02_fevrier", "03_mars",     "04_avril",
  "05_mai",     "06_juin",    "07_juillet",  "08_aout",
  "09_septembre","10_octobre","11_novembre", "12_decembre"
)

dir.create("figures/2_exploration", recursive = T)
# ------------------------------------------------------------------------
# 3. BOUCLE SUR LES 12 MOIS
# ------------------------------------------------------------------------
```

```
58
59 for (m in 1:12) {
60
61   df <- df_all %>% filter(month(Date) == m)
62
63   # Ignorer les mois sans données
64   if (nrow(df) == 0) {
65     cat("Mois", m, ": aucune donnée, ignoré.\n")
66     next
67   }
68
69   titre_mois <- format(df$Date[1], "%B %Y")
70
71   # -- Panneau a) : Pluie cumulée --
72   p_pluie <- ggplot(df, aes(x = Date, y = Pluie_mm)) +
73     geom_col(fill = "#4393c3", width = 60 * 15) +
74     scale_x_datetime(date_breaks = "1 week", date_labels = "%d %b") +
75     labs(title = "a) Pluie cumulée par pas de 15 min (mm)",
76          y = "Pluie (mm)") +
77     theme_visu
78
79   # -- Panneau b) : Intensité --
80   p_int <- ggplot(df, aes(x = Date, y = Int_mmh)) +
81     geom_col(fill = "#d6604d", width = 60 * 15) +
82     scale_x_datetime(date_breaks = "1 week", date_labels = "%d %b") +
83     labs(title = "b) Intensité de pluie (mm/h)",
84          y = "Intensité (mm/h)") +
85     theme_visu
86
87   # -- Panneau c) : Débit + seuil bypass --
88   p_debit <- ggplot(df, aes(x = Date, y = Qexona_m3h)) +
```

```
    geom_line(color = "#2166ac", linewidth = 0.4) +
    geom_hline(yintercept = 1000, color = "red",
               linetype = "dashed", linewidth = 0.7) +
    annotate("text", x = min(df$Date), y = max(df$Qexona_m3h, na.rm = TRUE) *
0.97,
             label = "Seuil bypass 1000 m³/h",
             color = "red", hjust = 0, size = 5) +
    scale_x_datetime(date_breaks = "1 week", date_labels = "%d %b") +
    labs(title = "c) Débit entrant STEP Exona (m³/h)",
         y = "Débit (m³/h)") +
    theme_visu

  # -- Assemblage --
  figure <- p_pluie / p_int / p_debit +
    plot_layout(heights = c(1.5, 1.5, 3))

  # -- Export --
  nom_fichier <- paste0("figures/2_exploration/exploration_", noms_mois[m],
".png")
  ggsave(
    filename = nom_fichier,
    plot     = figure,
    width    = 14,
    height   = 10,
    dpi      = 350,
    bg       = "white"
  )
  cat("Figure exportée :", nom_fichier, "\n")
}

cat("\nTerminé — 12 figures générées.\n")
```

Annexe 3.4 : Détection et synthèse des dépassements de consigne

Ce script détecte les épisodes de dépassement du seuil de bypass (1 000 m³/h) à la station d'Exona, calcule pour chacun sa durée, son débit maximal et son volume bypassé, puis produit une synthèse par saison ainsi que des graphiques en boîtes à moustaches comparant ces indicateurs entre saisons.

```
root_folder <- dirname(rstudioapi::getActiveDocumentContext()$path)
setwd(root_folder)
Sys.setenv(TZ = "UTC")

# ============================================================================
# Analyse des dépassements de consigne - STEP Exona
# Consigne bypass : Qexona_m3h > 1000 m3/h
# Pas de temps : 15 minutes
# ============================================================================

library(tidyverse)
library(readr)
library(lubridate)

# ----------------------------------------------------------------------------
# 1. Chargement des données
# ----------------------------------------------------------------------------

df <- read_delim(
  "data.csv",
  delim = ",",
  locale = locale(
    decimal_mark = ".",
    date_names  = "fr"
  ),
```

```
  show_col_types = FALSE
) %>%
  mutate(
    Date = dmy_hm(Date),   # format jj/mm/yyyy hh:mm
    .keep = "all"
  ) %>%
  arrange(Date)

# ------------------------------------------------------------------------
# 2. Paramètres
# ------------------------------------------------------------------------

CONSIGNE   <- 1000   # m3/h
DT_H       <- 15/60  # pas de temps en heures (15 min → 0.25 h)

# ------------------------------------------------------------------------
# 3. Détection et labellisation des épisodes de dépassement
# ------------------------------------------------------------------------

df <- df %>%
  mutate(
    depasse      = !is.na(Qexona_m3h) & Qexona_m3h > CONSIGNE,
    # Numéro d'épisode : incrémenté à chaque nouveau passage TRUE après FALSE
    episode_id   = cumsum(depasse & !lag(depasse, default = FALSE))
  ) %>%
  mutate(
    # On ne garde le numéro que sur les pas de temps en dépassement
    episode_id = if_else(depasse, episode_id, NA_integer_)
  )

# ------------------------------------------------------------------------
```

```
# 4. Fonction saison (hémisphère nord, calendrier météorologique)
# ------------------------------------------------------------------------------

get_saison <- function(date) {
  m <- month(date)
  case_when(
    m %in% c(12, 1, 2) ~ "Hiver",
    m %in% c(3, 4, 5)  ~ "Printemps",
    m %in% c(6, 7, 8)  ~ "Été",
    m %in% c(9, 10, 11) ~ "Automne"
  )
}

# ------------------------------------------------------------------------------
# 5. Résumé par épisode
# ------------------------------------------------------------------------------

resultats <- df %>%
  filter(depasse) %>%
  group_by(episode_id) %>%
  summarise(
    debut          = min(Date),
    fin            = max(Date),
    duree_h        = as.numeric(difftime(max(Date), min(Date), units = "hours")) +
DT_H,
    debit_max_m3h  = max(Qexona_m3h, na.rm = TRUE),
    # Volume bypassé = somme des débits résiduels (Q - consigne) × dt
    volume_bypasse_m3 = sum((Qexona_m3h - CONSIGNE) * DT_H, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  mutate(
```

```
    saison = get_saison(debut)
  ) %>%
  # Réordonnement des colonnes : saison en premier
  select(
    saison,
    episode_id,
    debut,
    fin,
    duree_h,
    debit_max_m3h,
    volume_bypasse_m3
  ) %>%
  arrange(debut)

# ------------------------------------------------------------------------------
# 6. Export CSV (encodage Excel / Windows)
# ------------------------------------------------------------------------------

write_excel_csv(
  resultats,
  file      = "tableaux/3_depassements_exona_2024.csv",
  delim     = ","   # séparateur point-virgule, compatible Excel FR
)

cat("\nFichier exporté : depassements_exona_2024.csv\n")


# ------------------------------------------------------------------------------
# Synthèse par saison
# ------------------------------------------------------------------------------

```

```
# Ordre calendaire des saisons (utile pour l'affichage et l'export)
ordre_saisons <- c("Hiver", "Printemps", "Été", "Automne")

synthese_saison <- resultats %>%
  group_by(saison) %>%
  summarise(
    nb_depassements        = n(),
    duree_mediane_h        = median(duree_h),
    debit_max_median_m3h   = median(debit_max_m3h),
    volume_total_bypasse_m3 = sum(volume_bypasse_m3),
    .groups = "drop"
  ) %>%
  mutate(saison = factor(saison, levels = ordre_saisons)) %>%
  arrange(saison)

write_excel_csv(
  synthese_saison %>% mutate(saison = as.character(saison)),
  file   = "tableaux/4_synthese_depassements_exona_2024.csv",
  delim  = ",",
  append = TRUE
)

# ============================================================================
# VISUALISATION DES DÉPASSEMENTS DE CONSIGNE PAR SAISON - STEP EXONA 2024
# ============================================================================
library(ggplot2)
library(patchwork)

resultats <- resultats %>%
  mutate(saison = factor(saison, levels = ordre_saisons))

```

```
couleurs_saisons <- c(
  "Hiver"     = "#92c5de",
  "Printemps" = "#4dac26",
  "Été"       = "#f4a582",
  "Automne"   = "#d6604d"
)

theme_depassements <- theme_bw(base_size = 12) +
  theme(
    panel.grid.minor   = element_blank(),
    panel.grid.major.x = element_blank(),
    axis.title.x       = element_blank(),
    axis.text          = element_text(size = 12, colour = "black"),
    plot.title         = element_text(face = "bold", size = 14, colour = "black"),
    legend.position    = "none"
  )

# Fonction générique : boxplot par saison + losange rouge pour la moyenne
make_panel <- function(data, var, ylab, titre, ylims = NULL) {
  p <- ggplot(data, aes(x = saison, y = .data[[var]], fill = saison)) +
    geom_boxplot(alpha = 0.7, outlier.shape = NA,
                 outlier.fill = "white", outlier.color = "grey40") +
    stat_summary(fun = mean, geom = "point",
                 shape = 23, size = 3,
                 fill = "red", color = "black") +
    scale_fill_manual(values = couleurs_saisons) +
    labs(title = titre, y = ylab) +
    theme_depassements

  if (!is.null(ylims)) {
    p <- p + scale_y_continuous(
```

```
      trans  = "sqrt",
      limits = ylims
    )
  } else {
    p <- p + scale_y_continuous(trans = "sqrt")
  }

  p
}

# ------------------------------------------------------------------------------
# CONSTRUCTION DES 3 PANNEAUX
# ------------------------------------------------------------------------------

p1 <- make_panel(
  data  = resultats,
  var   = "duree_h",
  ylab  = "Durée (h)",
  titre = "a) Durée des dépassements",
  ylim  = c(0,5)
)

p2 <- make_panel(
  data  = resultats,
  var   = "debit_max_m3h",
  ylab  = "Débit max (m³/h)",
  titre = "b) Débit maximum",
  ylim  = c(1000,2500)
)

p3 <- make_panel(
```

```
  data  = resultats,
  var   = "volume_bypasse_m3",
  ylab  = "Volume bypassé (m³)",
  titre = "c) Volume bypassé",
  ylim  = c(0,2500)
)

# ------------------------------------------------------------------------
# ASSEMBLAGE ET EXPORT
# ------------------------------------------------------------------------

figure <- p1 | p2 | p3

ggsave(
  filename = "figures/3_depassements_exona_boxplots.png",
  plot     = figure,
  width    = 14,
  height   = 5,
  dpi      = 350,
  bg       = "white"
)

cat("Figure exportée : figures/depassements_exona_boxplots.png\n")
```

Annexe 3.5 : Séparation du débit de base et des eaux parasites pluviales

Ce script sépare le débit entrant à Exona en deux composantes : le débit de base (Qbase_eff), estimé par une référence glissante médiane en période de temps sec (±21 jours, par créneau horaire et type de jour), et les eaux parasites pluviales (Qepp), calculées comme l'excédent de débit observé en période de pluie récente. Le bilan de masse Qexona = Qbase_eff + Qepp est vérifié, et des graphiques mensuels de validation sont générés pour les 12 mois de l'année.

```
root_folder <- dirname(rstudioapi::getActiveDocumentContext()$path)
setwd(root_folder)
Sys.setenv(TZ = "UTC")

# ============================================================================
# SÉPARATION QBASE_EFF / QEPP — VERSION FINALE
# ============================================================================
# Principe :
#   - Qbase    : référence glissante (médiane temps sec, ±21j, même créneau
#                horaire × type de jour) — estimateur du niveau attendu des
#                eaux usées en l'absence de pluie.
#   - Qepp     : eaux parasites pluviales
#                = pmax(Qexona - Qbase, 0)  si pluie récente
#                = 0                        sinon
#   - Qbase_eff: débit de base effectif = Qexona - Qepp
#                → absorbe tout Qexona en temps sec (Qepp = 0)
#                → absorbe la part sous Qbase en temps de pluie
#
# Bilan de masse garanti : Qexona = Qbase_eff + Qepp à tout instant.
# Qepp = 0 exactement hors épisodes pluvieux par construction.
#
# Paramètres à caler visuellement :
#   DELAI_RETOUR_H : durée (h) après laquelle Qepp revient à 0 sans pluie
#   SEUIL_PLUIE_MM : hauteur minimale (mm/pas) pour activer le masque
```

```
# ============================================================================

library(readr)
library(dplyr)
library(lubridate)
library(ggplot2)
library(patchwork)
library(zoo)

# ----------------------------------------------------------------------------
# 0. PARAMÈTRES
# ----------------------------------------------------------------------------

FENETRE_JOURS  <- 21    # demi-fenêtre glissante pour Qbase (±21 jours)
PAS_48H        <- 192   # 48h en nombre de pas de temps à 15 min (critère ts)
DELAI_RETOUR_H <- 2     # heures sans pluie avant que Qepp revienne à 0
# réduit à 2h : retour rapide au repos après un épisode
SEUIL_PLUIE_MM <- 0.5   # hauteur min (mm) pour qu'un pas soit considéré pluvieux
# relevé à 0.5mm : ignore les traces invisibles sur les graphiques

DT_MIN         <- 15    # pas de temps en minutes
pas_retour     <- as.integer(DELAI_RETOUR_H * 60 / DT_MIN)

# ----------------------------------------------------------------------------
# 1. CHARGEMENT ET PRÉPARATION
# ----------------------------------------------------------------------------

df <- read_csv(
  "data.csv",
  col_types = cols(
    Date       = col_character(),
```

```
    Pluie_mm   = col_double(),
    Int_mmh    = col_double(),
    Qexona_m3h = col_double()
  )
) %>%
  mutate(
    Date     = dmy_hm(Date),
    Pluie_mm = replace_na(Pluie_mm, 0),
    Int_mmh  = replace_na(Int_mmh,  0)
  ) %>%
  arrange(Date)

cat("Période couverte :", format(min(df$Date)), "→", format(max(df$Date)), "\n")
cat("Nombre de pas de temps :", nrow(df), "\n\n")

# ------------------------------------------------------------------------
# 2. IDENTIFICATION DES PÉRIODES TEMPS SEC
# ------------------------------------------------------------------------

df <- df %>%
  mutate(
    pluie_max_avant = rollapply(Pluie_mm, width = PAS_48H,
                                FUN = max, align = "right", fill = NA),
    pluie_max_apres = rollapply(Pluie_mm, width = PAS_48H,
                                FUN = max, align = "left",  fill = NA),
    temps_sec = (Pluie_mm == 0) &
      (!is.na(pluie_max_avant) & pluie_max_avant == 0) &
      (!is.na(pluie_max_apres) & pluie_max_apres == 0)
  )

cat("Pas de temps temps sec :", sum(df$temps_sec, na.rm = TRUE),
```

```
    "/", nrow(df), "(", round(mean(df$temps_sec, na.rm = TRUE) * 100, 1),
"%)\n\n")

# ------------------------------------------------------------------------
# 3. VARIABLES CALENDAIRES
# ------------------------------------------------------------------------

df <- df %>%
  mutate(
    heure_min = sprintf("%02d:%02d", hour(Date), minute(Date)),
    type_jour = if_else(wday(Date, week_start = 1) <= 5, "semaine", "weekend"),
    jour_num  = as.numeric(as.Date(Date))
  )

# ------------------------------------------------------------------------
# 4. CALCUL DE QBASE (RÉFÉRENCE GLISSANTE MÉDIANE)
# ------------------------------------------------------------------------

profil_jour <- df %>%
  filter(temps_sec, !is.na(Qexona_m3h)) %>%
  group_by(jour_num, heure_min, type_jour) %>%
  summarise(Q_med_jour = median(Qexona_m3h), .groups = "drop")

cat("Calcul de Qbase (référence glissante)...\n")

groupes  <- df %>% distinct(heure_min, type_jour)
df$Qbase <- NA_real_

for (i in seq_len(nrow(groupes))) {

  hm  <- groupes$heure_min[i]
```

```
  tj  <- groupes$type_jour[i]

  idx_df   <- which(df$heure_min == hm & df$type_jour == tj)
  jours_ts <- profil_jour %>% filter(heure_min == hm, type_jour == tj)

  if (nrow(jours_ts) == 0) next

  jours_ref <- df$jour_num[idx_df]

  df$Qbase[idx_df] <- sapply(jours_ref, function(jt) {
    vals <- jours_ts$Q_med_jour[abs(jours_ts$jour_num - jt) <= FENETRE_JOURS]
    if (length(vals) < 3) NA_real_ else median(vals)
  })
}

df <- df %>%
  mutate(Qbase = na.approx(Qbase, na.rm = FALSE))

cat("Qbase calculé.\n\n")

# ------------------------------------------------------------------------------
# 5. MASQUE PLUIE RÉCENTE
# ------------------------------------------------------------------------------
# pluie_recente = TRUE si au moins un pas de temps avec Pluie >= SEUIL_PLUIE_MM
# dans les DELAI_RETOUR_H dernières heures.
# Qepp est non nul UNIQUEMENT lorsque pluie_recente = TRUE.

df <- df %>%
  mutate(
    pluie_recente = rollapply(
      Pluie_mm,
```

```
      width = pas_retour,
      FUN   = function(x) any(x >= SEUIL_PLUIE_MM),
      align = "right",
      fill  = FALSE
    )
  )

pct_pluie_recente <- mean(df$pluie_recente, na.rm = TRUE) * 100
cat(sprintf("Pas de temps avec pluie récente : %.1f %%\n\n", pct_pluie_recente))

# ------------------------------------------------------------------------------
# 6. CALCUL DE QEPP ET QBASE_EFF
# ------------------------------------------------------------------------------
# Qepp      = pmax(Qexona - Qbase, 0)  si pluie_recente, sinon 0
# Qbase_eff = Qexona - Qepp            (absorbe tout Qexona hors pluie)
#
# Bilan garanti : Qexona = Qbase_eff + Qepp

df <- df %>%
  mutate(
    Qepp      = if_else(pluie_recente, pmax(Qexona_m3h - Qbase, 0), 0),
    Qbase_eff = Qexona_m3h - Qepp
  )

# Vérifications
erreur_max       <- max(abs(df$Qexona_m3h - df$Qbase_eff - df$Qepp), na.rm = TRUE)
pct_qepp_nul_sec <- mean(df$Qepp[!df$pluie_recente] == 0, na.rm = TRUE) * 100

cat(sprintf("Bilan de masse — erreur max        : %.2e m³/h\n",  erreur_max))
cat(sprintf("Qepp = 0 hors pluie récente        : %.1f %% (doit être 100%%)\n\n",
            pct_qepp_nul_sec))
```

```
179
180 # ------------------------------------------------------------------------
181 # 7. EXPORT
182 # ------------------------------------------------------------------------
183
184 df %>%
185   select(Date, Pluie_mm, Int_mmh, Qexona_m3h, Qbase_eff, Qepp) %>%
186   write_excel_csv("data_qbase_qepp.csv", delim = ",")
187
188 cat("Données exportées : data_qbase_qepp.csv\n\n")
189
190 # ------------------------------------------------------------------------
191 # 8. VISUALISATION DE VALIDATION
192 # ------------------------------------------------------------------------
193
194 theme_v <- theme_bw(base_size = 12) +
195   theme(
196     panel.grid.minor  = element_blank(),
197     axis.title.x      = element_blank(),
198     axis.text         = element_text(size = 12, colour = "black"),
199     axis.title.y      = element_text(size = 12, colour = "black")
200   )
201
202 valider_mois <- function(data, m, titre) {
203
204   d <- data %>% filter(month(Date) == m)
205
206   p_pluie <- ggplot(d, aes(x = Date, y = Pluie_mm)) +
207     geom_col(fill = "#4393c3", width = 60 * DT_MIN) +
208     scale_x_datetime(date_breaks = "1 week", date_labels = "%d %b") +
209     labs(title = "Pluie cumulée (mm)", y = "Pluie (mm)") +
```

```
    theme_v

  p_debit <- ggplot(d, aes(x = Date)) +
    geom_line(aes(y = Qexona_m3h), color = "#2166ac",
              linewidth = 0.35, alpha = 0.8) +
    geom_line(aes(y = Qbase_eff), color = "orange",
              linewidth = 0.7, linetype = "dashed") +
    geom_hline(yintercept = 1000, color = "red",
               linetype = "dashed", linewidth = 0.5) +
    scale_x_datetime(date_breaks = "1 week", date_labels = "%d %b") +
    labs(
      title = "Qexona (bleu) | Qbase (orange) | Consigne 1 000 m³/h (rouge)",
      y     = "Débit (m³/h)"
    ) +
    theme_v

  p_qepp <- ggplot(d, aes(x = Date)) +
    geom_hline(yintercept = 0, color = "grey40", linewidth = 0.4) +
    geom_col(aes(y = Qepp), fill = "#d6604d", alpha = 0.7, width = 60 * DT_MIN) +
    scale_x_datetime(date_breaks = "1 week", date_labels = "%d %b") +
    labs(
      title = sprintf("Qepp (rouge) - actif si pluie ≥ %.1f mm dans les %dh
précédentes",
                      SEUIL_PLUIE_MM, DELAI_RETOUR_H),
      y = "Qepp (m³/h)"
    ) +
    theme_v

  p_pluie / p_debit / p_qepp +
    plot_layout(heights = c(1, 2.5, 2)) +
    plot_annotation(
```

```
      title = titre,
      theme = theme(plot.title = element_text(face   = "bold",
                                              size   = 18,
                                              colour = "black",
                                              hjust  = 0))
    )
}

mois_validation <- list(
  list(m =  1, titre = "Janvier 2024 - Hiver"),
  list(m =  2, titre = "Février 2024 - Hiver"),
  list(m =  3, titre = "Mars 2024 - Printemps"),
  list(m =  4, titre = "Avril 2024 - Printemps"),
  list(m =  5, titre = "Mai 2024 - Printemps"),
  list(m =  6, titre = "Juin 2024 - Été"),
  list(m =  7, titre = "Juillet 2024 - Été"),
  list(m =  8, titre = "Août 2024 - Été"),
  list(m =  9, titre = "Septembre 2024 - Automne"),
  list(m = 10, titre = "Octobre 2024 - Automne"),
  list(m = 11, titre = "Novembre 2024 - Automne"),
  list(m = 12, titre = "Décembre 2024 - Hiver")
)

dir.create("figures/4_separation", showWarnings = FALSE, recursive = TRUE)

for (mv in mois_validation) {
  fig <- valider_mois(df, mv$m, mv$titre)
  nom <- paste0("figures/4_separation/validation_qbase_qepp_",
                sprintf("%02d", mv$m), ".png")
  ggsave(nom, fig, width = 14, height = 11, dpi = 350, bg = "white")
  cat("Figure exportée :", nom, "\n")
```

```
271 }
272
273 cat("\nTerminé.\n")
```

Annexe 3.6 : Modèle de prédiction du bypass (arbre de décision)

Ce script construit un arbre de décision permettant d'anticiper le risque de bypass à Exona à l'horizon d'une heure, à partir du débit moyen, du débit de l'heure précédente, de sa tendance et du cumul de pluie sur 2 et 6 heures. Après entraînement sur 80 % des données (split chronologique) et élagage de l'arbre, le script évalue le modèle sur le jeu de test (matrice de confusion, courbe précision-rappel, importance des variables), extrait les règles de décision sous forme lisible, et exporte l'ensemble des figures associées.

```
root_folder <- dirname(rstudioapi::getActiveDocumentContext()$path)
setwd(root_folder)
Sys.setenv(TZ = "UTC")

library(readr)
library(lubridate)
library(rpart)
library(caret)
library(zoo)
library(ggplot2)
library(patchwork)
library(dplyr)

CONSIGNE   <- 1000
SEUIL_PROB <- 0.5
CP_MIN     <- 0.0005
DEPTH      <- 5      # profondeur fixée — produit 8 feuilles
DT_MIN     <- 15
PAS_1H     <- 60 / DT_MIN

# ── 1. Chargement ──────────────────────────────────────────────────

df_raw <- read_csv("data_qbase_qepp.csv",
```

```
                    col_types = cols(.default = col_double(), Date =
col_character())) %>%
  mutate(Date = ymd_hms(Date)) %>% arrange(Date)

# ── 2. Agrégation horaire ──────────────────────────────────────────

df_1h <- df_raw %>%
  mutate(h = floor_date(Date, "hour")) %>%
  group_by(h) %>%
  summarise(
    Qexona_mean = mean(Qexona_m3h, na.rm = TRUE),
    Pluie_cumul = sum(Pluie_mm,    na.rm = TRUE),
    n_obs       = n(),
    .groups     = "drop"
  ) %>%
  filter(n_obs == 4) %>%
  arrange(h) %>%
  mutate(
    Date           = h,
    Pluie_cumul_2h = rollapply(Pluie_cumul, 2, sum, align = "right", fill = NA),
    Pluie_cumul_6h = rollapply(Pluie_cumul, 6, sum, align = "right", fill = NA)
  ) %>%
  select(-h)

cat("Heures completes :", nrow(df_1h), "\n\n")

# ── 3. Variables dérivées ──────────────────────────────────────────
# Note clé : heure est conservé en entier (0-23), PAS en facteur.
# Cela permet à rpart de produire des seuils horaires lisibles (ex. heure < 6)
# au lieu de codes de niveaux cryptiques (a, b, c...).

```

```
df_1h <- df_1h %>%
  mutate(
    heure        = hour(Date),          # entier 0-23
    mois         = factor(month(Date)), # facteur pour capter la saisonnalité
    Qexona_lag1  = lag(Qexona_mean, 1),
    dQ_dt        = Qexona_mean - lag(Qexona_mean, 1),
    Qexona_futur = lead(Qexona_mean, 1),
    bypass       = factor(if_else(Qexona_futur > CONSIGNE, 1L, 0L),
                          levels = c(0L, 1L), labels = c("Non", "Bypass"))
  )

features <- c("Qexona_mean", "Qexona_lag1", "dQ_dt",
              "Pluie_cumul_2h", "Pluie_cumul_6h",
              "heure", "mois")

df_model <- df_1h %>%
  select(Date, all_of(features), Qexona_futur, bypass) %>%
  filter(complete.cases(.))

cat("Observations :", nrow(df_model), "| Taux bypass :",
    round(mean(df_model$bypass == "Bypass") * 100, 1), "%\n\n")

# ── 4. Split chronologique 80/20 ──────────────────────────────────────────

n_train  <- floor(0.8 * nrow(df_model))
df_train <- df_model[1:n_train, ]
df_test  <- df_model[(n_train + 1):nrow(df_model), ]

cat(sprintf("Train : %d obs | Test : %d obs\n\n", nrow(df_train), nrow(df_test)))

# ── 5. Arbre de décision — depth fixé à 5 (8 feuilles) ────────────────────
```

```
85
86  cat(sprintf("Entrainement arbre de decision (depth=%d)...\n\n", DEPTH))
87
88  set.seed(42)
89  a_tmp <- rpart(
90    bypass ~ .,
91    data    = df_train[, c(features, "bypass")],
92    method  = "class",
93    control = rpart.control(cp = CP_MIN, maxdepth = DEPTH, minsplit = 20),
94    parms   = list(split = "gini")
95  )
96
97  cp_opt     <- min(a_tmp$cptable[which.min(a_tmp$cptable[, "xerror"]), "CP"],
0.005)
98  arbre_opt  <- prune(a_tmp, cp = cp_opt)
99  n_feuilles <- sum(arbre_opt$frame$var == "<leaf>")
100
101 cat(sprintf("Feuilles : %d | cp optimal : %.5f\n\n", n_feuilles, cp_opt))
102
103 prob_opt  <- predict(arbre_opt, newdata = df_test[, features], type = "prob")[,
"Bypass"]
104 class_opt <- factor(if_else(prob_opt >= SEUIL_PROB, "Bypass", "Non"),
105                     levels = c("Non", "Bypass"))
106 cm_opt    <- confusionMatrix(class_opt, df_test$bypass, positive = "Bypass")
107
108 cat("=== METRIQUES (jeu de test) ===\n")
109 print(cm_opt$table); cat("\n")
110 cat(sprintf("Accuracy  : %.3f\n", cm_opt$overall["Accuracy"]))
111 cat(sprintf("Rappel    : %.3f\n", cm_opt$byClass["Recall"]))
112 cat(sprintf("Precision : %.3f\n", cm_opt$byClass["Precision"]))
113 cat(sprintf("F1-score  : %.3f\n", cm_opt$byClass["F1"]))
114 cat(sprintf("Kappa     : %.3f\n\n", cm_opt$overall["Kappa"]))
```

```
115
116 # Courbe precision-rappel
117 seuils_tmp <- seq(0.05, 0.95, by = 0.05)
118 pr_opt <- lapply(seuils_tmp, function(s) {
119   pc   <- factor(if_else(prob_opt >= s, "Bypass", "Non"), levels = c("Non",
"Bypass"))
120   cm_s <- confusionMatrix(pc, df_test$bypass, positive = "Bypass")
121   data.frame(seuil     = s,
122              rappel    = as.numeric(cm_s$byClass["Recall"]),
123              precision = as.numeric(cm_s$byClass["Precision"]),
124              f1        = as.numeric(cm_s$byClass["F1"]))
125 }) %>% bind_rows()
126
127 write.csv(pr_opt, "tableaux/5_courbe_pr_arbre_1h.csv", row.names = FALSE)
128
129 # ── 6. Importance des variables ──────────────────────────────────────
130
131 vi     <- arbre_opt$variable.importance
132 imp_df <- data.frame(
133   variable       = names(vi),
134   importance     = as.numeric(vi),
135   importance_pct = as.numeric(vi) / sum(vi) * 100,
136   stringsAsFactors = FALSE
137 ) %>% arrange(desc(importance_pct))
138
139 cat("=== IMPORTANCE DES VARIABLES ===\n")
140 print(imp_df[, c("variable", "importance_pct")], row.names = FALSE, digits = 3)
141 cat("\n")
142
143 # ── 7. Règles de l'arbre – extraction lisible ────────────────────────
144 # heure étant numérique, rpart produit des seuils réels (ex. heure >= 5.5)
```

```
# que l'on arrondit à l'heure entière la plus proche pour la lisibilité.

extraire_regles <- function(arbre, seuil_prob = 0.5) {

  frame    <- arbre$frame
  splits   <- arbre$splits
  feuilles <- sort(as.integer(rownames(frame[frame$var == "<leaf>", ])))

  # Construire un mapping noeud -> index dans splits
  # Les splits sont listés dans l'ordre des noeuds non-terminaux de frame
  noeuds_internes <- as.integer(rownames(frame[frame$var != "<leaf>", ]))
  # Chaque noeud interne consomme frame$ncompete + frame$nsurrogate + 1 lignes de splits
  split_debut <- integer(nrow(frame))
  names(split_debut) <- rownames(frame)
  curseur <- 1L
  for (nd_char in rownames(frame)) {
    fi <- frame[nd_char, ]
    if (fi$var != "<leaf>") {
      split_debut[nd_char] <- curseur
      curseur <- curseur + fi$ncompete + fi$nsurrogate + 1L
    }
  }

  lignes <- c(
    "================================================================",
    "  REGLES DE DECISION — BYPASS EXONA (horizon t+1h, pas 1h)",
    "================================================================",
    ""
  )

```

```
175   for (nd in feuilles) {
176
177     conditions <- character(0)
178     cur <- nd
179
180     while (cur > 1) {
181       parent   <- floor(cur / 2)
182       par_char <- as.character(parent)
183       var_name <- as.character(frame[par_char, "var"])
184       droite   <- (cur %% 2 == 1)
185
186       # Index exact du split primaire de ce noeud parent
187       idx   <- split_debut[par_char]
188       seuil <- splits[idx, "index"]
189
190       if (var_name == "Qexona_mean") {
191         cond <- if (droite)
192           sprintf("Qexona >= %.0f m3/h", seuil)
193         else
194           sprintf("Qexona <  %.0f m3/h", seuil)
195
196       } else if (var_name == "Qexona_lag1") {
197         cond <- if (droite)
198           sprintf("Qexona(t-1h) >= %.0f m3/h", seuil)
199         else
200           sprintf("Qexona(t-1h) <  %.0f m3/h", seuil)
201
202       } else if (var_name == "dQ_dt") {
203         cond <- if (droite)
204           sprintf("dQ/dt >= %.0f m3/h", seuil)
205         else
```

```
206             sprintf("dQ/dt <  %.0f m3/h", seuil)
207
208         } else if (var_name == "heure") {
209           h_seuil <- ceiling(seuil)
210           cond <- if (droite)
211             sprintf("heure  >= %02dh00", h_seuil)
212           else
213             sprintf("heure  <  %02dh00", h_seuil)
214
215         } else {
216           cond <- if (droite)
217             sprintf("%s >= %.2f", var_name, seuil)
218           else
219             sprintf("%s <  %.2f", var_name, seuil)
220         }
221
222         conditions <- c(cond, conditions)
223         cur <- parent
224       }
225
226       fi    <- frame[as.character(nd), ]
227       n_tot <- fi$n
228       n_byp <- fi$yval2[2]
229       p_byp <- round(n_byp / n_tot * 100, 1)
230
231       if (fi$yval == 2) {
232         entete <- sprintf("Regle %d — >>> BYPASS predit  (P=%.1f%%, n=%d)", nd,
p_byp, n_tot)
233       } else {
234         entete <- sprintf("Regle %d —     Non-bypass      (P=%.1f%%, n=%d)", nd,
p_byp, n_tot)
235       }
```

```
236
237     lignes <- c(lignes, entete,
238                 sprintf("  SI   %s", conditions[1]))
239     if (length(conditions) > 1) {
240       for (cond in conditions[-1]) {
241         lignes <- c(lignes, sprintf("  ET   %s", cond))
242       }
243     }
244     lignes <- c(lignes, "")
245   }
246
247   lignes <- c(lignes,
248               "----------------------------------------------------------------",
249               sprintf("Seuil de classification : P(bypass) >= %.2f", seuil_prob),
250               "----------------------------------------------------------------"
251   )
252   lignes
253 }
254
255 regles <- extraire_regles(arbre_opt, SEUIL_PROB)
256 cat(paste(regles, collapse = "\n"), "\n\n")
257
258 writeLines(regles, "tableaux/6_regles_arbre_1h.txt")
259 cat("Regles exportees : regles_arbre_1h.txt\n\n")
260
261 # ── 8. Figures ──────────────────────────────────────────────────────
262
263 dir.create("figures/5_modele_oracle", showWarnings = FALSE, recursive = TRUE)
264
265 theme_v <- theme_bw(base_size = 12) +
266   theme(panel.grid.minor = element_blank(),
```

```
        axis.text  = element_text(size = 12, colour = "black"),
        axis.title = element_text(size = 12, colour = "black"),
        plot.title = element_text(face = "bold", size = 14, colour = "black"))

# Arbre
png("figures/5_modele_oracle/arbre_1h_simple.png",
    width = 3200, height = 2000, res = 350, bg = "white")
par(mar = c(1, 1, 3, 1))
plot(arbre_opt, uniform = TRUE, compress = TRUE, margin = 0.1,
     main = paste0("Arbre de décision"))
text(arbre_opt, use.n = TRUE, cex = 0.7)
dev.off()
cat("Figure : arbre_1h_simple.png\n")

# Matrice de confusion
cm_df <- as.data.frame(cm_opt$table)
colnames(cm_df) <- c("Predit", "Reel", "n")
cm_df <- cm_df %>% group_by(Reel) %>% mutate(pct = n / sum(n) * 100)
p_cm <- ggplot(cm_df, aes(x = Predit, y = Reel, fill = pct)) +
  geom_tile(color = "white", linewidth = 1) +
  geom_text(aes(label = sprintf("%d\n(%.1f%%)", n, pct)), size = 5) +
  scale_fill_gradient(low = "#f7f7f7", high = "#4393c3", name = "% par classe") +
  labs(title = "Matrice de confusion",
       x = "Classe prédite", y = "Classe réelle") + theme_v
ggsave("figures/5_modele_oracle/matrice_confusion_arbre_1h.png",
       p_cm, width = 6, height = 5, dpi = 350, bg = "white")
cat("Figure : matrice_confusion_arbre_1h.png\n")

# Courbe precision-rappel
pr_plot <- pr_opt %>%
```

```
  filter(!is.na(rappel), !is.na(precision), is.finite(rappel), is.finite(precision))

p_pr <- ggplot(pr_plot, aes(x = rappel, y = precision)) +
  geom_path(colour = "#4393c3", linewidth = 1) +
  geom_point(data = pr_plot %>% filter(abs(seuil - SEUIL_PROB) < 0.03),
             shape = 21, size = 4, fill = "orange", colour = "black") +
  scale_x_continuous(labels = scales::percent_format()) +
  scale_y_continuous(labels = scales::percent_format()) +
  labs(title = "Courbe precision-rappel",
       subtitle = sprintf("Seuil actuel : %.2f (cercle orange)", SEUIL_PROB),
       x = "Rappel", y = "Précision") + theme_v
ggsave("figures/5_modele_oracle/courbe_pr_arbre_1h.png",
       p_pr, width = 8, height = 6, dpi = 350, bg = "white")
cat("Figure : courbe_pr_arbre_1h.png\n")

# Importance des variables
p_imp <- ggplot(imp_df, aes(x = reorder(variable, importance_pct), y = importance_pct)) +
  geom_col(fill = "#4393c3", alpha = 0.85) +
  geom_text(aes(label = sprintf("%.1f%%", importance_pct)), hjust = -0.1, size = 4) +
  coord_flip() +
  ylim(c(0, 55)) +
  labs(title = "Importance des variables",
       x = NULL, y = "Importance (%)") + theme_v
ggsave("figures/5_modele_oracle/importance_arbre_1h.png",
       p_imp, width = 8, height = 5, dpi = 350, bg = "white")
cat("Figure : importance_arbre_1h.png\n")

# Serie temporelle
df_tp <- df_test %>% mutate(prob_bypass = prob_opt)
```

```
p_deb <- ggplot(df_tp, aes(x = Date)) +
  geom_line(aes(y = Qexona_mean), color = "#2166ac", linewidth = 0.4, alpha = 0.8) +
  geom_point(data = df_tp %>% filter(bypass == "Bypass"),
             aes(y = Qexona_futur), color = "#d6604d", shape = 16, size = 1, alpha = 0.6) +
  geom_hline(yintercept = CONSIGNE, linetype = "dashed", colour = "red", linewidth = 0.5) +
  scale_x_datetime(date_breaks = "1 month", date_labels = "%b") +
  labs(title = "Qexona (bleu) | bypass reels (rouge) | consigne (tiret)",
       x = NULL, y = "m3/h") + theme_v
p_pro <- ggplot(df_tp, aes(x = Date, y = prob_bypass)) +
  geom_line(colour = "#d6604d", linewidth = 0.4) +
  geom_hline(yintercept = SEUIL_PROB, linetype = "dashed", colour = "grey30") +
  scale_x_datetime(date_breaks = "1 month", date_labels = "%b") +
  labs(title = sprintf("P(bypass) seuil %.2f", SEUIL_PROB),
       x = NULL, y = "P(bypass)") + theme_v
ggsave("figures/5_modele_oracle/serie_temporelle_arbre_1h.png",
       p_deb / p_pro + plot_layout(heights = c(2, 1)),
       width = 14, height = 8, dpi = 350, bg = "white")
cat("Figure : serie_temporelle_arbre_1h.png\n")

cat("\nTermine.\n")

```

Annexe 4 : Liste des épisodes pluvieux identifiés en 2024

N°	Saison	Début	Fin	Durée (h)	Pluie cumulée (mm)	Int. moy. (mm/h)	Int. max. (mm/h)
1	Hiver	25/01/2024 13:15	25/01/2024 15:30	2.5	0.8	0.32	0
2	Hiver	25/01/2024 22:30	26/01/2024 13:15	15	4.6	0.31	4.8
3	Hiver	01/02/2024 06:45	01/02/2024 08:00	1.5	1.6	1.07	4.8
4	Hiver	07/02/2024 06:45	07/02/2024 22:15	15.75	3.6	0.23	1.2
5	Hiver	08/02/2024 05:00	09/02/2024 00:15	19.5	6.2	0.32	3.6
6	Hiver	10/02/2024 10:15	10/02/2024 19:15	9.25	8.8	0.95	10.4
7	Hiver	11/02/2024 11:45	11/02/2024 15:00	3.5	0.4	0.11	0
8	Hiver	12/02/2024 17:15	13/02/2024 01:15	8.25	1.8	0.22	15.6
9	Hiver	16/02/2024 12:00	16/02/2024 12:30	0.75	0.4	0.53	0
10	Hiver	18/02/2024 11:00	18/02/2024 17:15	6.5	6.2	0.95	11.8
11	Hiver	21/02/2024 15:00	22/02/2024 17:45	27	25.8	0.96	105.6
12	Hiver	23/02/2024 14:00	23/02/2024 18:45	5	6.6	1.32	25.4
13	Hiver	24/02/2024 17:45	24/02/2024 18:45	1.25	5.2	4.16	45.8

14	Hiver	25/02/2024 16:45	26/02/2024 22:45	30.25	15.6	0.52	13.6
15	Hiver	29/02/2024 10:45	29/02/2024 15:00	4.5	1	0.22	0
16	Hiver	29/02/2024 23:15	29/02/2024 23:45	0.75	0.6	0.8	4.6
17	Printemps	01/03/2024 20:00	01/03/2024 22:00	2.25	6	2.67	13
18	Printemps	02/03/2024 10:00	03/03/2024 05:00	19.25	11.8	0.61	8.4
19	Printemps	05/03/2024 13:30	05/03/2024 15:45	2.5	2.8	1.12	22.4
20	Printemps	05/03/2024 22:30	06/03/2024 09:30	11.25	1.8	0.16	1.6
21	Printemps	09/03/2024 16:45	09/03/2024 23:15	6.75	10.4	1.54	8
22	Printemps	11/03/2024 15:15	11/03/2024 15:30	0.5	0.4	0.8	1.4
23	Printemps	12/03/2024 11:15	12/03/2024 22:00	11	5.4	0.49	2
24	Printemps	15/03/2024 10:30	15/03/2024 13:45	3.5	1.4	0.4	11.2
25	Printemps	17/03/2024 06:00	17/03/2024 11:30	5.75	2.2	0.38	2.4
26	Printemps	20/03/2024 23:45	21/03/2024 00:00	0.5	0.4	0.8	0.8
27	Printemps	22/03/2024 23:00	23/03/2024 00:45	2	3.6	1.8	7
28	Printemps	23/03/2024 13:30	23/03/2024 19:00	5.75	5.4	0.94	31.6

29	Printemps	26/03/2024 06:15	26/03/2024 19:45	13.75	5.6	0.41	3.6
30	Printemps	28/03/2024 01:30	28/03/2024 04:00	2.75	3.2	1.16	5
31	Printemps	29/03/2024 02:15	29/03/2024 07:15	5.25	0.6	0.11	0
32	Printemps	29/03/2024 18:15	30/03/2024 08:00	14	19.4	1.39	11.2
33	Printemps	31/03/2024 15:30	31/03/2024 16:00	0.75	4.4	5.87	88.6
34	Printemps	01/04/2024 13:45	01/04/2024 16:45	3.25	1.6	0.49	29
35	Printemps	02/04/2024 18:00	03/04/2024 10:30	16.75	9	0.54	3.2
36	Printemps	04/04/2024 05:45	04/04/2024 09:30	4	7	1.75	11.6
37	Printemps	08/04/2024 00:15	08/04/2024 02:15	2.25	3	1.33	3.8
38	Printemps	08/04/2024 09:45	08/04/2024 10:00	0.5	1.2	2.4	3.6
39	Printemps	09/04/2024 00:00	09/04/2024 01:15	1.5	3	2	8.2
40	Printemps	09/04/2024 14:30	09/04/2024 16:45	2.5	3.4	1.36	33.2
41	Printemps	15/04/2024 17:15	15/04/2024 23:45	6.75	3	0.44	26.4
42	Printemps	16/04/2024 12:30	16/04/2024 19:00	6.75	5.4	0.8	37.2
43	Printemps	17/04/2024 15:45	17/04/2024 16:00	0.5	1.4	2.8	7.2

44	Printemps	18/04/2024 04:15	18/04/2024 06:00	2	1.4	0.7	5.2
45	Printemps	19/04/2024 19:45	19/04/2024 22:00	2.5	2.2	0.88	16.8
46	Printemps	24/04/2024 04:45	24/04/2024 05:30	1	0.6	0.6	2.4
47	Printemps	27/04/2024 00:00	27/04/2024 06:45	7	2.2	0.31	4.4
48	Printemps	27/04/2024 14:00	27/04/2024 23:45	10	8.4	0.84	5.8
49	Printemps	30/04/2024 07:30	30/04/2024 09:00	1.75	0.8	0.46	0
50	Printemps	01/05/2024 22:15	02/05/2024 00:00	2	5.2	2.6	6.6
51	Printemps	02/05/2024 15:45	02/05/2024 20:45	5.25	6.8	1.3	4
52	Printemps	04/05/2024 06:45	04/05/2024 09:30	3	0.8	0.27	0
53	Printemps	05/05/2024 07:00	05/05/2024 09:30	2.75	1	0.36	2.2
54	Printemps	05/05/2024 16:00	05/05/2024 17:45	2	0.8	0.4	0
55	Printemps	06/05/2024 01:45	06/05/2024 03:00	1.5	3	2	25.4
56	Printemps	14/05/2024 01:15	14/05/2024 19:15	18.25	14.6	0.8	14
57	Printemps	15/05/2024 17:45	15/05/2024 20:00	2.5	11	4.4	89.4
58	Printemps	16/05/2024 02:15	16/05/2024 06:30	4.5	0.4	0.09	0

59	Printemps	18/05/2024 17:30	18/05/2024 18:45	1.5	9.8	6.53	53.4
60	Printemps	19/05/2024 13:15	19/05/2024 13:30	0.5	1.8	3.6	15.8
61	Printemps	21/05/2024 17:15	21/05/2024 18:15	1.25	1.8	1.44	15.2
62	Printemps	22/05/2024 21:30	22/05/2024 21:45	0.5	0.6	1.2	3
63	Printemps	23/05/2024 17:00	23/05/2024 18:00	1.25	2.2	1.76	3.8
64	Printemps	24/05/2024 18:15	25/05/2024 03:00	9	15	1.67	192
65	Printemps	25/05/2024 21:30	25/05/2024 21:45	0.5	0.8	1.6	13.6
66	Printemps	26/05/2024 21:30	26/05/2024 22:15	1	1	1	6.6
67	Printemps	28/05/2024 12:45	28/05/2024 19:00	6.5	1.8	0.28	11.4
68	Printemps	29/05/2024 05:15	29/05/2024 21:15	16.25	22	1.35	11.4
69	Printemps	30/05/2024 13:45	30/05/2024 20:15	6.75	1.6	0.24	30
70	Printemps	31/05/2024 14:45	01/06/2024 03:45	13.25	7.2	0.54	33.8
71	Été	01/06/2024 14:30	01/06/2024 19:30	5.25	1.2	0.23	1
72	Été	10/06/2024 15:30	10/06/2024 16:00	0.75	3.4	4.53	30.6
73	Été	11/06/2024 17:15	11/06/2024 18:45	1.75	2.8	1.6	17.4

74	Été	14/06/2024 03:15	14/06/2024 06:45	3.75	0.6	0.16	0.8
75	Été	16/06/2024 06:30	16/06/2024 10:30	4.25	5.6	1.32	6.4
76	Été	16/06/2024 17:15	17/06/2024 01:30	8.5	9	1.06	103.8
77	Été	17/06/2024 19:30	17/06/2024 19:45	0.5	2.8	5.6	30
78	Été	18/06/2024 07:00	18/06/2024 10:30	3.75	9	2.4	15.4
79	Été	21/06/2024 07:30	21/06/2024 14:00	6.75	6.6	0.98	7.2
80	Été	22/06/2024 00:00	22/06/2024 13:00	13.25	5	0.38	42
81	Été	29/06/2024 18:45	29/06/2024 23:00	4.5	1.4	0.31	1.2
82	Été	30/06/2024 06:30	30/06/2024 07:15	1	0.6	0.6	0
83	Été	01/07/2024 03:00	01/07/2024 05:00	2.25	2.2	0.98	28.2
84	Été	04/07/2024 03:15	04/07/2024 06:45	3.75	3	0.8	39.4
85	Été	09/07/2024 15:45	09/07/2024 16:30	1	0.8	0.8	4.6
86	Été	11/07/2024 23:45	12/07/2024 08:45	9.25	11.8	1.28	48.2
87	Été	13/07/2024 21:45	14/07/2024 01:30	4	1.4	0.35	37.2
88	Été	16/07/2024 03:30	16/07/2024 03:45	0.5	1	2	6.2

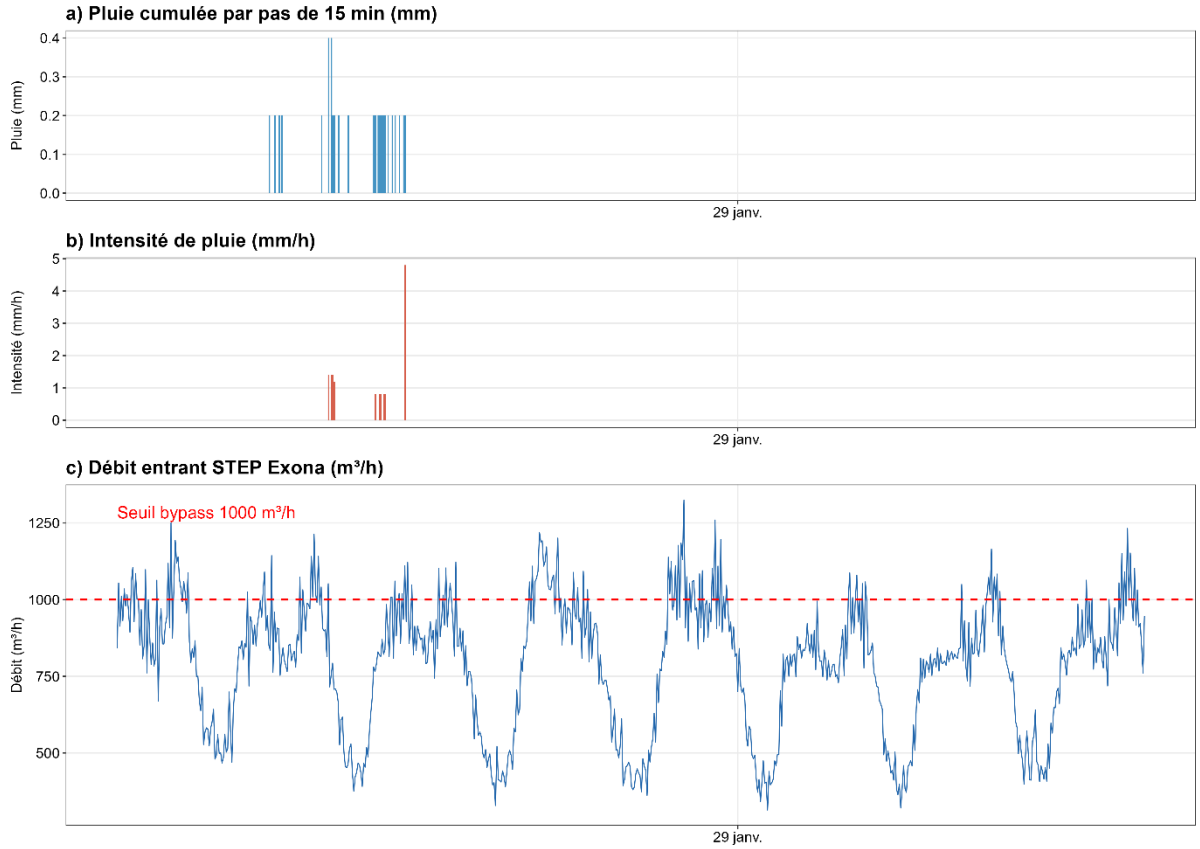
89	Été	21/07/2024 01:15	21/07/2024 04:00	3	2	0.67	10.6
90	Été	23/07/2024 05:30	23/07/2024 09:15	4	2.4	0.6	10
91	Été	26/07/2024 03:45	26/07/2024 11:45	8.25	2.4	0.29	4.2
92	Été	26/07/2024 18:15	27/07/2024 20:15	26.25	30.6	1.17	21.2
93	Été	31/07/2024 04:00	31/07/2024 09:45	6	6.4	1.07	4.6
94	Été	01/08/2024 06:30	01/08/2024 07:00	0.75	3.4	4.53	52.2
95	Été	01/08/2024 18:00	02/08/2024 04:00	10.25	19.6	1.91	45.4
96	Été	14/08/2024 06:45	14/08/2024 07:00	0.5	0.6	1.2	8
97	Été	17/08/2024 08:45	17/08/2024 23:45	15.25	33	2.16	73.8
98	Été	20/08/2024 18:00	20/08/2024 20:00	2.25	5.2	2.31	54
99	Été	24/08/2024 18:30	24/08/2024 19:45	1.5	1.8	1.2	57.8
100	Été	30/08/2024 07:45	30/08/2024 18:00	10.5	12.8	1.22	85.4
101	Automne	05/09/2024 05:30	05/09/2024 14:45	9.5	22.2	2.34	27.8
102	Automne	07/09/2024 03:30	07/09/2024 04:30	1.25	2.4	1.92	58.8
103	Automne	07/09/2024 12:15	07/09/2024 17:30	5.5	0.8	0.15	1.2

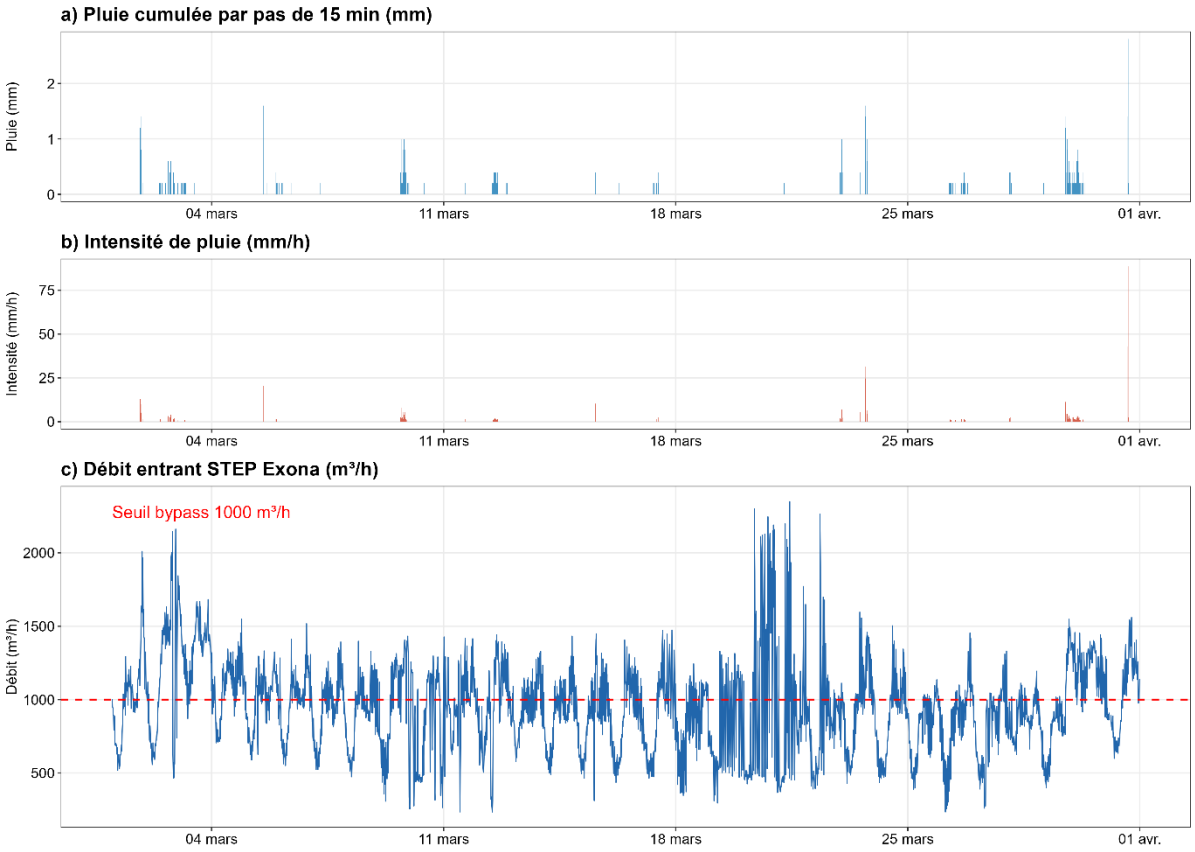
104	Automne	08/09/2024 16:00	09/09/2024 02:00	10.25	6.4	0.62	11.2
105	Automne	09/09/2024 08:45	09/09/2024 11:45	3.25	0.4	0.12	0
106	Automne	11/09/2024 02:15	11/09/2024 07:15	5.25	3	0.57	12.8
107	Automne	12/09/2024 13:30	12/09/2024 19:45	6.5	4.6	0.71	73
108	Automne	13/09/2024 18:45	13/09/2024 19:00	0.5	1.4	2.8	11
109	Automne	21/09/2024 00:00	21/09/2024 05:30	5.75	6.2	1.08	54.4
110	Automne	22/09/2024 01:00	22/09/2024 04:15	3.5	2.2	0.63	9
111	Automne	22/09/2024 20:00	22/09/2024 22:00	2.25	6.4	2.84	33.4
112	Automne	24/09/2024 04:15	24/09/2024 09:15	5.25	1.6	0.3	8.4
113	Automne	25/09/2024 09:45	26/09/2024 03:15	17.75	30.8	1.74	78.4
114	Automne	26/09/2024 10:30	26/09/2024 17:15	7	27.8	3.97	121.2
115	Automne	27/09/2024 19:00	28/09/2024 00:00	5.25	1	0.19	7.4
116	Automne	30/09/2024 11:45	30/09/2024 13:30	2	0.6	0.3	1.8
117	Automne	30/09/2024 20:45	30/09/2024 21:45	1.25	1.8	1.44	5.4
118	Automne	02/10/2024 12:30	02/10/2024 15:00	2.75	1.6	0.58	4.2

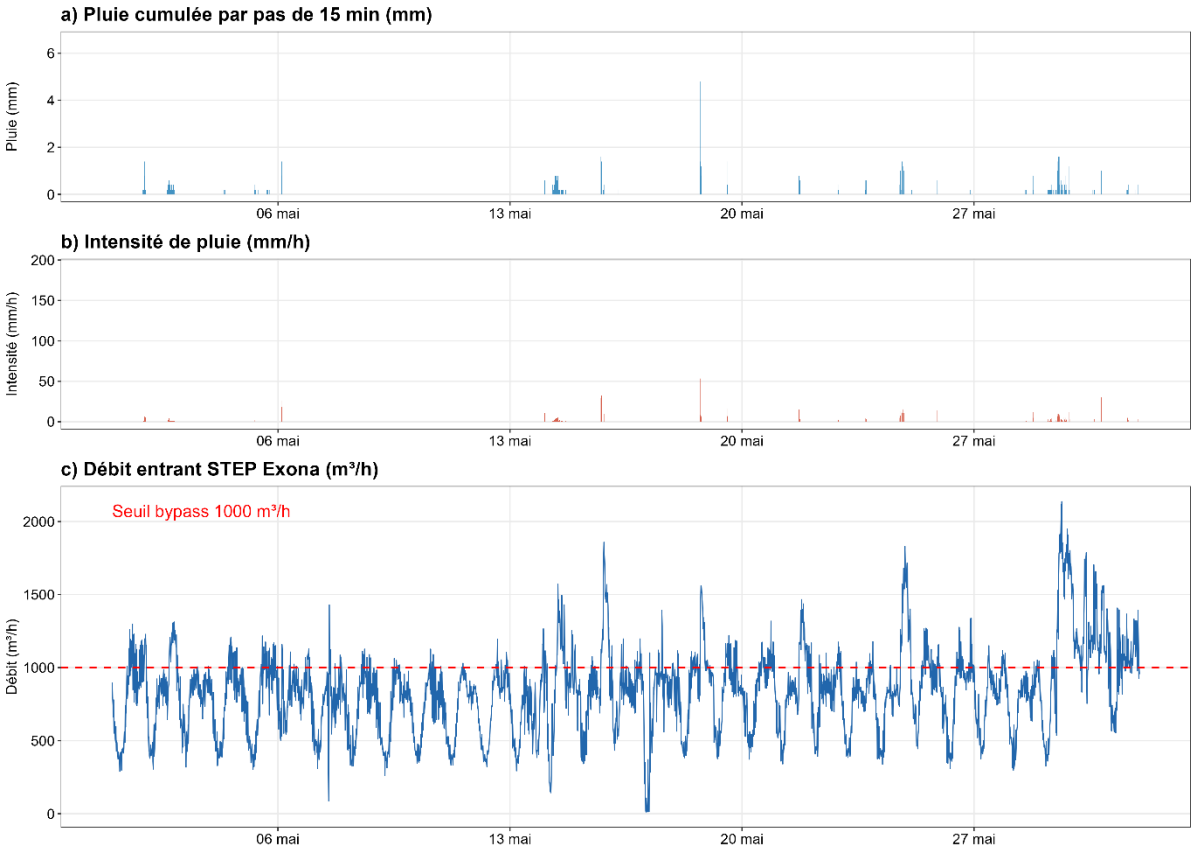
119	Automne	06/10/2024 10:30	06/10/2024 18:30	8.25	2	0.24	3.4
120	Automne	07/10/2024 12:45	07/10/2024 22:30	10	11.8	1.18	61
121	Automne	08/10/2024 19:30	09/10/2024 23:45	28.5	46.8	1.64	19
122	Automne	10/10/2024 06:30	10/10/2024 13:00	6.75	2.4	0.36	4
123	Automne	16/10/2024 21:30	17/10/2024 00:30	3.25	6.8	2.09	122.6
124	Automne	17/10/2024 11:00	17/10/2024 17:15	6.5	1.6	0.25	17.2
125	Automne	18/10/2024 00:30	18/10/2024 04:30	4.25	1.4	0.33	0.8
126	Automne	19/10/2024 11:45	19/10/2024 12:00	0.5	0.4	0.8	1
127	Automne	19/10/2024 18:30	19/10/2024 19:00	0.75	1	1.33	2.6
128	Automne	25/10/2024 18:30	26/10/2024 02:45	8.5	1.4	0.16	1
129	Automne	18/11/2024 12:45	19/11/2024 02:15	13.75	6	0.44	1.8
130	Automne	19/11/2024 11:15	19/11/2024 21:15	10.25	21.8	2.13	46
131	Automne	21/11/2024 11:15	21/11/2024 23:00	12	9.2	0.77	2.2
132	Automne	22/11/2024 09:30	22/11/2024 11:30	2.25	1.4	0.62	0.8
133	Automne	25/11/2024 08:45	25/11/2024 14:15	5.75	9.6	1.67	5.4

134	Automne	27/11/2024 08:30	27/11/2024 19:45	11.5	3.2	0.28	20
135	Hiver	05/12/2024 08:45	05/12/2024 13:30	5	4.4	0.88	4.6
136	Hiver	05/12/2024 21:45	06/12/2024 01:15	3.75	6.4	1.71	64.8
137	Hiver	07/12/2024 03:00	07/12/2024 06:45	4	7.4	1.85	5.8
138	Hiver	07/12/2024 15:00	08/12/2024 02:45	12	3.4	0.28	3.2
139	Hiver	08/12/2024 09:45	08/12/2024 18:00	8.5	4	0.47	5
140	Hiver	17/12/2024 05:30	17/12/2024 08:00	2.75	0.4	0.15	0
141	Hiver	18/12/2024 22:45	19/12/2024 09:15	10.75	4	0.37	19.8
142	Hiver	21/12/2024 00:00	21/12/2024 04:30	4.75	0.6	0.13	0.8
143	Hiver	21/12/2024 15:15	22/12/2024 16:30	25.5	8.4	0.33	131
144	Hiver	24/12/2024 03:00	24/12/2024 07:15	4.5	0.8	0.18	0.8

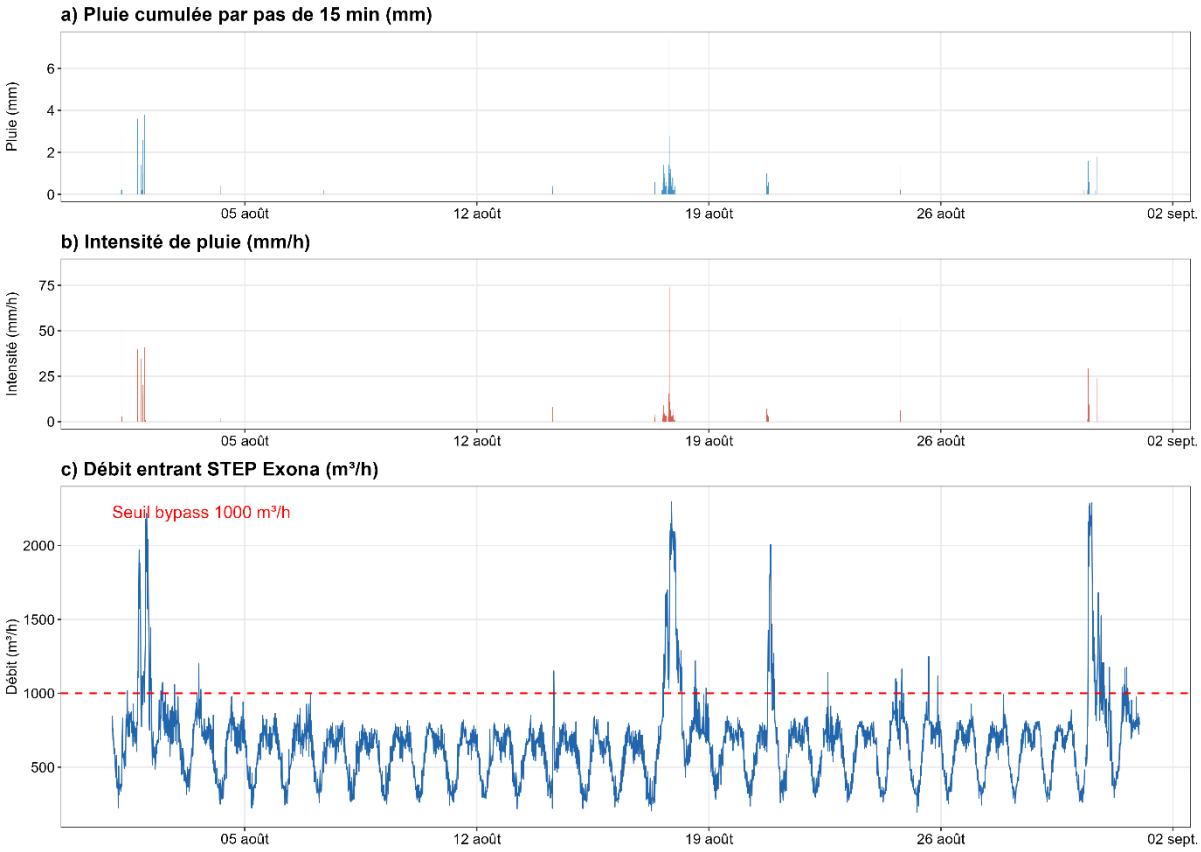
Annexe 5 : Graphiques de l'évaluation temporelle de la relation pluie-débit entrant à la station Exona pour les mois de janvier, mars, mai, juin, août, septembre, novembre, décembre

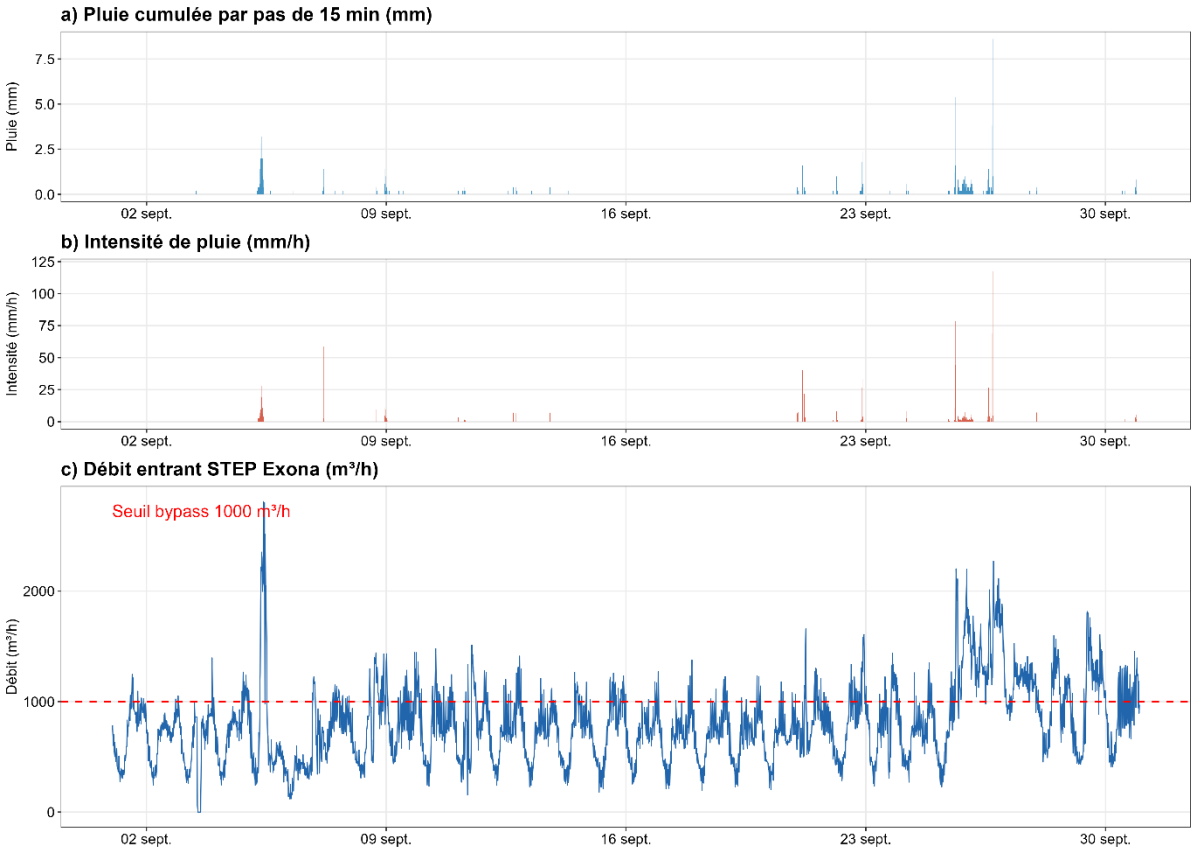


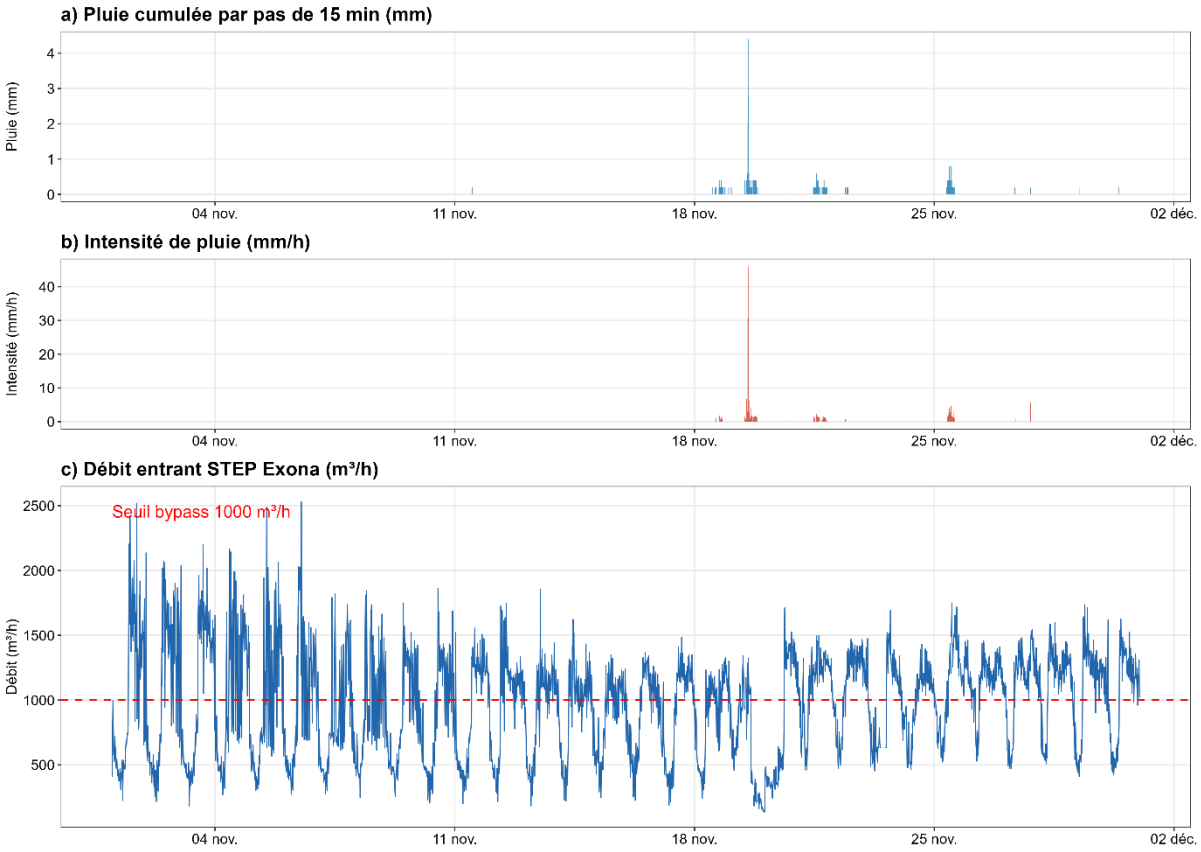

a) Pluie cumulée par pas de 15 min (mm)
Pluie (mm)
2
1
0
04 mars
11 mars
18 mars
25 mars
01 avr.
b) Intensité de pluie (mm/h)
Intensité (mm/h)
75
50
25
0
04 mars
11 mars
18 mars
25 mars
01 avr.
c) Débit entrant STEP Exona (m³/h)
Seuil bypass 1000 m³/h
Débit (m³/h)
2000
1500
1000
500
04 mars
11 mars
18 mars
25 mars
01 avr.


a) Pluie cumulée par pas de 15 min (mm)
Pluie (mm)
0
2
4
6
06 mai
13 mai
20 mai
27 mai
b) Intensité de pluie (mm/h)
Intensité (mm/h)
0
50
100
150
200
06 mai
13 mai
20 mai
27 mai
c) Débit entrant STEP Exona (m³/h)
Seuil bypass 1000 m³/h
Débit (m³/h)
0
500
1000
1500
2000
06 mai
13 mai
20 mai
27 mai


a) Pluie cumulée par pas de 15 min (mm)
Pluie (mm)
0
1
2
3
4
03 juin
10 juin
17 juin
24 juin
01 juil.
b) Intensité de pluie (mm/h)
Intensité (mm/h)
0
25
50
75
100
03 juin
10 juin
17 juin
24 juin
01 juil.
c) Débit entrant STEP Exona (m³/h)
Seuil bypass 1000 m³/h
Débit (m³/h)
0
500
1000
1500
2000
03 juin
10 juin
17 juin
24 juin
01 juil.

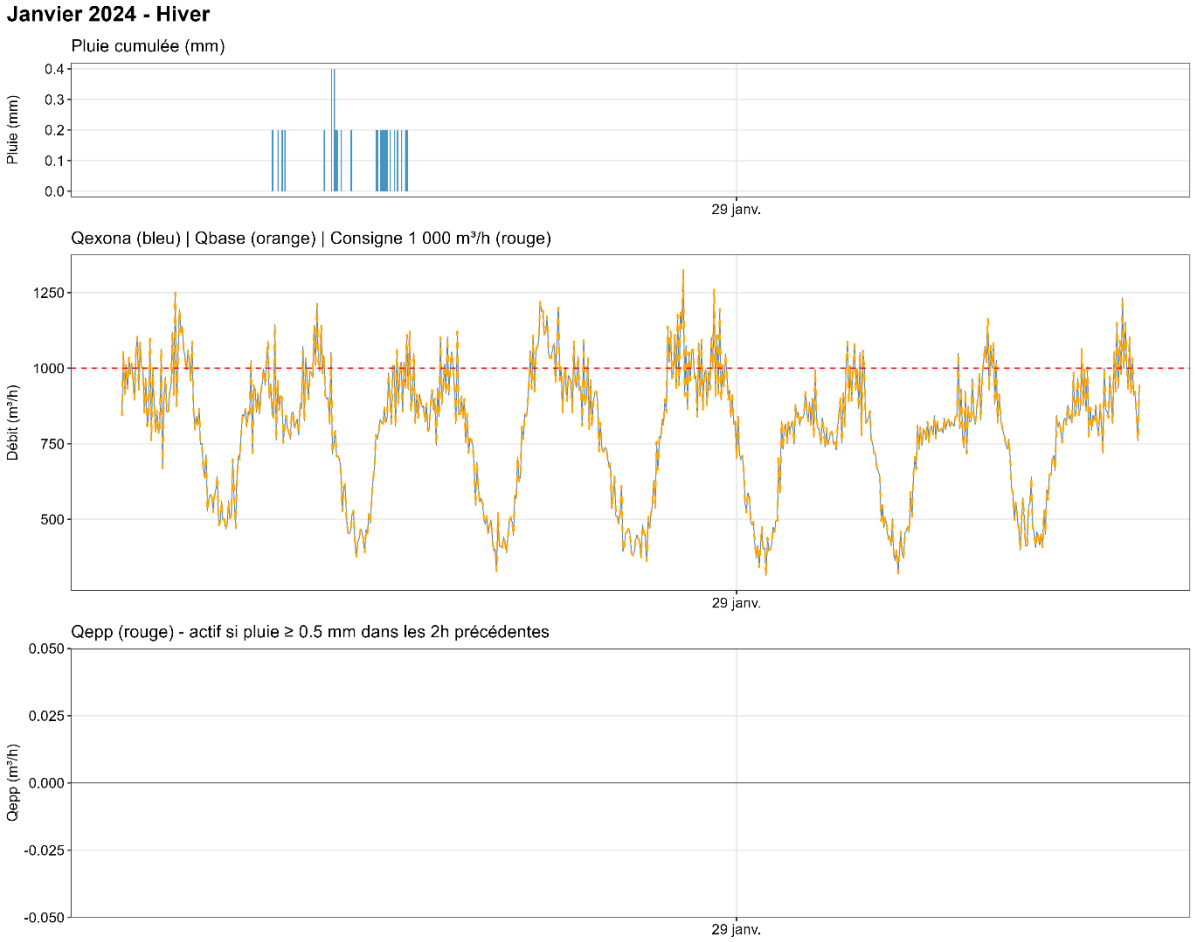

a) Pluie cumulée par pas de 15 min (mm)
Pluie (mm)
0
2
4
6
05 août
12 août
19 août
26 août
02 sept.
b) Intensité de pluie (mm/h)
Intensité (mm/h)
0
25
50
75
05 août
12 août
19 août
26 août
02 sept.
c) Débit entrant STEP Exona (m³/h)
Seuil bypass 1000 m³/h
Débit (m³/h)
500
1000
1500
2000
05 août
12 août
19 août
26 août
02 sept.

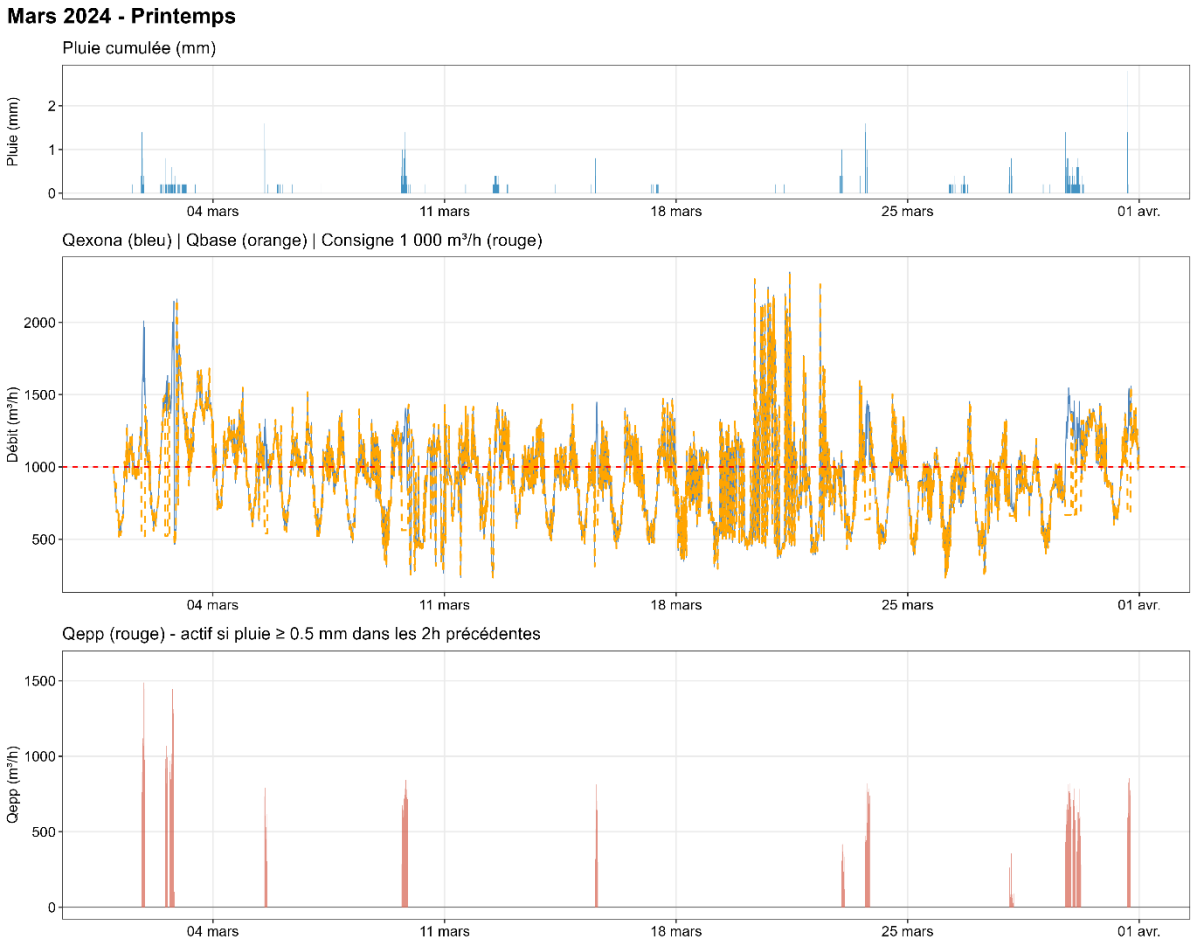

a) Pluie cumulée par pas de 15 min (mm)
Pluie (mm)
7.5
5.0
2.5
0.0
02 sept.
09 sept.
16 sept.
23 sept.
30 sept.
b) Intensité de pluie (mm/h)
Intensité (mm/h)
125
100
75
50
25
0
02 sept.
09 sept.
16 sept.
23 sept.
30 sept.
c) Débit entrant STEP Exona (m³/h)
Seuil bypass 1000 m³/h
Débit (m³/h)
2000
1000
0
02 sept.
09 sept.
16 sept.
23 sept.
30 sept.


a) Pluie cumulée par pas de 15 min (mm)
Pluie (mm)
0
1
2
3
4
04 nov.
11 nov.
18 nov.
25 nov.
02 déc.
b) Intensité de pluie (mm/h)
Intensité (mm/h)
0
10
20
30
40
04 nov.
11 nov.
18 nov.
25 nov.
02 déc.
c) Débit entrant STEP Exona (m³/h)
Seuil bypass 1000 m³/h
Débit (m³/h)
500
1000
1500
2000
2500
04 nov.
11 nov.
18 nov.
25 nov.
02 déc.


a) Pluie cumulée par pas de 15 min (mm)
Pluie (mm)
3
2
1
0
02 déc.
09 déc.
16 déc.
23 déc.
30 déc.
b) Intensité de pluie (mm/h)
Intensité (mm/h)
100
50
0
02 déc.
09 déc.
16 déc.
23 déc.
30 déc.
c) Débit entrant STEP Exona (m³/h)
Seuil bypass 1000 m³/h
Débit (m³/h)
2000
1500
1000
500
0
02 déc.
09 déc.
16 déc.
23 déc.
30 déc.

Annexe 6 : Graphiques de l'évaluation temporelle de la séparation du débit de base (Qbase) et des eaux parasites pluviales (Qepp) à la station d'Exona pour les mois de janvier, mars, mai, juin, août, septembre, novembre, décembre




Mars 2024 - Printemps
Pluie cumulée (mm)
Pluie (mm)
0
1
2
04 mars
11 mars
18 mars
25 mars
01 avr.
Qexona (bleu) | Qbase (orange) | Consigne 1 000 m³/h (rouge)
Débit (m³/h)
500
1000
1500
2000
04 mars
11 mars
18 mars
25 mars
01 avr.
Qepp (rouge) - actif si pluie ≥ 0.5 mm dans les 2h précédentes
Qepp (m³/h)
0
500
1000
1500
04 mars
11 mars
18 mars
25 mars
01 avr.

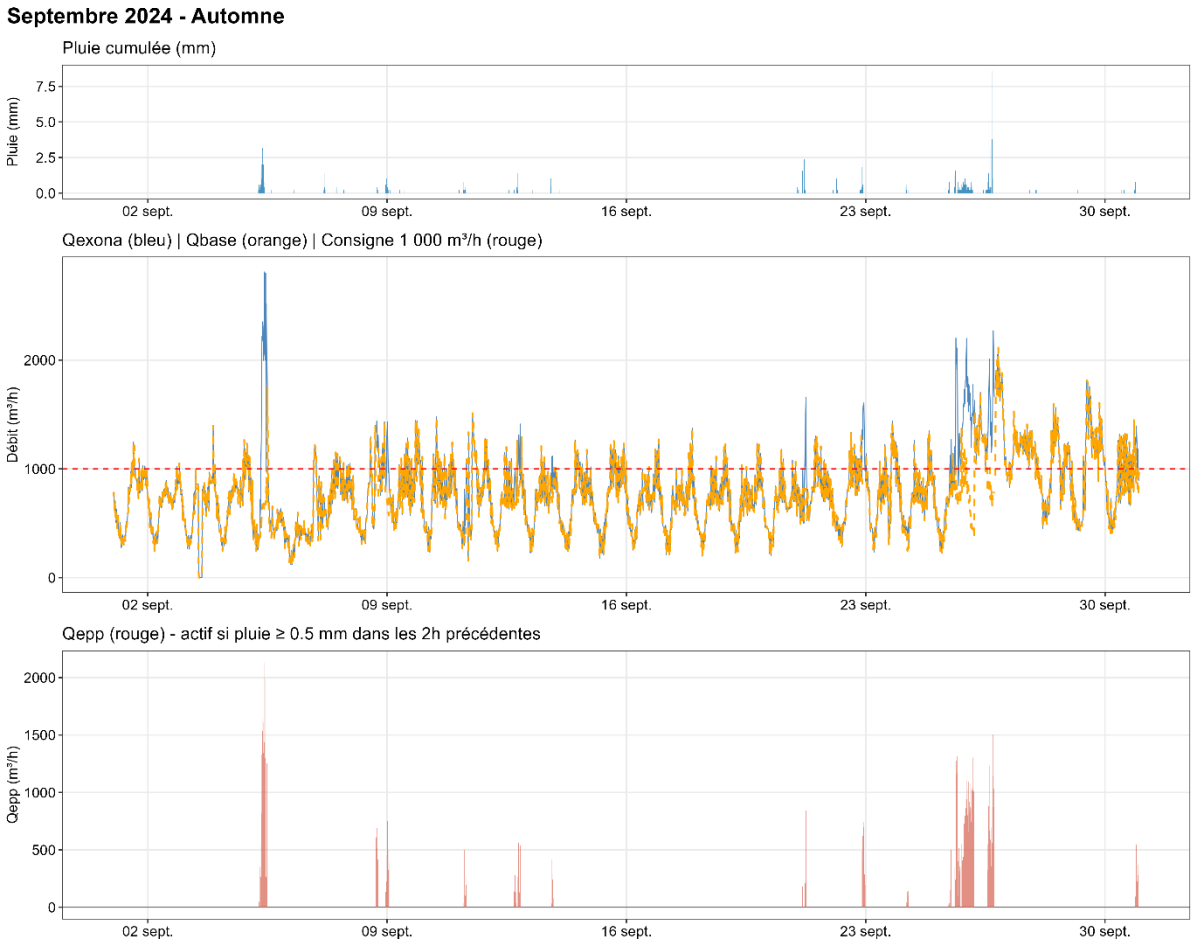

Mai 2024 - Printemps
Pluie cumulée (mm)
Pluie (mm)
0
2
4
6
06 mai
13 mai
20 mai
27 mai
Qexona (bleu) | Qbase (orange) | Consigne 1 000 m³/h (rouge)
Débit (m³/h)
0
500
1000
1500
2000
06 mai
13 mai
20 mai
27 mai
Qepp (rouge) - actif si pluie ≥ 0.5 mm dans les 2h précédentes
Qepp (m³/h)
0
250
500
750
1000
1250
06 mai
13 mai
20 mai
27 mai

Juin 2024 - Été

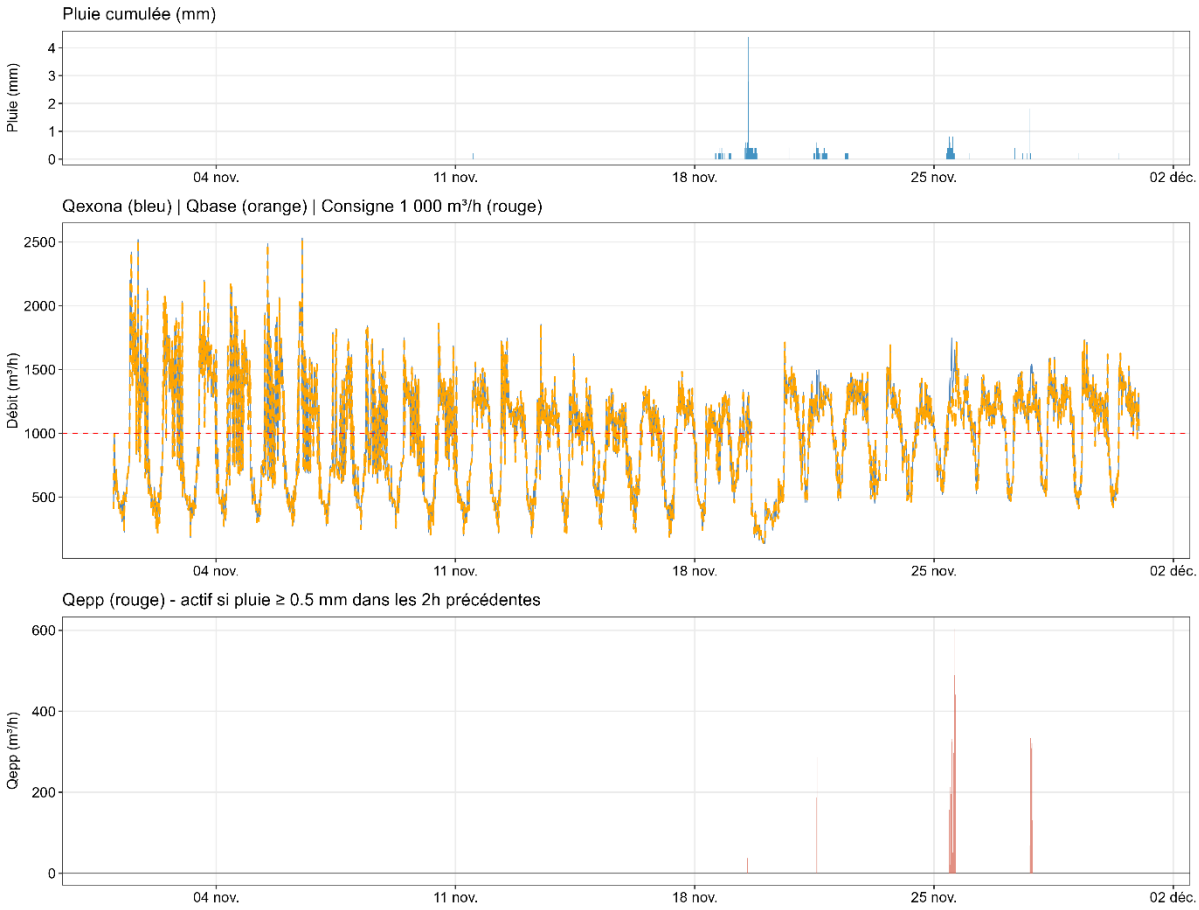


Août 2024 - Été




Septembre 2024 - Automne
Pluie cumulée (mm)
Pluie (mm)
7.5
5.0
2.5
0.0
02 sept.
09 sept.
16 sept.
23 sept.
30 sept.
Qexona (bleu) | Qbase (orange) | Consigne 1 000 m³/h (rouge)
Débit (m³/h)
2000
1000
0
02 sept.
09 sept.
16 sept.
23 sept.
30 sept.
Qepp (rouge) - actif si pluie ≥ 0.5 mm dans les 2h précédentes
Qepp (m³/h)
2000
1500
1000
500
0
02 sept.
09 sept.
16 sept.
23 sept.
30 sept.

Novembre 2024 - Automne



Décembre 2024 - Hiver

