



**OUTILS AUTOMATIQUES D'ÉVALUATION DE LA  
QUALITÉ DES DONNÉES  
SUIVI EN TEMPS RÉEL DE LA QUALITÉ DE L'EAU D'UN  
RUISSEAU URBAIN**

**Mémoire**

PASCAL POIRIER

Maîtrise en Génie des eaux  
Maître en science (M.Sc.)

Québec, Canada  
© Pascal Poirier, 2015



## Résumé

Lors d'un suivi en temps réel de la qualité de l'eau d'une rivière, une quantité innombrable de données peut être récoltée permettant ainsi une bonne description des dynamiques changeant rapidement. Cependant, des questions s'imposent. Quelle est la qualité des données mesurées en temps réel? Quels sont les effets de données de mauvaise qualité sur les objectifs associés à leur utilisation? Ce mémoire tente de résoudre cette problématique en présentant des outils automatiques d'évaluation de la qualité des données. Ces outils sont généraux et simples, en ce sens qu'ils s'appliquent à différents emplacements (rivière, STEP, égout, etc.) et qu'ils sont basés sur des méthodes univariées faciles à implanter. Suite à l'application de ces outils, la qualité des données est connue et accrue favorisant ainsi l'atteinte des objectifs reliés à un suivi en temps réel. Leur application est démontrée sur des données mesurées lors d'un suivi en temps réel de la qualité de l'eau d'un ruisseau urbain effectué à l'aide de la station *monEAU*.



## **Abstract**

During real-time water quality monitoring in river systems, a huge amount of data is produced which allows a better description of the rapidly changing dynamics compared to spot sampling methods. However, what is the quality of the measured data and what are the impacts of poor data quality on the objectives pursued during real-time water quality monitoring? Therefore, this master thesis proposes automatic data quality assessment tools to tackle this issue. These tools are general and simple, which means that they can be applied to any type of water and that they are based on easy-to-implement univariate methods. Following the application of these tools, data quality will be improved and known as such and will allow to better achieve the objectives of a real-time water quality monitoring project. The application of these tools is demonstrated on data collected during real-time water quality monitoring of an urban river performed using a *monEAU* station.



## Avant-Propos

Le présent projet a été réalisé à l'Université Laval. Un projet de maîtrise est d'une telle ampleur que généralement beaucoup de gens y sont impliqués directement et indirectement à son aboutissement. C'est pourquoi je tiens à souligner l'importance de chacun dans le bon déroulement de ce projet et à les remercier.

J'aimerais particulièrement remercier mon directeur de recherche, Peter A. Vanrolleghem, pour sa supervision, ainsi que pour la confiance qu'il m'a accordée durant toute la durée du projet. Peter m'a permis d'améliorer et de développer certaines compétences qui sont un atout pour se distinguer dans ce domaine.

Je tiens également à souligner la collaboration de Peter A. Vanrolleghem en tant que coauteur de l'article présenté dans ce mémoire s'intitulant *Hardware and software tools to increase data quality in automated water quality measuring stations*. Il a apporté de nombreuses suggestions constructives tant au niveau des méthodes d'évaluation de la qualité des données qu'au niveau de la présentation de l'article. De mon côté, j'ai développé, programmé et testé les méthodes d'évaluation de la qualité des données.

Je tiens également à remercier tous les membres du groupe de recherche *modelEAU* pour leurs conseils et leur aide en lien avec le projet, mais plus particulièrement pour ces deux belles années passées à vous côtoyer. Merci à vous tous!

Je souhaite remercier la compagnie Primodal qui a construit et développé la station *monEAU* et qui était toujours disponible pour répondre à mes questions et résoudre mes problèmes. Merci John et Lina!

Je remercie la ville de L'Ancienne-Lorette pour avoir rendu possible la mise en place de la station *monEAU*.

Je veux exprimer ma reconnaissance envers Mme et M. Hamel pour avoir facilité la mise en place de la station *monEAU* en donnant accès à leur terrain sur lequel était installée la station *monEAU*. Je tiens à les remercier aussi pour tous les services qu'ils m'ont rendus lors d'imprévus. Merci Mme et M. Hamel.

Je remercie Maria Pujadas Ferrer, Sylvie Leduc, Étienne Gaborit, Cyril Garneau et Mathieu Beaupré pour leur aide diverse. Merci à vous tous!

Je veux souligner la participation spéciale de Tyler L. Miller et de Jonathan Rochon lors des échantillonnages aux petites heures du matin et lors de l'installation des capteurs durant une belle après-midi d'été. Et je remercie également Tyler pour ses judicieux conseils sur la langue de Shakespeare! Merci à vous deux.

Finalement, je veux souligner l'importance de mes parents dans la réalisation de ce projet. Tout au long de ma maîtrise, ils m'ont encouragé, poussé et ils ont cru en moi. Merci infiniment!

Rire, c'est risquer de paraître fou.  
Pleurer, c'est risquer de paraître sentimental.  
Rechercher les autres, c'est s'exposer aux complications.  
Dévoiler ses sentiments, c'est risquer de montrer sa vraie nature.  
Traîner ses idées, ses rêves devant la foule, c'est risquer de les perdre.

Aimer, c'est risquer de ne pas être aimé en retour.  
Espérer, c'est risquer le désespoir.  
Essayer, c'est risquer l'échec.  
Mais il faut prendre des risques car ne rien risquer c'est plus hasardeux.  
Celui qui ne risque rien, ne fait rien, n'est rien.

Il peut éviter la souffrance et la tristesse  
mais il ne peut apprendre le vrai sens des sentiments,  
du renouvellement, de la sublimation, de l'amour de la vie.  
Enchaîné par ses certitudes, il est esclave, il a abandonné la liberté.  
Seul celui qui risque et se risque est libre...

(Albert Coccoz, guide de haute montagne, disparu dans une avalanche, avec sa femme et sept autres personnes.)



# Table des matières

Résumé .....	iii
Abstract .....	iv
Avant-Propos .....	v
Table des matières .....	viii
Liste des tableaux .....	x
Liste des figures .....	xi
1. Introduction .....	1
1.1 Mise en contexte.....	1
1.2 Objectifs et démarche.....	2
1.3 Plan du mémoire.....	3
2. Revue de littérature .....	5
2.1 Suivi en temps réel de la qualité de l'eau.....	5
2.2 Facteurs influençant la qualité des données .....	10
2.3 Évaluation de la qualité des données.....	15
2.3.1 Les méthodes hors ligne.....	20
2.3.2 Les méthodes en ligne.....	25
3. Matériels et méthodologie.....	31
3.1 Description du bassin versant.....	31
3.2 La station monEAU .....	32
3.3 Capteurs et paramètres mesurés .....	37
3.4 Jaugeage du débit .....	42
4. Étude préliminaire .....	49
4.1 Objectifs de l'étude préliminaire.....	49
4.2 Description des points d'échantillonnage .....	49
4.3 Déroulement de l'étude préliminaire.....	51
4.4 Résultats et analyses de l'étude préliminaire .....	52
4.4.1 Débit.....	52
4.4.2 Sonde YSI 6600EDS (température, conductivité et turbidité).....	54
4.4.3 MeS .....	59
4.4.4 Hach TNT (ammonium et nitrate).....	62
5. Suivi en temps réel .....	68
5.1 Objectifs du suivi en temps réel .....	68
5.2 Description des sites de mesure .....	68
5.3 Matériels et installation .....	73
5.3.1 Installation des capteurs .....	73
5.3.2 Installation et protection du matériel.....	75
5.3.3 Systèmes de nettoyage automatiques .....	82
5.4 Entretien des capteurs.....	84
5.5 Amélioration du logiciel monEAU.....	88
5.6 Résultats et analyses du suivi en temps réel.....	89
6. ADQATs .....	96
6.1 Résumé.....	97
6.2 Hardware and software tools to increase data quality in automated water quality measuring stations.....	98

6.2.1	Abstract .....	98
6.2.2	Keywords .....	98
6.2.3	Introduction.....	99
6.2.4	Case study .....	100
6.2.5	Hardware: installation, cleaning-system and protection.....	100
6.2.6	ADQAT .....	105
6.2.7	Conclusion .....	115
7.	Discussion sur les ADQATs .....	116
7.1	Importance de la fréquence de mesure .....	116
7.2	Génération et interprétation des scores .....	117
7.3	Diagnostic des capteurs .....	121
8.	Conclusion .....	122
	L'étape suivante au développement des ADQATs serait l'intégration d'un algorithme permettant la génération et l'interprétation des scores pour attribuer un score final de validation aux données. Ainsi, la qualité des données mesurées pourrait être déterminée selon leur score final. Un algorithme qui pourrait également être intégré aux ADQATs est le diagnostic des capteurs basé sur les caractéristiques des données mesurées et des fautes détectées. Ceci permettrait d'identifier les causes possibles influençant la qualité des données (ex. affaiblissement de la batterie) et donc de déterminer l'action à poser pour y remédier. ....	124
	Bibliographie .....	126
	ANNEXE 1. Fiche de jaugeage des vitesses d'écoulement avec un micro-moulinet .....	132
	ANNEXE 2. Fiche d'échantillonnage de l'étude préliminaire .....	134
	ANNEXE 3. Matériels et méthodologie d'une journée typique de l'étude préliminaire ...	136
	ANNEXE 4. Résultats de l'étude préliminaire (Suite) .....	140
	ANNEXE 5. Jaugeage du débit durant l'étude préliminaire .....	150
	ANNEXE 6. Jaugeage du débit durant le suivi en temps réel .....	158
	ANNEXE 7. Chartes de contrôle .....	162
	ANNEXE 8. Méthodes en ligne de détection de fautes.....	166

## Liste des tableaux

Tableau 1 : Synthèse des différents principes de mesure et des facteurs influençant la qualité des données mesurées (Beaupré, 2010).....	14
Tableau 2 : Taxonomie des fautes, définitions et causes possibles.....	25
Tableau 3 : Superficie des sous-bassins versants et leur pourcentage d'imperméabilité.....	32
Tableau 4 : Paramètres analysés lors de l'étude préliminaire et leur méthode d'analyse....	38
Tableau 5 : Paramètres analysés lors du suivi en temps réel et leur méthode d'analyse .....	38
Tableau 6 : Capteurs utilisés lors du suivi en temps réel et leurs paramètres de mesure.....	39
Tableau 7: Bathymétrie de la section transversale au point d'échantillonnage amont .....	151
Tableau 8: Bathymétrie de la section transversale au point d'échantillonnage aval .....	151
Tableau 9: Étalonnage des micro-moulinets .....	152
Tableau 10: Points conservés pour la détermination des droites d'étalonnage.....	152
Tableau 11: Exemple de fiche de mesure des vitesses d'écoulement avec un micro-moulinet .....	154
Tableau 12 : Débits jaugés au point d'échantillonnage amont.....	155
Tableau 13 : Débits jaugés au point d'échantillonnage aval .....	155
Tableau 14 : Bathymétrie de la section transversale au point de mesure aval .....	159
Tableau 15 : Étalonnage du micro-moulinet 172-112-5 .....	159
Tableau 16 : Débits jaugés au point de mesure aval .....	161
Tableau 17 : Valeurs des constantes des modèles et du lisseur pour différents intervalles de mesure et pour les paramètres SS* et conductivité .....	167



## Liste des figures

Figure 1: Séries d'hydrogrammes étalés sur une période de 24 heures et mesurés à la station WK de la rivière Waiakeakua (Tomlinson et De Carlo, 2003) .....	6
Figure 2 : Schéma de la station de mesure (van Griensven et al. 2000) .....	8
Figure 3 : La cabine de mesure et le support des capteurs remonté pour un entretien .....	9
Figure 4 : Schéma d'un capteur et de ses composantes principales (Ni et al., 2009).....	11
Figure 5 : Réponse $G_s$ d'un capteur en fonction de la valeur $G_e$ du mesurande .....	11
Figure 6 : Illustration de l'erreur d'« offset » sur un capteur.....	12
Figure 7 : Illustration de l'erreur de sensibilité sur un capteur .....	13
Figure 8 : Illustration d'une erreur due aux grandeurs d'influence et/ou aux erreurs simultanées d'« offset » et de sensibilité (Bertrand-Krajewski et al., 2000).....	13
Figure 9 : Concept d'évaluation univariée/multivariée de la qualité des données .....	16
Figure 10 : Procédure de validation des données de drainage urbain.....	18
Figure 11 : Génération des scores de validation .....	19
Figure 12 : Effets de dérive, de saut ou de valeur aberrante (Thomann et al., 2002).....	21
Figure 13 : Approche de suivi pour les capteurs en ligne. Les flèches en gras indiquent la séquence des actions à effectuer si le processus est présumé en contrôle. ....	22
Figure 14 : Charte de contrôle « Shewart » avec des critères de « out-of-control » plus sévères pour la phase d'avertissement et ceux normalement utilisés pour la phase d'alarme .....	24
Figure 15 : Schéma de l'utilisation d'un modèle de prédiction pour détecter des données aberrantes (Krajewski et Krajewski, 1989).....	27
Figure 16 : Bassin versant du ruisseau Notre-Dame.....	31
Figure 17 : Agrandissement du bassin versant du ruisseau Notre-Dame .....	32
Figure 18 : Montage du réseau de suivi de la qualité de l'eau, le système monEAU .....	33
Figure 19 : Composantes matérielles du prototype 3 de la station monEAU .....	36
Figure 20 : Montage pour jauger les vitesses selon la méthode a) point par point;.....	43
Figure 21 : Types d'hélices pour le premier montage .....	45
Figure 22 : Installation pour l'étalonnage des hélices du premier montage .....	45
Figure 23 : Jaugeage du débit selon la méthode point par point.....	46
Figure 24 : Intégration des vitesses ponctuelles à la section d'écoulement .....	46
Figure 25 : Jaugeage du débit selon la méthode verticale par verticale .....	47
Figure 26 : Carte indiquant les points d'échantillonnage amont et aval.....	50
Figure 27 : Point d'échantillonnage amont .....	50
Figure 28 : Point d'échantillonnage aval.....	51
Figure 29 : Débits mesurés et simulés en amont au mois d'octobre 2009.....	53
Figure 30 : Débits mesurés et simulés en aval au mois d'octobre 2009 .....	54
Figure 31 : Températures en amont et aval au mois d'octobre 2009 .....	55
Figure 32 : Conductivités en amont et en aval au mois d'octobre 2009.....	56
Figure 33 : Conductivités en amont et en aval le 19 octobre 2009.....	57
Figure 34 : Turbidités en amont et en aval au mois d'octobre 2009 .....	58
Figure 35: Agrandissement de la figure 34.....	58
Figure 36 : Turbidité en amont le 19 octobre 2009 .....	59
Figure 37 : MeS en amont et en aval au mois d'octobre 2009 .....	60
Figure 38 : MeS en amont et en aval le 1er octobre 2009 .....	61

Figure 39 : MeS en amont et en aval le 19 octobre 2009.....	61
Figure 40 : Concentrations d'ammonium en amont et en aval au mois d'octobre 2009.....	63
Figure 41: Concentrations d'ammonium en amont et en aval le 1er octobre 2009 .....	64
Figure 42: Concentrations d'ammonium en amont et en aval le 19 octobre 2009 .....	65
Figure 43: Concentrations de nitrate en amont et en aval au mois d'octobre 2009 .....	66
Figure 44: Concentrations de nitrate en amont et en aval le 1 <sup>er</sup> octobre 2009 .....	67
Figure 45: Concentrations de nitrate en amont et en aval le 19 octobre 2009 .....	67
Figure 47 : Sites d'installation des deux stations du suivi en temps réel et points d'échantillonnage .....	69
Figure 48 : Point d'installation aval des capteurs mesurant la hauteur et la vitesse.....	70
Figure 49 : Point d'installation aval des capteurs mesurant la qualité de l'eau.....	71
Figure 50: Égout pluvial en amont de la station de mesure .....	71
Figure 51 : Point d'installation amont des capteurs mesurant la hauteur et la vitesse.....	72
Figure 52 : Point d'installation amont des capteurs mesurant la qualité de l'eau.....	73
Figure 53 : Protection de la station monEAU.....	76
Figure 54 : Installation des capteurs dans la première cage de protection.....	77
Figure 55 : Effet d'obstruction des débris par la cage de protection (5 août 2010).....	78
Figure 56 : Protection des capteurs contre les débris .....	78
Figure 57: Retrait de la sous-section de la première cage de protection lors d'un entretien et localisation des capteurs.....	80
Figure 58 : Deuxième cage de protection installée dans le ruisseau .....	81
Figure 59 : Retrait des sous-sections de la deuxième cage de protection lors d'un entretien .....	81
Figure 60 : Boîte d'insonorisation du compresseur .....	83
Figure 61 : Vérification du fonctionnement des capteurs avec des solutions standards .....	85
Figure 62 : Suivi en temps réel de l'évènement pluvial du 2010-08-02 .....	91
Figure 63 : Suivi en temps réel de l'évènement pluvial du 2010-09-13 .....	92
Figure 64 : Storm sewer outlet just upstream the monEAU station .....	101
Figure 65 : SS* and conductivity smoothed time series during a rain event .....	101
Figure 66 : Protective cage blocks debris.....	102
Figure 67 : Sensors rose out of the water but remain protected.....	103
Figure 68 : Cage with metal front plate.....	104
Figure 69 : Turbidity sensor with wiper.....	104
Figure 70 : Micro-fouling effect during intense rain event.....	105
Figure 71 : Manual cleaning, Filled symbols: accepted,.....	107
Figure 72 : Standard solution control charts Filled symbols: accepted,.....	107
Figure 73: Scheme of the integrated on-line data validation methods proposed in the paper .....	109
Figure 74: Illustration of the detection method for outliers (indicated with crosses) on SS* time series.....	111
Figure 75: On-line data validation methods on a SS* time series .....	114
Figure 76: Méthode proposée de génération de scores .....	118
Figure 77: Concentrations d'azote total en amont et en aval au mois d'octobre 2009 .....	140
Figure 78: Concentrations d'azote total en amont et en aval le 1 <sup>er</sup> octobre 2009.....	141
Figure 79: Concentrations d'azote total en amont et en aval le 19 octobre 2009.....	142
Figure 80: Concentrations d'E.coli en amont et en aval au mois d'octobre 2009 .....	144
Figure 81: Concentrations d'E.coli en amont et en aval le 19 octobre 2009 .....	145

Figure 82: Concentrations de coliformes totaux en amont et en aval au mois d'octobre 2009 .....	146
Figure 83: Coliformes totaux en amont et en aval le 1 octobre 2009 .....	147
Figure 84: Concentrations de coliformes totaux en amont et en aval le 19 octobre 2009 ..	147
Figure 85: Concentrations de manganèse total en amont et en aval au mois d'octobre 2009 .....	148
Figure 86: Concentrations de zinc total en amont et en aval au mois d'octobre 2009 .....	149
Figure 87: Droite d'étalonnage du micro-moulinet 172-112-5 .....	153
Figure 88: Droite d'étalonnage du micro-moulinet 111987-6.....	153
Figure 89: Droite d'étalonnage du micro-moulinet 111134 .....	153
Figure 90: Courbe de tarage au point d'échantillonnage amont.....	156
Figure 91: Courbe de tarage au point d'échantillonnage aval .....	156
Figure 92: Étalonage du micro-moulinet 172-112-5 .....	160
Figure 93 : Courbe de tarage au point de mesure aval.....	161
Figure 94: « Standard solution control chart » du capteur bulle à bulle .....	163
Figure 95 : « Standard solution control chart» du capteur 3700 sc Conductimètre inductif .....	163
Figure 96 : « Standard solution control chart » du capteur pH sc différentiel numérique (pH7 : marqueur en losange, pH 10 : marqueur carré) .....	164
Figure 97 : « Standard solution control chart » du capteur LDO .....	164
Figure 98 : « Standard solution control charts » pour la mesure de la température des capteurs 3700 sc Conductimètre inductif (marque en losange), pH sc différentiel numérique (marqueur carré) et LDO (marqueur en triangle) .....	165
Figure 99: Méthode de détection des données aberrantes sur les SS* .....	168
Figure 100 : Méthode de détection des données aberrantes sur les SS* .....	168
Figure 101 : Méthode de lissage des données sur les SS* .....	169
Figure 102 : Méthode de lissage des données sur les SS* .....	169
Figure 103 : Méthode de détection des données aberrantes sur les SS* .....	171
Figure 104 : Méthode de détection des données aberrantes sur les SS*, agrandissement de la figure 100 ( $\alpha x=0,1$ ; $\alpha e=0,1$ ; $K=5$ ; $\Delta t=5$ sec.) .....	171
Figure 105 : Méthodes de détection de fautes sur les SS* .....	172
Figure 106 : Méthodes de détection de fautes sur les SS*, agrandissement de la figure 102 .....	172
Figure 107 : Méthode de détection des données aberrantes sur les SS* .....	174
Figure 108 : Méthode de détection des données aberrantes sur les SS*, agrandissement de la figure 104 ( $\alpha x=0,2$ ; $\alpha e=0,1$ ; $K=5$ ; $\Delta t=15$ sec.) .....	174
Figure 109 : Méthodes de détection de fautes sur les SS* .....	175
Figure 110 : Méthodes de détection de fautes sur les SS*, agrandissement de la figure 106 .....	175
Figure 111 : Méthode de détection des données aberrantes sur les SS* .....	177
Figure 112 : Méthode de détection des données aberrantes sur les SS*, agrandissement du premier « loss-of-control » de la figure 108 ( $\alpha x=0,25$ ; $\alpha e=0,1$ ; $K=5$ ; $\Delta t=30$ sec.).....	177
Figure 113 : Méthode de détection des données aberrantes sur les SS*, agrandissement du deuxième « loss-of-control » de la figure 108 ( $\alpha x=0,25$ ; $\alpha e=0,1$ ; $K=5$ ; $\Delta t=30$ sec.) ..	178
Figure 114 : Méthodes de détection de fautes sur les SS* .....	178
Figure 115 : Méthode de détection des données aberrantes sur les SS* .....	180

Figure 116 : Méthode de détection des données aberrantes sur les SS*, agrandissement du « loss-of-control » de la figure 112 ( $\alpha_x=0,35$ ; $\alpha_e=0,1$ ; $K=5$ ; $\Delta t=1$ min.).....	180
Figure 117 : Méthodes de détection de fautes sur les SS* .....	181
Figure 118 : Méthode de détection des données aberrantes sur les SS* .....	182
Figure 119 : Méthode de détection des données aberrantes sur les SS*, agrandissement du « loss-of-control » de la figure 115 ( $\alpha_x=0,55$ ; $\alpha_e=0,1$ ; $K=5$ ; $\Delta t=2$ min.) .....	182
Figure 120 : Méthodes de détection de fautes sur les SS* .....	183
Figure 121 : Méthode de détection des données aberrantes sur la conductivité .....	184
Figure 122 : Méthode de détection des données aberrantes sur la conductivité, agrandissement d'un « loss-of-control » de la figure 118.....	185
Figure 123 : Méthode de détection des données aberrantes sur la conductivité, agrandissement de l'encadré de la figure 118 .....	185
Figure 124 : Méthode de détection des fautes sur une série de conductivité .....	186

# 1. Introduction

## 1.1 Mise en contexte

De nos jours, les exigences liées au suivi de la qualité de l'eau sont de plus en plus sévères. Ceci est attribuable, particulièrement, à la complexification des stratégies de contrôle et à la nécessité d'un suivi continu de la qualité de l'eau à l'effluent des stations d'épuration, à l'évolution des méthodes de gestion de la ressource en eau (protection des eaux de surfaces), etc. Ces exigences sont principalement axées sur un meilleur suivi des variations de la qualité de l'eau. Par exemple, dans certains bassins versants, les processus hydrologiques et de la qualité de l'eau ont des vitesses de variation de l'ordre de la minute et de l'heure (Tomlinson et De Carlo, 2003). Si tel est le cas, un suivi traditionnel avec des échantillonnages ponctuels est inapproprié. Un suivi en temps réel effectué à l'aide d'une station de mesure et de capteurs est donc nécessaire pour bien mesurer les variations de la qualité de l'eau lorsque cette dernière varie rapidement.

Par ailleurs, les avancements technologiques favorisant l'utilisation des stations de mesure fournissent de nouvelles capacités de lecture des capteurs, des moyens de communications à longue distance et d'interprétation des données, un format livrable via le web, etc. Les capteurs permettent parfois d'effectuer un suivi de la qualité de l'eau nécessitant peu d'entretien et fournissant une précision acceptable (Jeppsson et al., 2002, Vanrolleghem et Lee, 2003).

Cependant, contrairement aux analyses en laboratoire, lesquelles ont une précision et une exactitude connues, la mesure en temps réel peut être propice à la collecte de données de mauvaise qualité à cause des conditions difficiles dans lesquelles les capteurs sont installés. Les séries de données mesurées peuvent également être influencées par du bruit, des données aberrantes et des données manquantes. Il y a donc une nécessité d'avoir des outils automatiques d'évaluation de la qualité des données (ADQATs qui est l'acronyme de « Automatic Data Quality Assessment Tools »). Ils doivent être automatisés, car un suivi en temps réel génère une grande quantité de données et une évaluation manuelle de la qualité des données est pratiquement impossible. Si la qualité des données n'est pas évaluée, les données pourraient avoir un effet néfaste sur le résultat de leur utilisation. Par exemple,

l'utilisation d'un modèle de la qualité de l'eau à des fins de gestion qui serait construit à partir de données de mauvaise qualité pourrait biaiser les décisions basées sur les résultats obtenus par ce modèle. Alors, un suivi en temps réel de la qualité de l'eau doit absolument posséder des ADQATs.

## 1.2 Objectifs et démarche

L'objectif principal de ce projet consiste à développer des ADQATs permettant d'e mesurer en temps réel des données de qualité connue et accrue. Leur utilisation favorisera ainsi l'atteinte des objectifs (modélisation, stratégies de contrôle, gestion, etc.) reliés à un suivi en temps réel.

Les ADQATs présentés dans ce mémoire tentent de répondre à la vision du système monEAU quant à l'évaluation de la qualité des données. Cette vision se définit par l'exploitation d'un concept général d'évaluation de la qualité des données pouvant être appliqué à différents emplacements (rivière, STEP, égout, etc.). Particulièrement, ces ADQATs consistent au niveau 1 d'évaluation de la qualité des données du système monEAU, soit l'exploitation de méthodes univariées (Rieger et Vanrolleghem, 2008). Ils n'utilisent donc aucune méthode multivariée et aucune connaissance du processus et d'un expert. Ils permettent une évaluation efficace de la qualité des données grâce à l'exploitation de plusieurs méthodes d'évaluation basées sur différentes caractéristiques des données et du système. Les ADQATs développés dans le cadre de ce projet seront implantés ultérieurement dans le logiciel monEAU.

Les ADQATs développés dans le cadre de ce projet ont été évalués sur des données récoltées durant un suivi en temps réel de la qualité de l'eau d'un ruisseau urbain. Ce suivi a été réalisé à l'aide d'une station de mesure, la station monEAU.

Antérieurement au suivi en temps réel, une campagne d'échantillonnage a été effectuée afin de connaître les gammes de concentration de différents paramètre et les dynamiques générales de la qualité de l'eau lors d'évènements pluviaux. Les connaissances acquises lors de cette campagne ont été très utiles à la mise en branle du suivi en temps réel.

### 1.3 Plan du mémoire

Le présent mémoire présente d'abord une revue de littérature portant sur différents aspects d'un suivi en temps réel de la qualité de l'eau, les facteurs influençant la qualité des données et les méthodes d'évaluation de la qualité des données. Ensuite, ce mémoire décrit le matériel et la méthodologie empruntés, l'étude préliminaire et le suivi en temps réel de la qualité de l'eau du ruisseau à l'étude. Finalement, l'article *Hardware and software tools to increase data quality in automated water quality measuring stations* est présenté ainsi qu'une discussion sur la génération et l'interprétation des scores, le diagnostic des capteurs et l'importance d'une haute fréquence de mesure dans l'utilisation des ADQATs. Les objectifs de l'étude préliminaire, du suivi en temps réel et des ADQATs sont également rappelés dans les sections correspondantes et certains éléments additionnels de méthodologie y sont également présentés. Cette présentation facilitera la compréhension des grandes étapes de ce mémoire dans l'atteinte de l'objectif principal qui consiste à développer des ADQATs permettant de mesurer en temps réel des données de qualité connue et accrue.



## **2. Revue de littérature**

### **2.1 Suivi en temps réel de la qualité de l'eau**

Afin de démontrer l'importance de la mesure en temps réel dans le suivi de la qualité de l'eau, cette section explique premièrement le besoin d'une haute fréquence de mesure qui n'est possible qu'avec les stations de mesure. Ensuite, deux cas réels d'utilisation des capteurs en rivière sont présentés, ainsi que les problématiques en lien avec les stations de mesure.

Une haute fréquence de lecture pour le suivi de la qualité de l'eau avec l'aide de capteurs est reconnue comme un besoin par la communauté scientifique (Horsburgh et al., 2010). En effet, tel qu'illustré à la figure 1, la qualité de l'eau dans un cours d'eau suivant un évènement pluvial peut parfois changer très rapidement selon l'intensité de la pluie, les conditions précédant l'évènement, etc. (Kirchner et al., 2004). Par exemple, Tomlinson et De Carlo (2003) ont montré que pour certains paramètres, un intervalle de lecture de 5 minutes était essentiel afin de bien représenter les changements rapides de la qualité de l'eau dans le cours d'eau à l'étude. Durant leur étude, ils ont mesuré des variations dans la turbidité passant de 42 NTU à 1257 NTU en 5 minutes. D'ailleurs, une fréquence de mesure cohérente avec le temps de réponse hydrologique du bassin versant est nécessaire afin de comprendre les interactions entre les processus hydrologiques et hydrochimiques (Kirchner et al., 2004).

Un suivi en temps réel réduit l'incertitude sur les paramètres d'un modèle et, par conséquent, augmente sa fiabilité. En fait, ceci est étroitement lié à la quantité et à la qualité des données disponibles pour la calibration (Vandenberghe et al., 2002). La fréquence de mesure doit aussi être suffisamment élevée pour que la fréquence maximale d'intérêt se situe entre le tiers et le quart de la fréquence de mesure. Ceci permet d'éviter l'« aliasing effect » qui se définit par l'introduction d'erreurs causées lorsqu'un signal est échantillonné à un taux qui ne permet pas une analyse appropriée des hautes fréquences (Olsson et Newell, 1999).

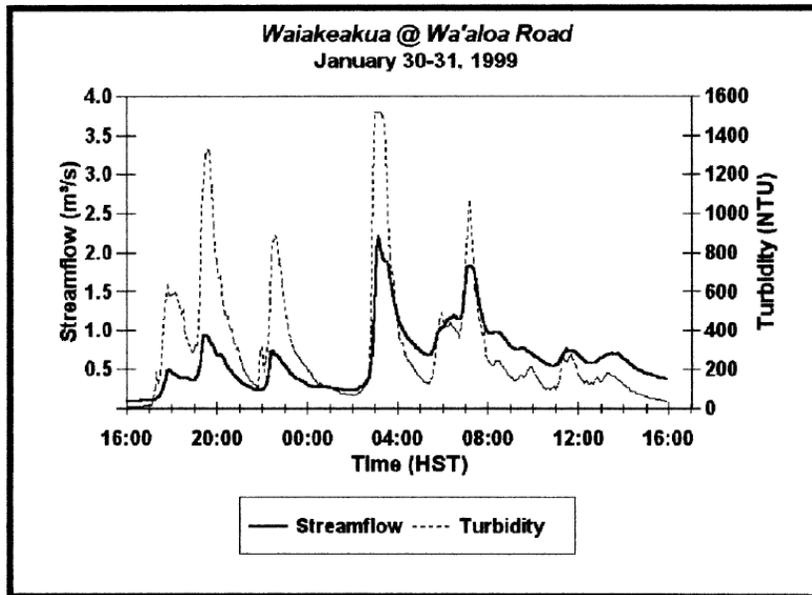


Figure 1: Séries d'hydrogrammes étalés sur une période de 24 heures et mesurés à la station WK de la rivière Waiakeakua (Tomlinson et De Carlo, 2003)

Un suivi en temps réel permet également une gestion et un contrôle plus efficaces des procédés tels que dans les stations d'épuration des eaux et de production d'eau potable. Entre autre, le débit d'air insufflé dans un bassin de boues activées peut être ajusté selon l'oxygène dissous mesuré en temps réel. Une modification rapide du débit d'air insufflé aux variations d'oxygène dissous permet de réduire les coûts reliés à une surconsommation électrique.

Par ailleurs, une haute fréquence de mesure permet l'exploitation de nouvelles méthodes d'analyses des données dans le domaine de la qualité de l'eau telles que l'analyse spectrale, l'analyse en ondelette, etc. (Kirchner et al., 2004).

Le suivi en temps réel est largement répandu dans les usines de traitement des eaux usées (Vanrolleghem et Lee, 2003) et dans les réseaux d'assainissement (Lombard et al., 2010), comparativement aux eaux de surface où seulement quelques cas d'étude ont été sujets d'un suivi en temps réel (Vandenberghe et al., 2005). Parmi ces cas, les types de suivis en temps réel sont très variés, en ce sens que les capteurs utilisés et les paramètres de mesure sont différents d'un projet à l'autre. Notamment, il y a des études qui se spécialisent dans la mesure en temps réel d'un seul paramètre (Wang et Thongngamdee, 2003), tandis que d'autres mesurent plusieurs paramètres à la fois pour acquérir des connaissances dans les processus aquatiques (Beck et al., 1998). Aussi, certaines études promeuvent l'utilisation

d'un montage simple nécessitant peu d'entretien (Edwards, 1998, Johnson, 1998), tandis que d'autres utilisent un système complexe (Beck et al., 1998, Du Preez et al., 1998).

Concernant le suivi de la qualité de l'eau en rivières, la règle de l'art d'un suivi en temps réel a grandement évolué, passant d'un suivi avec une boucle de déviation (suivi en ligne), à un suivi in situ où les capteurs sont installés directement dans le milieu (Winkler et al., 2008). Notamment, van Griensven et al. (2000) et Pressl et al. (2004) ont réalisé respectivement un suivi en temps réel en ligne et in situ. La figure 2 et la figure 3 illustrent leur montage. Chaque type de suivi comporte ses avantages et ses inconvénients et ceux-ci sont discutés ci-après.

van Griensven et al. (2000) mentionnent que les données récoltées par un suivi en ligne sont propices aux erreurs de mesures qui sont principalement causées par le colmatage du système et par l'encrassement des membranes des électrodes. Un rinçage régulier du système était donc nécessaire pour pallier ces problèmes. Cependant, cela perturbait les données après le rinçage pour une durée égale au temps de réponse du système. De plus, la température de l'eau mesurée était biaisée par la température présente à l'intérieur de la cabine dans laquelle était installé le système. Conséquemment, la température mesurée servant à compenser l'influence de la température de l'eau sur les paramètres mesurés par les électrodes avait un effet direct sur la qualité des données. Le principal avantage de ce type de suivi est que les capteurs sont facilement accessibles pour un entretien régulier.

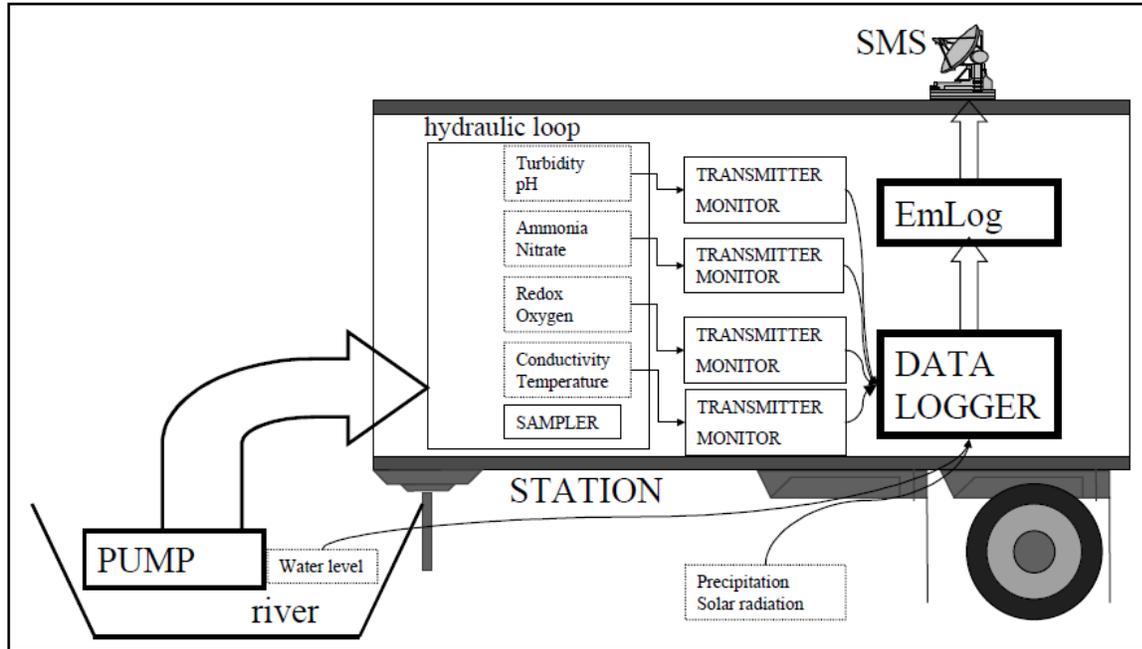


Figure 2 : Schéma de la station de mesure (van Griensven et al. 2000)

Le suivi in situ de la qualité de l'eau comporte aussi des inconvénients, toutefois ceux-ci sont différents de ceux qui peuvent être rencontrés lors d'un suivi en ligne. Tout d'abord, l'installation doit tenir compte des changements rapides de la hauteur d'eau, de la protection contre les débris transportés par le courant, du mouvement des pierres, etc. De plus, les capteurs doivent être sortis du milieu pour être calibrés, sinon la calibration doit être effectuée in situ, ce qui n'est pas toujours possible. Par contre, ce type de suivi a l'avantages de nécessiter peu de matériels comparativement à un suivi en ligne (pompe, unité de filtration, cabine, etc.) et il est rapidement mis en fonction. Contrairement à van Griensven et al. (2000) qui mentionnent qu'il est difficile d'obtenir des données fiables et représentatives de la qualité de l'eau en rivière, Pressl et al. (2004) affirment que l'installation in situ des capteurs dans une rivière est bien adaptée pour le suivi de la qualité de l'eau et que cela permet la mesure de données de qualité. Ces deux cas illustrent bien l'évolution de la règle d'art du suivi en temps réel de la qualité de l'eau en rivière.

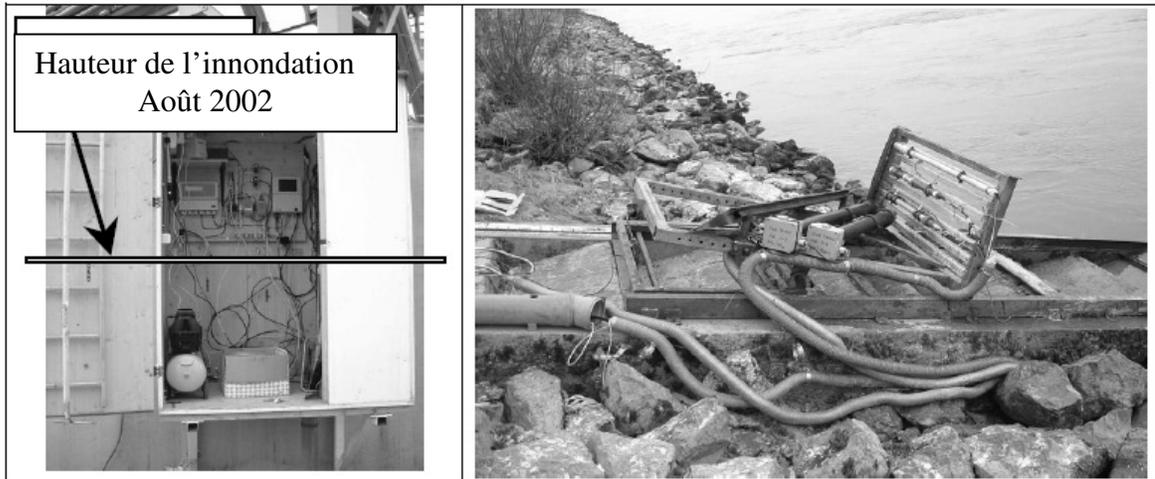


Figure 3 : La cabine de mesure et le support des capteurs remonté pour un entretien (Pressl et al., 2004)

Les stations de mesures sont généralement développées afin de répondre à un objectif spécifique, par exemple pour le contrôle dans des stations d'épuration des eaux et production d'eau potable, la compréhension de processus hydrologiques et hydrochimiques, etc.). Cela engendre un manque au niveau de leur standardisation. De plus, les problèmes liés à la qualité des données mesurées et à la flexibilité insuffisante des stations (connexion de nouveaux capteurs, nouveaux objectifs, etc.) limitent leurs utilisations. Ces problèmes sont causés respectivement par un manque d'outils d'évaluation de la qualité des données et par la conception initiale de la station basée sur des objectifs particuliers et des spécifications de conceptions individuelles (Rieger et Vanrolleghem, 2008). Bien que le suivi en temps réel soit souvent perçu comme un objectif spécifique, des études se concentrent sur le développement de réseaux de suivi de la qualité de l'eau intégrant l'information de différents emplacements sur un bassin versant (Strobl et al., 2006). Cela crée de nouveaux besoins pour l'échange bidirectionnel de données tels que la transmission des données par différentes options de télémétrie, la sécurité et l'accessibilité des données (Rieger et Vanrolleghem, 2008).

Afin de résoudre les problématiques énumérées précédemment, un projet en cours impliquant des groupes de recherche, des organisations publiques et des compagnies privées développent une nouvelle génération de réseaux de suivi de la qualité de l'eau, le système monEAU. Le système monEAU combine les technologies les plus avancées et permet la plus haute

flexibilité possible en terme de connexion de capteurs, d'emplacement et d'objectifs de suivi (Rieger et Vanrolleghem, 2008).

## **2.2 Facteurs influençant la qualité des données**

Comme mentionné précédemment, les capteurs sont propices à mesurer des données de mauvaise qualité à cause des conditions difficiles dans lesquelles ils sont installés. C'est pourquoi cette section présente différentes approches empruntées par certains chercheurs pour discuter des facteurs influençant la qualité des données mesurées par des capteurs.

Ni et al. (2009) discutent des fautes possibles pouvant être observées dans les données mesurées par les capteurs en général. Celles-ci sont des indicateurs de la qualité des données. Ils définissent une faute comme étant une donnée anormale déviant du comportement normal prévu. Les fautes sont causées par les caractéristiques environnementales et par le système de mesure en question et elles se reflètent dans les caractéristiques des données mesurées telles que la pente, la variance, etc. Par exemple, van Griensven et al (2000) racontent que la haute concentration de fines particules en suspension, qui est une caractéristique environnementale, causait des problèmes reliés à l'encrassement des électrodes. Cet encrassement affectait directement la qualité des données mesurées. D'ailleurs, deux types d'encrassement influençant surtout la qualité des données mesurées par les capteurs optiques caractérisent les eaux, soit le micro-encrassement et le macro-encrassement. Ils ont une influence distincte sur la qualité des données et ils sont définis respectivement par l'augmentation du signal par salissure des éléments optiques et par l'obstruction du capteur par des déchets en suspension. Le micro-encrassement peut être pallié par un système automatique de nettoyage, tandis que le macro-encrassement exige généralement un entretien manuel (Bertrand-Krajewski et al., 2000).

Concernant les caractéristiques du système de mesure, la figure 4 illustre un capteur typique ainsi que le parcours des données au travers des composantes principales d'un système de mesure. Ces composantes sont toutes des endroits où une faute peut se produire.

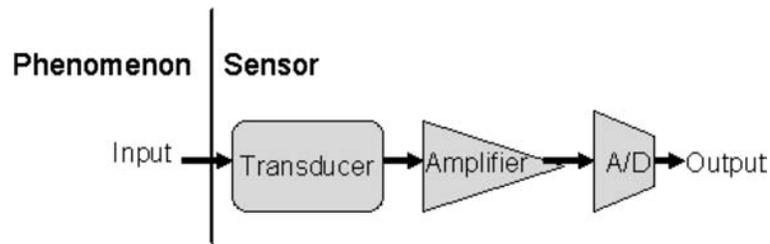


Figure 4 : Schéma d'un capteur et de ses composants principales (Ni et al., 2009)

Dans la section 2.3 Évaluation de la qualité des données, le tableau 2 présente une liste de fautes pouvant être observées lors d'un suivi en temps réel ainsi que leurs causes possibles.

Bertrand-Krajewski et al. (2000) décrivent les erreurs instrumentales telles que les erreurs systématiques et aléatoires. Ces dernières sont des erreurs inévitables et elles ne sont pas discutées ici. Les erreurs systématiques conduisent à des écarts entre les valeurs réelles et les valeurs mesurées. Pour un capteur idéal, le signal de sortie  $G_S$  (réponse) est proportionnel au signal d'entrée  $G_e$  sur l'étendue de mesure et n'est pas nécessairement linéaire en dehors de cette étendue. L'équation suivante représente cette relation où  $S_e$  et  $Z_0$  sont respectivement la sensibilité et l'« offset ». La figure 5 représente ces notions.

$$G_S = S_e G_e + Z_0$$

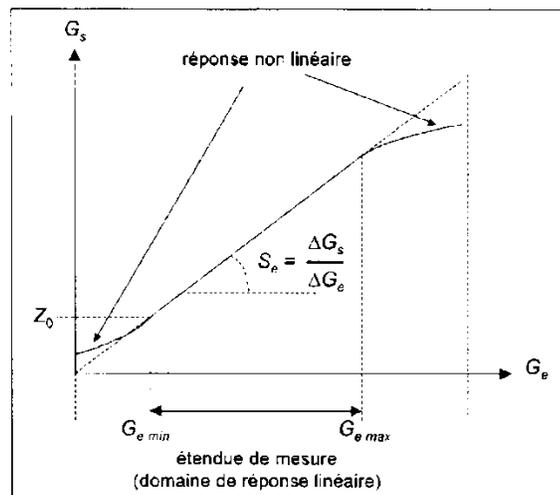


Figure 5 : Réponse  $G_S$  d'un capteur en fonction de la valeur  $G_e$  du mesurande (Bertrand-Krajewski et al., 2000)

Cependant, les capteurs ne sont pas idéaux et des erreurs systématiques peuvent être introduites telles que l'erreur d'« offset », l'erreur de sensibilité, l'erreur liée aux grandeurs d'influence et l'erreur de linéarité. Ces erreurs produisent un signal de sortie biaisé et affectent donc la qualité des données. Les erreurs d'« offset » et de sensibilité sont causées respectivement par un décalage de l'« offset » et à une pente ou sensibilité incorrecte, et ces erreurs sont représentées à la figure 6 et la figure 7. L'erreur liée aux grandeurs d'influence, telle qu'illustrée à la figure 8, est causée par un paramètre influençant le signal de sortie et il peut avoir un effet simultané sur l'« offset » et la sensibilité. Par exemple, van Griensven et al. (2000) ont observé que la mesure de l'oxygène dissout était influencée par la variation de pression à l'intérieur de la boucle de déviation. Finalement, l'erreur de linéarité, telle qu'observée aux limites de l'étendue de mesure à la figure 5, est définie par une non-linéarité dans la relation entre les signaux d'entrée et de sortie.

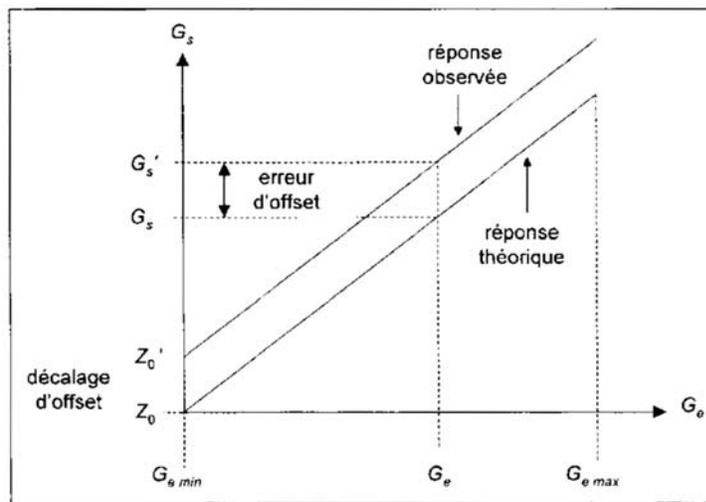


Figure 6 : Illustration de l'erreur d'« offset » sur un capteur (Bertrand-Krajewski et al., 2000)

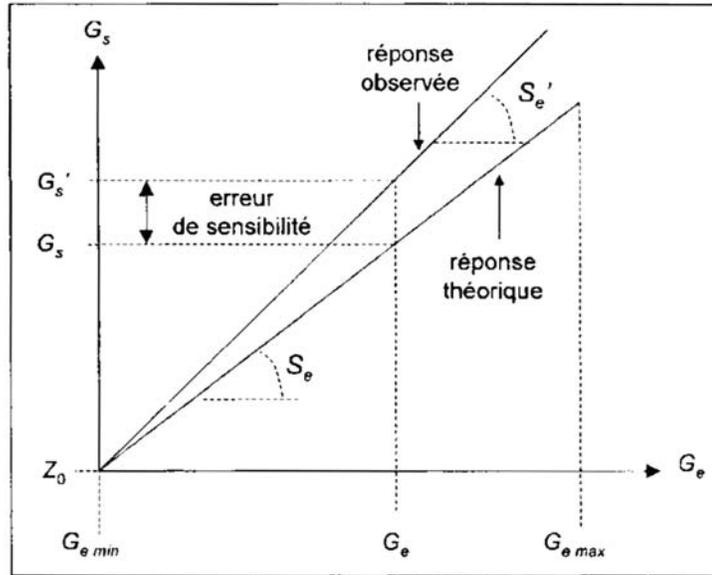


Figure 7 : Illustration de l'erreur de sensibilité sur un capteur (Bertrand-Krajewski et al., 2000)

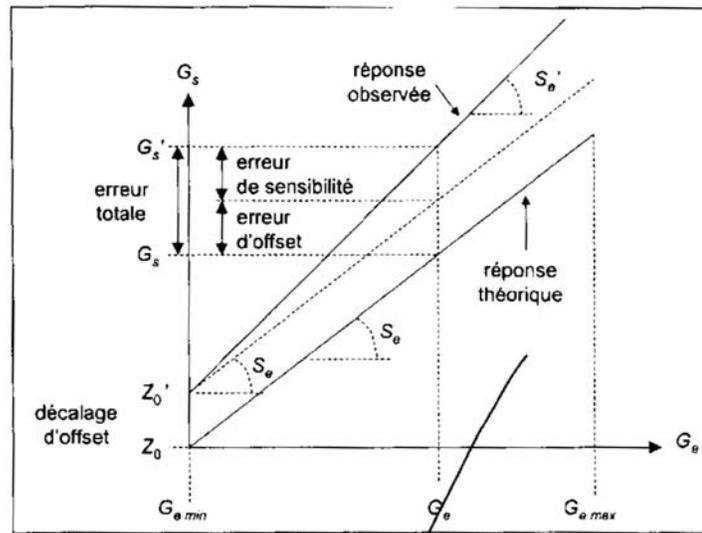


Figure 8 : Illustration d'une erreur due aux grandeurs d'influence et/ou aux erreurs simultanées d'« offset » et de sensibilité (Bertrand-Krajewski et al., 2000)

Une multitude de facteurs peut causer des erreurs systématiques. Beaupré (2010) présente une synthèse des différents principes de mesure des capteurs ainsi que quelques facteurs causant des erreurs systématiques. Le tableau 1 présente quelques éléments de cette synthèse.

Ce tableau démontre que peu importe le type de capteur, la qualité des données peut être influencée par un facteur quelconque faisant partie soit des caractéristiques environnementales ou du système.

Tableau 1 : Synthèse des différents principes de mesure et des facteurs influençant la qualité des données mesurées (Beaupré, 2010)

Principe de mesure	Facteurs influençant la qualité des données
Électrode sensible au gaz	<ul style="list-style-type: none"> <li>• Colmatage et contamination de l'électrode par des ions d'hydroxyde (Aspegren et al., 1993)</li> <li>• Dérive de l'électrode (Patry et Takács, 1995)</li> <li>• Bulle de gaz prise sous l'embout de l'électrode (Andersen et Wagner, 1990)</li> </ul>
Électrode sélective d'ions	<ul style="list-style-type: none"> <li>• Contamination de l'électrode (Wacheux et al., 1993, Sikow et Pursiainen, 1995)</li> <li>• Dérive de l'électrode (Wacheux et al., 1993)</li> <li>• Interférence dans la mesure du nitrate avec du <math>\text{HCO}_3^-</math> (Sikow et Pursiainen, 1995), du chlore (APHA, 1992), du bromure (Rieger et al., 2002, Winkler et al., 2004) et de l'iode (Winkler et al., 2004)</li> <li>• Interférence dans la mesure de l'ammonium avec des amines, du mercure et de l'argent (APHA, 2005), ainsi que du potassium et du sodium (Rieger et al., 2002, Winkler et al., 2004)</li> </ul>
Absorbance et diffusion de la lumière	<ul style="list-style-type: none"> <li>• Interférence dans la mesure de la turbidité par des bulles d'air et par l'encrassement optique (Vanrolleghem et Lee, 2003)</li> <li>• Influence de la variation de la grosseur des particules et autres propriétés du milieu dans la relation entre les MeS et la turbidité (Pfannkuche et Schmidt, 2003)</li> </ul>
Colorimétrie	<ul style="list-style-type: none"> <li>• Sensible aux variations de température (Wacheux et al., 1996)</li> </ul>
Sonde diélectrique	<ul style="list-style-type: none"> <li>• Influence de la variation de la capacité de l'eau ou de la structure, de la composition et de la perméabilité de la membrane cellulaire (Vanrolleghem et Lee, 2003)</li> </ul>

## 2.3 Évaluation de la qualité des données

La section précédente a démontré que la qualité des données mesurées lors d'un suivi en temps réel peut être influencée par de nombreux facteurs. Il est donc impératif d'évaluer la qualité des données. Pour ce faire, il existe une multitude de méthodes.

Rieger et Vanrolleghem (2008) proposent un concept d'évaluation de la qualité des données basée sur trois types d'information, soit des échantillons de référence qui sont comparés avec les mesures du capteur (méthode hors ligne), l'état du capteur (ex. mode calibration, nettoyage automatique, etc.) et l'information des séries temporelles qui est utilisée pour des méthodes univariées et multivariées (méthode en ligne). Dans les conditions normales, un échantillon de référence devrait être pris à un intervalle d'environ deux semaines en combinaison avec des méthodes d'évaluation en ligne. Chaque type d'information peut engendrer un avertissement, lequel demande une fréquence d'échantillonnage plus élevée. Une alarme est déclenchée seulement si l'avertissement est confirmé avec des échantillons additionnels. La figure 9 illustre ce concept.

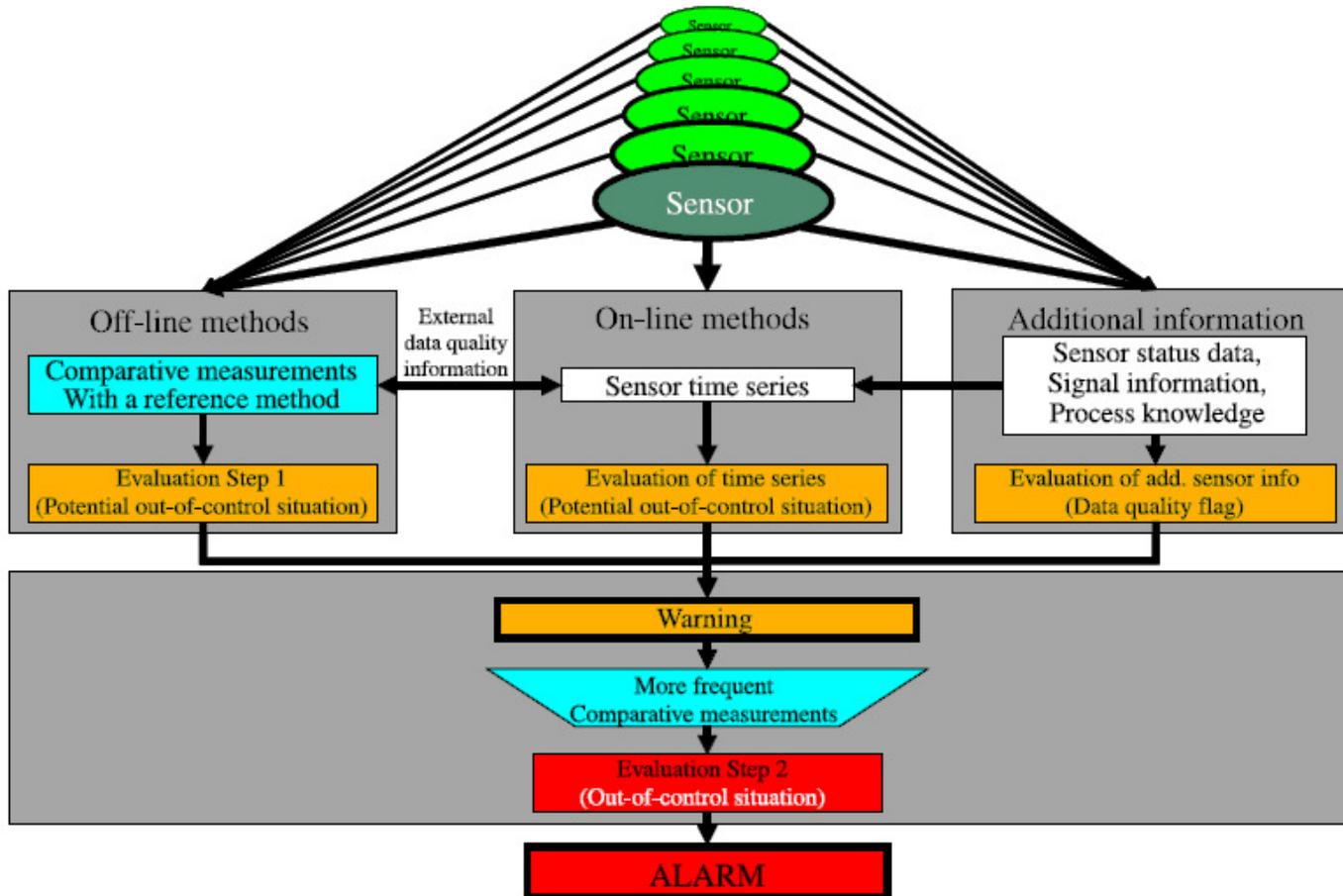


Figure 9 : Concept d'évaluation univariée/multivarié de la qualité des données (Rieger et Vanrolleghem, 2008)

Branisavljevic et al. (2010) proposent une procédure d'évaluation de la qualité des données relativement similaire à celle suggérée par Rieger et Vanrolleghem (2008). En ce sens qu'elle est aussi basée sur une combinaison de méthodes en ligne et hors ligne appliquées à une même série de données afin d'en augmenter les performances. Cependant, contrairement à Rieger et Vanrolleghem (2008), cette procédure n'exploite pas un concept d'entretien proactif. Cette dernière est schématisée à la figure 10. Brièvement, cette procédure d'évaluation de la qualité des données se divise en trois étapes principales, soit la préparation et la distribution des données aux différents tests, la génération des scores de validation et leur interprétation. Chaque paramètre de mesure peut posséder sa propre procédure d'évaluation afin qu'elle soit optimale au paramètre en question.

La procédure débute avec la collecte et l'assimilation de données provenant de différentes sources. Un traitement de base est effectué sur les données (ex. une vérification simple) et ces dernières sont stockées dans une base de données temporaire. Le gestionnaire des données distribue les données et les métadonnées sur demande de la base de données temporaire aux modules de validation des données. La validation des données est précédée d'un traitement des données qui ajuste les données selon les besoins des méthodes de validations, de configurations et de réglages.

Dans l'étape de la génération d'un score de validation, les données passent au travers d'une série de méthodes de validation selon un ordre spécifique, comme illustré à la figure 11 . Il existe trois catégories de méthodes de validation, soit les méthodes liées qui sont liées à d'autres méthodes selon un ordre d'exécution précis, les méthodes dépendantes qui peuvent être exécutées en parallèle, mais qui dépendent des résultats d'autres méthodes, et les méthodes indépendantes. Le score de validation de chaque méthode est attaché à la donnée en tant que métadonnée et est stocké dans la base de données.

Finalement, les scores de validation sont interprétés pour produire un score final qui indique la validité de la donnée mesurée.

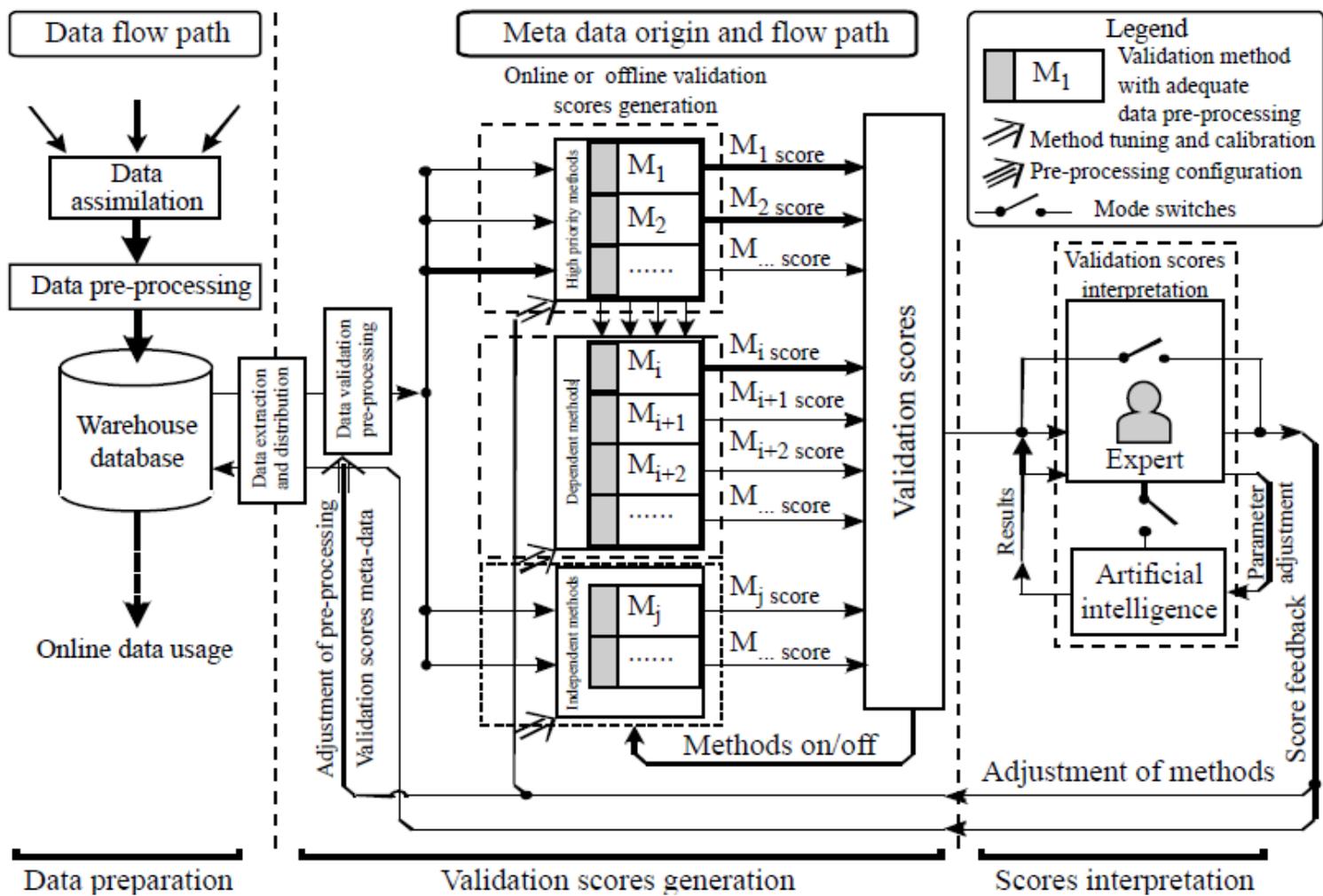


Figure 10 : Procédure de validation des données de drainage urbain  
(Branisavljevic et al., 2010)

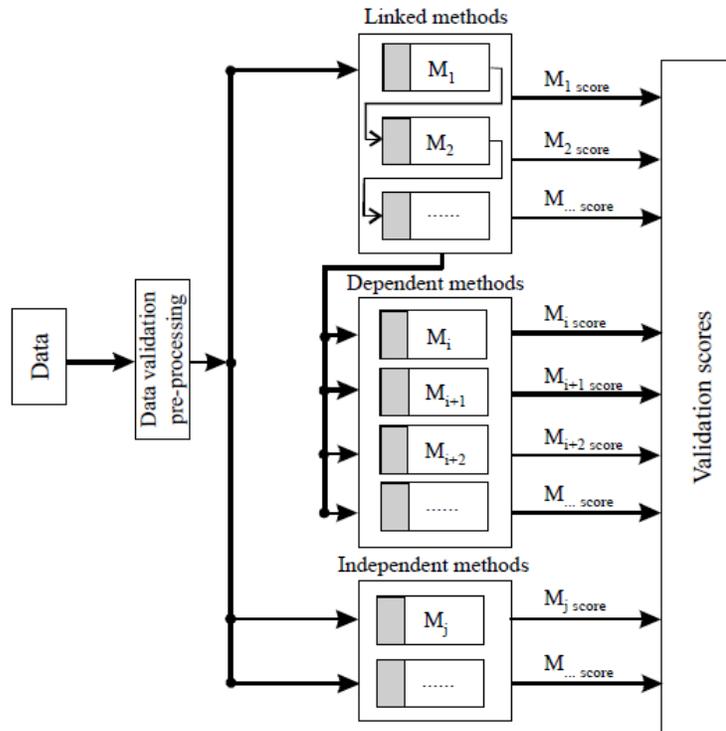


Figure 11 : Génération des scores de validation  
(Branisavljevic et al., 2010)

L'utilisation d'un ensemble de méthodes, telle que décrit précédemment, pour évaluer la qualité des données est préconisée par plusieurs chercheurs. Par exemple, certains suggèrent un ensemble de méthodes d'évaluation appliqué aux données de débit dont une méthode est basée sur un modèle physique en temps réel (Krajewski et Krajewski, 1989) appliqué aux données de drainage urbain avec quelques méthodes multivariées basées sur des mesures redondantes (Mourad et Bertrand-Krajewski, 2002, Bertrand-Krajewski et al., 2000) ou basée sur l'analyse en composantes principales (Branisavljevic et al., 2010), etc.

Les sections suivantes présentent, en premier lieu, des méthodes hors ligne et, en deuxième lieu, des méthodes en ligne, qui sont définies respectivement par l'auteur de ce mémoire comme des méthodes nécessitant une intervention de l'utilisateur et utilisant seulement l'information contenue dans les séries temporelles.

### 2.3.1 Les méthodes hors ligne

Les méthodes hors ligne d'évaluation de la qualité des données rencontrées dans le domaine de la qualité de l'eau sont principalement les chartes de contrôle. Cependant, les hypothèses permettant leur utilisation ne sont généralement pas respectées pour des séries temporelles environnementales. Ces hypothèses sont que les erreurs sur les données mesurées soient indépendantes et aléatoires, et possèdent une moyenne et une variance constantes. Il est donc nécessaire de modifier les chartes de contrôle rencontrées dans le domaine industriel telles que les chartes « Shewart », « cumulative sum », « moving average » et « exponentially weighted moving average » (EWMA), afin de pouvoir les employer dans le domaine de la qualité de l'eau. Ces modifications consistent principalement à utiliser les résidus entre les valeurs mesurées par un capteur et les valeurs obtenues par une méthode redondante quelconque au lieu d'utiliser directement les valeurs mesurées.

Le concept de base permettant l'exploitation de chartes de contrôle est que tout processus possède une variabilité. Cette variabilité est composée d'une variabilité résultante des causes normales du système communément appelée le bruit, et d'une variabilité identifiable qui est produite par une cause spéciale. Le processus est dit en équilibre statistique lorsque les causes de variabilité identifiable ont été détectées, identifiées et éliminées (Nelson, 1984). De plus, l'hypothèse que les observations futures possèdent les mêmes propriétés statistiques que les observations passées doit être respectée. Cette hypothèse permet l'exploitation de limites de contrôle basées sur la performance passée du processus pour évaluer si les conditions futures sont en équilibres statistiques. Ainsi, à l'aide de charte de contrôle, il est possible de détecter un changement dans la moyenne (valeur cible) ou la variabilité du processus (Berthouex, 1989). Une description détaillée des différents types de chartes de contrôle se trouve dans Montgomery (2001).

Deux types de charte de contrôle rencontrés dans le domaine du suivi en temps réel de la qualité de l'eau sont les chartes « Shewart » comparant des valeurs d'échantillons de référence analysés en laboratoire (Thomann et al., 2002) ou des valeurs de solutions standards (NMKL, 1990) (valeurs de référence) à leur mesure avec un capteur. Ainsi, si le processus est en équilibre statistique, les résidus entre les mesures du capteur et les valeurs

de référence respectent les hypothèses énumérées précédemment, c'est-à-dire que les résidus sont indépendants, aléatoires et ont une moyenne et une variance constantes.

Thomann et al. (2002) exploitent une méthodologie basée sur la charte de contrôle « Shewart » permettant de détecter des effets de dérive, de saut ou de valeur aberrante dans les données tels qu'illustrés à la figure 12. De plus, ils emploient une analyse de régression des échantillons de référence et de leur mesure avec un capteur afin de détecter une mauvaise courbe de calibration. Ce concept est divisé en plusieurs phases et est illustré à la figure 13.

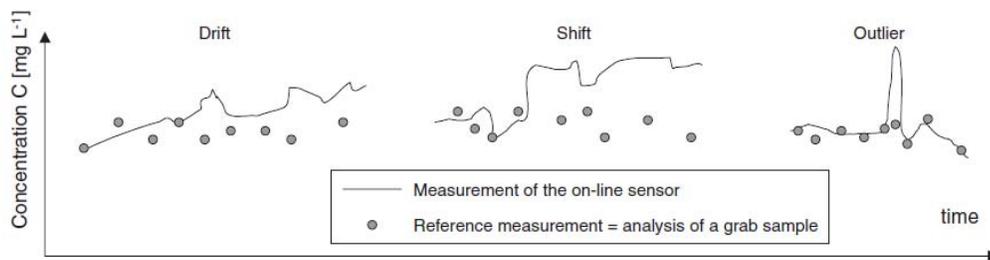


Figure 12 : Effets de dérive, de saut ou de valeur aberrante (Thomann et al., 2002)

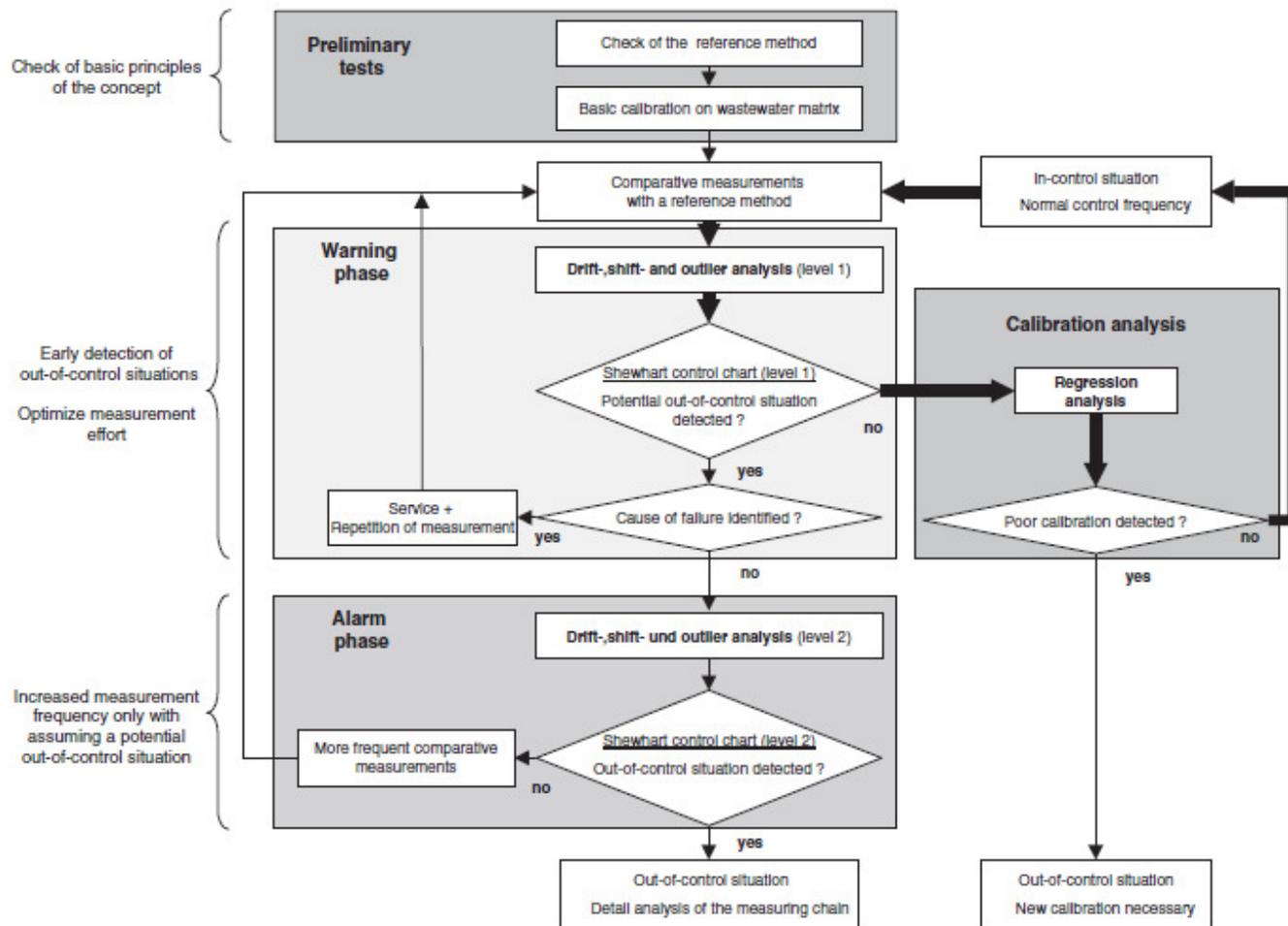


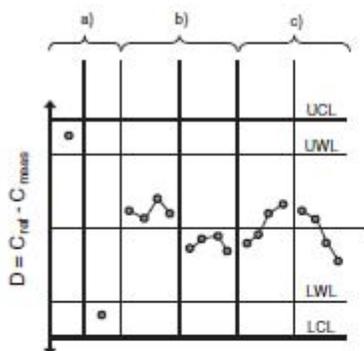
Figure 13 : Approche de suivi pour les capteurs en ligne. Les flèches en gras indiquent la séquence des actions à effectuer si le processus est présumé en contrôle. (Thomann et al., 2002)

Dans la phase des tests préliminaires, une analyse des principes de base tels que la précision et la calibration des capteurs, etc. est tout d'abord effectuée. Ensuite, dans la phase d'avertissement, les résidus entre les échantillons de référence et leur mesure par un capteur sont étudiés à l'aide d'une charte de contrôle « Shewart » et de critères d'« out-of-control ». Si ces critères ne sont pas respectés, la cause de ce non-respect est recherchée et un entretien est effectué, par exemple un nettoyage du capteur. Un échantillon de référence est immédiatement pris après cet entretien.

Si la situation d'avertissement demeure, il y a une situation possible d'« out-of-control » et la phase d'alarme est enclenchée. La fréquence d'échantillonnage est donc augmentée et ne change pas tant qu'il n'y a pas une situation d'« out-of-control » de détectée ou que les critères d'« out-of-control » soient respectés à nouveau. D'ailleurs, si ces critères sont respectés, ce qui signifie qu'il n'y a pas d'effets de dérive, de saut ou de valeur aberrante dans les données, une analyse de régression est effectuée sur les échantillons de référence et leur mesure par un capteur afin de déterminer si la courbe de calibration est bonne. La figure 14 montre les phases d'avertissement et d'alarme ainsi que leurs critères d'« out-of-control ». Ce concept à deux phases permet de détecter une situation d'« out-of-control » rapidement et d'optimiser l'effort investit (Thomann et al., 2002).

### warning phase

- a) 1 point outside of the warning limits
- b) 4 consecutive points on one side of the center line
- c) 4 points in a row steadily increasing or decreasing



### alarm phase

- a) 1 point outside of the control limits
- b) 2 of 3 consecutive points outside the warning limits but still inside the control limits
- c) 8 consecutive points on one side of the center line
- d) 6 points in a row steadily increasing or decreasing

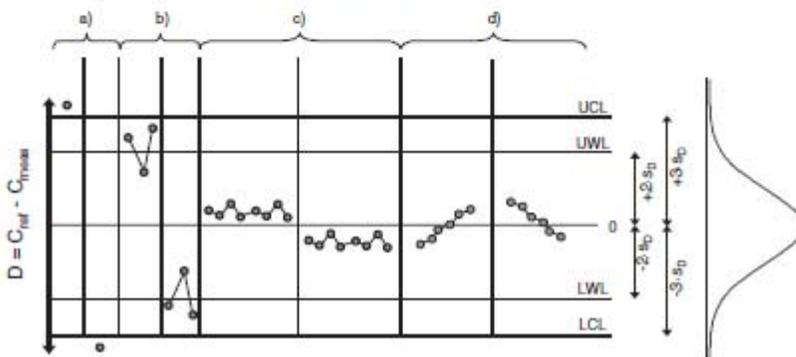


Figure 14 : Charte de contrôle « Shewart » avec des critères de « out-of-control » plus sévères pour la phase d'avertissement et ceux normalement utilisés pour la phase d'alarme (Thomann et al., 2002)

La charte de contrôle proposée par NMKL (1990) est similaire à celle décrite précédemment sauf que la variable de contrôle est la mesure d'une solution standard. Cela a donc les désavantages d'exclure les effets des grandeurs d'influence et de nécessiter la sortie des capteurs hors du milieu pour effectuer la mesure. Ce type de charte de contrôle permet aussi de détecter des erreurs systématiques et d'identifier si un entretien ou une calibration est nécessaire. Cependant, contrairement au concept décrit précédemment, la méthodologie proposée par NMKL (1990) n'est pas divisée en deux phases principales.

Berthouex (1989) propose de construire des chartes de contrôle (les chartes « Shewart », « cumulative sum », « moving average » et EWMA) à partir des résidus calculés entre les valeurs prédites par un modèle et les valeurs mesurées pour évaluer les performances d'une usine de traitement des eaux usées. Même si Berthouex (1989) a démontré l'efficacité de cette méthode à l'aide d'échantillons analysés en laboratoire, elle pourrait être appliquée aux données mesurées en temps réel pour détecter des fautes. En effet, si les propriétés statistiques des résidus changent, cela signifie qu'il y a un comportement anormal dans les données mesurées. Ce comportement anormal pourrait être causé par un facteur influençant la qualité des données tel que l'encrassement du capteur.

Le modèle que Berthouex (1989) préconise pour modéliser les séries temporelles est un modèle autorégressif à moyenne mobile intégrée (ARIMA). Il faut tout d'abord identifier le modèle ARIMA approprié afin que les résidus respectent les hypothèses nécessaires à l'utilisation des chartes de contrôle.

Cette méthode aurait pu être décrite dans la section 2.3.2 Les méthodes en ligne, étant donné qu'elle pourrait être appliquée en temps réel. D'ailleurs, des méthodes relativement similaires à celle-ci y sont discutées.

### **2.3.2 Les méthodes en ligne**

Les méthodes en ligne utilisent l'information contenue dans les séries temporelles. Il y a donc une quantité innombrable de méthodes permettant d'évaluer la qualité des données. Comme mentionné précédemment, les fautes sont des indicateurs de la qualité des données et elles se reflètent dans les caractéristiques des données mesurées. Les caractéristiques des données telles que la moyenne, la pente, la variance, la distance à une autre lecture (méthode redondante), etc. peuvent donc être utilisées pour détecter des fautes contenues dans les séries de données. Le tableau 2 présente des fautes pouvant être observées lors d'un suivi en temps réel, ainsi que les indicateurs permettant de les détecter et leurs causes possibles.

Tableau 2 : Taxonomie des fautes, définitions et causes possibles

(traduction française de Ni et al., 2009)

<b>Faute</b>	<b>Définition</b>	<b>Indications et causes possibles</b>
Donnée aberrante	Donnée isolée ou anormalement distante des modèles.	La distance entre d'autres mesures est plus grande que prévue. Le gradient change considérablement lorsque la donnée aberrante est incluse. Les causes sont généralement inconnues à moins qu'un logiciel soit inséré dans l'enregistreur de données.
Pointe	Plusieurs données avec un gradient plus élevé que prévu.	Un changement soudain dans le gradient lequel est plus élevé que prévu. Faible corrélation temporelle entre les données historiques et la pointe. Les causes fréquentes incluent une défaillance dans les batteries, le matériel et les connexions.
«Coincé à»	Les données mesurées ne possèdent aucune variation pour une période de temps indéterminée.	La variance se situe près de zéro ou est égale à zéro. Une corrélation spatiale peut être utilisée pour déterminer si un «Coincé à» est une faute. La cause de cette faute est fréquemment un mauvais fonctionnement des composantes du capteur.
Bruit ou variance élevé	Les données montrent une variance ou un bruit anormalement élevé.	La variance est plus élevée que prévue ou que démontrée par les modèles historiques. Une corrélation spatiale peut être utilisée pour évaluer si la variation est due à l'environnement. Ceci peut être causé par une défaillance dans le matériel, un environnement en dehors de la gamme de mesure, ou un faible niveau d'alimentation de la batterie.
Calibration	Les données mesurées sont décalées des valeurs réelles.	L'erreur de calibration et la dérive du capteur est la principale cause de cette faute. Le capteur peut être décalé ou avoir un gain différent des valeurs réelles. La différence peut dériver dans le temps.
Connexion et matériel	Un mauvais fonctionnement dans les composantes du capteur qui cause la collecte de données inexactes.	Le comportement des données est dépendant du matériel. Les caractéristiques communes incluent des données anormalement faibles ou élevées, dépassement fréquent de la gamme prévue. Les perturbations environnementales et l'âge du capteur peut engendrer une plus grande probabilité d'erreur. Les autres causes incluent un court-circuit ou la perte d'une connexion.
Batterie faible	Baisse dans le voltage des batteries à un niveau que le capteur ne peut pas mesurer des données exactes.	L'état de la batterie est un indicateur des performances d'un système. Le comportement général inclue un gradient imprévu suivi par un manque de données ou une variance égale à zéro. Il peut avoir un bruit excessif.
Environnement hors-gamme	L'environnement dépasse la gamme de mesure du capteur.	Il peut y avoir un bruit plus élevé ou un aplatissement des données. Cela peut être aussi un indice d'une mauvaise calibration.
Coupure	Le capteur atteint le minimum ou le maximum aux limites du convertisseur analogique-numérique (ADC)	Les données montrent un minimum ou un maximum aux limites. Ceci est causé par un environnement qui dépasse la gamme du ADC.

Selon le tableau 2, une des caractéristiques de données qui peut être exploitée pour détecter des données aberrantes (« outliers ») est la distance entre la lecture d'un capteur à une autre

lecture. Par ailleurs, Hill et Minsker (2010) utilisent le terme anomalie plutôt que donnée aberrante et ils définissent une anomalie comme étant une donnée qui dévie des tendances historiques. Tel qu'illustré à la figure 15, une approche empruntée par certains chercheurs pour détecter des données aberrantes est de comparer les données mesurées à des données prédites à l'aide d'un modèle. Si les données mesurées se situent à l'extérieur d'un intervalle de prédiction, elles sont considérées comme étant aberrantes.

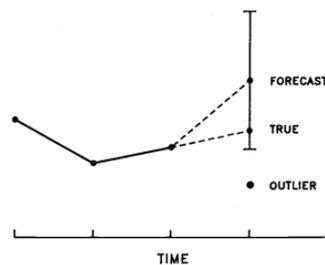


Figure 15 : Schéma de l'utilisation d'un modèle de prédiction pour détecter des données aberrantes (Krajewski et Krajewski, 1989)

Par exemples, Krajewski et Krajewski (1989) exploitent un modèle physique, tandis que Hill et Minsker (2010) préconisent des modèles autorégressifs univariés. Dans les deux cas, l'intervalle de prédiction est défini par l'écart-type de l'erreur. Clement et al. (2007) suggèrent l'utilisation d'un modèle non paramétrique multivarié additif, nommé modèle additif, et d'un intervalle de prédiction défini à l'aide d'un « bootstrap » sur l'erreur de prédiction. Ce type d'intervalle de prédiction ne nécessite pas que l'erreur de prédiction suive une distribution Gaussienne et a donc comme avantage d'être plus robuste face à une déviation de l'erreur par rapport à une distribution normale.

Le modèle autorégressif univarié est intéressant sous plusieurs aspects. Contrairement au modèle physique, il ne nécessite qu'une seule variable d'entrée et il est rapidement employable. À l'inverse du modèle multivarié, le modèle univarié permet d'éviter les complications causées par des fréquences de mesure différentes dans les séries de données provenant de différents capteurs. De plus, lorsqu'il y a des données manquantes ou aberrantes dans les séries temporelles, cela complique la prédiction des données. L'utilisation d'un

modèle autorégressif univarié diminue l'effet de ces données manquantes ou aberrantes en exigeant seulement une série temporelle.

D'ailleurs, Hill et Minsker (2010) étudient deux possibilités soit d'introduire les données aberrantes (anomalies) dans le modèle ou de les remplacer par les données prédites pour calculer l'intervalle de prédiction au pas suivant.

D'autres méthodes permettent de détecter des anomalies dans les données telles que le « minimum volume ellipsoid » (Rousseeuw et Leroy, 1996), le « convex peeling » (Rousseeuw et Leroy, 1996), le « nearest neighbor » (Tang et al., 2002, Ramaswamy et al., 2000), le « clustering » (Bolton et Hand, 2001), le « neural network classifier » (Kozuma et al., 1994), le « support vector machine classifier » (Bulut et al., 2005), et le « decision tree » (John, 1995). Cependant, ces méthodes ont le désavantage d'être inutilisable en temps réel, soit qu'elles nécessitent que toutes les données soient récoltées avant d'être appliquées, qu'elles exigent un temps de calcul énorme ou qu'elles nécessitent une pré-classification des données (anormal/normal) laquelle caractérise les types d'anomalie pouvant être détectés.

Les fautes sont aussi détectées à l'aide de modèles prédisant le comportement prévu des données en définissant un intervalle dans lequel une caractéristique des données calculée (pente, variance, corrélation) devrait se situer (intervalle d'acceptabilité). Il est possible de prédire un intervalle d'acceptabilité pour une caractéristique à l'aide de la moyenne et de la variance d'un échantillon de données, de la médiane et des quartiles, de modèles de régression ou d'autres modèles plus complexes (Ni et al., 2009). Certains chercheurs définissent simplement un intervalle d'acceptabilité en choisissant des percentiles (ex. 0,01<sup>th</sup> et 99,99<sup>th</sup>) des caractéristiques calculées sur des séries de données historiques (Mourad et Bertrand-Krajewski, 2002, Krajewski et Krajewski, 1989).

Mourad et Bertrand-Krajewski (2002) ont développé une méthode permettant de détecter des pentes anormales dans les séries de données causées par la présence de données aberrantes. Cette méthode consiste à filtrer les séries à l'aide d'une moyenne à fenêtre mobile et de calculer les résidus entre les séries originales et filtrées. Le test de Page-Hinkley (Ragot et al., 1990) est ensuite utilisé pour détecter les valeurs de résidu anormalement élevées.

De nombreux chercheurs exploitent les analyses multivariées pour évaluer la qualité des données mesurées. Par exemple, Mourad et Bertrand-Krajewski (2002) proposent d'utiliser

deux capteurs en redondance pour détecter des tendances ou des écarts anormales dans les données. Les résidus entre les données de ces deux capteurs servent d'indicateur pour détecter des fautes. Les résidus peuvent être comparés à une valeur seuil ou des méthodes comme l'algorithme Page-Hinkley peuvent être employées pour détecter des valeurs anormalement élevées (Brunet et al., 1990). Cependant, il peut être difficile de déterminer quel capteur parmi les deux mesure des données de mauvaise qualité. L'utilisation de trois capteurs en redondance permet donc une détection de fautes de façon plus efficace (Menendez Martinez et al., 2001). Une méthode similaire à la précédente est d'exploiter des paramètres de mesure qui sont corrélés, par exemple, la hauteur d'eau et la vitesse d'écoulement (Mourad et Bertrand-Krajewski, 2002).

Outre ces méthodes multivariées relativement simples, il en existe des plus complexes. Ces méthodes ne sont pas décrites dans cette section, car elles ne correspondent pas à un des objectifs de ce mémoire qui est de construire des ADQATs basés sur des méthodes univariées.



### 3. Matériels et méthodologie

#### 3.1 Description du bassin versant

Le bassin versant à l'étude choisi pour l'installation de la station de mesure monEAU est celui du ruisseau Notre-Dame situé dans la municipalité de L'Ancienne-Lorette. La figure 16 montre une portion du bassin versant du ruisseau Notre-Dame et la figure 17 est un agrandissement du rectangle noir de la figure 16. La flèche noire sur la figure 17 indique le ruisseau Notre-Dame. Ce bassin versant se divise principalement en deux sous-bassins versants, soit le sous-bassin versant de l'aéroport Jean-Lesage, en amont du ruisseau, et le sous-bassin versant urbain situé le long du ruisseau. Le tableau 3 présente leur superficie ainsi que leur pourcentage d'imperméabilité. Les données présentées dans cette section ont été obtenues par la compagnie BPR (Québec, Qc, Canada).

La section du ruisseau présentée sur la figure 17 possède une longueur d'environ 2900 m. Ce ruisseau est fortement influencé par un réseau d'égout pluvial drainant l'eau de ruissellement du sous-bassin versant urbain et il est un affluent de la rivière Lorette. Il est important de mentionner la présence d'un cimetière d'automobile en amont près de la route de l'Aéroport.

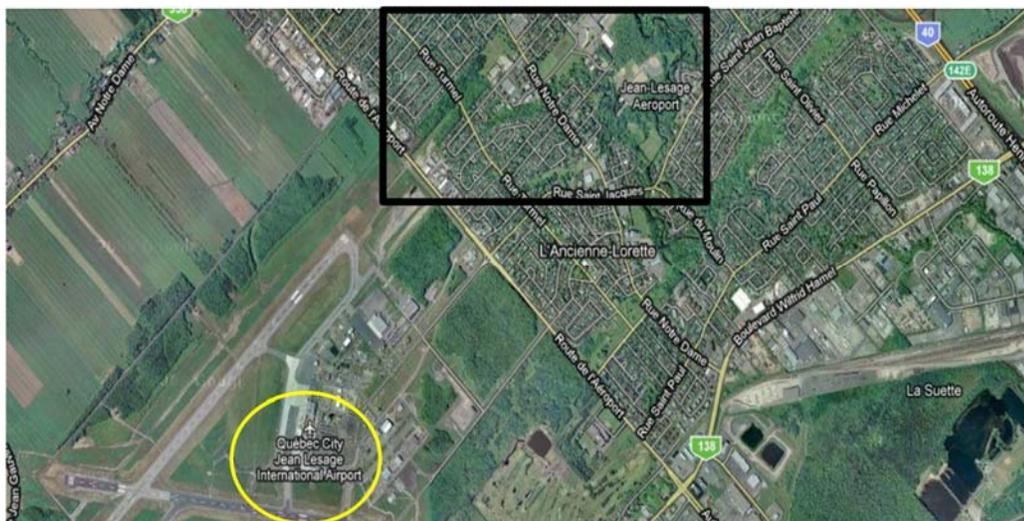


Figure 16 : Bassin versant du ruisseau Notre-Dame

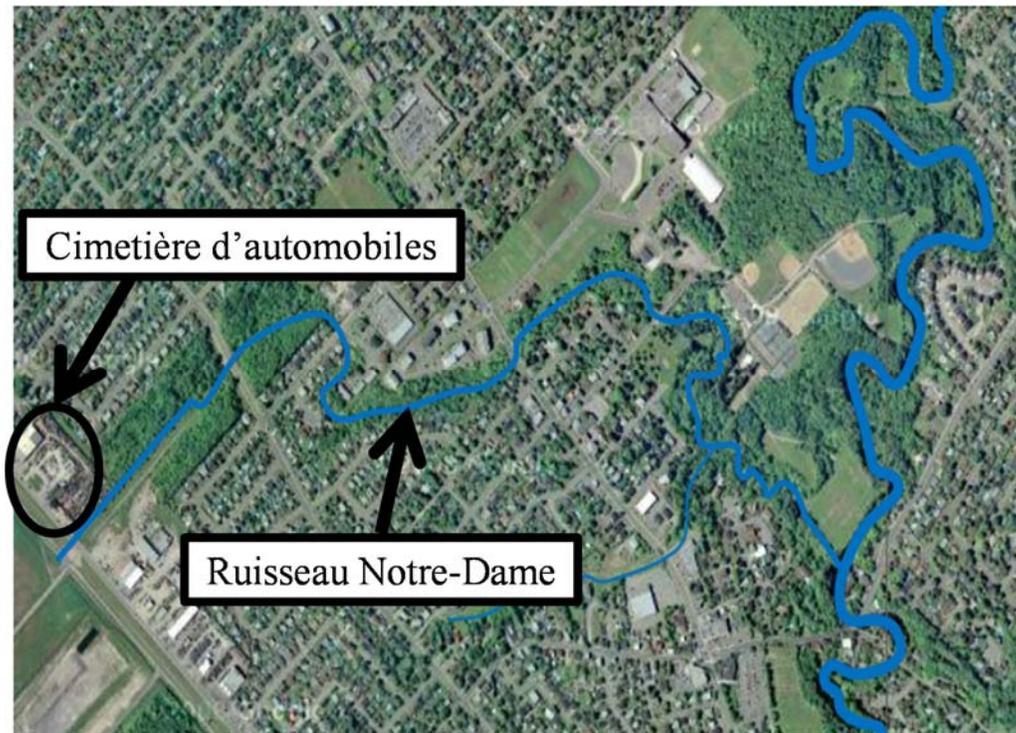


Figure 17 : Agrandissement du bassin versant du ruisseau Notre-Dame

Tableau 3 : Superficie des sous-bassins versants et leur pourcentage d'imperméabilité

Sous-bassin versant	Aire (ha)	Imperméabilité (%)
Total	561	-
De l'aéroport	275	8,5
Urbain	239	42
Parc	47	1

### 3.2 La station monEAU

Un prototype de la station monEAU, comme discutée dans le chapitre 2 Revue de littérature, a été utilisé lors du suivi en temps réel. Cette section décrit donc le concept du système monEAU expliqué par Rieger et Vanrolleghem (2008), ainsi que les composantes du prototype de la station monEAU. La figure 18 illustre ce concept et représente un réseau de suivi de la qualité de l'eau. Il est important de souligner que la station monEAU est commercialisée sous le nom RSM30 par la compagnie Primodal Inc. (Hamilton, Ontario, Canada) et qu'elle est en constant développement et amélioration dans le but de répondre aux objectifs de Rieger et Vanrolleghem (2008), qui sont discutés ci-après.

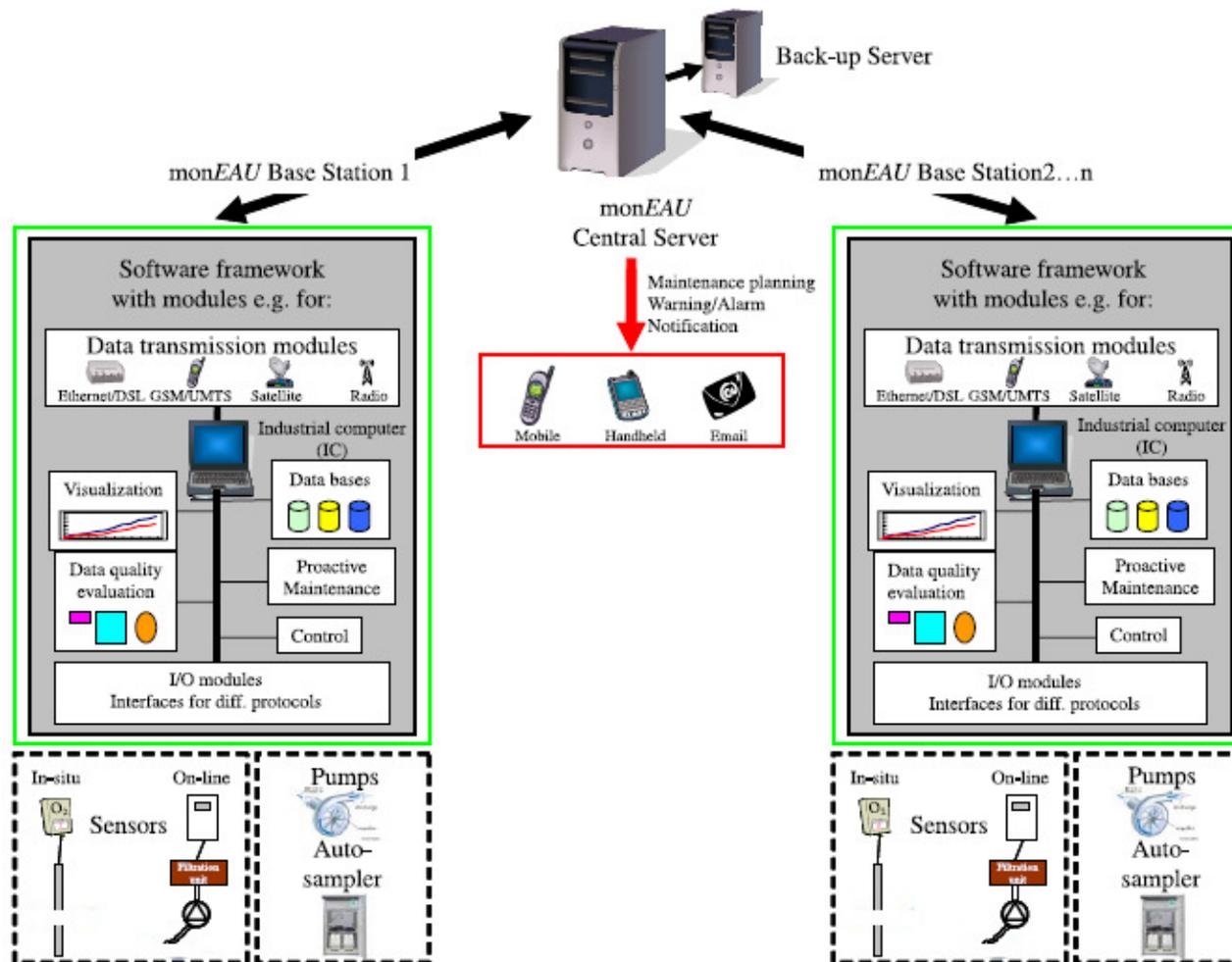


Figure 18 : Montage du réseau de suivi de la qualité de l'eau, le système monEAU (Rieger et Vanrolleghem, 2008)

Le système monEAU possède une flexibilité permettant son utilisation pour différents objectifs de suivi et de recherche, à différents emplacements (rivière, STEP, égout, etc.) et avec différents types de capteurs et de méthodes de mesure (in situ, en ligne et hors-ligne). Plusieurs protocoles de communication sont disponibles pour les capteurs de la qualité de l'eau (Profibus, HART, 4–20 mA, USB, etc.) et la transmission des données entre les capteurs et la station se fait via un protocole bus permettant l'exploitation de toute l'information provenant des capteurs (métadonnées) et une communication bidirectionnelle. Des connexions 4-20mA sont aussi fournies comme solution de remplacement.

Le système monEAU permet une utilisation à distance étant donné que les stations ne sont pas nécessairement installées à des endroits faciles d'accès. Sa conception considère donc la fréquence des exigences d'entretien, la minimisation de la demande d'énergie en combinant différentes options d'alimentation (ex. énergie solaire), différentes options de télémétrie (ligne téléphonique, DSL, câble TV, SMS, etc.) et un accès à distance des capteurs et du fonctionnement de la station de mesure.

Le logiciel monEAU est ouvert et modulaire, c'est-à-dire qu'il est composé d'un logiciel principal possédant des fonctionnalités de bases, où il peut y être ajouté des modules par l'entremise d'un API spécifique (« Application Programming Interface »). Le système peut être adapté aux demandes spéciales telles que l'ajout d'outils d'évaluation et de visualisation des données, etc. Ces modules peuvent être écrits dans un langage standard de programmation et, par conséquent, il est facile pour l'utilisateur d'augmenter les fonctionnalités d'une station. Le logiciel monEAU est orienté vers l'utilisateur, c'est-à-dire que l'information disponible dépend du niveau de l'utilisateur (opérateur, expert, etc.) et de l'emplacement (à la station ou au serveur central).

L'exportation des données se fait à l'aide du format standard « IEEE 754 floating point » garantissant la lecture des données par différents logiciels et permettant l'utilisation ultérieure des données. Différentes mesures de sécurité sont mises en place afin de garantir la sécurité des données tant au niveau de la station qu'au niveau du serveur central. En effet, un réseau redondant de disques indépendants au serveur central (RAID niveau 5) et un disque dur avec une mémoire suffisante à la station sont utilisés pour combler les ruptures de communication entre la station et le serveur central et pour stocker les données brutes, traitées

(évaluation de la qualité des données), condensées (moyennes) et dérivées (charge, temps de rétention, etc.).

Le système *monEAU* fournit des outils automatiques d'évaluation de la qualité des données, ce qui ajoute un aspect de qualité à la base de données. Ces outils sont basés sur trois sources d'information soit des échantillons de référence qui sont comparés aux mesures des capteurs, des données sur l'état et les diagnostics des capteurs et de l'information contenue dans les séries temporelles. Plus précisément, le système *monEAU* exploite un concept général d'évaluation de la qualité des données pouvant être appliqué à différents emplacements (rivière, STEP, égout, etc.). Il permet aussi la possibilité d'intégrer des méthodes spécifiques à un emplacement basées sur le savoir des experts. Trois niveaux d'évaluation de la qualité des données peuvent être sélectionnés dans le logiciel *monEAU*:

- Niveau 1 : Méthodes univariées;
- Niveau 2 : Méthodes multivariées d'analyses des séries temporelles sans connaissance du processus ou d'un expert;
- Niveau 3 : Méthodes d'évaluation avancées incluant les connaissances du processus et d'un expert.

Les méthodes d'évaluation de la qualité des données développées dans ce mémoire consistent au niveau 1 et elles sont basées principalement sur la troisième source d'information.

La base du concept *monEAU* est que le corps de la station soit le même pour toutes les stations. Cependant, la configuration finale peut varier d'une station à une autre attribuable aux contraintes d'espace, d'énergie et aux conditions environnementales. C'est-à-dire que les options d'alimentation électrique, de transmission des données, etc. peuvent différer. Les composantes du prototype de la station *monEAU* qui a été utilisé lors du suivi en temps réel de la qualité de l'eau du ruisseau Notre-Dame sont donc décrites ci-après.

La station *monEAU* qui a été installée pour le suivi en temps réel du ruisseau Notre-Dame est le troisième prototype développé par la compagnie Primodal. La figure 19 est une photographie des composantes matérielles du prototype 3 de la station *monEAU*. La description du logiciel *monEAU* présentée ci-après n'est plus totalement représentative, car des modifications y ont été apportées. D'ailleurs, certaines modifications pour améliorer le

logiciel monEAU ont été suggérées par l'auteur de ce mémoire et sont présentées à la section 5.5 Amélioration du logiciel monEAU.

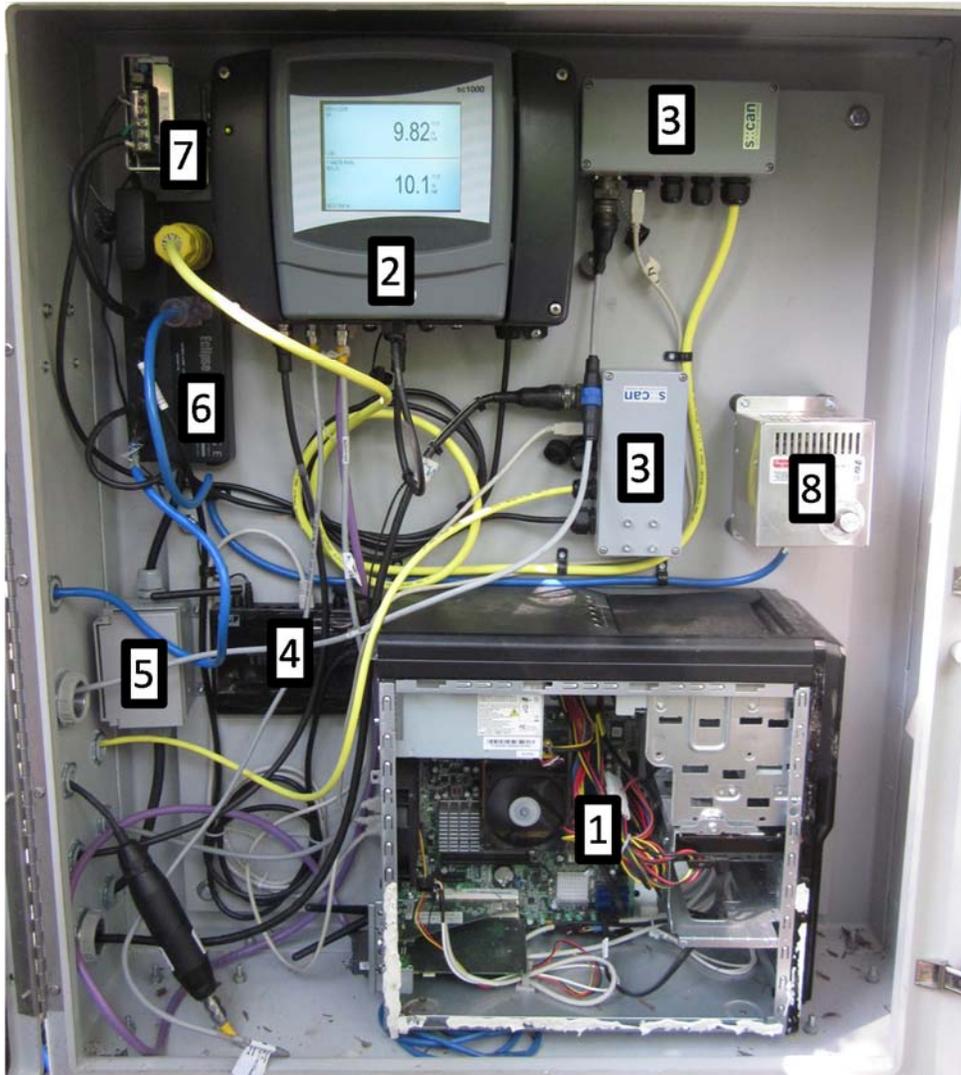


Figure 19 : Composantes matérielles du prototype 3 de la station monEAU

Les pièces numérotées sur la figure 19 sont:

- 1) Ordinateur de la station monEAU;
- 2) Transmetteur sc1000 de Hach;
- 3) Terminal con::nect de s::can;
- 4) Routeur ;
- 5) Disjoncteur GFCI;
- 6) Barre d'alimentation;
- 7) Transformateur 120V-12V pour le transmetteur sc1000 de Hach;
- 8) Appareil de chauffage.

Le boîtier de protection du prototype 3 de la station *monEAU* contenant toutes les composantes matérielles est composé de fibre de verre et est hermétique afin que l'humidité n'affecte pas le fonctionnement de la station de mesure. Le type d'alimentation électrique de la station est une source de courant électrique de 120V. Les capteurs qui peuvent y être connectés sont des capteurs Hach et *s::can*. Cependant, les capteurs *s::can* ne sont pas encore pris en charge par le logiciel *monEAU*. Il faut donc utiliser le logiciel fourni par la compagnie *s::can*. Le téléchargement des données et l'accès à distance de la station se font via un réseau sans fil. La communication bidirectionnelle entre les capteurs et la station n'est pas encore exécutable, en ce sens que le transfert des données se fait seulement du capteur au logiciel *monEAU* et qu'il n'est pas possible, par exemple, de calibrer le capteur à partir du logiciel *monEAU*.

Le serveur central peut être représenté par l'ordinateur portable de l'utilisateur. Aucun RAID n'est exploité pour garantir la sécurité des données. La mémoire du disque dur de la station, qui est de 698 GB, est suffisamment élevée pour stocker une quantité innombrable de données. Par exemple, durant le suivi en temps réel, seulement 234 MB ont servis à stocker les données mesurées. De plus, les données peuvent être sauvegardées dans des fichiers CSV, XML, Microsoft Excel Workbook (.xls) et Microsoft Excel Spreadsheet (.xml). Un module de visualisation des données permet d'afficher un paramètre ou plusieurs paramètres simultanément. La mesure de chaque paramètre se fait à un intervalle fixe et ce dernier peut être différent pour chaque paramètre. Lors du suivi en temps réel, le logiciel *monEAU* ne possédait pas encore de modules d'évaluation de la qualité des données et de filtration des données.

### **3.3 Capteurs et paramètres mesurés**

Une étude préliminaire et un suivi en temps réel ont été réalisés dans le cadre de ce mémoire. Brièvement, l'étude préliminaire consistait en une campagne d'échantillonnage du ruisseau Notre-Dame et le suivi en temps réel consistait à mesurer la qualité de l'eau du ruisseau Notre-Dame à l'aide de capteurs. Les chapitres 4 Étude préliminaire et 5 Suivi en temps réel décrivent en détails ces deux études.

Durant l'étude préliminaire et le suivi en temps réel, certains paramètres ont été analysés en laboratoire. Ces données ont permis principalement d'avoir une idée générale du comportement de la qualité de l'eau lors de l'étude préliminaire et de calibrer certains capteurs lors du suivi en temps réel. Le tableau 4 et le tableau 5 présentent respectivement les paramètres analysés lors de l'étude préliminaire et du suivi en temps réel ainsi que leur méthode d'analyse correspondante.

Tableau 4 : Paramètres analysés lors de l'étude préliminaire et leur méthode d'analyse

Paramètre	Unité	Méthode d'analyse
MeS	mg/L	Méthode 2540 D du STANDARD METHODS
Pathogènes (E. coli et coliformes totaux)	UFC/100mL	Méthode Coliscan Easygel
Ammonium	mg/L NH <sub>3</sub> -N	Hach TNT plus 830 Méthode 10205
Nitrate	mg/L NO <sub>3</sub> -N	Hach TNT plus 835 Méthode 10206
Azote total	mg/L N	Hach TNT plus 826 Méthode 10208
Orthophosphate	mg/L PO <sub>4</sub> <sup>3-</sup>	Hach TNT plus 844 Méthode 10209 Reactive
Phosphore total	mg/L PO <sub>4</sub> <sup>3-</sup>	Hach TNT plus 844 Méthode 10210 Total
pH	-log[H <sup>+</sup> ]	pH/ionmètre
Métaux lourds totaux (Cu, Mn et Zn)	mg/L	ILQ-MÉTAUX-Détermination des métaux dans l'eau potable et l'eau usée

Tableau 5 : Paramètres analysés lors du suivi en temps réel et leur méthode d'analyse

Paramètre	Unité	Méthode d'analyse
MeS	mg/L	Méthode 2540 D du STANDARD METHODS
Nitrate	mg/L NO <sub>3</sub> -N	Hach TNT plus 835 Méthode 10206
Potassium	mg/L K <sup>+</sup>	Hach Powder pillows Méthode 8049
Ammonium	mg/L NH <sub>3</sub> -N	Hach TNT plus 830 Méthode 10205

Une sonde YSI 6600EDS mesurant la température, la conductivité et la turbidité in situ a été utilisée lors de l'étude préliminaire. Par ailleurs, le pH a été analysé en laboratoire, cependant il est suggéré de mesurer le pH in situ avec un capteur. En effet, la température et la

concentration de gaz dissous influençant l'acidité peuvent varier entre l'échantillonnage et l'analyse de l'échantillon en laboratoire (APHA, 2005).

Pour déterminer le débit durant l'étude préliminaire, la hauteur d'eau a été mesurée à chaque échantillonnage avec une règle graduée. Elle était mesurée à partir d'un point fixe, une brique ancrée sur le lit du ruisseau. La méthodologie permettant de jauger le débit est décrite à la section 3.4 Jaugeage du débit.

Les capteurs installés lors du suivi en temps réel ont permis de mesurer plusieurs paramètres simultanément avec une haute fréquence de mesure. Le tableau 6 présente ces capteurs ainsi que leurs paramètres de mesure.

Tableau 6 : Capteurs utilisés lors du suivi en temps réel et leurs paramètres de mesure

Capteur	Paramètres mesurés avec capteur
Débitmètre sigma950 (Capteurs bulle à bulle et de vitesse Doppler) - Hach	Hauteur d'eau et vitesse d'écoulement
Spectro::lyser UV-Vis – s::can	Turbidité et nitrate
Ammo::lyser – s::can	Température, pH, potassium et ammonium
Solitax sc - Hach	MeS
3700 sc Conductimètre inductif - Hach	Conductivité et température
LDO - Hach	Oxygène dissous et température
Capteur pHD sc différentiel numérique - Hach	pH et température

Les principes de mesure des capteurs indiqués au tableau 6 sont décrits brièvement ci-après. Les méthodes de calibration des capteurs ne sont pas expliquées dans cette section. Elles le sont à la section 5.4 Entretien des capteurs, étant donné qu'il existe plusieurs méthodes de calibration pour un même capteur et que certaines solutions de calibration ont été déterminées à la suite de l'étude préliminaire. Pour plus de détails sur les capteurs, il faut se référer à leur manuel d'utilisateur.

Le débitmètre sigma950 permet la connexion de plusieurs capteurs simultanément comme le capteur bulle à bulle et le capteur de vitesse Doppler. Le capteur bulle à bulle mesure la

hauteur d'eau à un endroit en poussant un jet d'air continu au travers d'un tube de plastique. La pression dans le tube est proportionnelle à la hauteur du liquide à l'endroit où l'extrémité du tube est installée. Le débitmètre sigma950 mesure cette pression et la convertit en une hauteur de liquide. Le capteur de vitesse Doppler exploite l'effet Doppler pour mesurer la vitesse d'écoulement. Des ondes sonores à haute fréquence sont transmises dans le milieu et sont réfléchies par les particules en mouvement vers le capteur. La fréquence de retour est modifiée et elle est proportionnelle à la vitesse des particules. Le débitmètre sigma950 utilise la fréquence de retour pour calculer la vitesse d'écoulement. Aucune calibration n'est nécessaire pour ce capteur, car la fréquence d'émission est fixée avec un générateur de fréquence très précis contrôlé par un cristal de quartz. Ce capteur ne mesure pas une vitesse ponctuelle, mais une vitesse moyenne située entre le nez de la sonde et la surface du liquide en suivant un angle d'une vingtaine de degrés (Hach, 2007, 2008).

Le Spectro::lyser UV-Vis mesure l'absorbance dans les spectres ultra-violet et visible. Une lampe émet un faisceau lumineux et, à la suite du contact avec le milieu, ce faisceau est capté par un détecteur mesurant son intensité situé à 180° de l'émetteur. Chaque molécule d'une substance dissoute contenue dans le milieu absorbe des radiations à une certaine longueur d'onde et la concentration de substances contenue dans le milieu augmente l'intensité d'absorption. Cette relation permet donc de déterminer la concentration de différentes substances. De plus, ce capteur compense l'effet de la turbidité sur l'absorbance mesurée pour la détermination d'une substance en particulier (s::can, 2007).

L'Ammo::lyser détermine la concentration d'ammonium total dissous en mesurant le potentiel d'oxydo-réduction de substances précises à l'aide d'électrodes sélectives d'ions ( $H^+$ ,  $K^+$ ,  $NH_4^+$ ). L'équilibre entre l'ion d'ammonium et l'ammoniac complique la mesure de l'ammonium total dissout, car cet équilibre est dépendant du pH et de la température. De plus, le potassium possède des propriétés similaires à l'ammonium, de sorte que la membrane sélective d'ion de l'électrode de l'ammonium est partiellement perméable aux ions de potassiums. La température, le pH et la concentration de potassium doivent donc être mesurés pour compenser leur influence sur la mesure de l'ammonium total dissout. Chaque électrode est munie d'une membrane sélective spécifique à un ion et la combinaison d'une électrode et d'un électrolyte permet de mesurer le potentiel d'oxydo-réduction de l'ion en question. L'ajout d'une électrode de référence permet de convertir ce potentiel en une concentration

ionique à l'aide de l'équation de Nernst et, ainsi, il est possible de déterminer le pH, la concentration de potassium et d'ammonium (s::can, 2007).

Le capteur Solitax sc possède deux options de mesure indépendantes de la couleur du milieu soit la turbidité ou les MeS. La dernière option qui est celle choisie lors du suivi en temps réel exploite la méthode de mesure par lumière infrarouge diffusée à deux rayons. Cette option fournit une mesure de MeS par défaut résultante d'une pré-calibration effectuée par la compagnie Hach (Hach, 2004).

Le capteur 3700 sc Conductimètre inductif mesure la conductivité électrique en envoyant un courant alternatif à travers une bobine toroïdale qui induit un courant dans le milieu. Ensuite, le courant induit produit dans une deuxième bobine toroïdale un courant proportionnel à la conductivité électrique du milieu (Hach, 2004).

Le principe de mesure du capteur LDO est basé sur une méthode optique qui mesure la durée de la luminescence après une impulsion d'excitation. Le capuchon du capteur est enduit d'un matériel luminescent et une lampe DEL émettant de la lumière bleue excite instantanément ces composantes chimiques luminescentes. Lorsqu'elles reviennent à un état d'équilibre, une lumière rouge est produite et est détectée par une photodiode. Le temps pour que les composantes chimiques reviennent à l'équilibre est influencé par la concentration en oxygène et ce temps est inversement proportionnel à la concentration en oxygène (Hach, 2006).

Le capteur pHD sc différentiel numérique mesure le pH à l'aide d'une électrode de verre et d'une électrode de référence. L'électrode de verre transforme l'énergie chimique provenant des ions d'hydrogène du milieu en énergie électrique. Ensuite, le flux ionique entre la solution de référence et le milieu ferme le circuit électrique. Une tension est formée et permet de déterminer le pH du milieu à l'aide de l'équation de Nernst (Hach, 2006).

### 3.4 Jaugeage du débit

Comme mentionné précédemment, la hauteur d'eau a été mesurée soit avec une règle lors de l'étude préliminaire ou avec un capteur bulle à bulle lors du suivi en temps réel. Ainsi, il a été possible de déterminer le débit au point de mesure à l'aide d'une courbe de tarage. La courbe de tarage est développée en établissant la relation entre la hauteur d'eau et le débit.

Par ailleurs, il est important de mentionner que le jaugeage du débit est un aspect que l'auteur de ce mémoire n'a pas étudié en profondeur et, donc, cet aspect peut présenter quelques lacunes. Par exemple, la présence d'un contrôle hydraulique, sa permanence dans le temps et dans l'espace influencent directement le tracé de la courbe de tarage et auraient dû être étudiées. En effet, la courbe de tarage peut être non univoque (deux ou plus de deux valeurs de débits peuvent correspondre à une même hauteur d'eau) si la permanence d'un contrôle hydraulique n'est pas respectée (Jaccon, 1986). Aussi, les sites choisis ne permettent pas le jaugeage de tous les débits, en ce sens qu'il est dangereux de jauger un débit élevé à cause de la force du courant.

Le mode opératoire adopté pour réaliser le jaugeage des débits a été la mesure à gué étant donné que la profondeur et la vitesse d'écoulement du ruisseau Notre-Dame sont généralement inférieures à 1m et 1 m/s (Aldegheri, 1979). Ce mode consiste à ce que l'opérateur, équipé de bottes, mesure des vitesses d'écoulement en se déplaçant directement dans le cours d'eau.

La bathymétrie de la section où l'on veut estimer le débit doit être tout d'abord relevée. La méthode qui a été sélectionnée par l'auteur consiste à installer une corde horizontale de niveau traversant le cours d'eau qui est le niveau de référence et de mesurer la profondeur du lit par rapport à ce niveau à des distances connues à partir d'un point origine. Aussi, il est bien important de mesurer la profondeur du lit au point où la hauteur d'eau est mesurée pour élaborer la courbe de tarage.

Ensuite, le débit doit être jaugé pour des hauteurs d'eau couvrant la plus grande plage possible. Deux méthodes ont été utilisées pour jauger le débit avec une hélice soit en jaugeant les vitesses point par point ou verticale par verticale (Jaccon, 1986). La méthode de jaugeage des vitesses point par point est la méthode employée durant l'étude préliminaire.

Lors du suivi en temps réel, elle a été remplacée par la méthode verticale par verticale suite à un bris dans le montage utilisé pour la première méthode. Le premier montage utilisé possède un pied qui se pose directement sur le lit du ruisseau, tandis que le deuxième montage n'a pas de pied et demeure suspendu lors du jaugeage des vitesses. Il est donc difficile avec le deuxième montage de positionner l'hélice de façon stable par rapport au lit du ruisseau pour jauger une vitesse ponctuelle. C'est pourquoi la méthode verticale par verticale a été adoptée avec ce montage. Les montages utilisés pour jauger les vitesses selon les méthodes point par point et verticale par verticale sont montrés respectivement à la figure 20 a) et b). Le texte ci-après explique la méthode point par point et la méthode verticale par verticale effectuées respectivement avec le premier et le deuxième montage.

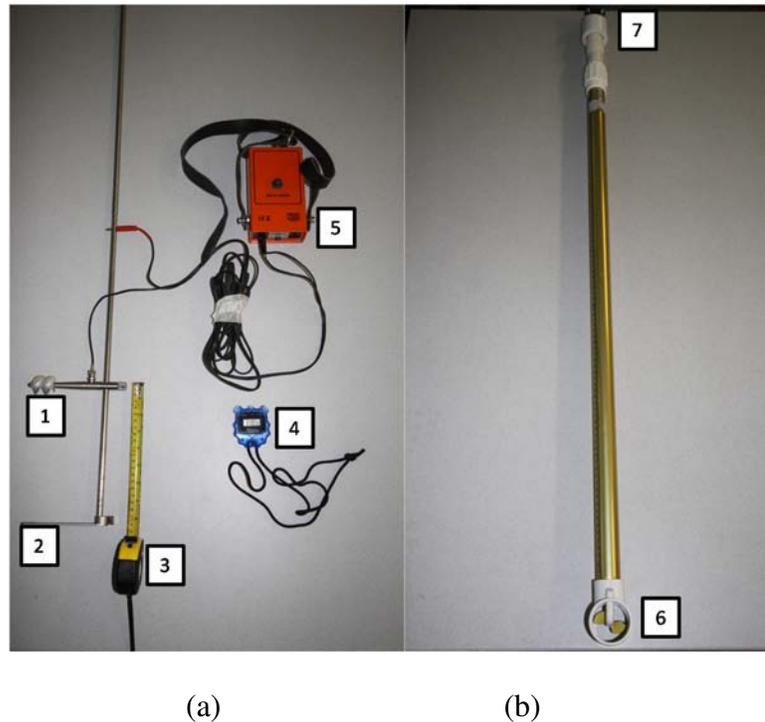


Figure 20 : Montage pour jauger les vitesses selon la méthode a) point par point;  
b) verticale par verticale

Les pièces numérotées sur la figure 20 sont:

- 1) Hélice;
- 2) Pied;
- 3) Gallon à mesurer;
- 4) Chronomètre;
- 5) Compteur de tour;
- 6) Hélice;
- 7) Écran digital.

Pour le premier montage, il faut premièrement déterminer les coefficients de la formule de tarage de l'hélice qui est la relation entre la vitesse d'écoulement de l'eau et le nombre de tours de l'hélice dans un intervalle de temps. La formule générale est :

$$V = an + b$$

Où :

- V est la vitesse du courant;
- n est le nombre de tours de l'hélice par intervalle de temps;
- a est le pas réel de l'hélice;
- b est la vitesse dite de frottement.

Les coefficients a et b sont déterminés expérimentalement pour chaque hélice (Aldegheri, 1979). En effet, le premier montage peut changer d'hélice selon la grandeur des vitesses présentes dans le cours d'eau. La figure 21 montre trois types d'hélices utilisées pour le jaugeage du débit lors de l'étude préliminaire et dont leur sensibilité augmente de la gauche vers la droite. Contrairement à la méthode décrite dans Aldegheri (1979) où les hélices sont étalonnées par translation rectiligne dans un canal où l'eau est parfaitement immobile, la méthode adoptée par l'auteur de ce mémoire compare les vitesses de rotation des hélices aux vitesses d'écoulement mesurées avec un appareil de référence, un ADV exploitant l'effet Doppler. En comparant ainsi plusieurs vitesses de rotation et d'écoulement, on peut déterminer la valeur des coefficients a et b. La figure 22 présente l'installation employée lors de l'étalonnage des hélices du premier montage.



Figure 21 : Types d'hélices pour le premier montage

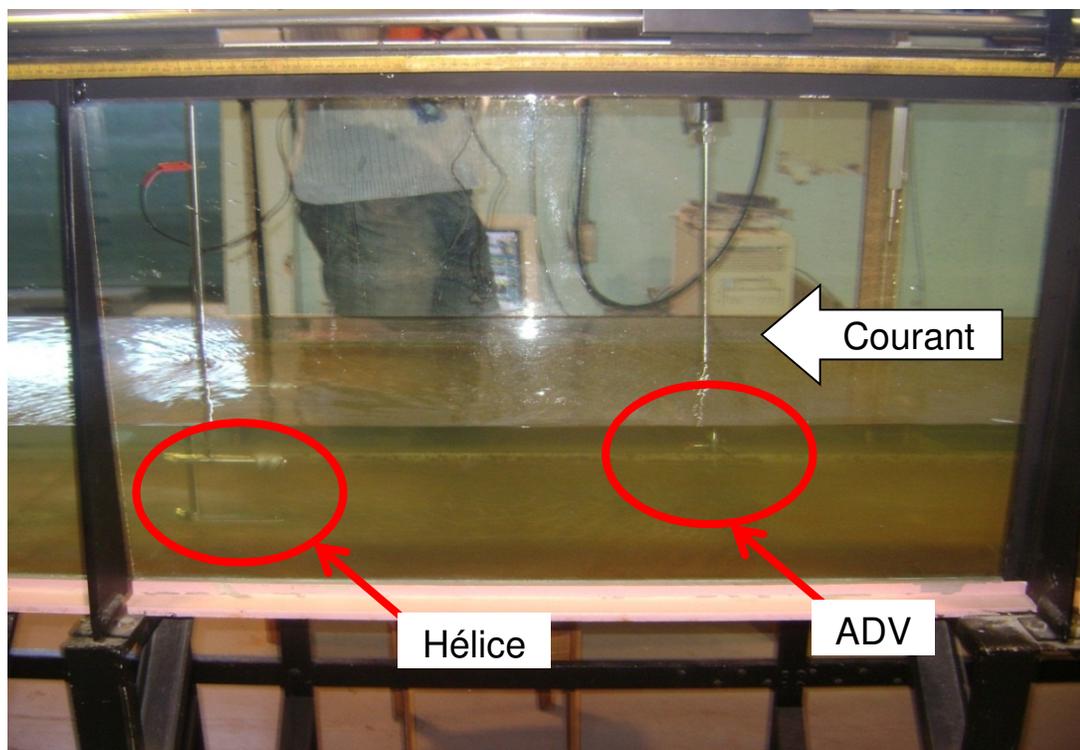


Figure 22 : Installation pour l'étalonnage des hélices du premier montage

Sur le terrain, le jaugeage du débit s'effectue en jaugeant plusieurs vitesses à des positions connues par rapport au lit du cours d'eau et au point d'origine. La figure 23 illustre cette méthodologie. Sur cette figure, le niveau de référence qui est la corde horizontale est équivalent à la valeur 0 sur l'axe des ordonnées. La hauteur d'eau au point de mesure est de 48 cm et les vitesses ont été mesurées à trois hauteurs par rapport au lit du cours d'eau soit à 10, 20 et 30 cm. À l'annexe 1 se trouve la fiche qui a été utilisée pour le jaugeage des vitesses.

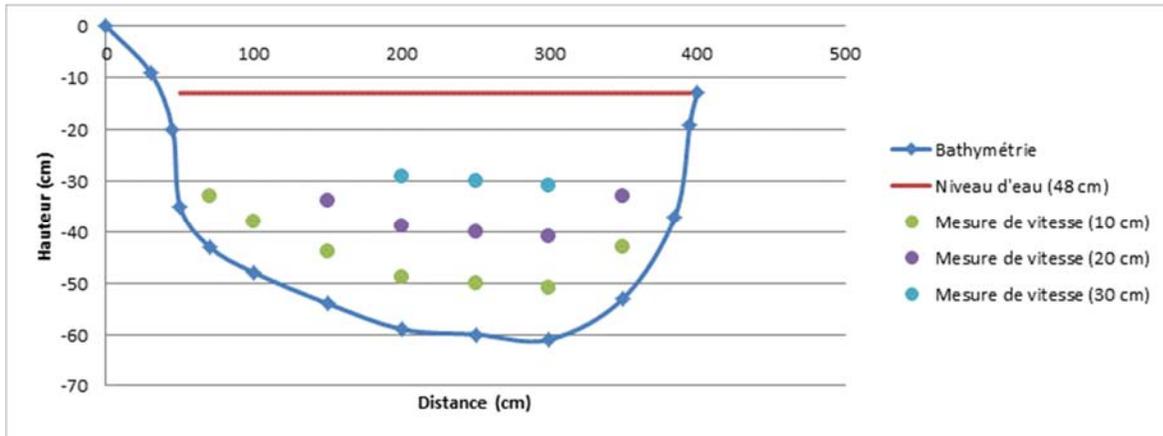


Figure 23 : Jaugeage du débit selon la méthode point par point

Lorsque plusieurs vitesses d'écoulement ponctuelles ont été jaugées, il suffit de les intégrer à la section d'écoulement pour obtenir une valeur de débit, comme illustré à la figure 24 (Aldegheri, 1979).

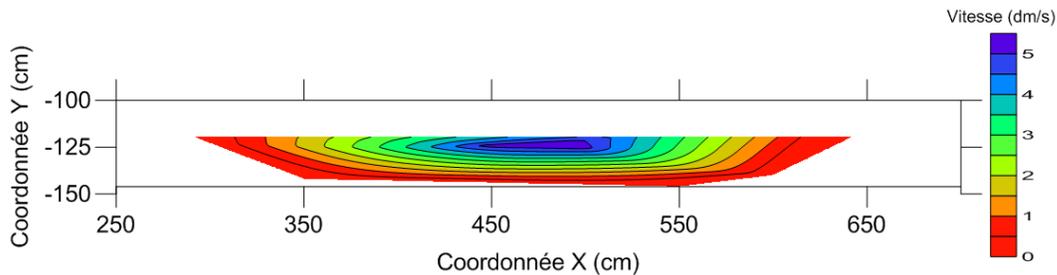


Figure 24 : Intégration des vitesses ponctuelles à la section d'écoulement avec le logiciel Surfer Plot

La méthode verticale par verticale consiste à jauger à plusieurs endroits la vitesse moyenne verticale en effectuant un mouvement continu du haut vers le bas avec une hélice (Aldegheri, 1979). Cette méthode a été utilisée avec le deuxième montage qui possède une hélice déjà étalonnée par la compagnie fournisseur. La vitesse d'écoulement jaugée s'affiche sur un écran digital et est en mètre par seconde.

Ensuite, le débit est calculé selon la méthode de la section médiane qui attribue une aire à chaque vitesse moyenne verticale dont les limites sont définies par la moitié de la distance entre deux mesures (Aldegheri. 1979). La figure 25 illustre cette méthodologie. Sur cette figure, le niveau d'eau est de 48 cm et les vitesses moyennes verticales ont été jaugées à une distance de 75, 100, 150, 200, 250, 300 et 350 cm par rapport au point d'origine. Deux aspects intéressants de cette méthode sont que le jaugeage des vitesses est rapide et que le calcul du débit est simple.

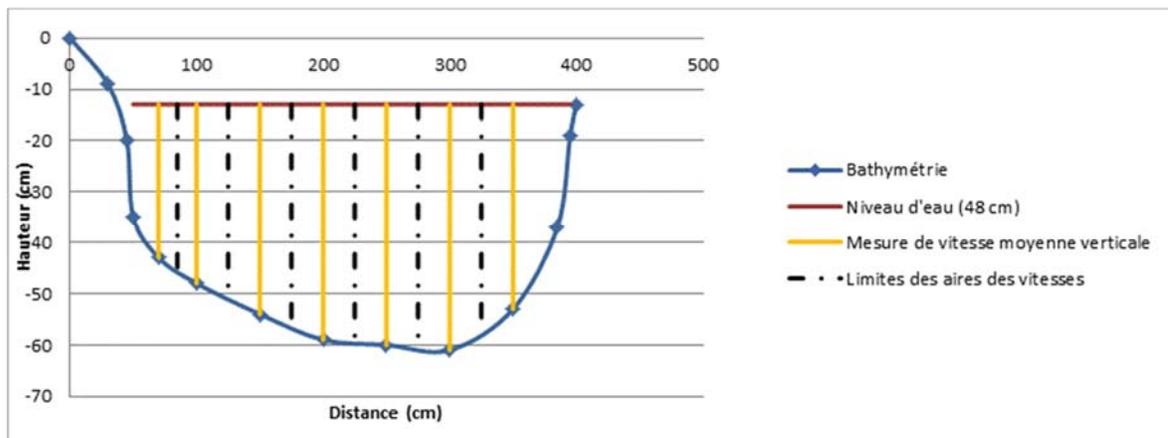


Figure 25 : Jaugeage du débit selon la méthode verticale par verticale

Finalement, après avoir jaugé le débit pour plusieurs hauteurs d'eau soit avec la méthode point par point ou verticale par verticale, il est possible de déterminer une courbe de tarage qui décrit la relation entre le débit et la hauteur d'eau. Cette relation permet par la suite de déterminer le débit lorsque seulement la hauteur d'eau est mesurée (Jaccon, 1986). Les équations les plus couramment utilisées pour décrire cette relation sont les formes exponentielle et polynômiale (Jaccon, 1986) représentées respectivement par :

$$Q = a(h - h_o)^n$$

$$Q = a + a_1h + a_2h^2 + \dots$$

Où :

h est la hauteur à l'échelle pour le débit Q;

$h_0$  est la hauteur à l'échelle pour le débit  $Q_0$ ;  
 $a$  et  $n$  sont des constantes si le lit est très stable et le profil en travers régulier.

Par ailleurs, étant donné qu'il n'était pas possible de jauger le débit lorsqu'il était élevé, la méthode d'extrapolation employée pour les hautes eaux consiste simplement à l'exploitation de l'équation décrivant la relation entre la hauteur d'eau et le débit. Le critère d'utilisation de cette méthode est que la section transversale doit être très régulière (Jacon, 1986).

## **4. Étude préliminaire**

### **4.1 Objectifs de l'étude préliminaire**

Les objectifs principaux de l'étude préliminaire du ruisseau Notre-Dame effectuée en octobre 2009 étaient de connaître les gammes et les dynamiques générales de la qualité de l'eau lors d'évènements pluviaux. Les connaissances acquises lors de cette étude ont été exploitées pour la mise en branle du suivi en temps réel. En particulier, ces connaissances ont guidé le choix de certaines solutions standards à utiliser pour la calibration des capteurs. Dans le cadre du projet *rivEAU*, les données récoltées durant cette étude ont servi à la modélisation hydrologique et à la modélisation de la qualité de l'eau et elles ont permis d'étudier l'effet du sous-bassin versant urbain sur la qualité de l'eau de ce ruisseau. De plus, lors de l'étude préliminaire, l'auteur s'est familiarisé avec les processus d'échantillonnage, de préservation et d'analyse en laboratoire, ainsi que la méthode de jaugeage du débit. Ce qui s'avéra utile pour le suivi en temps réel.

### **4.2 Description des points d'échantillonnage**

Deux points d'échantillonnage, où était aussi jaugé le débit, ont été choisis afin d'évaluer l'effet du sous-bassin versant urbain sur la qualité de l'eau. En effet, le point d'échantillonnage amont est peu influencé par le sous-bassin versant urbain tandis que le point d'échantillonnage aval est fortement influencé par ce dernier. Les critères sur lesquels le choix des points d'échantillonnage était principalement basé sont une section dont l'écoulement est parallèle aux rives (sans tourbillons importants et remous) et ne possédant pas de zone d'eau morte. Aussi, les points choisis doivent posséder un écoulement lent dû au mode opératoire pour le jaugeage du débit qui est la mesure à gué. Le respect de ces critères facilite le jaugeage du débit (Aldegheri, 1979). Comme indiqué à la figure 26, le point d'échantillonnage amont est situé quelques mètres en aval du ponceau drainant le sous-bassin versant de l'aéroport situé sous la route de l'Aéroport et le point d'échantillonnage aval est situé quelques mètres en amont du pont pour les piétons liant le terrain de soccer du site des loisirs à la rue St-Pierre.

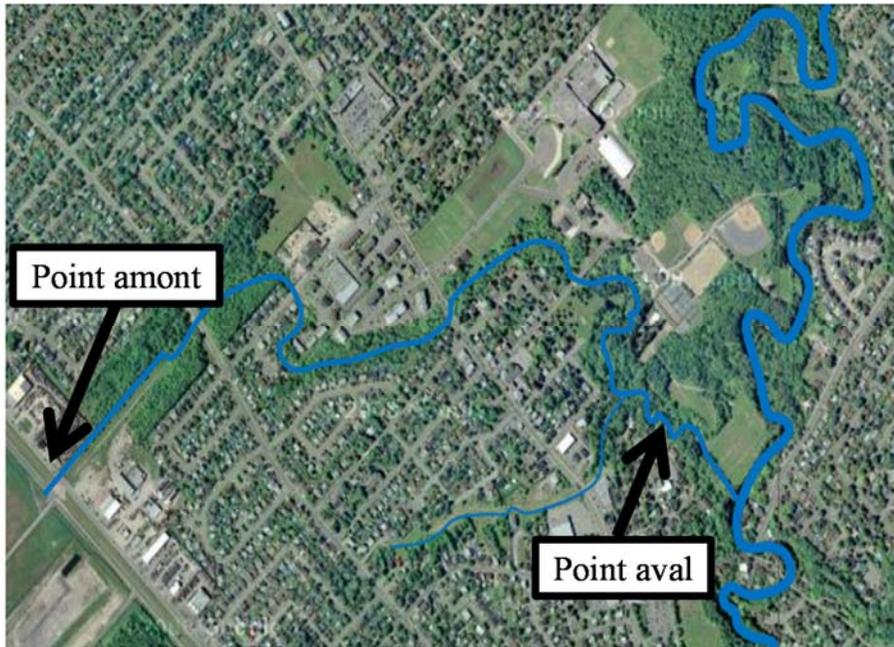


Figure 26 : Carte indiquant les points d'échantillonnage amont et aval

La figure 27 et la figure 28 sont respectivement des photographiques des points d'échantillonnage amont et aval.



Figure 27 : Point d'échantillonnage amont



Figure 28 : Point d'échantillonnage aval

### **4.3 Déroulement de l'étude préliminaire**

L'étude préliminaire du ruisseau Notre-Dame s'est étalée du 1<sup>er</sup> octobre au 31 octobre 2009. Tous les jours, des échantillons ont été pris le matin vers 10 heures et le soir vers 18 heures, aux deux points d'échantillonnage. De plus, des échantillons ont été pris à un intervalle de 4 heures durant une journée après une longue période d'évènements pluviaux, le 1<sup>er</sup> octobre 2009 et, également, durant une journée après quelques jours sans précipitation, le 19 octobre 2009.

Durant cette campagne d'échantillonnage, les analyses ou la préservation à long terme (acidification et congélation) pour des analyses ultérieures ont été faites tous les après-midi sur les échantillons pris à 18 heures de la journée précédente et à 10 heures de la journée même. Pour la préservation de moins de 24 heures, les échantillons étaient préservés dans une chambre froide à 4°C. Les échantillons ont donc été analysés ou préservés à long terme dans un délai de moins de 24 heures. À l'annexe 2 et à l'annexe 3 se trouvent respectivement la fiche d'échantillonnage et la procédure employées lors de l'étude préliminaire.

## **4.4 Résultats et analyses de l'étude préliminaire**

Cette section présente les résultats de certains paramètres mesurés lors de l'étude préliminaire ainsi qu'une brève discussion sur les dynamiques observées. Les paramètres présentés ci-après sont ceux qui ont aussi été mesurés durant le suivi en temps réel. Le débit est tout d'abord discuté, puisque les paramètres de la qualité de l'eau sont illustrés en fonction du débit. L'intensité des précipitations mesurée au pluviomètre situé à l'aéroport Jean-Lesage est également montrée afin de comprendre les variations des débits. Ensuite, cette section présente les paramètres mesurés avec la sonde YSI 6600EDS, soit la température, la conductivité et la turbidité, les MeS, ainsi que l'ammonium et le nitrate mesurés avec les tests Hach TNT. Même si le pH a été mesuré durant le suivi en temps réel, il n'est pas présenté ci-après. En effet, comme discuté dans la section 3.3 Capteurs et paramètres mesurés, les valeurs mesurées lors de l'étude préliminaire sont biaisées. L'annexe 4 montre les résultats des paramètres de la qualité de l'eau qui ne sont pas présentés dans cette section.

Deux types de figures sont principalement présentées dans cette section, soit les résultats des échantillonnages de 10 heures et de 18 heures du mois d'octobre et ceux pris à un intervalle de 4 heures durant les journées du 1<sup>er</sup> et du 19 octobre. Pour le premier type de figures, les paramètres de la qualité de l'eau sont présentés selon les heures approximatives d'échantillonnage, soit 10 heures et 18 heures, tandis que, pour le deuxième type de figures, ils sont présentés selon les heures exactes d'échantillonnage. Pour le deuxième type de figures, les débits simulés amont et aval sont présentés par la même ligne, car ils sont presque identiques.

### **4.4.1 Débit**

Cette section présente les débits déterminés à l'aide d'une courbe de tarage, ainsi que les débits simulés avec le logiciel SWMM aux points d'échantillonnage amont et aval.

L'élaboration des courbes de tarage est présentée à l'annexe 5. La simulation des débits a été réalisée par Étienne Gaborit, étudiant au doctorat à l'Université Laval. Pour ce faire, il a effectué une calibration des paramètres originaux du modèle SWMM fournis par la compagnie BPR à l'aide d'une partie des débits mesurés par cette compagnie durant une étude effectuée en 2006. La validation du modèle a été effectuée à l'aide de la seconde partie des débits mesurés par la compagnie BPR en 2006 et des débits mesurés durant le mois d'octobre lors de l'étude préliminaire.

La figure 29 et la figure 30 présentent respectivement les débits mesurés et simulés aux points d'échantillonnage amont et aval du mois d'octobre. Il est à noter que l'échelle du débit n'est pas la même sur ces deux figures. Il se peut que le modèle simule moins bien les débits aux environs du 23 et du 24 octobre, car il y a eu des précipitations sous forme de neige. Étant donné que les débits simulés en amont et en aval correspondent relativement bien aux débits mesurés, les paramètres de la qualité de l'eau sont illustrés en fonction des débits simulés.

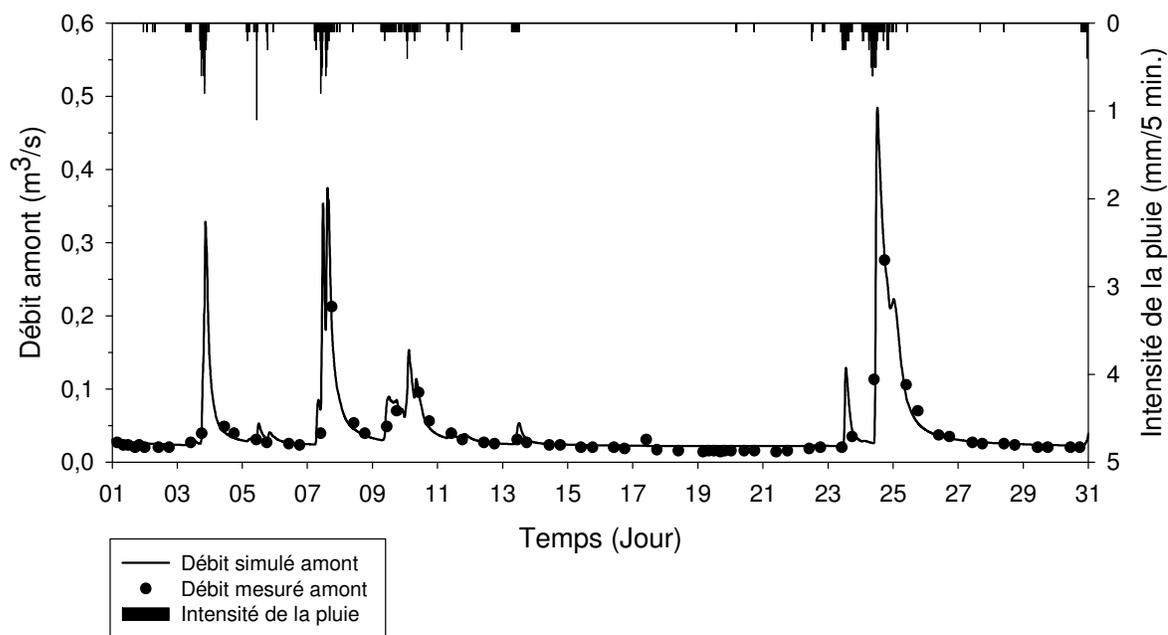


Figure 29 : Débits mesurés et simulés en amont au mois d'octobre 2009

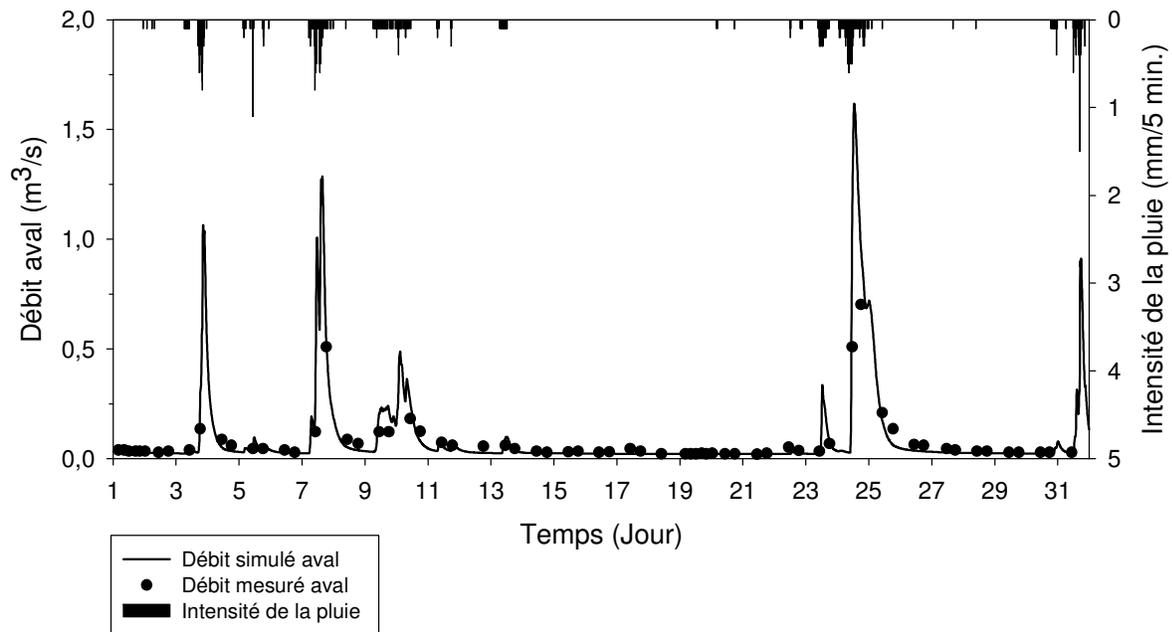


Figure 30 : Débits mesurés et simulés en aval au mois d’octobre 2009

#### 4.4.2 Sonde YSI 6600EDS (température, conductivité et turbidité)

Cette section présente les paramètres mesurés par la sonde YSI 6600EDS, soit la température, la conductivité et la turbidité. En général, ces paramètres ont été mesurés à chaque échantillonnage en laissant la sonde 3 minutes dans le ruisseau avec un intervalle de mesure de 1 minute. Ce sont les moyennes de ces trois mesures qui sont illustrées sur les figures suivantes. Aussi, deux sondes YSI 6600EDS ont été installées dans le ruisseau aux points d’échantillonnage amont et aval durant environ 24 heures lors de la journée du 19 octobre, soit la deuxième journée où les échantillons ont été pris toutes les 4 heures. Par ailleurs, il y a beaucoup de données manquantes, ceci a été causé par des problèmes de calibration, de mesure, de sauvegarde ou la sonde était hors d’usage. Les données manquantes sont indiquées par un espace vide sur les figures.

##### 4.4.2.1 Température

La figure 31 montre les températures mesurée en amont et en aval durant le mois d’octobre. Sur cette figure, on observe que les températures de l’eau en amont et en aval suivent un

cycle journalier, c'est-à-dire qu'elles diminuent la nuit et qu'elles augmentent le jour. De plus, la variation de la température en aval est de moins grande amplitude que celle en amont. Par ailleurs, il a été démontré que les surfaces imperméables augmentent généralement la température des cours d'eau dans lesquels l'eau de ruissèlement se jette à la suite d'une précipitation estivale (EPA, 2003). Cependant, avec une telle fréquence de mesure, il est impossible d'observer un changement rapide de la qualité de l'eau. Cela sera donc confirmé avec une plus haute fréquence de mesure lors du suivi en temps réel.

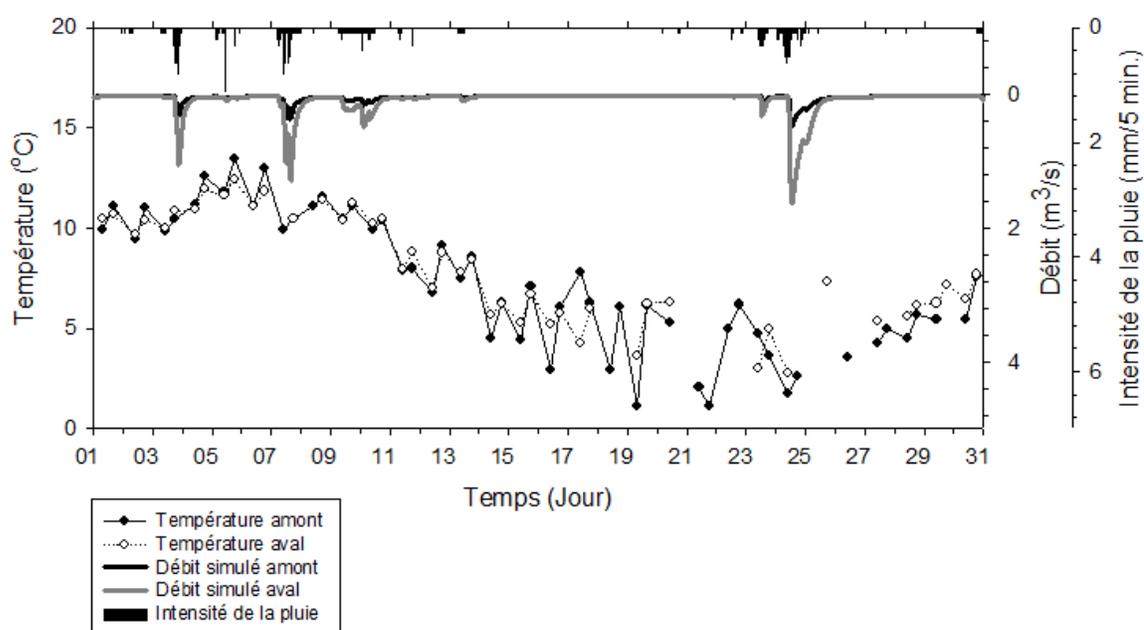


Figure 31 : Températures en amont et aval au mois d'octobre 2009

#### 4.4.2.2 Conductivité

La figure 32 présente les conductivités mesurées en amont et en aval durant le mois d'octobre. Il y a quelques valeurs de conductivité manquantes à cause de problèmes avec la sonde YSI 6600EDS. Selon la figure 32, les conductivités amont et aval diminuent généralement lorsque le débit augmente. Cependant, les valeurs de conductivité les plus élevées en aval surviennent lorsque le débit augmente après une longue période sans précipitation. Aussi, en période de temps sec, lorsque le débit est composé principalement

d'eau souterraine, elles augmentent jusqu'à l'atteinte d'un plateau ayant une valeur se situant entre 1,1 et 1,2 mS/cm.

La figure 33 présente les conductivités mesurée aux points d'échantillonnage amont et aval pour la journée du 19 octobre. Le débit simulé est constant tout au long de la journée, car cette journée correspond à une période de temps sec. D'après la figure 33, on constate que la conductivité demeure relativement constante tout au long de la journée, à l'exception d'une légère variation mesurée en amont aux alentours de 14 heures.

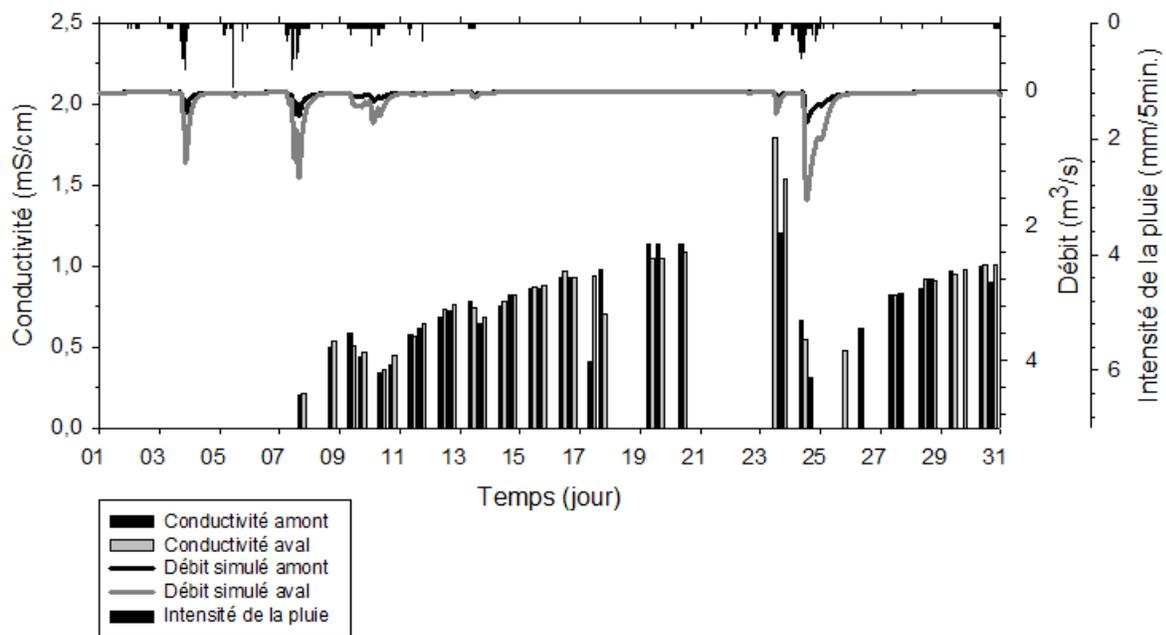


Figure 32 : Conductivités en amont et en aval au mois d'octobre 2009

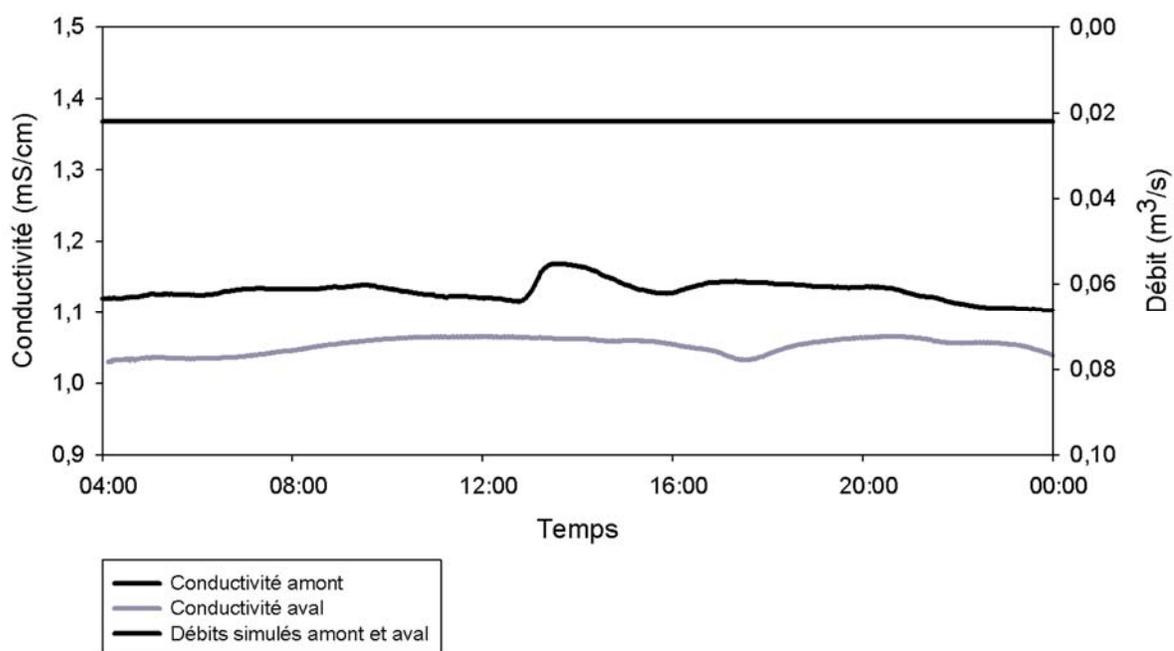


Figure 33 : Conductivités en amont et en aval le 19 octobre 2009

#### 4.4.2.3 Turbidité

La

montre les turbidités mesurées aux points d'échantillonnage amont et aval durant le mois d'octobre et la figure 35 est un agrandissement de l'encadré de la

. Certaines valeurs prises à un intervalle d'une minute présentaient une très grande variation possiblement causée par une suspension de sédiment lors de l'installation de la sonde dans le ruisseau. Dans ce cas-ci, la valeur de turbidité qui semblait la plus plausible a été choisie au lieu de prendre la moyenne des trois valeurs. La

révèle que la variation des turbidités amont et aval est très grande. Elles augmentent lorsque le débit augmente et, en temps sec, elles diminuent pour atteindre une valeur relativement constante. Il y a des pointes de turbidité qui ne s'expliquent pas par la présence de précipitation. Cependant, ces pointes pourraient être expliquées par le fait qu'elles se produisent lorsque la température de l'air baisse sous 0°C durant la nuit, ce qui pourrait diminuer la cohésion des particules de sol en période de dégel et ainsi augmenter la turbidité. Ce processus se nomme décimentation.

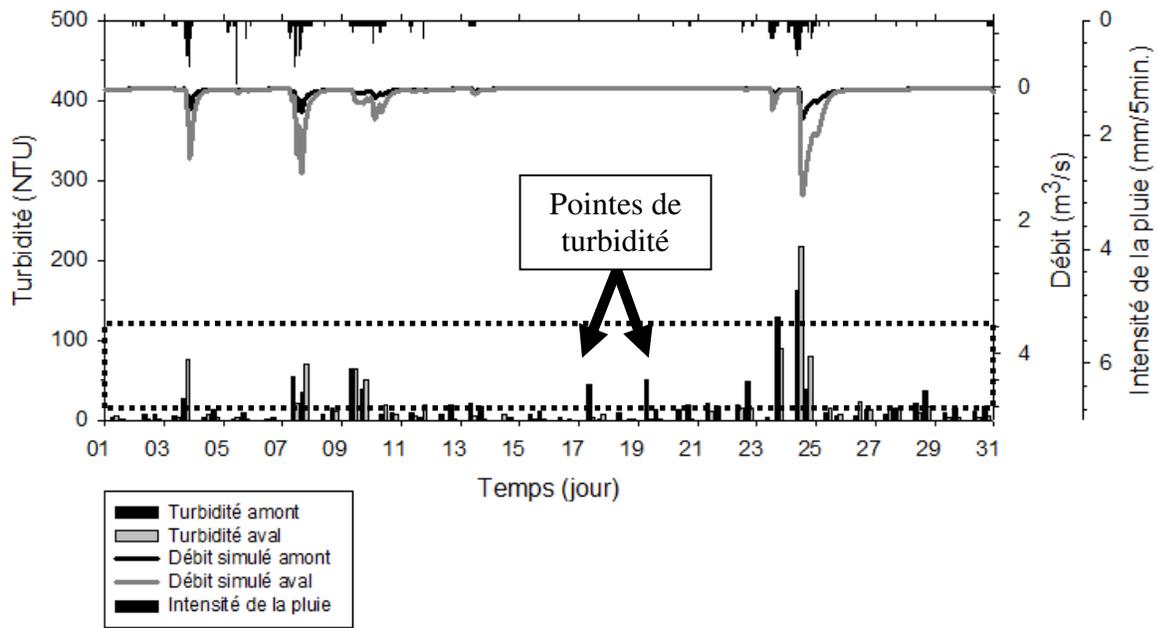


Figure 34 : Turbidités en amont et en aval au mois d'octobre 2009

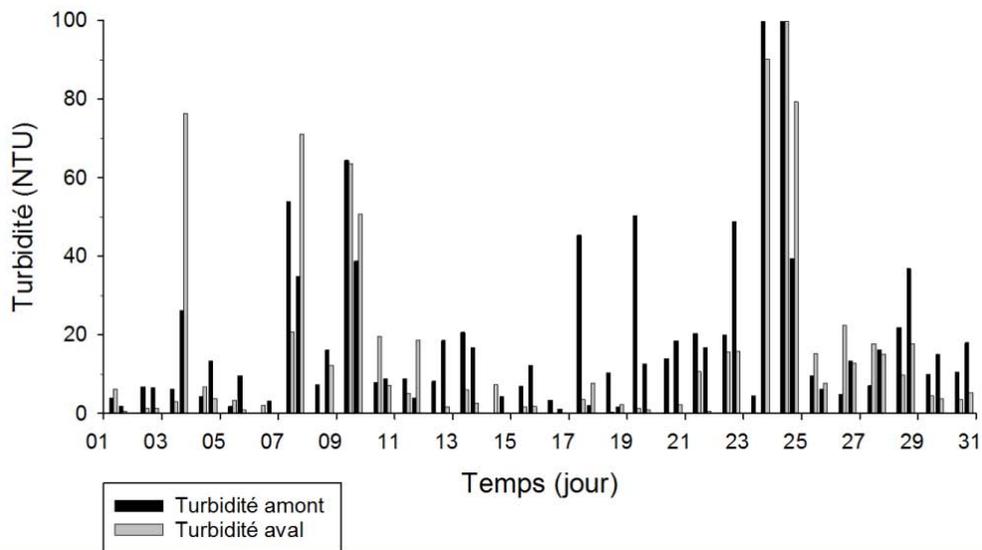


Figure 35: Agrandissement de la

La figure 36 présente la turbidité en amont pour la journée du 19 octobre. Cette journée correspond à une pointe de turbidité indiquée sur la figure 34. La turbidité en aval pour cette journée n'a pas été mesurée à cause d'un problème de calibration de la sonde YSI 6600EDS. Cette figure montre que la turbidité en amont varie beaucoup pour la journée du 19 octobre, même si le débit demeure constant. Comme mentionné précédemment, cette variation pourrait être expliquées par le fait que la température de l'air baisse sous 0°C durant la nuit, ce qui pourrait diminuer la cohésion des particules de sol en période de dégel et ainsi augmenter la turbidité.

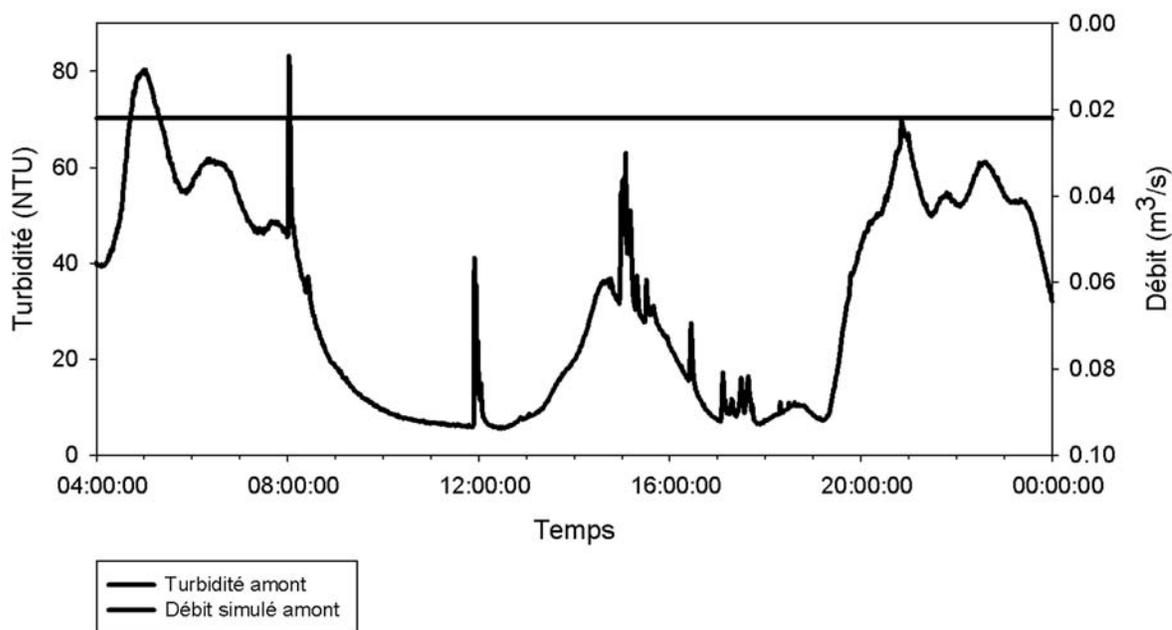


Figure 36 : Turbidité en amont le 19 octobre 2009

#### 4.4.3 MeS

La figure 37 illustre les concentrations de MeS mesurées en amont et en aval durant le mois d'octobre. Pour certaines analyses, des résidus de filtre restaient collés sur l'entonnoir de filtration engendrant ainsi des mesures de MeS plus faibles que les valeurs réelles. Les concentrations de MeS négatives causées par cette perte de filtre ne sont pas présentées sur

la figure 37, tandis que les concentrations de MeS positives influencées par cette perte de filtre sont quand même présentées sur cette figure. En observant la figure 37, on constate que la concentration de MeS varie énormément et qu'elle se comporte de façon similaire à la turbidité. En période sèche, la concentration de MeS est faible et relativement constante et, lors d'un événement pluvial, elle augmente. Comme pour la turbidité, il y a des pointes de MeS qui ne s'expliquent pas par des précipitations, mais qui coïncident avec des températures de l'air sous 0°C.

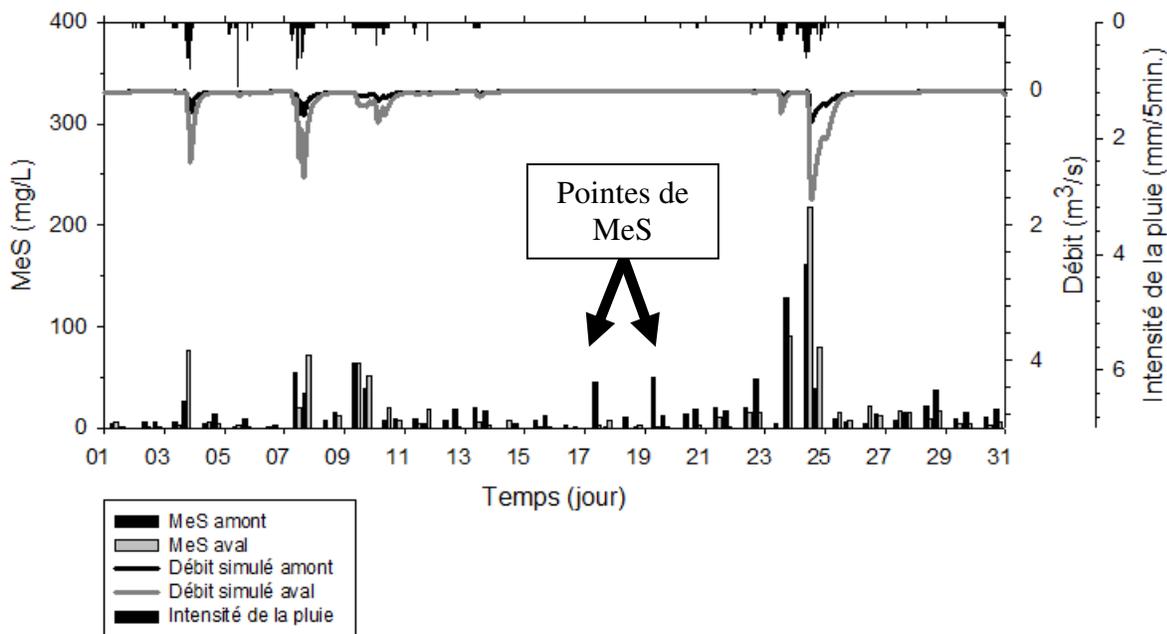


Figure 37 : MeS en amont et en aval au mois d'octobre 2009

La figure 38 et la figure 39 montrent respectivement les concentrations de MeS en amont et en aval pour les journées du 1<sup>er</sup> et du 19 octobre. Deux valeurs ne sont pas présentées sur la figure 38, car elles sont négatives. Selon ces deux figures, la concentration de MeS en amont varie beaucoup et est élevée comparativement à la concentration de MeS en aval. Les causes de cette variation et de cette différence entre les valeurs amont et aval sont inconnues et ne s'expliquent pas par une variation de débit, car, pour ces deux journées, les débits amont et aval sont relativement constants.

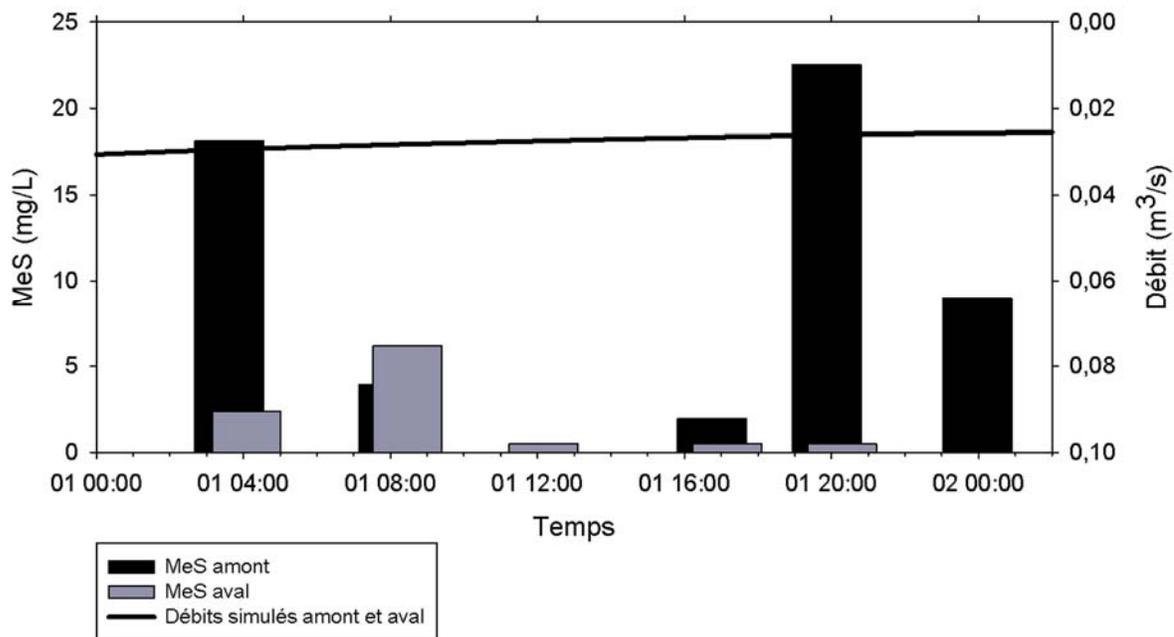


Figure 38 : MeS en amont et en aval le 1er octobre 2009

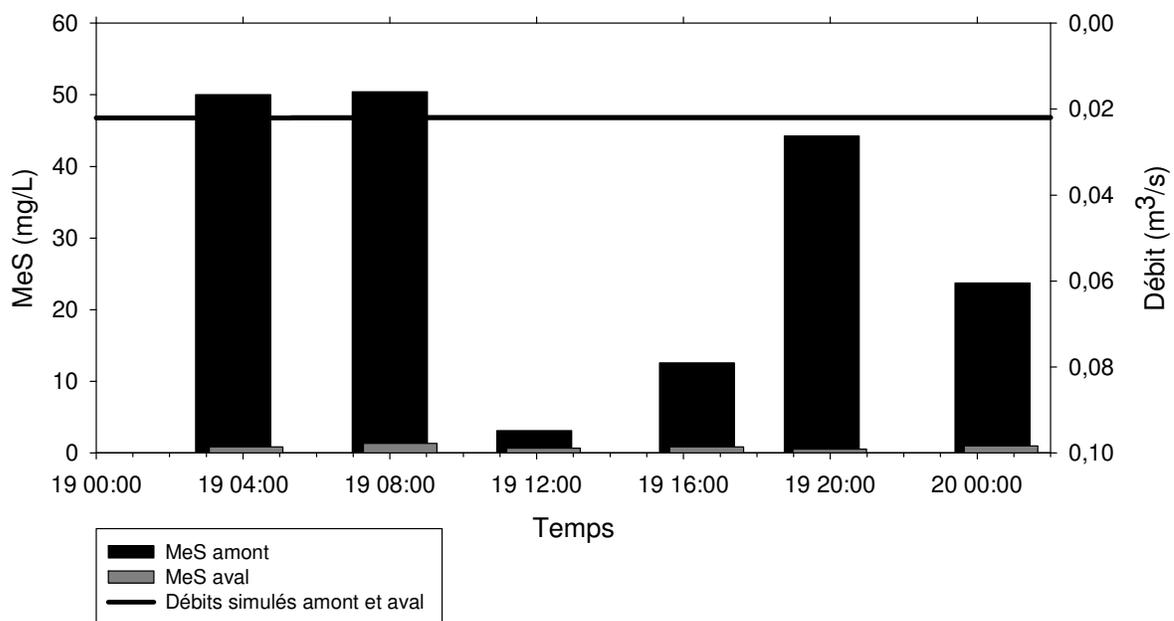


Figure 39 : MeS en amont et en aval le 19 octobre 2009

#### 4.4.4 Hach TNT (ammonium et nitrate)

Quelques problèmes reliés à l'utilisation des tests Hach TNT sont survenues lors de l'étude préliminaire. Tout d'abord, il y a eu un manque de tests Hach TNT pour effectuer les analyses. Certains échantillons ont donc été congelés ou acidifiés pour être préservés à long terme et analysés ultérieurement. L'acidification est la méthode de préservation recommandée par Hach pour l'ammonium et le nitrate à l'exception de l'orthophosphate qui ne se préserve pas au-delà de 48 heures (Hach, 2008).

Parfois, la mesure sur le spectromètre indiquait le message suivant : « out of range ». Selon la compagnie Hach, les valeurs correspondantes à ce message ne sont pas fiables. Sur les figures suivantes, les limites supérieure et inférieure des gammes de mesure des tests Hach TNT sont donc représentées respectivement par des lignes à long pointillé et à court pointillé. Par ailleurs, il est important de mentionner qu'il y a de nombreux facteurs qui peuvent influencer la mesure tels que la température, le temps entre l'ajout du réactif et la mesure, les éléments chimiques présents dans l'échantillon, etc. (Hach, 2008). Cela explique peut-être les valeurs qui semblent aberrantes. De plus, lorsque l'échantillon est très turbide ou coloré, cela peut interférer et augmenter significativement les mesures de l'ammonium, du nitrate et de l'orthophosphate (Hach, 2008). Les échantillons auraient dû être préalablement filtrés pour enlever les particules et des échantillons témoins, « Sample Blank », auraient dû être mesurés pour tenir compte de l'interférence causée par la turbidité de l'échantillon. Cependant, ceci n'a pas été effectué.

##### 4.4.4.1 Ammonium

La **Error! Reference source not found.** présente les concentrations d'ammonium mesurées en amont et en aval durant le mois d'octobre. Cette figure révèle que la concentration en aval est généralement plus faible que la concentration en amont et qu'elles suivent la même tendance. Elles diminuent légèrement lorsque le débit augmente et elles augmentent graduellement tout au long du mois d'octobre. Cette augmentation coïncide avec la période de tombée des feuilles d'arbre et s'explique peut-être par cette dernière. Aussi, il y a quelques

pointes dans les concentrations d'ammonium qui sortent de la tendance et qui semblent être des valeurs aberrantes.

Il est également important de mentionner que la proportion en pourcentage d'ammoniac du couple ammoniac/ammonium est fonction de la température et du pH. Lorsque la température ou le pH augmente, la proportion d'ammoniac augmente. Cependant, pour les gammes de variation de température et de pH mesurées durant le mois d'octobre, cette variation est négligeable par rapport à celle mesurée et ne peut expliquer l'augmentation en ammonium vers la fin du mois d'octobre.

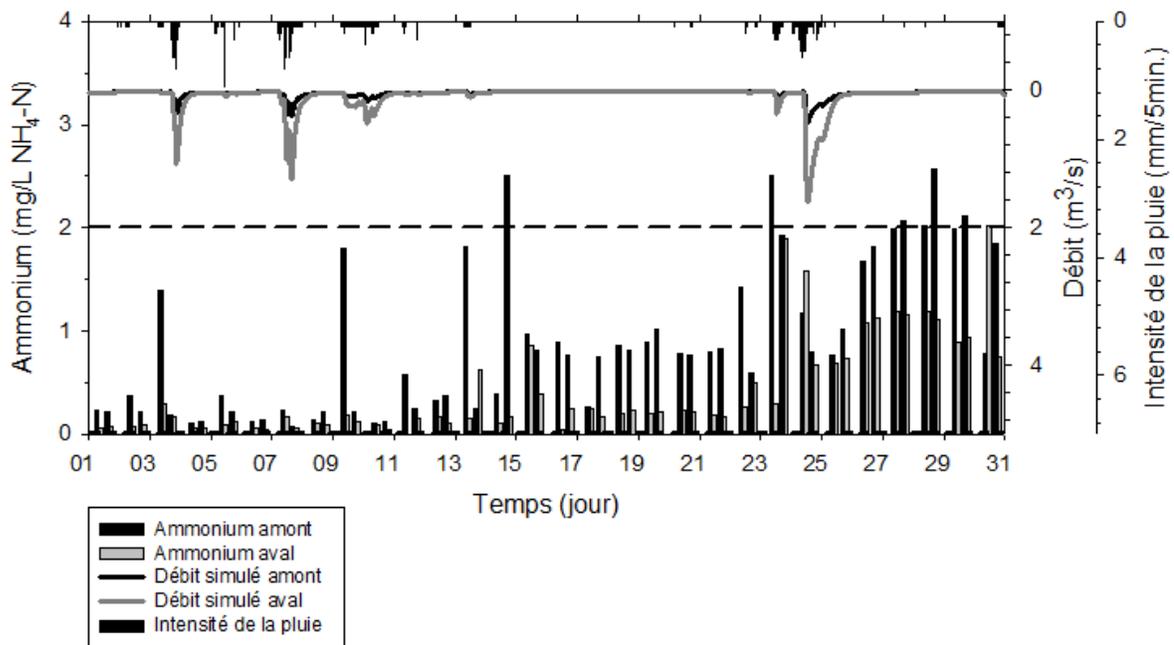


Figure 40 : Concentrations d'ammonium en amont et en aval au mois d'octobre 2009 (ligne à long pointillée : limite supérieure, ligne à court pointillée : limite inférieure)

La Figure 41 et la Figure 42 montrent respectivement les concentrations d'ammonium mesurées en amont et en aval pour les journées du 1<sup>er</sup> et du 19 octobre. On observe que la concentration d'ammonium en amont est plus élevée qu'en aval. Ceci pourrait s'expliquer par des processus de nitrification, ainsi que par la dilution causée par l'augmentation du débit vers l'aval. Il est important de rappeler que le débit présenté sur ces figures est le débit simulé

et il est le même en amont qu'en aval. Par contre, lors du jaugeage du débit en période de temps sec, le débit amont était plus faible que le débit aval.

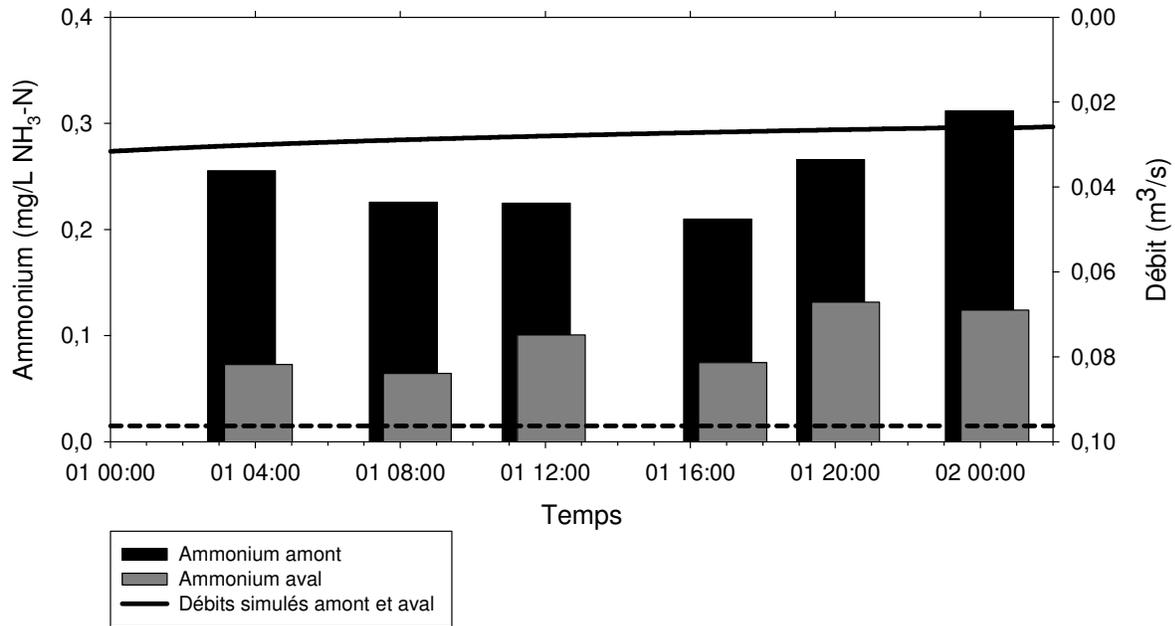


Figure 41: Concentrations d'ammonium en amont et en aval le 1er octobre 2009  
(ligne à court pointillée : limite inférieure)

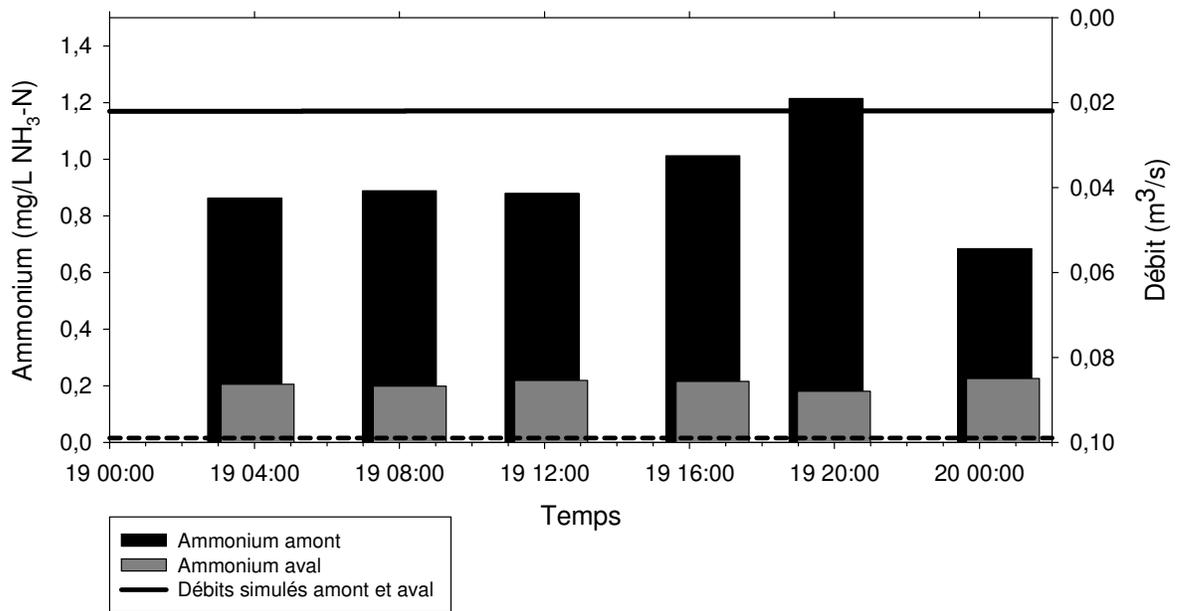


Figure 42: Concentrations d'ammonium en amont et en aval le 19 octobre 2009  
(ligne à court pointillée : limite inférieure)

#### 4.4.4.2 Nitrate

La Figure 43 **Error! Reference source not found.** illustre les concentrations de nitrate mesurées en amont et en aval durant le mois d'octobre. Cette figure indique que les concentrations de nitrate diminuent lorsque le débit augmente et que la concentration en amont est généralement plus élevée que celle en aval. Contrairement à l'ammonium, il n'y a pas d'augmentation graduelle tout au long du mois d'octobre. L'augmentation de la concentration de nitrate à la fin du mois d'octobre semble être causée par le début d'une période de temps sec.

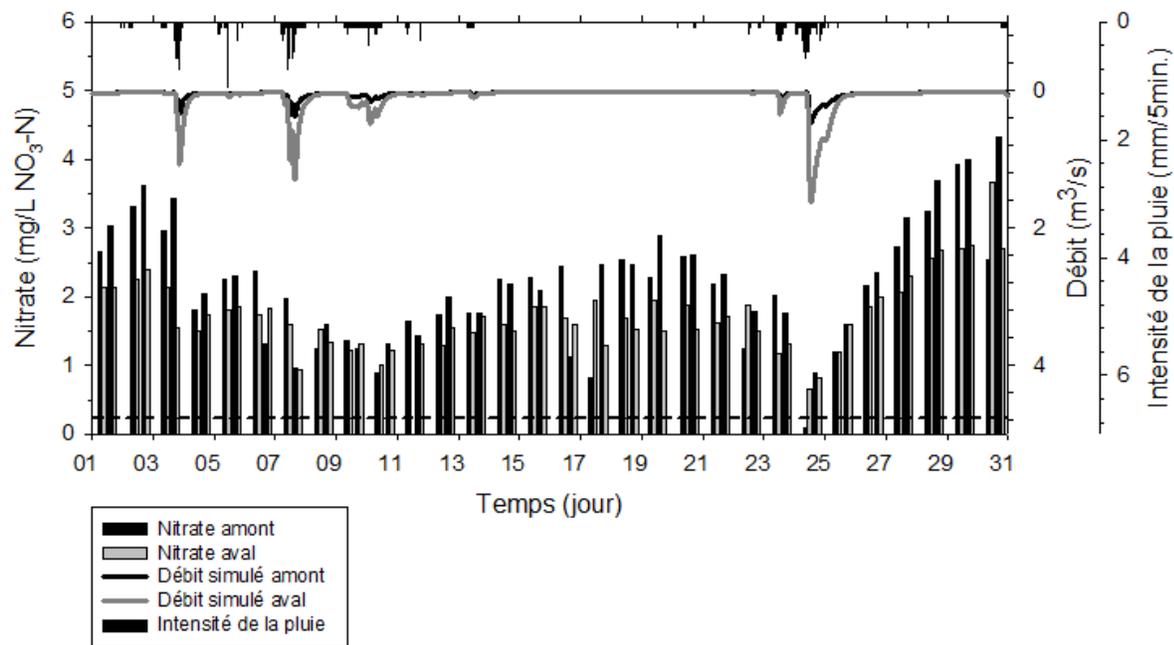


Figure 43: Concentrations de nitrate en amont et en aval au mois d'octobre 2009  
(ligne à court pointillée : limite inférieur)

La Figure 44 et la Figure 45 montrent respectivement les concentrations de nitrate mesurées en amont et en aval pour les journées du 1 et du 19 octobre. Comme pour les concentrations d'ammonium, les concentrations de nitrate sont plus faibles en aval qu'en amont et ceci pourrait s'expliquer par la dilution causée par l'augmentation du débit vers l'aval.

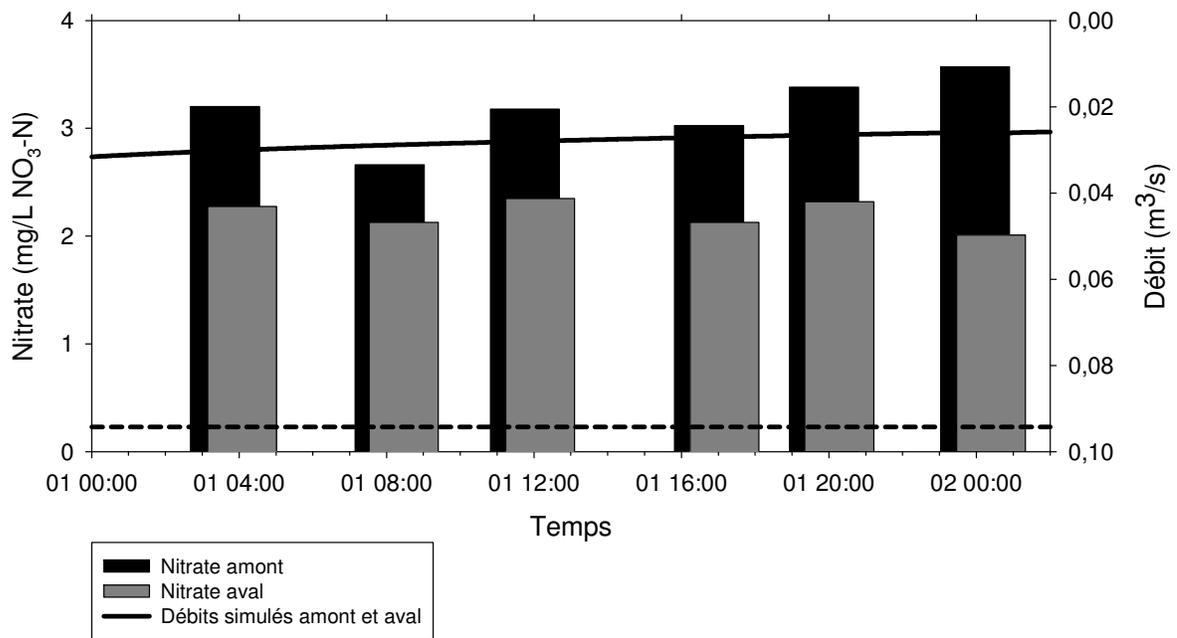


Figure 44: Concentrations de nitrate en amont et en aval le 1<sup>er</sup> octobre 2009  
(ligne à court pointillée : limite inférieur)

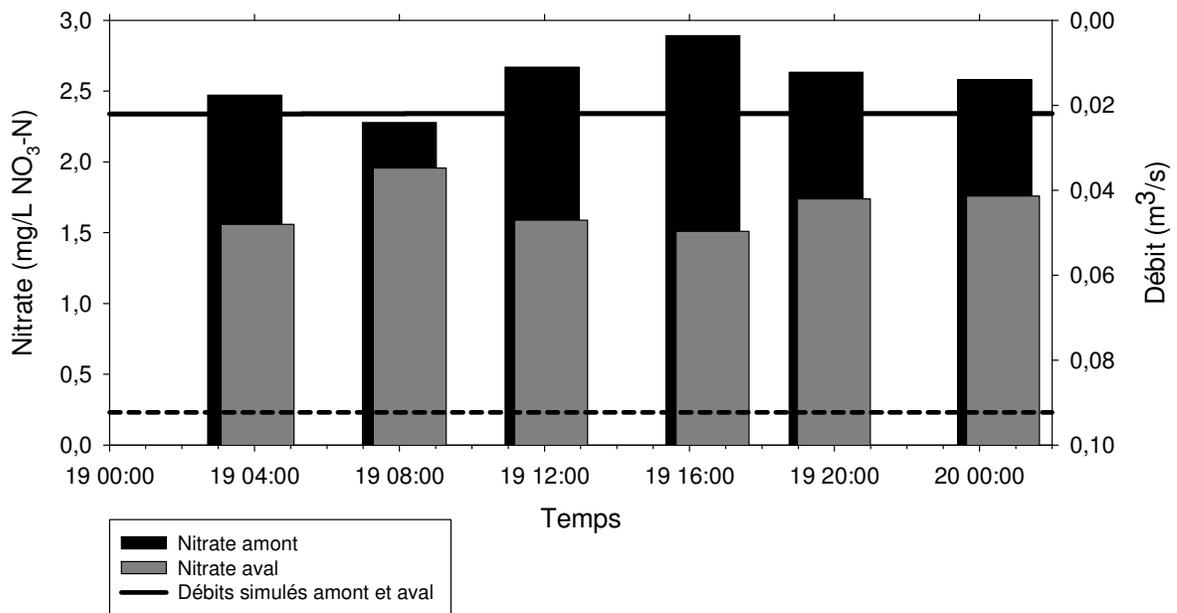


Figure 45: Concentrations de nitrate en amont et en aval le 19 octobre 2009  
(ligne à court pointillée : limite inférieur)



## **5. Suivi en temps réel**

### **5.1 Objectifs du suivi en temps réel**

Le suivi en temps réel du ruisseau Notre-Dame consistait à mesurer les dynamiques de la qualité de l'eau lors d'évènements pluviaux. Pour ce faire, l'intervalle de mesure des capteurs a été établi à 2 minutes, à l'exception du Solitax sc dont l'intervalle a été ajusté à 5 secondes pour des fins expérimentales. Les données récoltées ont principalement permis d'évaluer les ADQATs qui ont été développées dans le cadre de ce projet (voir la chapitre 6 ADQATs). De plus, comme pour l'étude préliminaire, les données récoltées ont également été utilisées dans le cadre du projet rivEAU à des fins de modélisation hydrologique et de modélisation de la qualité de l'eau.

Au départ, deux stations de mesure devaient être installées soit en amont et en aval du sous-bassin versant urbain du ruisseau Notre-Dame. Cependant, seulement une station a été installée en aval de ce ruisseau en raison de la charge de travail que nécessitait l'installation initiale d'une station de mesure prototype. Les deux stations auraient permis d'évaluer l'influence d'un quartier urbain sur la qualité de l'eau dans le cadre du projet rivEAU.

### **5.2 Description des sites de mesure**

Même si seulement une station de mesure a été installée au site aval durant le suivi en temps réel, cette section décrit les sites de mesure amont et aval qui ont été choisis pour l'installation des deux stations. L'auteur de ce mémoire a choisi de décrire le site amont, car il a déterminé ce site suite à une évaluation et inspection du ruisseau. Le site amont a été utilisé dans un suivi ultérieur réalisé par une chercheure postdoctorale, Janelcy Alferes. Les sites de suivi amont et aval sont indiqués à la figure 46 et le site amont ne correspond pas au même site que le point d'échantillonnage amont de l'étude préliminaire.

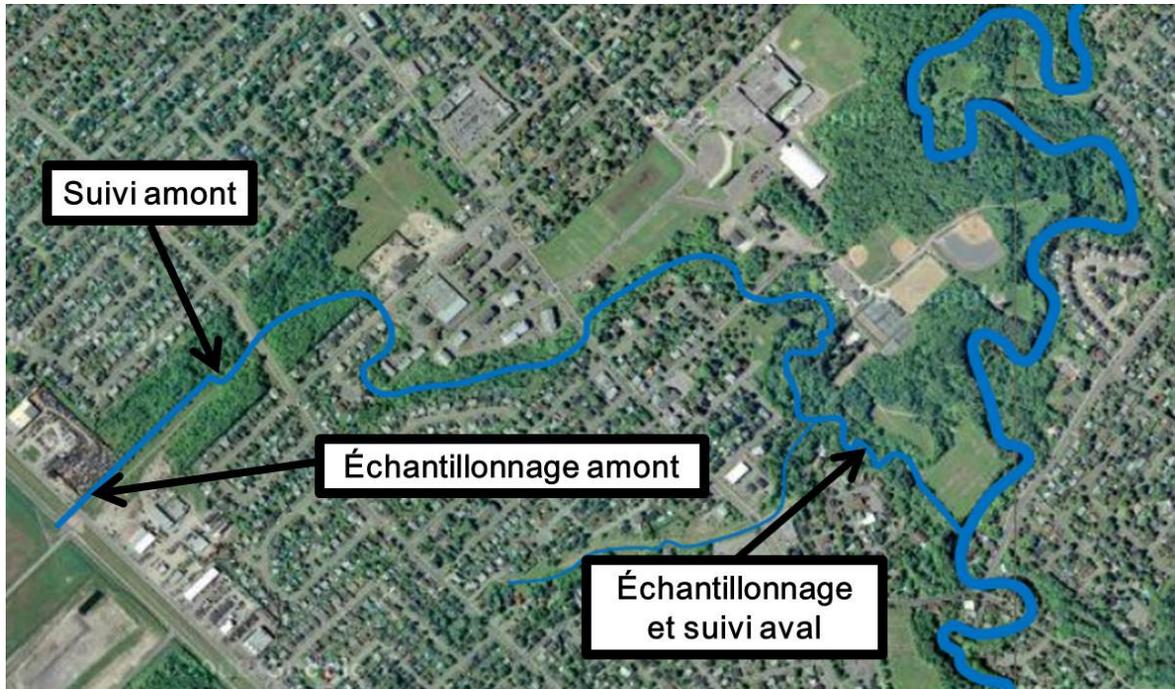


Figure 46 : Sites d'installation des deux stations du suivi en temps réel et points d'échantillonnage

Le site d'installation de la station aval est près du point d'échantillonnage aval de l'étude préliminaire. Les capteurs bulle à bulle et de vitesse Doppler ont été installés à un endroit différent des capteurs mesurant la qualité de l'eau à cause des conditions favorables distinctes pour le jaugeage du débit et la mesure de la qualité de l'eau. La figure 47 montre le point d'installation aval, encadré en noir, des capteurs mesurant la hauteur d'eau et la vitesse d'écoulement. Le choix du point d'installation de ces capteurs est basé sur les mêmes critères que ceux décrits dans la section 4.2 Description des points d'échantillonnage, soit un écoulement parallèle aux rives, l'absence de zone d'eau morte et des vitesses lentes (Aldegheri, 1979, Hach, 2008). Ces capteurs ont été vissés sur une planche de bois qui a été vissée sur une dalle de béton déjà présente sur le lit du ruisseau. L'aire de la section d'écoulement à ce point est constante favorisant donc la stabilité de la relation hauteur-débit (Jacon, 1987). De plus, un aspect intéressant avec cette dalle de béton est que la hauteur d'eau était mesurée à un point stable. Ainsi, la courbe de tarage développée durant ce suivi peut être réutilisée et améliorée durant un suivi ultérieur. À titre d'indication, le point d'échantillonnage aval de l'étude préliminaire est quelques mètres en aval de ce point.

Concernant les capteurs de la qualité de l'eau, ils ont été installés à quelques mètres en amont du point de mesure de la hauteur d'eau et de la vitesse d'écoulement. Un déversoir naturel, tel que présenté à la figure 48, garde une hauteur d'eau relativement haute même en débit d'étiage immergeant donc toujours les capteurs. Sur la figure 48, les capteurs ne sont pas encore installés. De plus, les conditions d'écoulement à ce point sont favorables, c'est-à-dire que la turbulence est faible et les vitesses se situent généralement entre 0,1 et 3 m/s (s::can, 2007). Par contre, comme indiqué à la figure 49, il y a un égout pluvial, situé quelques mètres en amont des capteurs, qui a un impact sur la représentativité des données mesurées. Les capteurs ont quand même été installés à cet endroit, malgré ce fait. L'influence de l'égout pluvial sur les données mesurées est discutée dans la section ADQATs.



Figure 47 : Point d'installation aval des capteurs mesurant la hauteur et la vitesse



Figure 48 : Point d'installation aval des capteurs mesurant la qualité de l'eau



Figure 49: Égout pluvial en amont de la station de mesure

Comme pour la station en aval, les points d'installation choisis en amont du ruisseau pour les capteurs servant au jaugeage du débit et mesurant la qualité de l'eau se situent à des endroits différents. Le point choisi pour les capteurs bulle à bulle et de vitesse Doppler sont l'entrée du ponceau situé sous la rue Turmel. Ce ponceau possède une section circulaire uniforme

permettant l'exploitation de la hauteur et de la vitesse mesurées en continu pour jauger directement le débit (Bertrand-Krajewski et al., 2000). De façon visuelle, l'écoulement y semble être uniforme et peu turbulent, deux conditions favorables à l'installation de ces capteurs (Hach, 2008). La figure 50 est une photographie de l'entrée du ponceau où les capteurs bulle à bulle et de vitesse Doppler auraient été installés.

Le point d'installation amont choisi pour les capteurs mesurant la qualité de l'eau se situe quelques mètres en amont du point de jaugeage du débit. Le choix de cet endroit est basé sur les mêmes critères que ceux décrits pour le point d'installation aval des capteurs mesurant la qualité de l'eau. La figure 51 est une photographie du point d'installation amont des capteurs mesurant la qualité de l'eau.



Figure 50 : Point d'installation amont des capteurs mesurant la hauteur et la vitesse



Figure 51 : Point d'installation amont des capteurs mesurant la qualité de l'eau

## **5.3 Matériels et installation**

### **5.3.1 Installation des capteurs**

La section précédente décrit les points d'installation des capteurs choisis pour effectuer le suivi, tandis que cette section discute des spécifications d'installation de chaque capteur qui sont indépendantes du site (angle, sens, etc.). Seulement les renseignements qui se sont avérés utiles à l'installation des capteurs durant ce suivi y sont discutés. Pour plus de détails, il faut se référer au manuel de l'utilisateur de chaque capteur. D'ailleurs, certains manuels ne présentent aucun renseignement sur l'installation des capteurs. Il est important de mentionner que les capteurs ont été installés in situ, c'est-à-dire qu'ils ont été installés directement dans le ruisseau.

- L'extrémité du capteur bulle à bulle doit être située au moins 2,5 à 5 cm sous la plus petite hauteur d'eau prévue et elle doit être perpendiculaire au sens du courant (Hach, 2008).

- Le capteur de vitesse Doppler doit mesurer la vitesse dans le sens contraire du courant. Une flèche sur le capteur indique dans quelle direction le positionner (Hach, 2008).
- Il est recommandé d'installer le Spectro::lyser dans une position horizontale, parallèle au courant et avec la face plane de la section de mesure orientée verticalement pour des vitesses inférieures à 1 m/s, ce qui est généralement le cas au point d'installation aval. Cette position assure qu'il n'y a pas de sédiments ou de bulles d'air qui s'accumulent sur les fenêtres optiques (enva). De plus, le capteur doit être installé à une hauteur minimum de 10 cm par rapport au lit du ruisseau afin d'éviter la suspension des sédiments lors du fonctionnement du système de nettoyage automatique par insufflation d'air (s::can, 2007).
- Les électrodes de l'Ammono::lyser doivent être dirigées vers le bas avec un angle d'inclinaison par rapport à l'horizontale d'au moins 10°. Si l'Ammono::lyser n'est pas installé verticalement, il doit être orienté vers le sens du courant. Ceci évite de réduire la durée de vie des électrodes causé par l'impact des particules transportées par le courant sur les électrodes. Comme pour le Spectro::lyser, il doit être installé à une hauteur minimum de 10 cm par rapport au lit du ruisseau (s::can, 2007).
- Les fenêtres optiques du Solitax sc doivent être orientées vers le sens du courant afin de ne pas les endommager par la collision de particules abrasives (Hach, 2004).
- Dans le manuel du capteur 3700 sc Conductimètre inductif, il n'y a pas de renseignements pertinents quant à son installation dans le ruisseau. Cependant, selon le soutien technique de la compagnie Hach, ce capteur doit être installé loin des autres capteurs, car ces derniers peuvent causer une interférence dans le signal. Aussi, il a été observé que le signal est affecté par la présence d'objets métalliques, tel que la cage de protection des capteurs.

### **5.3.2 Installation et protection du matériel**

Les conditions environnementales et le vandalisme sont des aspects qui doivent être considérés lors de l'installation et de la protection du matériel (Pressl et al., 2004). Durant le suivi en temps réel, tout le matériel était cadenassé soit la station mon*EAU*, le compresseur du système de nettoyage automatique par insufflation d'air et la cage de protection des capteurs. De plus, une clôture de protection était installée alentour de la station et du compresseur, comme montré à la figure 52. Pour protéger le matériel électronique de la station mon*EAU* contre l'humidité et la pénétration d'eau, la station a été construite de façon hermétique et un disjoncteur GFCI a été installé. Une planche de bois a quand même été posée au-dessus de la station pour assurer que l'eau n'y pénètre pas à l'intérieur lorsqu'il pleut.

Par ailleurs, le câble électrique connectant la source de courant qui provenait d'une résidence à la station mon*EAU* nécessitait un diamètre plus élevé que celui des câbles électriques standards. En effet, avec un diamètre standard, la distance entre la source de courant et la station (>50 m) aurait causé une diminution du voltage réduisant donc la durée de vie des appareils électroniques. Le diamètre utilisé pour éviter cette diminution était de 14 AWG.



Figure 52 : Protection de la station monEAU

Afin de protéger les capteurs contre le vandalisme et les débris transportés par le courant, ils étaient installés dans une cage de protection qui est présentée à la figure 53. Les parois de la cage ont été construites avec des treillis métalliques possédant des mailles carrées de deux pouces de côté. Cela a l'avantage de laisser l'eau circuler librement tout en protégeant les capteurs contre les gros débris. Cependant, la cage était installée directement dans le ruisseau et elle avait l'effet d'obstruer les débris transportés par le courant. Parfois, lorsque des débris

demeuraient pris entre les mailles du treillis métallique, l'effet d'obstruction augmentait et encore plus de débris y demeuraient pris. Cet effet est représenté à la figure 54 qui est une photographie prise après un évènement pluvial d'une période de retour de 5 à 10 ans dont l'intensité des précipitations a été de 37,7 mm en deux heures. Néanmoins, les capteurs demeuraient protégés, comme montrés à la figure 55, à l'exception d'une fois où l'impact d'une branche qui est passée au travers d'une maille a brisé une des électrodes de l'Ammonyser.

Pour que la cage demeure en place, elle était attachée à des arbres avec deux cordes. De plus, les câbles des capteurs étaient attachés sur toute leur longueur à un câble plus rigide et plus tendu. Ils étaient tous installés à l'intérieur d'un tuyau protecteur comme indiqué à la figure 53 par une flèche. Ainsi, si un débris avait heurté le tuyau protecteur, c'est le câble plus rigide et plus tendu qui aurait reçu l'impact.



Figure 53 : Installation des capteurs dans la première cage de protection



Figure 54 : Effet d'obstruction des débris par la cage de protection (5 août 2010)



Figure 55 : Protection des capteurs contre les débris

Pour effectuer un entretien sur les capteurs, il fallait sortir une sous-section de la cage de protection. La figure 56 illustre cette action et les différents capteurs y sont également localisés. Le capteur LDO mesurant l'oxygène dissous et la température n'était pas installé lors de la prise de cette photographie. Le retrait de la sous-section de la cage de protection pouvait être parfois difficile lorsqu'il y avait beaucoup de débris coincés entre les mailles de la cage. À la suite de l'apprentissage avec cette première cage de protection, une deuxième cage a été construite et est présentée à la figure 57. La sous-section de la cage a été remplacée par deux sous-sections. Lors d'un entretien sur un capteur avec la deuxième cage de protection, il faut simplement basculer le dessus de la cage et retirer la sous-section correspondante au capteur, comme montré à la figure 58. Ainsi, les autres capteurs demeurent en place et fonctionnels. Les sous-sections sont plus petites et plus légères que celle de la première cage. Aussi, elles n'ont pas tendance à accrocher des débris, car elles sont situées au milieu de la cage. De plus, le treillis métallique à l'avant de la cage a été remplacé par des plaques métalliques inclinées. Ces plaques diminuent l'effet d'obstruction de la cage en favorisant le glissement des débris sur les côtés, mais elles favorisent l'accumulation de sédiments à l'aval de la cage.

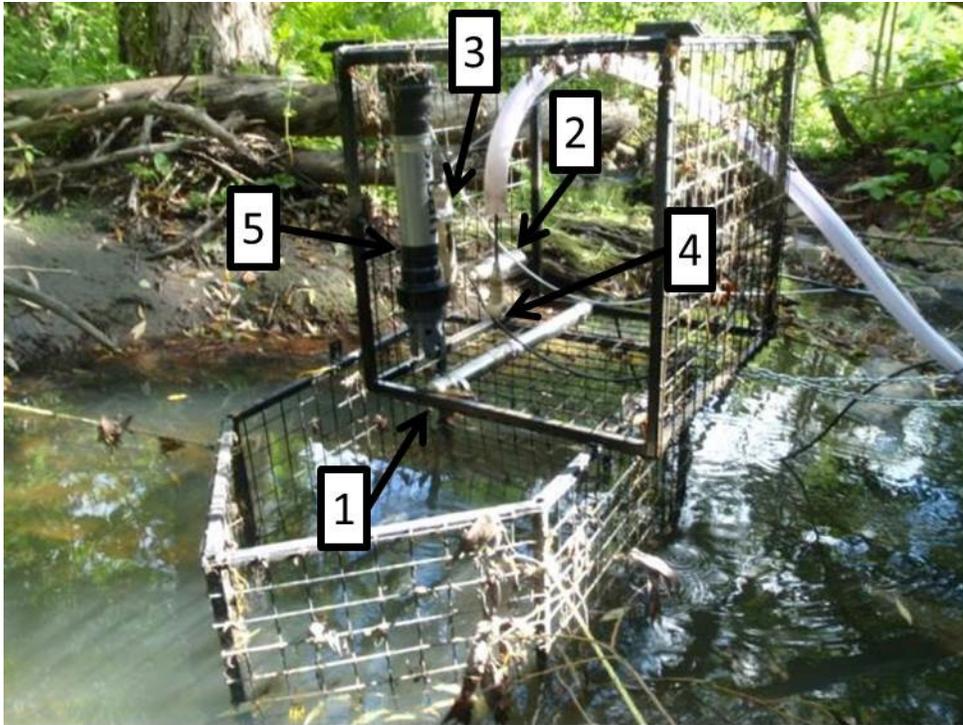


Figure 56: Retrait de la sous-section de la première cage de protection lors d'un entretien et localisation des capteurs

Numérotation des capteurs sur la figure 56 :

- 1 : Spectro::lyser (Nitrate et turbidité)
- 2 : Solitax sc (MeS)
- 3 : pHD sc différentiel numérique (pH et température)
- 4 : 3700 sc Conductimètre inductif (Conductivité et température)
- 5 : Ammo::lyser (Ammonium, potassium, pH et température)



Figure 57 : Deuxième cage de protection installée dans le ruisseau



Figure 58 : Retrait des sous-sections de la deuxième cage de protection lors d'un entretien

### 5.3.3 Systèmes de nettoyage automatiques

Comme mentionné précédemment dans le chapitre 2 Revue de littérature, le micro-encrassement des capteurs, principalement des capteurs optiques, cause une dérive dans le signal. Les systèmes de nettoyage automatiques remédient particulièrement à ce type d'encrassement (Bertrand-Krajewski et al., 2000). D'ailleurs, il a été observé que le ruisseau Notre-Dame a un pouvoir de micro-encrassement très fort. Il était donc nécessaire d'avoir un tel système pour les capteurs optiques. Dans ce cas-ci, le Solitax sc et le Spectro::lyser, deux capteurs optiques, possédaient respectivement un système de nettoyage automatique par raclage et par insufflation d'air. L'Ammono::lyser possédait aussi un système de nettoyage automatique par insufflation d'air qui était connecté avec celui du Spectro::lyser.

Le compresseur du système de nettoyage automatique par insufflation d'air a été installé dans une boîte d'insonorisation. Cette boîte réduisait considérablement le bruit émis par le compresseur et, donc, le bruit ne dérangeait pas les résidents habitant aux alentours de la station de mesure. La figure 59 est une photographie de la boîte d'insonorisation avec le compresseur à l'intérieur. Cette boîte a été construite tout simplement avec un bac de plastique et des planches d'isolant en mousse de polystyrène. Par ailleurs, durant le suivi, le moteur du compresseur a brisé et ceci pourrait être attribuable au fait que l'air ne circulait pas assez facilement à l'intérieur de la boîte.



Figure 59 : Boîte d'insonorisation du compresseur

Les deux systèmes de nettoyage automatiques permettent certains ajustements afin de les adapter aux conditions dans lesquelles les capteurs sont installés. Durant le suivi en temps réel, le système de nettoyage par raclage balayait les lentilles optiques à un intervalle de 15 minutes et le système par insufflation d'air injectait un jet d'air à un intervalle de 6 minutes, d'une durée d'une seconde et à une pression de six bars. La pression d'air avait été ajustée à six bars, car, selon le soutien technique de s::can, c'est la pression maximale acceptable par l'Ammono::lyser sans endommager les électrodes. Comme le Spectro::lyser et l'Ammono::lyser partageaient le même compresseur, la pression a donc été réglée à 6 bars. Cependant, si seulement le Spectro::lyser avait été connecté au système, la pression d'air aurait pu être de 10 bars qui est la pression maximale acceptable par ce capteur.

## 5.4 Entretien des capteurs

Durant un suivi en temps réel, un entretien régulier des capteurs est nécessaire afin d'assurer une bonne qualité des données mesurées (Pressl et al., 2004, van Griensven et al., 2000). Cette section décrit donc la méthodologie développée par l'auteur de ce mémoire pour l'entretien de chacun des capteurs installés lors du suivi en temps réel, excepté pour l'Ammonolyser qui n'a pas fonctionné. L'entretien consistait principalement en un nettoyage manuel et une vérification de leur fonctionnement. Par ailleurs, durant ce suivi, tous les entretiens effectués sur les capteurs et toutes les observations pouvant influencer la qualité des données ont été inscrits dans une fiche de vie comme suggéré par Bertrand-Krajewski et al. (2000).

Les capteurs ont chacun leurs techniques de nettoyage selon la persistance des résidus et ces techniques sont généralement décrites dans leur manuel de l'utilisateur. Par exemple, pour le Spectrolyser, selon un ordre croissant de la persistance des résidus, l'utilisateur devrait le nettoyer avec de l'eau, de l'éthanol, de l'agent nettoyant s::can ou de l'acide hydrochlorique à une concentration de 3% (s::can, 2007). D'après une des méthodes hors ligne d'évaluation de la qualité des données développées au chapitre 6 ADQATs, les capteurs doivent être nettoyés selon la technique enlevant les résidus les plus persistants. En effet, selon cette méthode, les mesures in situ avant et après un nettoyage manuel sont comparées afin d'évaluer l'effet de l'encrassement sur les mesures. Si des résidus influençant la qualité des données demeurent sur les capteurs, cette méthode sera biaisée. C'est pourquoi il faut nettoyer parfaitement les capteurs dès la première fois. Toutefois, il faut s'assurer que les techniques de nettoyage ne réduisent pas la durée de vie des capteurs.

Outre le nettoyage manuel, la vérification du fonctionnement (calibration ou fonctionnement général) des capteurs fait partie de l'entretien à effectuer. Cette vérification se fait à l'aide d'une charte de contrôle exploitant la mesure d'une solution standard avec un capteur (NMKL, 1990). Pour certains capteurs, une méthode redondante a été utilisée au lieu d'une solution standard. Brièvement, s'il n'y a pas de différence significative entre la valeur mesurée par un capteur et la valeur de la solution standard ou de la méthode redondante, cela signifie qu'il n'y a pas de mauvais fonctionnement détecté. Il est important de mentionner que le capteur doit être parfaitement nettoyé selon la méthode décrite précédemment avant

d'effectuer cet entretien. Ainsi, s'il y a une différence significative entre la valeur mesurée et la valeur de la solution standard ou de la méthode redondante, celle-ci n'est pas causée par l'encrassement du capteur. La figure 60 est une photographie prise durant la vérification du fonctionnement des capteurs avec des solutions standards. Comme le nettoyage manuel, la vérification du fonctionnement à l'aide d'une solution standard ou d'une méthode redondante est exploitée par une des méthodes hors ligne d'évaluation de la qualité des données développée au chapitre 6 ADQATs.

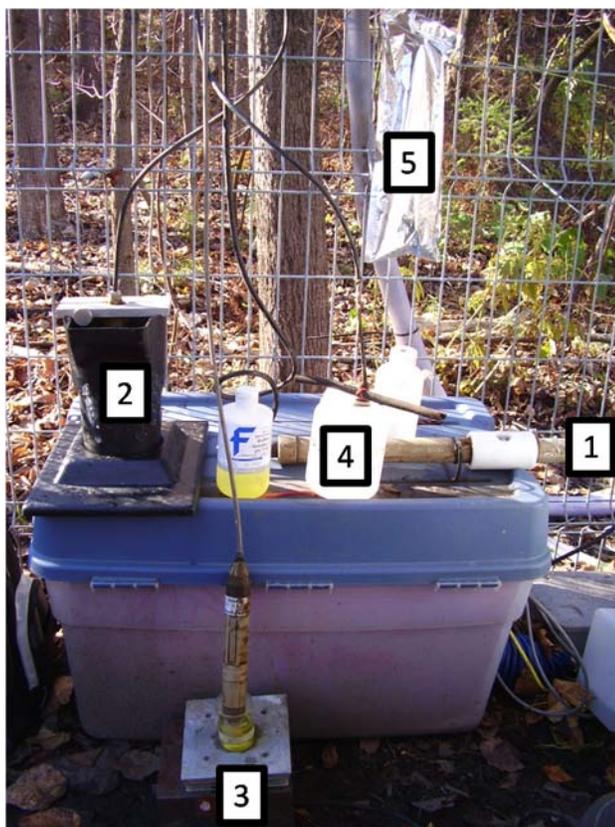


Figure 60 : Vérification du fonctionnement des capteurs avec des solutions standards  
(Voir texte ci-dessous pour plus de détails)

Les solutions standards et les méthodes redondantes utilisées pour la vérification du fonctionnement des capteurs sont dépendantes du paramètre mesuré et, parfois, de la gamme des valeurs rencontrées dans le milieu étudié.

- Le fonctionnement des capteurs optiques Spectro::lyser et Solitax sc a été vérifié avec de l'eau ultrapure (Voir les numéros 1 et 2 sur la figure 60).
- Le fonctionnement du capteur pHD sc différentiel numérique a été vérifié avec deux solutions tampons de pH 7 et 10 qui étaient cohérentes avec la gamme de pH du ruisseau (Voir le numéro 3 sur la figure 60).
- Le fonctionnement du capteur 3700 sc Conductimètre inductif a été vérifié avec une solution de NaCl (Voir le numéro 4 sur la figure 60) dont la conductivité se situait entre 1000 et 1100  $\mu\text{S}/\text{cm}$ , ce qui correspondait à la conductivité de l'eau du ruisseau en période de temps sec.
- Le fonctionnement du capteur LDO a été vérifié avec un sac d'air saturé en eau (Voir le numéro 5 sur la figure 60).
- Le fonctionnement des capteurs pHD sc différentiel numérique, LDO et 3700 sc Conductimètre inductif a aussi été vérifié en comparant la température mesurée in situ par ces capteurs à celle mesurée par un thermomètre (méthode redondante). Cependant, il n'est pas possible de calibrer la température du capteur LDO.
- Le fonctionnement du capteur bulle à bulle a été vérifié en comparant la hauteur d'eau mesurée par ce capteur à celle mesurée avec une règle graduée (méthode redondante).
- Le fonctionnement du capteur de vitesse Doppler n'a pas été vérifié. Cependant, ce capteur ne nécessite aucune calibration (Hach, 2007).

Les valeurs des solutions standards utilisées pour la vérification du fonctionnement des capteurs pHD sc différentiel numérique et 3700 sc Conductimètre inductif ont été déterminées à la suite de l'étude préliminaire. Cette étude a également démontré que les gammes de mesure de nitrate et de turbidité du Spectro::lyser disponible pour le suivi en temps réel n'étaient pas cohérentes avec la concentration de nitrate et la turbidité rencontrées dans le ruisseau. Un réducteur de la longueur du chemin optique a donc été ajouté au Spectro::lyser pour remédier à ce problème.

À la suite d'une vérification du fonctionnement, s'il y a une différence significative entre la valeur mesurée par le capteur et celle de la solution standard ou de la méthode redondante, cela indique qu'il y a possiblement un mauvais fonctionnement du capteur. Cependant, avant de calibrer le capteur en question, il est important de le nettoyer une deuxième fois et de le vérifier à nouveau avec la solution standard ou la méthode redondante. En effet, cette différence peut être causée par un mauvais nettoyage du capteur. Dans le cas d'une vérification avec une solution standard, s'il y a toujours une différence à la suite du deuxième nettoyage, il faut revérifier avec une nouvelle solution standard. Cette différence peut être causée par une contamination de la solution standard. Finalement, si cette différence est toujours mesurée, la calibration du capteur en est possiblement la cause. Le capteur doit donc être calibré avec la solution standard ou la méthode redondante qui a servi à la vérification de son fonctionnement.

Selon la méthodologie empruntée par l'auteur de ce mémoire pour la calibration du Spectro::lyser et du Solitax sc, cette dernière n'est possible qu'à la fin du suivi à l'aide de tous les échantillons analysés en laboratoire. La vérification avec une solution standard a permis seulement de détecter un micro-encrassement des fenêtres optiques ou un mal fonctionnement général tel que la présence de rainures sur les fenêtres optiques. Leur méthode de calibration n'est pas discutée ici. Par contre, Bertrand-Krajewski et al. (2007) proposent une méthode pour déterminer une régression polynomiale entre les MeS d'échantillons mesurés en laboratoire et leur turbidité mesurée par un capteur. Ce type de calibration est applicable avec le Solitax sc.

Par ailleurs, durant ce suivi, aucune règle n'a été utilisée pour déterminer si la différence entre la valeur mesurée par le capteur et la valeur de la solution standard ou de la méthode redondante était significative. Ceci était fait subjectivement, car il y a eu un manque de temps pour effectuer les tests préliminaires nécessaires à l'utilisation d'une méthode objective telle que décrite par NMKL (1990) et par Thomann et al. (2002), Aussi, un point qui a été oublié lors de la calibration a été de noter dans la fiche de vie l'intercepte et la pente de la courbe de calibration. Ces informations permettent de suivre l'évolution de la réponse de l'électrode indiquant le niveau d'usure de cette dernière (van Griensven et al., 2000).

## 5.5 Amélioration du logiciel monEAU

À la suite de l'utilisation du prototype du logiciel monEAU permettant l'affichage, la sauvegarde et le téléchargement des données mesurées en temps réel, l'auteur de ce mémoire a émis des suggestions à la compagnie Primodal Inc. afin d'améliorer ce logiciel. Ces suggestions sont présentées ci-dessous.

- Il faudrait pouvoir ajouter des commentaires qui figureraient sur l'interface affichant les données mesurées ou ajouter un code qui serait présent dans les métadonnées, afin d'indiquer une action ou une observation influençant les données mesurées. Par exemple, si le capteur a été calibré, nettoyé ou s'il était hors de l'eau, le fait de l'indiquer permettrait de comprendre le comportement de la série de données à ce moment.
- Lorsque les données sont exportées de la station à un ordinateur, le fichier contenant les données devrait contenir des informations telles que les unités des paramètres mesurés, l'emplacement du point de mesure, le nom du projet et les personnes impliquées dans le projet. Ainsi, tous les fichiers pourraient être retracés et utilisés par quelqu'un d'externe au projet.
- Lorsqu'un entretien est effectué sur un capteur, la sauvegarde des données est arrêtée afin d'éviter d'avoir des données non représentatives dans la base de données. Étant donné que les capteurs nécessitent une fréquence d'entretien différente, le logiciel devrait permettre l'arrêt de la sauvegarde d'un seul capteur à la fois. Ainsi, durant un entretien d'un capteur, les autres capteurs pourraient demeurer en fonction.
- Lors d'un entretien des capteurs, les données devraient continuer d'être affichées sur l'interface même si la sauvegarde des données est arrêtée, car celles-ci peuvent être nécessaires durant l'entretien. Par exemple, lors de la vérification d'un capteur avec une solution standard, il est important de voir la valeur mesurée même si celle-ci n'est pas sauvegardée.
- Lorsque des données de plusieurs paramètres sont exportées dans un fichier, toutes les données des différents paramètres sont classées dans la même colonne. Les paramètres devraient plutôt être classés par colonne, c'est-à-dire la première colonne

contiendrait le temps à laquelle les mesures sont prises et, ensuite, une colonne pour chaque paramètre contiendrait les valeurs mesurées. Il serait ainsi plus facile de travailler avec les données.

- Après avoir fait un rafraîchissement des données (les nouvelles données mesurées sont ajoutées à l'interface), il faut redimensionner l'échelle de vision des données. Cela peut être relativement accablant si, par exemple, quelques mois de données ont été récoltés et que seulement les données mesurées durant la dernière heure doivent être affichées. Le logiciel devrait afficher la même échelle et les mêmes données après un rafraîchissement.
- Il faudrait aussi pouvoir redimensionner l'échelle de vision des données et afficher les données voulues en délimitant un rectangle alentour de ces dernières. Cela faciliterait leur affichage.
- Un filtre simple soit la moyenne ou la médiane pourrait être appliqué à une série de données afin de réduire les effets du bruit et des données aberrantes. Ainsi, la tendance d'une série de données serait plus facilement observable.
- Le logiciel devrait permettre le choix d'un intervalle de mesure et, aussi, d'un nombre de mesures par intervalle. Par exemple, l'utilisateur devrait pouvoir choisir un intervalle de mesure d'une heure et que dix mesures soient prises à un intervalle de 5 secondes chaque heure. Cela permettrait de calculer des statistiques et de détecter les données aberrantes à chaque intervalle, tout en conservant un intervalle cohérent avec le processus.

Jusqu'à présent, seulement la suggestion du filtre simple a été ajoutée au logiciel. Cependant, la compagnie Primodal Inc. prévoit améliorer le logiciel et y inclure ces suggestions.

## **5.6 Résultats et analyses du suivi en temps réel**

Étant donnée la quantité innombrable de données récoltées durant le suivi en temps réel du ruisseau Notre-Dame, seulement le suivi de deux événements pluviaux est présenté ci-après.

La figure 61 et la figure 62 représentent respectivement les données brutes mesurées lors des deux évènements pluviaux datant du 2 août et 13 septembre 2010. Ces deux évènements pluviaux ont été choisis de façon à élaborer l'analyse. Les débits simulé et mesuré sont présentés sur ces figures et l'élaboration de la courbe de tarage est décrite à l'annexe 6. Les paramètres mesurés par l'Analyser ne sont pas présentés, car ce dernier n'a pas fonctionné durant tout le suivi. De plus, les paramètres mesurés par le capteur LDO ne sont pas présents à la figure 61, car il n'était pas installé à ce moment. Sur ces figures les températures identifiées Température 1, Température 2 et Température 3 correspondent respectivement aux températures mesurées par les capteurs pHD sc différentiel numérique, 3700 sc conductimètre inductif et LDO. Comme mentionné précédemment dans la section 3.3 Capteurs et paramètres mesurés, les MeS présentées sur ces figures proviennent d'une mesure de MeS par défaut résultante d'une précalibration effectuée par la compagnie Hach (Hach, 2004).

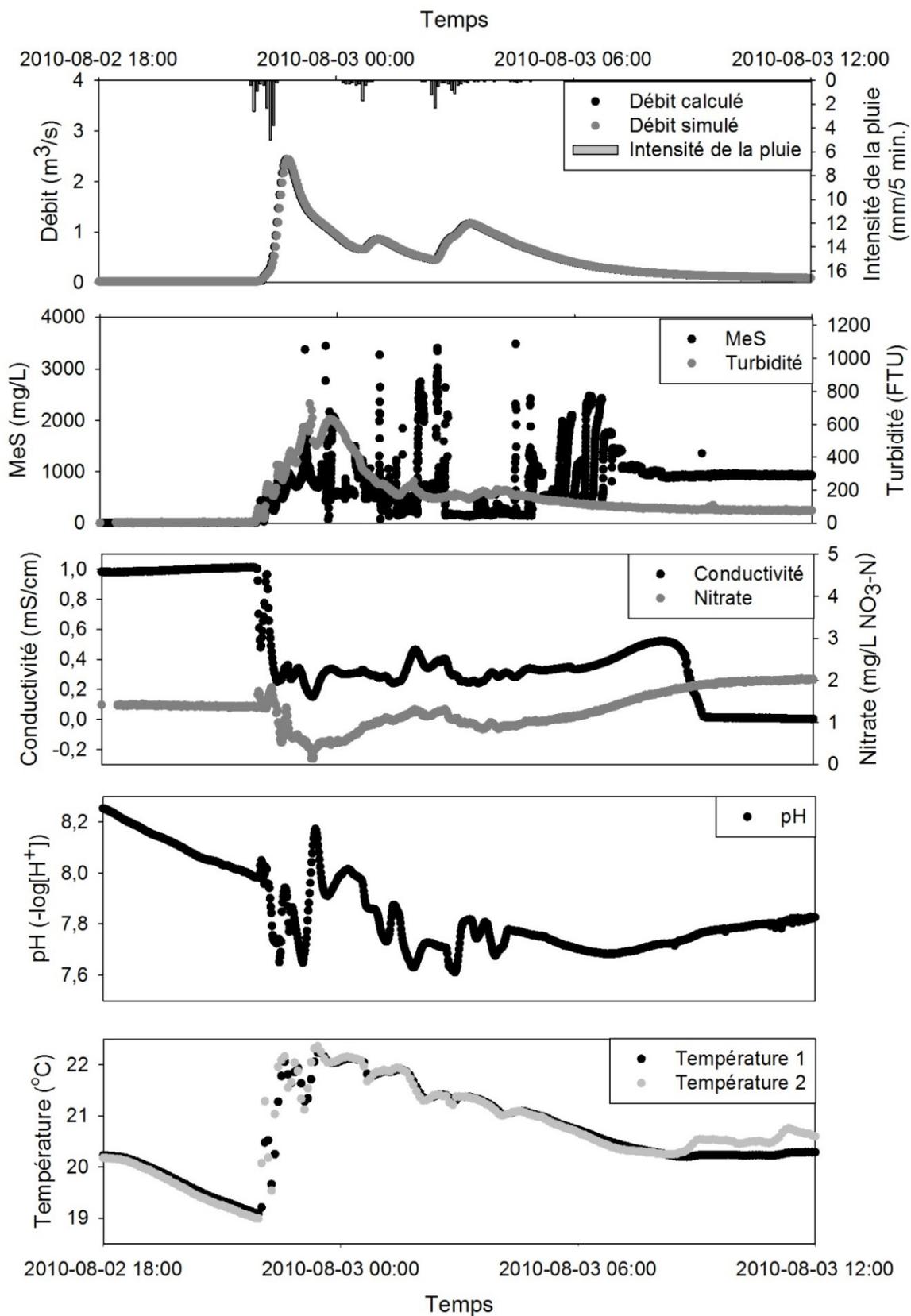


Figure 61 : Suivi en temps réel de l'évènement pluvial du 2010-08-02

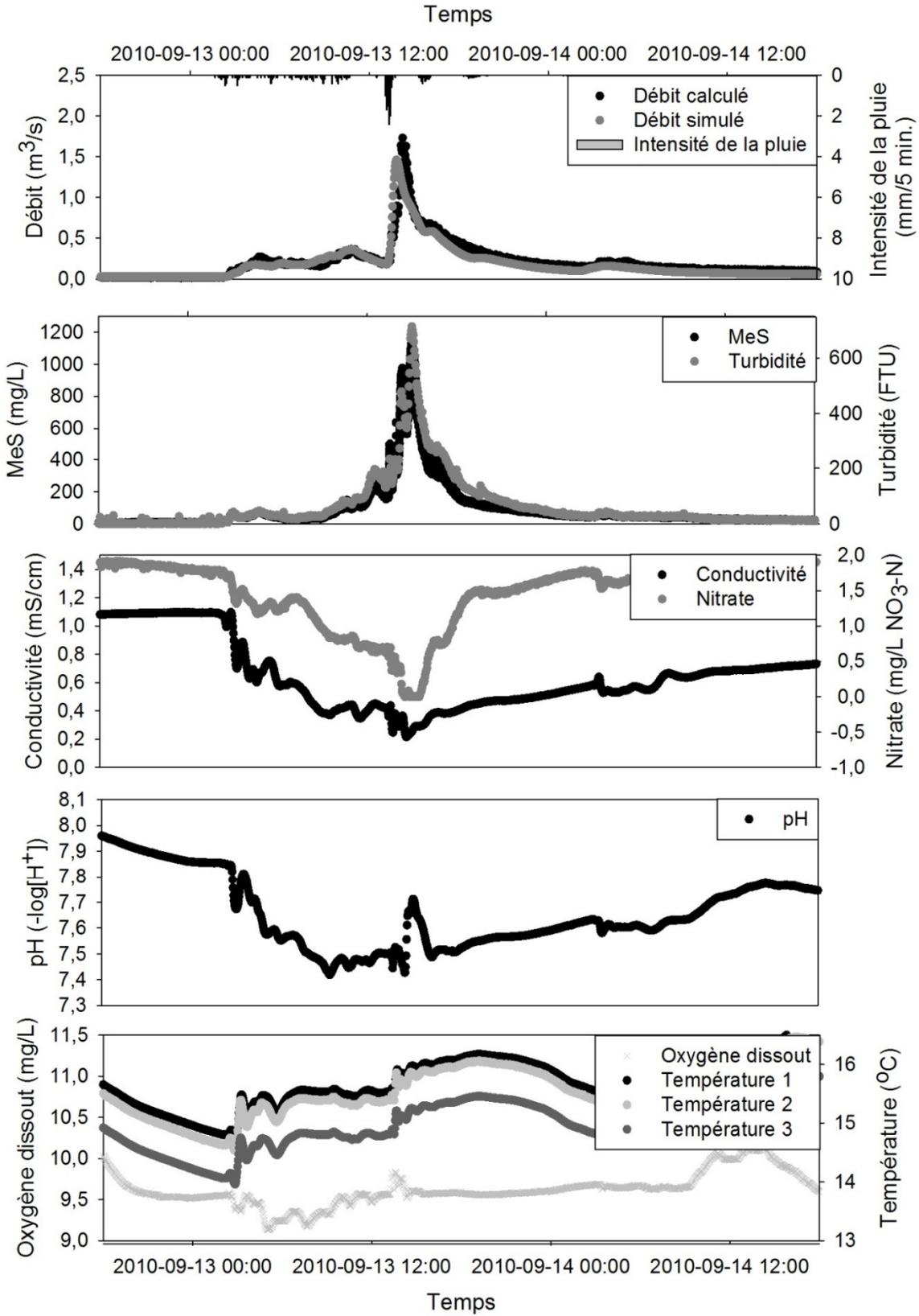


Figure 62 : Suivi en temps réel de l'évènement pluvial du 2010-09-13

La figure 61 et la figure 62 démontrent l'importance d'une haute fréquence de mesure. D'après ces figures, on remarque que le débit simulé représente très bien le débit calculé. D'ailleurs, la fiabilité du modèle hydrologique est augmentée grâce à la grande quantité de données mesurées comparativement à l'étude préliminaire. Cependant, les données doivent aussi être de bonne qualité pour que le modèle soit plus fiable (Vandenberghe et al., 2002), d'où l'importance des ADQATs.

Comparativement à l'étude préliminaire, le suivi en temps réel permet de mieux déterminer les dynamiques de la qualité de l'eau lors d'un évènement pluvial. Selon la figure 62, les MeS et la turbidité sont corrélées positivement avec le débit, tandis que la conductivité et la concentration de nitrate sont corrélées négativement avec le débit. Cependant, ces corrélations sont difficilement observables à la figure 61, car les données mesurées sont de mauvaise qualité. Ceci est justifié plus loin dans cette section. La figure 62 montre que, lorsque le débit redescend à sa valeur en temps sec, les MeS, la turbidité, la conductivité et la concentration de nitrate redeviennent aussi à leur valeur en temps sec. Concernant la concentration d'oxygène dissout, le pH et la température, ces paramètres ne semblent pas être corrélés avec le débit, toutefois ils varient lors d'un évènement pluvial. Par ailleurs, ces trois paramètres suivent un cycle journalier, mais ce dernier n'est pas observable sur les figures précédentes à cause d'une échelle de temps trop petite.

Les variations présentes sur la figure 61 et la figure 62 seraient difficilement mesurables avec une méthode basée sur l'échantillonnage. Par exemple, on remarque qu'au début des deux crues tous les paramètres varient selon différentes vitesses de variation. Contrairement à l'étude préliminaire, le suivi en temps réel confirme que, durant la saison estivale, les surfaces imperméables augmentent la température des cours d'eau dans lesquels l'eau de ruissellement se jette à la suite d'une précipitation. En effet, au début des deux crues, on observe une augmentation rapide de la température.

En observant la figure 61 et la figure 62, on peut détecter visuellement des comportements anormaux dans les séries de données qui sont possiblement des indicateurs de fautes. Tout d'abord, à la figure 61, on observe des sauts et des pentes abruptes dans les séries de MeS et de turbidité probablement causés par les nettoyages automatiques et l'encrassement des

lentilles. De plus, la série de MeS ne suit pas la même tendance que la turbidité et, pourtant, ces deux paramètres sont corrélés positivement au débit selon la figure 62.

On remarque aussi à la figure 61 que l'écart entre les températures 1 et 2 augmente rapidement vers la fin des séries de données, contrairement à la figure 62 où les écarts entre les températures 1, 2 et 3 demeurent relativement constants. Le comportement de l'écart entre les températures 1 et 2 à la figure 61 indique que la calibration n'en est probablement pas la cause. En se référant à la fiche de vie complétée tout au long du suivi, il est noté que, le 5 août 2010, le capteur de conductivité (Température 2) était hors de l'eau suite au déplacement de la cage par des débris transportés lors de la crue. Ceci explique donc l'écart entre les séries de température à la figure 61. De plus, la conductivité mesurée à ce moment est nulle, ce qui est caractéristique de ce capteur lorsqu'il est hors de l'eau. Le comportement des écarts entre les températures 1, 2 et 3 à la figure 62 indique qu'une mauvaise calibration en est probablement la cause. D'ailleurs, la température 3 est mesurée par le capteur LDO et sa calibration n'est pas possible.

Les exemples précédents démontrent que l'utilisation de trois capteurs est plus efficace pour détecter quel capteur est fautif (Menendez Martinez et al., 2001). En effet, il est difficile de déterminer visuellement quelle série est fautive à la figure 61, contrairement à la figure 62 où la série de la température 3 est sûrement fautive étant donné son écart plus grand par rapport aux deux autres séries de température.

Les analyses précédentes démontrent qu'une multitude de facteurs peuvent influencer la qualité des données mesurées durant un suivi en temps réel. Ce type de suivi doit donc être accompagné d'ADQATs afin d'obtenir des données de qualité connue et accrue. Ces analyses donnent également un aperçu de certaines méthodes de détection de fautes et de diagnostic des capteurs.



## 6. ADQATs

L'Article *Hardware and software tools to increase data quality in automated water quality measuring stations* a été soumis à la conférence AutMoNet 2011. Toutefois, cette conférence n'a pas eu lieu et l'article n'a jamais été publié. Il est présenté ci-après tel qu'il a été soumis à la conférence AutMoNet 2011. Aucune modification n'a été effectuée pour son insertion dans ce mémoire, malgré le fait que certains aspects ont déjà été discutés dans les sections précédentes.

Des détails supplémentaires sur l'application des méthodes hors ligne et en ligne se trouvent respectivement à l'annexe 7 et l'annexe 8.

## 6.1 Résumé

Un suivi en temps réel d'un ruisseau urbain a été effectué à l'aide de la station monEAU. Ce suivi a permis de bien décrire les dynamiques de la qualité de l'eau de ce ruisseau présentes durant les évènements pluviaux grâce à une haute fréquence de mesure. Cependant, une telle fréquence de mesure produit une quantité innombrable de données et leur qualité peut être incertaine ou mauvaise. Des outils automatiques d'évaluation de la qualité des données sont donc indispensables pour valider les données mesurées. L'article suivant présente des méthodes univariées hors ligne et en ligne d'évaluation de la qualité des données. Les méthodes hors ligne utilisent des chartes de contrôle basées sur la comparaison de mesures. Les méthodes en ligne détectent et éliminent, premièrement, les données aberrantes et, ensuite, lissent les données pour calculer certaines caractéristiques. La détection de fautes s'effectue en imposant des limites d'acceptabilité à ces caractéristiques. Outre les outils automatiques d'évaluation de la qualité des données, cet article discute aussi de l'influence du matériel utilisé lors d'un suivi sur la qualité des données.

## 6.2 Hardware and software tools to increase data quality in automated water quality measuring stations

Pascal Poirier<sup>1</sup>, Peter A. Vanrolleghem<sup>1</sup>

<sup>1</sup> modelEAU, Département de génie civil et de génie des eaux, Université Laval, 1065 Av. de la Médecine, G1V 0A6, Québec (Québec), Canada (Email: [pascal.poirier.1@ulaval.ca](mailto:pascal.poirier.1@ulaval.ca); [peter.vanrolleghem@gci.ulaval.ca](mailto:peter.vanrolleghem@gci.ulaval.ca))

### 6.2.1 Abstract

An automated water quality monitoring station, *monEAU*, has been installed to in-situ measure the water quality of an urban river observed during rain events. The measurement campaign spanned from July to November 2010. A high measuring frequency has allowed to well describe the dynamics in the reactive water body. However, such high measuring frequency also has the drawback of producing a large amount of data with an uncertain quality. Therefore, automatic data quality assessment tools (ADQATs) are necessary for data validation. In this paper, off-line and on-line methods for data validation are presented using univariate information. They are easily implemented and require little process knowledge. The off-line methods use control charts based on comparative measurements. The on-line methods firstly detect and replace outliers, and then calculate smoothed data which are used to compute data features allowing fault detection. Faults are detected by comparing these features with acceptability limits. Besides the ADQATs, the influence of the hardware used for collecting reliable data sets is discussed based on the experience learned from this measurement campaign.

### 6.2.2 Keywords

Automated monitoring station; Data quality assessment; Data smoothing; Data validation; Fault detection; Outlier detection.

### 6.2.3 Introduction

Nowadays, on-line sensors are largely used for water quality monitoring of different water bodies to reach different goals. Data acquisition systems can provide high measuring frequencies which allows better description of the dynamics in reactive water bodies (river, combined sewer, etc.) compared to grab or composite samples. Consequently, a large amount of data is produced with uncertain quality. Manual data validation is very tedious. Therefore, automatic data quality assessment tools (ADQATs) are necessary to validate time series and to use them for their meant application. In that sense, poor quality data could drastically affect the results of their application, namely water quality models for river basin management, WWTP control rules, etc. According to the *monEAU* vision (Rieger and Vanrolleghem, 2008), ADQATs should be based on reference samples, sensor status/diagnosis data and time series information. Some authors have proposed validation methods using several tests applied to streamflow data with one test using a physically-based real-time model (Krajewski and Krajewski, 1989), to urban drainage data with some multivariate tests using material and analytical redundancy (Mourad and Bertrand-Kralewski, 2002, Bertrand-Kralewski et al., 2000) or Principal Component Analysis (Branisavljevic et al., 2010).

This paper presents simple ADQATs using univariate methods which depend only to a small content on process knowledge. They have been tested on river water quality time series obtained from in-situ measurement using the *monEAU* monitoring station. As suggested by Branisavljevic et al. (2010), the ADQATs presented in this study are a combination of several validation methods applied on the same data set in order to improve their validation performance. The finally produced sensor status/diagnosis data information and the validation score used for data quality assessment and diagnosis are not described in this paper, for lack of space.

Besides ADQATs, the choice of the hardware is another important aspect of a water quality monitoring system in order to collect reliable data sets. Vandenberghe et al. (2005) and Pressl et al. (2004) who have performed respectively on-line and in-situ river water quality monitoring have faced different problems related to the chosen hardware. Thus, this paper also discusses the influence of the hardware as experienced by the authors, namely

installation, cleaning system and protection of the in-situ river water quality monitoring system.

#### **6.2.4 Case study**

The ADQATs have been applied to an in-situ monitoring station, *monEAU*, measuring the water quality dynamics during rain events affecting a small urban river located in Ancienne-Lorette (Québec, Canada). The measurement campaign spanned from July to November 2010 and several sensors monitored different water quality parameters. Here, a turbidity sensor is used to demonstrate the application of the ADQATs. The reading interval was five seconds. The time series consists of the turbidity sensor suspended solids default output which is represented by  $SS^*$ . The data are non-stationary and have heteroscedastic noise characteristics (the noise generally increases with the reading value).

#### **6.2.5 Hardware: installation, cleaning-system and protection**

Hardware is an important contributor to data quality, i.e. the data quality is affected considerably by the installation, the cleaning system and the protection of the sensors to floating debris. When a sensor is installed, the location chosen has to meet certain criteria (water depth, flow velocity, mixing characteristics, pressure, etc.) which are usually specified in the user's manual of the sensor. During the measurement campaign, a storm sewer outlet was present just upstream of the *monEAU* station's location (Figure 63) which due to a too short mixing distance leads sometimes to unrepresentative readings. However, before the sensor installation, the authors thought that the mixing distance was sufficient. When run-off is suddenly discharged from the storm sewer outlet due to rain, one observes in the data that the water quality of the river and of the outlet are very different because the river water quality reacts slower than the outlet water quality. In fact, depending on the flow paths, the sensors measure either the storm run-off or the upstream river water quality. Consequently, the user has to be careful when using such time series. ADQATs have to be able to deal with such high temporal variability. Figure 63 and Figure 64 show respectively the storm sewer outlet upstream of the sensors and a typical on-line  $SS^*$  and conductivity smoothed time

series calculated with the ADQATs developed in this article and the rain intensity to better understand the dynamics of the water quality. The zoom window presented on Figure 64 shows a high temporal variability caused by the insufficient mixing distance. After a while, the storm run-off stops and the river water quality dynamics become smoother.



Figure 63 : Storm sewer outlet just upstream the monEAU station

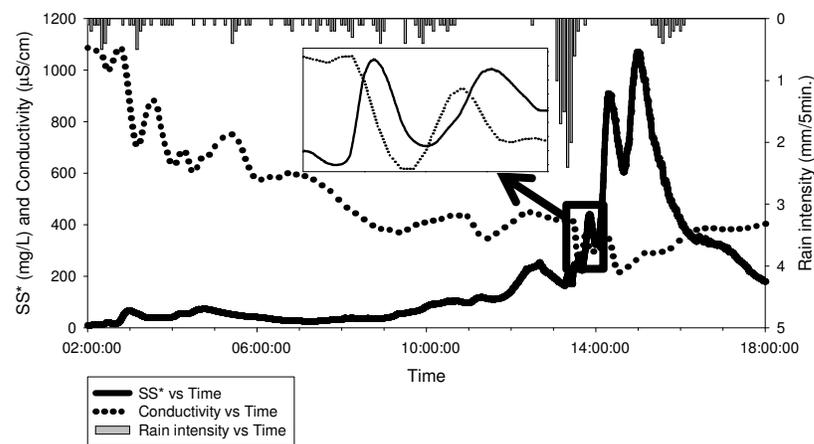


Figure 64 : SS\* and conductivity smoothed time series during a rain event

Furthermore, the installation must consider floating materials, vandalism, macro-fouling and micro-fouling (Bertrand-Krajewski et al., 2000), and a steep rise of water level during rain

events. Pressl et al. (2004) faced similar problems, but, considering their much larger river, the adopted solution is different. In this small river, the way that was chosen to install and protect the in-situ sensors consists of a secured cage containing the sensors dropped on the river bed and attached with cables. The first prototype cage was made of a metal grid which reduced the macro-fouling effect by avoiding the contact of debris (branches, leaves, etc) with the sensors. However, this setup suffered from obstruction by floating materials. As can be observed on Figure 65, the cage collected an increasing amount of larger and larger branches during rain events. As a result, the cage with the sensors rose out of the water. Still the sensors remained safely protected (Figure 66). ADQATs are easily able to detect the exact time when the sensors lost the contact with the water.



Figure 65 : Protective cage blocks debris



Figure 66 : Sensors rose out of the water but remain protected

This type of problems has occurred during a number of intense rain events. Therefore, a metal front plate has been added to the second prototype cage to reduce the obstructing effect of the metal grid and to facilitate floating material to flow by (Figure 67). It has been demonstrated that the effect of the metal front plate is negligible on the representativeness of the measured data.

Besides the installation and the protection of the sensors, an automatic cleaning system seems indispensable when using optical sensors to maintain long-term reliable measurement. Of course the cleaning intensity will depend on the fouling characteristics of the water (Bertrand-Krajewski et al., 2000). In this case, it was necessary to have an automatic cleaning system to reduce micro-fouling effect. Here, the turbidity sensor is provided with a wiper which was set up to clean every 15 minutes (Figure 68). One can observe that the sensor is very fouled except where the wiper passes by. Figure 69 demonstrates the micro-fouling effect during a rain event which causes a fast drift in the on-line SS\* data. Every 15 minutes there is a shift which corresponds to the time interval of the automatic cleaning. Even with an automatic

cleaning system, ADQATs obviously have to be implemented to detect when data are affected by such micro-fouling effects. The methods mentioned below easily detect erratic behavior as presented in Figure 69. To still get good data under such conditions consists of increasing the wiper's frequency.



Figure 67 : Cage with metal front plate



Figure 68 : Turbidity sensor with wiper

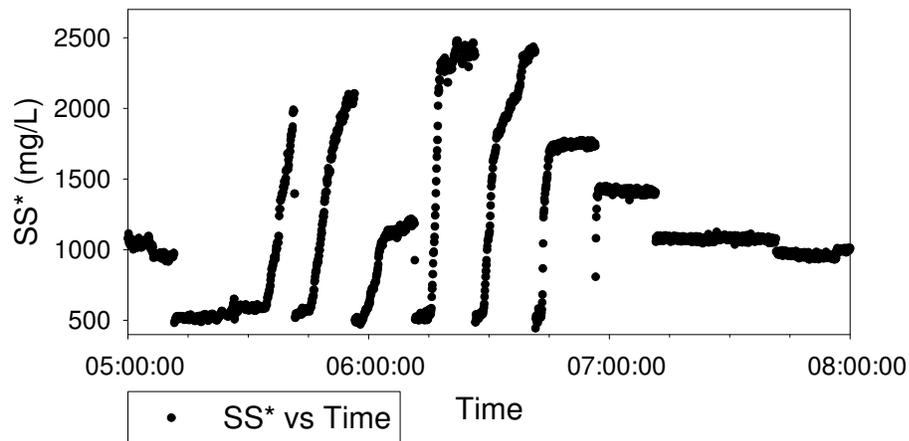


Figure 69 : Micro-fouling effect during intense rain event

## 6.2.6 ADQAT

The developed ADQATs can be divided in two main parts: off-line methods using control charts that use comparative measurements and on-line methods using time series information. The on-line methods firstly detect and replace outliers by using univariate autoregressive models which can be seen as an analytical redundancy method (Hill and Minsker, 2010). The resulting time series is called the accepted data time series. Then, they compute information by using a smoothed data series that is compared with acceptability limits. The data features computed by the on-line methods are the locally realistic range (Mourad and Bertrand-Krajewski, 2002), rate of change, variance and distance from other readings (Ni et al., 2009). Moreover, the autocorrelation of the residuals between accepted data and smoothed data are computed for fault detection. It is important to mention that a fault is defined as an anomalous data that exceeds normal expected behavior (Ni et al., 2009).

### 6.2.6.1 Off-line methods

The off-line methods use comparative measurements, but since the turbidity sensor  $SS^*$  output are obtained from the factory calibration without local calibration (an option provided with the sensor (Hach, 2004), the sensor measurements could not be compared to laboratory analysis results. The Shewart chart proposed by Thomman et al. (2002) is therefore not

applicable in this case. Thus, a control chart which compares a standard solution value and its sensor measurement, called the standard solution control chart, seemed appropriate to detect drift, shift or outlier effects. Also, a control chart which compares the in-situ measurements before and after manual cleaning, called the cleaning control chart, was used to detect fouling effects. Note that an automatic cleaning system increases the time interval before a manual cleaning is necessary (Bertrand-Krajewski et al., 2000).

For the turbidity sensor, the standard solution control chart used distilled water as standard and the measurements obtained should be less than 1 mg/L for good quality data (Hach, 2004).

The cleaning control charts should be evaluated when the water quality is stable, ideally during dry weather periods. This ensures that the difference between the measurements before and after cleaning is due to the fouling effect and not due to the natural variation of the water quality. To detect whether there is a significant difference between the in-situ measurements before and after the manual cleaning, a one-sided (fouling cannot have a negative effect on readings) Student's t-test with a confidence interval of 99% considering two different population variances is used. Note that the variance before and after the cleaning may be different and, therefore, a t-test with two population variance needs to be used (Montgomery, 2001).

Figure 70 presents the manual cleaning control chart and Figure 71 the standard solution control chart obtained during the campaign. The square closed and the round open markers mean respectively that the fault remains undetected and fault is being detected. Both control charts show more faults towards the end of the measurement campaign. This is probably due to the degradation of the optics and of the automatic cleaning system (e.g. the wiper needs replacement).

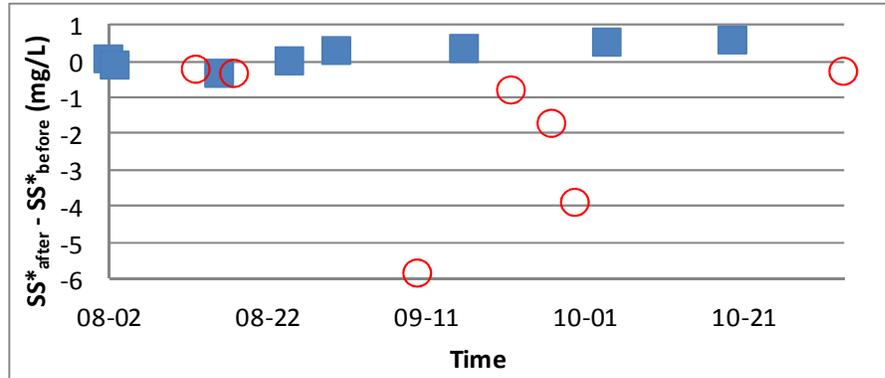


Figure 70 : Manual cleaning, Filled symbols: accepted, Open symbols: failed test and fault

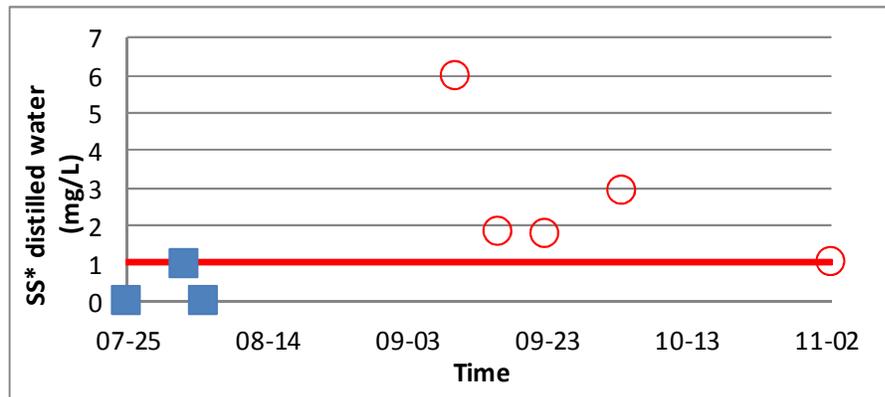


Figure 71 : Standard solution control charts Filled symbols: accepted, Open symbols: failed test and fault

### 6.2.6.2 On-line methods

The on-line data validation methods are based on time series information. The proposed integrated set of methods can be divided in two main steps, outlier detection and fault detection. An outlier is considered a fault too but it has to be detected first and replaced because it significantly skews the mean, variance, gradient, and other data features (Ni et al., 2009). Outlier detection is carried out by using univariate autoregressive models which compare measured values with forecast values. Once outliers have been detected and replaced by the forecast value, the resulting time series are smoothed. Faults are detected by defining acceptable limits to data features. Two of the data features used for fault detection are the locally realistic range and the rate of change, slope between two data, which are

computed from the smoothed data. The smoothed data are calculated using a central weighted average over a window of a certain length. Using a smooth to calculate the data features has the advantage of avoiding the corruption by the signal noise. Also, the autocorrelation of the residuals (calculated as the difference between the accepted data and the smoothed data) is computed for fault detection. It gives information on the smoother and the noise of the accepted data, i.e. if there is a residual autocorrelation this means that either the smoothed data is not representative of the accepted data or the signal noise is not randomly distributed. The signal noise variance of the time series is estimated by calculating the residuals standard deviation (RSD). Moreover, before the RSD is calculated, the data have to be transformed because the signal noise is heteroscedastic (the noise generally increases with the reading value) which, if the data are not transformed, do not allow the use of fixed RSD acceptable limits for fault detection as developed in this report. Figure 72 illustrates the integrated on-line data validation methods. All steps are explained in details below.

#### *6.2.6.2.1 Outlier detection*

Olsson and Newell (1999) proposed to fit a model to the data to detect outliers and, then to test the residuals. This outlier detection method is based on univariate autoregressive models which determine a prediction interval. The prediction interval is defined by adding or subtracting a multiple of the standard deviation of the forecast error to the forecast value. If the one-step-ahead sensor data falls outside the prediction interval, it is considered as an outlier and the models, forecast value model and forecast error model, are not updated with the new data to calculate the next one-step-ahead predictive interval. The outlier is then eliminated from the time series and it is replaced by the expected value, forecast value, which will eventually be used in the smoother (see Figure 73). In case the one-step-ahead sensor data falls inside the prediction interval, the models are updated and the next one-step-ahead prediction interval is calculated and it goes on so forth. Hill and Minsker (2010) described a similar procedure applied to different data-driven univariate autoregressive models and discussed why these models simplify the outlier detection process.

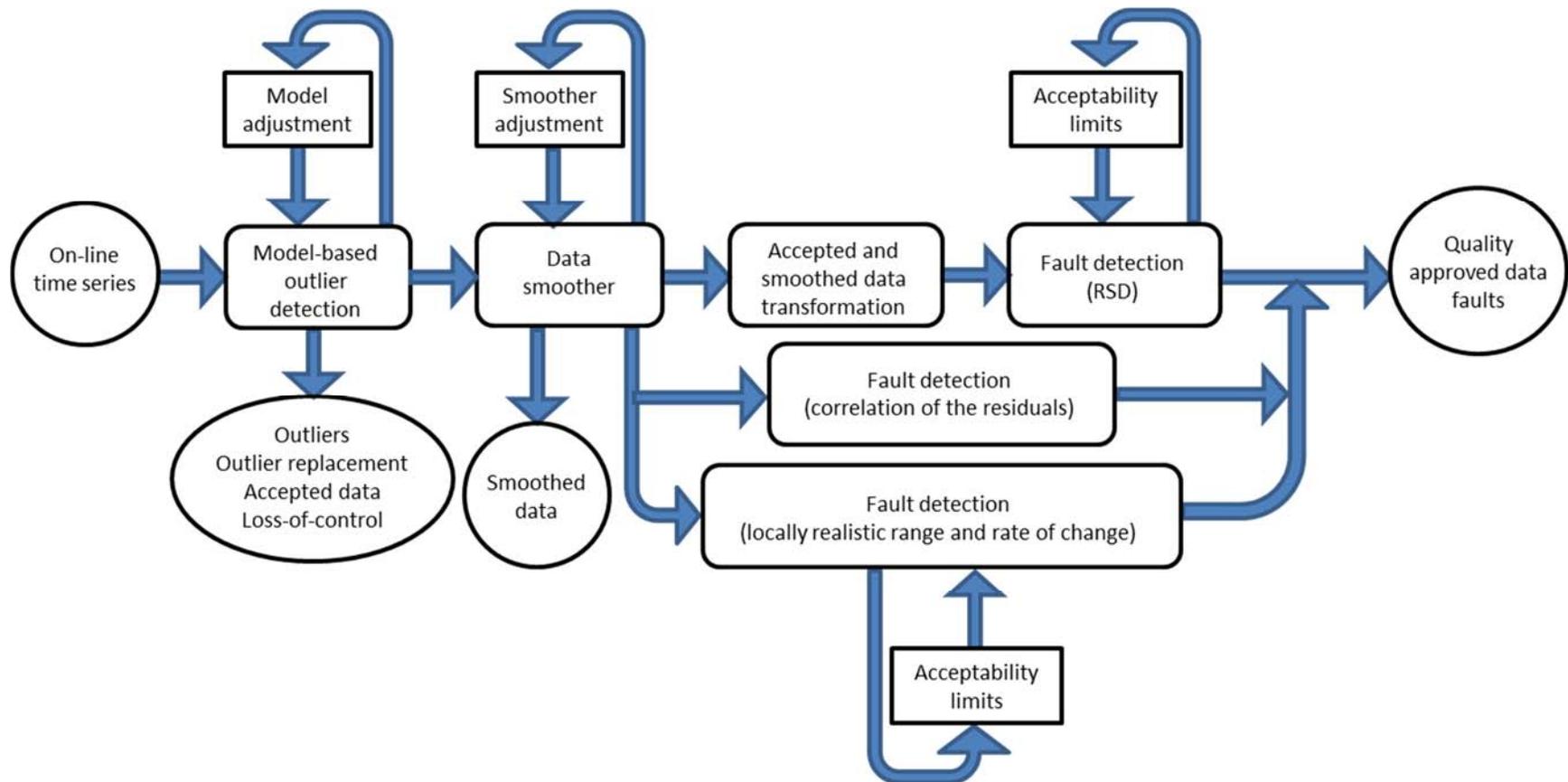


Figure 72: Scheme of the integrated on-line data validation methods proposed in the paper (RSD= residuals standard deviation)

After evaluation of different models, the ones retained to predict the expected value and the standard deviation of the forecast error are respectively a third-order exponential smoothing and a simple exponential smoothing. The prediction interval is defined by

$$x'_{T+1}(T) \pm K\sigma'_{e_1}(T)$$

where  $x'_{T+1}(T)$  and  $\sigma'_{e_1}(T)$  are the one-step-ahead forecast value and standard deviation of the forecast error, and  $K$  is a factor which multiplies the one-step-ahead standard deviation of the forecast error.  $K$  can be adjusted to make the outlier detection more or less restrictive.

These models are simple and easy to tune, i.e. only one smoothing constant ( $\alpha$ ) needs to be determined for each model. Sequence of trials were conducted on a set of data using several values for  $\alpha$ , and the value of  $\alpha$  that minimizes the root mean square of the difference between on-line and forecasted values is chosen. The method used to detect outliers is mainly described in Montgomery (1990). With the outlier detection method, a quality index is given to each data, i.e. if the data fall inside the prediction interval the quality index is 1, if not it is 0. When there is a sudden change (step) or an abnormal abrupt slope in the time series, this may lead to many consecutive data falling outside the prediction interval called here a loss-of-control. A method has been implemented to minimise the data loss. When a loss-of-control occurs, a reinitialization of the forecasting models is done  $n$  data after the first outlier and the outlier detection method is applied backward on the time series. When the outlier detection method has returned to the point where the loss-of-control occurred, the method is applied forward, starting halfway between the loss-of-control point and the reinitialization point. Thus, data set is obtained where the outlier detection has been applied backward and forward. The first half of this data set used the backward application and the second half used the forward application to detect outliers. Thus, the used of unrepresentative prediction interval due to initial conditions is fixed because it lets the prediction interval to get stabilized before it is used to detect outliers. A quality index of 0 is given to the data which belong to the loss-of-control as well. An interesting aspect is that the SS\* time series noise is heteroscedastic and the prediction interval must therefore adapt to the noise variance changes. Figure 73

presents an application of the outlier detection method with a loss-of-control situation. On this time series, the loss-of-control is caused by a sudden change occurring at the first vertical black bar. The reinitialization occurs 40 readings later at the second vertical black bar. Full recovery has been achieved thanks to the reverse application of the outlier detection method.

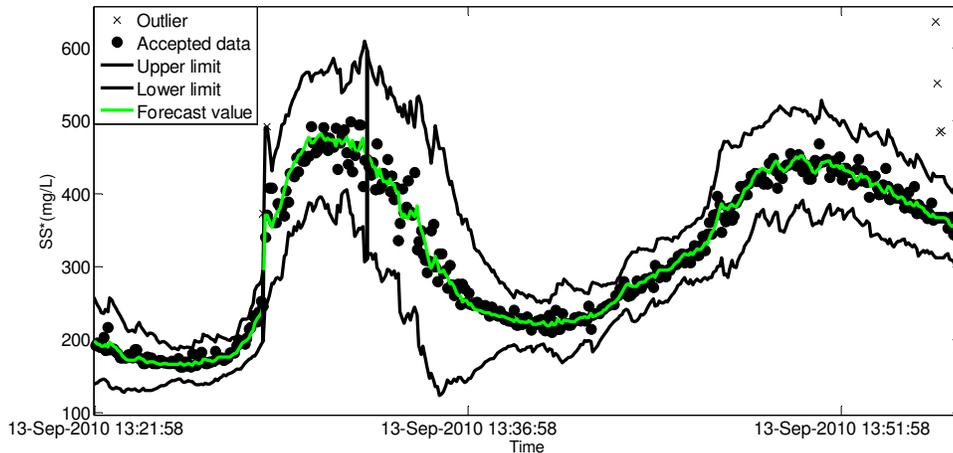


Figure 73: Illustration of the detection method for outliers (indicated with crosses) on SS\* time series

The user must check what may have caused the loss-of-control. If it is due to a high rate of change as presented in Figure 73, the fault will be detected later by the implemented fault detection methods.

#### 6.2.6.2.2 Data smoothing

Smoothing the time series is an important step in the developed on-line data validation methods, because from the smoothed data a number of features are calculated that allows data validation, namely locally realistic range, rate of change, residuals standard deviation and autocorrelation of the residuals. Using the smoothed data to calculate these features avoids data corruption by, for instance, noise. Here, the Nadaraya-Watson kernel estimator which is nonparametric regression smoother is used to smooth the time series (Schimek, 2000). Kernel smoothing belongs to the class of weighted moving averages which is suitable

for time series smoothing as it assigns a weighting factor to each value in the data series according to its age. The most recent data gets the greatest weight and each data gets a smaller weight as we move backward and forward in the series. To determine the bandwidth ( $h$ ) of the kernel estimator, cross-validation was used (Schimek, 2000).

The approach used to check the independence of the residuals is a runs test with a confidence interval of 95% carried out on a moving window (Dochain and Vanrolleghem, 2001). The window size chosen here is 30 which leads to a rapid detection of dependence of residuals. If the runs test fails, abnormal behaviour is detected (non-random noise) or the bandwidth of the kernel estimator not suitable for this time series. As mentioned above, outliers are replaced by forecast values which are then used further by the smoother. The fraction of forecast values used to smooth the data at time  $t$  is calculated to give an indication on how good the smoothed data and the deduced data features are at this time.

#### *6.2.6.2.3 Determination of the model and smoother parameters and defining the acceptability limits*

The model and smoother parameters, and the acceptability limits of the data features are determined on the basis of time series which have been accepted by the user to be of good quality. During the measurement campaign, three rain events have been chosen to assess normal behaviour. The limits of the locally realistic range, the rate of change and the RSD are determined by taking certain percentiles. Krajewski and Krajewski (1989) used a similar procedure to determine the acceptability limits of data features. Experience with the application of on-line data validation methods on a number of actual time series may induce modification of the limits to become more or less restrictive. The model and smoother parameters can be changed as well to improve them.

#### *6.2.6.2.4 Data transformation*

As the measurement noise is heteroscedastic, a transformation of the data has to be carried out in order to stabilize it and to detect faults based on the RSD. The three rain events used to set the acceptability limits of RSD have shown that a natural logarithmic transformation is most appropriate. The logarithmic transformation is performed separately on the smoothed data and the accepted data time series and it largely stabilizes the RSD. However, based on the Levene's test, it remains heteroscedastic. The window size used to calculate the RSD is 30 data.

#### *6.2.6.2.5 Fault detection*

This section presents the application of the on-line data validation methods on a real time series. Figure 74 shows a one day SS\* time series which contains many faults. In order to render Figure 74 readable, the prediction interval is not shown but outliers and accepted data are indicated respectively by cross and square markers. The horizontal lines presented on the time series, slope and RSD figures are the acceptability limits determined from the three "normal" rain events and those presented on the runs test figure represent the 95% of confidence interval. One can observe that many outliers are detected and some loss-of-control occurs. Between 21:00 hrs and 24:00 hrs September 30<sup>th</sup>, the on-line methods detect abnormally increasing and decreasing slopes which are probably due to poor mixing of the storm sewer flow outlet as explained before. At the same time, the runs test indicates a correlation in the residuals and the RSD is abnormally high due to the fact that the kernel smoother is averaging the peak to such an extent that a series of same-sign residuals is created. Furthermore, in the last quarter of this time series, abnormal behavior occurs which leads to the use of a high percentage of forecast data in the smoother.

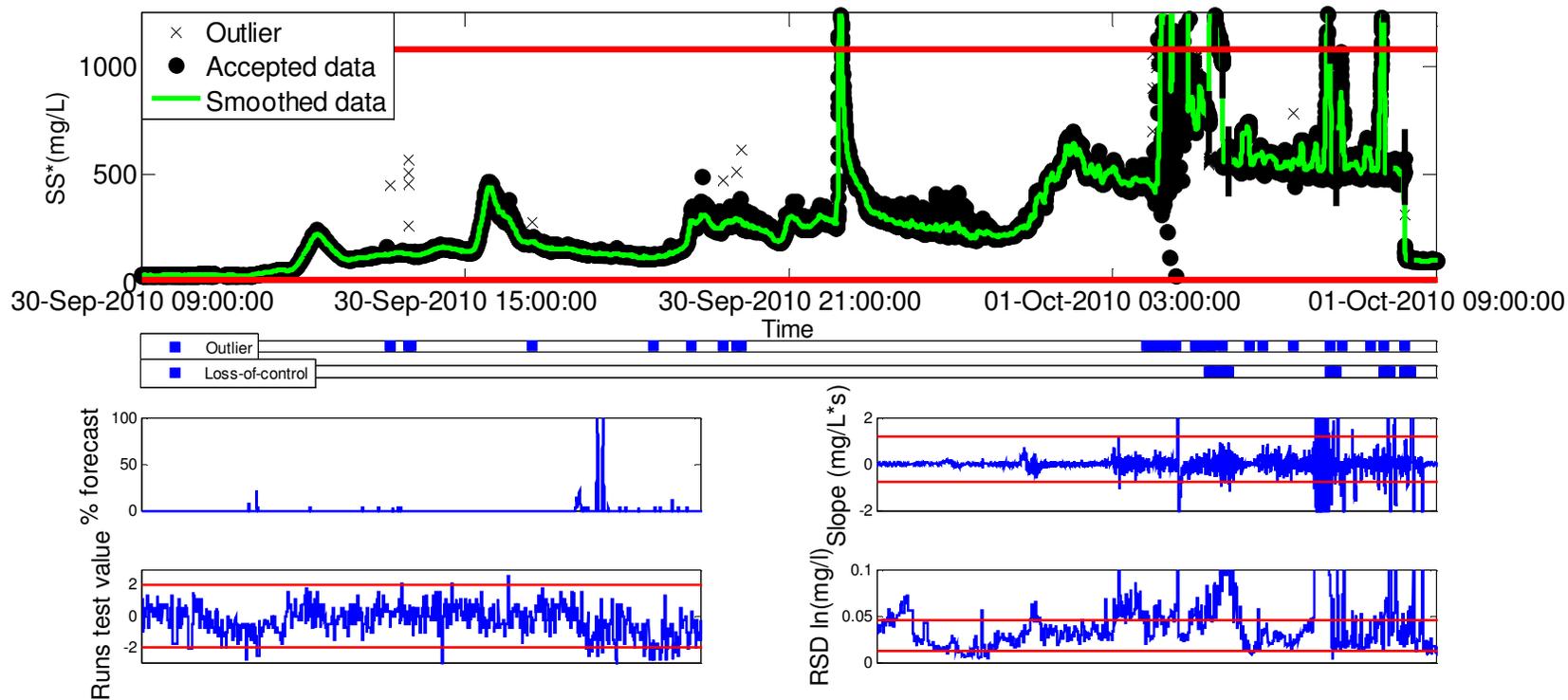


Figure 74: On-line data validation methods on a SS\* time series

## 6.2.7 Conclusion

The ADQATs presented in this study have been demonstrated to be very efficient fault detection tools. Moreover, they are easily implementable and require little process knowledge.

The off-line methods have revealed faults caused by drift/shift effects in the standard solution control chart and fouling effects in the manual cleaning control chart probably due to the deterioration of the optics and the wiper.

The on-line two-step data validation methods (outlier detection and fault detection), seem to be appropriate for this type of time series. Removing outliers first ensures that they do not have an effect on subsequent data smoothing which could lead to ill-calculated data features. Also, using smoothed data avoids noise corruption. Computing several data features (locally realistic range, rate of change and residuals standard deviation) and models and smoother information (outliers, loss-of-control, % of forecast and runs test) has the advantage of giving the user a better understanding of the time series behavior and the performance of the models and the smoother.

Besides the ADQATs, using and installing hardware suitable to the environmental conditions is important for reliable use of automated water quality measuring stations. According to the experience of the authors, consideration of the mixing characteristics, macro-fouling and micro-fouling are crucial to obtain representative data. These characteristics have to be considered in the choice of the installation, the cleaning system and the protection.

## 7. Discussion sur les ADQATs

### 7.1 Importance de la fréquence de mesure

L'application des ADQATs présentés dans ce mémoire a démontré que l'utilisation d'une haute fréquence de mesure durant un suivi en temps réel présente plusieurs avantages dans l'évaluation de la qualité des données. Par exemple, la méthode hors ligne utilisant le « manual cleaning control chart » nécessite que le nettoyage des capteurs s'effectue lorsque la qualité de l'eau est relativement stable, afin que la différence détectée entre les valeurs mesurées avant et après le nettoyage soit causée par l'encrassement du capteur et non par la variation de la qualité de l'eau. Cependant, une haute fréquence de mesure permet de réduire l'effet de cette variation en diminuant l'intervalle de temps requis pour obtenir un nombre suffisant de données nécessaires au test de Student.

Des essais avec différentes fréquences de mesure ont été réalisés avec les méthodes en ligne et sont présentés à l'annexe 8. Ces essais ont démontré que les méthodes en ligne fonctionnent relativement bien pour les séries de SS\* et de conductivité possédant un intervalle de mesure de 2 minutes. Cependant, leurs performances sont grandement améliorées si la fréquence de mesure est plus élevée. En effet, une haute fréquence de mesure permet :

- d'avoir des données lissées représentatives des données mesurées, c'est-à-dire qu'aucune corrélation significative dans les résidus n'est détectée;
- de détecter rapidement et de mieux localiser une faute avec le « runs test » et le RSD tout en utilisant un nombre élevé de données pour le calcul;
- d'augmenter la performance du modèle de prédiction de la valeur et, par conséquent, de restreindre l'intervalle de prédiction de la méthode de détection des données aberrantes;
- d'améliorer la réinitialisation de la méthode de détection des données aberrantes et, donc, de récupérer la majeure partie des données détectées comme aberrantes lors d'un « loss-of-control ».

## 7.2 Génération et interprétation des scores

D'après la méthodologie proposée par Branisavljevic et al. (2010), les prochaines étapes suivant l'application des ADQATs seraient la génération et l'interprétation des scores de validation. Tout d'abord, un score devrait être généré pour chaque méthode d'évaluation. Ensuite, les différents scores devraient être combinés pour produire un score final de validation et, ainsi, déterminer la validité des données mesurées. Les scores peuvent être de type binaire (0 ou 1), continu (0 à 100%) ou descriptif (bon, incertain et mauvais). L'auteur de ce mémoire préfère un score descriptif pour les ADQATs développés dans le cadre de ce projet. Un score continu peut être difficile à interpréter par l'utilisateur, tandis qu'un score binaire n'est pas assez flexible dans la classification des données. D'après l'expérience de l'auteur de ce mémoire, la classification incertaine peut être souvent nécessaire.

La vision de l'auteur de ce mémoire quant à la génération et l'interprétation des scores est que, selon la méthode d'évaluation, le score peut être bon ou mauvais, ou bon ou incertain. Aussi, le score est lié soit aux données brutes, aux données lissées ou aux deux. Ceci se justifie par le fait que même si certaines données brutes peuvent être de mauvaise qualité, les données lissées qui leur sont associées peuvent être de bonne qualité lorsque les mesures sont prises à une haute fréquence. La méthodologie décrite ci-après exclut autant que possible l'intervention humaine dans l'interprétation du score afin qu'elle puisse être appliquée automatiquement. Cependant, lorsque le score d'une méthode équivaut à incertain, l'intervention humaine peut parfois permettre de déterminer si le score est bon ou mauvais. La figure 75 est un schéma illustrant la méthode proposée pour la génération de scores.

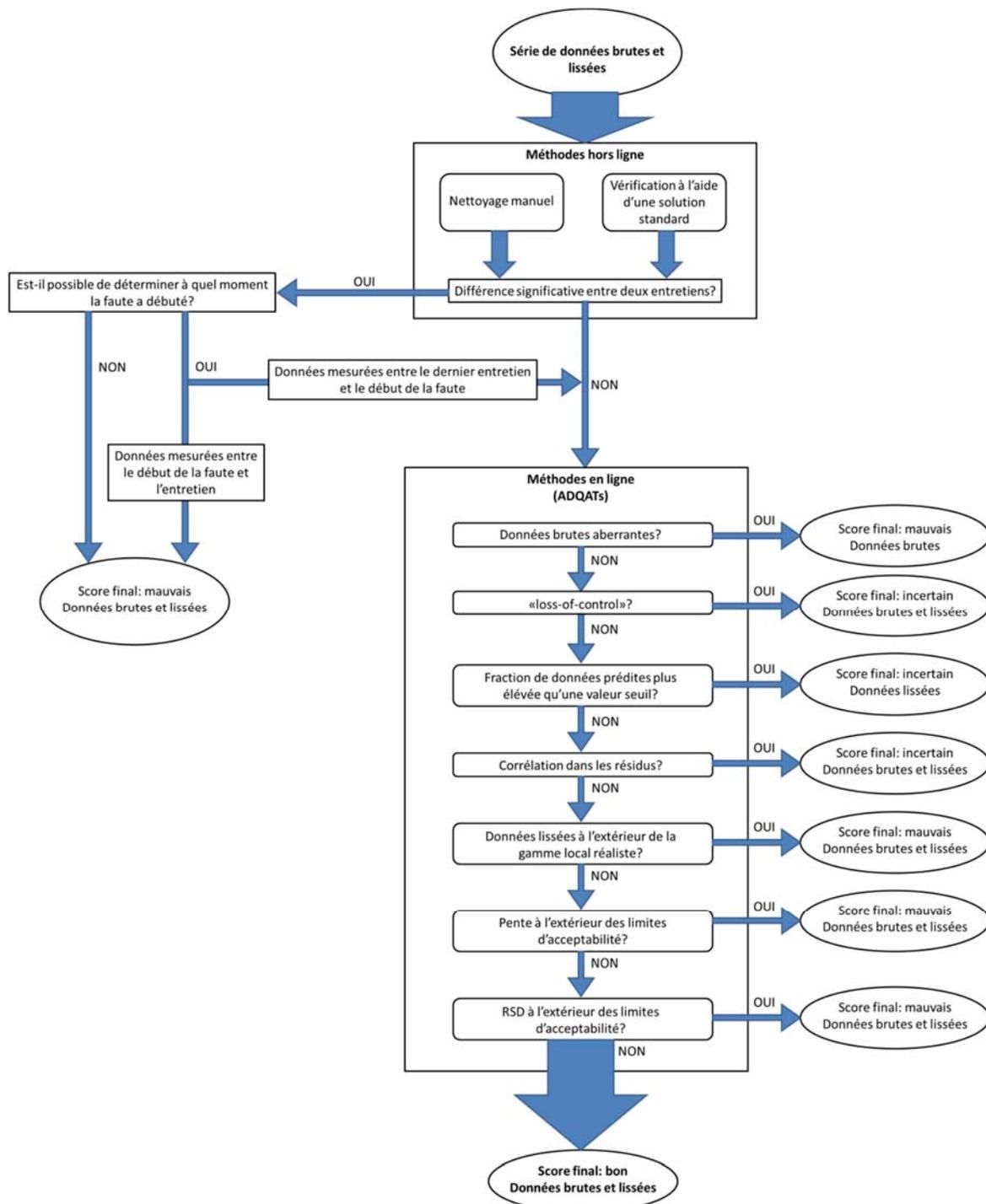


Figure 75: Méthode proposée de génération de scores

Le texte ci-dessous explique la méthode proposée de génération et d'interprétation de scores telle que présentée à la figure 75. Tout d'abord, la séquence proposée commence par la génération et l'interprétation des scores des méthodes hors lignes, et ce fait comme suit:

- Lorsqu'une différence significative est détectée entre les données mesurées avant et après un nettoyage manuel ou entre la valeur d'une solution standard et sa mesure, un score final de validation équivalant à mauvais est attribué à toutes les données brutes et lissées mesurées depuis le dernier entretien (nettoyage manuel ou vérification à l'aide d'une solution standard). En effet, s'il est impossible de déterminer à quel moment la faute a débuté, il est sécuritaire de considérer toutes les données entre les deux entretiens comme étant mauvaises. Par contre, s'il est possible de déterminer à quel moment la faute a débuté, ce score est attribué aux données mesurées entre le début de la faute et l'entretien. Si aucune différence significative n'est détectée entre les deux entretiens, le score des données brutes et lissées équivaut à bon.

Si les méthodes hors lignes attribuent un score de bon aux données brutes et lissées, ces dernières sont ensuite évaluées par les méthodes en ligne. La génération et l'interprétation des scores des méthodes en ligne se fait comme suit:

- Lorsqu'une donnée brute est détectée comme étant aberrante, le score final pour cette donnée brute équivaut à mauvais. Sinon, le score équivaut à bon.
- Lorsqu'il y a un « loss-of-control », les données brutes et les données lissées correspondantes au « loss-of-control » ont un score final équivalant à incertain. Sinon leur score équivaut à bon. S'il y a un « loss-of-control », cela signifie que le modèle de prédiction n'a pas pu prédire le comportement des données brutes. Il faut donc vérifier ce qui a causé le « loss-of-control ».
- Lorsque les données sont lissées à l'aide d'une fraction de données prédites plus élevée qu'une valeur seuil déterminée par l'utilisateur, le score final de ces données lissées équivaut à incertain. En effet, si la fraction des données prédites utilisées dans le lissage est élevée, il y a un plus grand risque que les données lissées ne soient pas

représentatives des données réelles. Si la fraction est sous la valeur seuil, le score des données lissées équivaut à bon.

- Lorsque le « runs test » détecte une corrélation dans les résidus, les données brutes et lissées correspondantes ont un score final équivalant à incertain. En effet, la corrélation des résidus peut être attribuable soit au bruit dans les des données brutes qui n'est pas aléatoire ou aux données lissées qui ne sont pas représentatives des données brutes. Une vérification est donc nécessaire pour déterminer quelle est la cause de cette corrélation et quelles données en sont affectées (brutes ou lissées). Si aucune corrélation n'est détectée dans les résidus, le score des données brutes et lissées équivaut à bon.
- Lorsque les données lissées sont à l'extérieur de la gamme locale réaliste, le score final des données brutes et lissées correspondantes équivaut à mauvais. Sinon, le score équivaut à bon.
- Lorsque la pente calculée à partir des données lissées est à l'extérieur des limites d'acceptabilité, le score final des données brutes et lissées correspondantes équivaut à mauvais. Sinon, le score équivaut à bon.
- Lorsque l'écart-type des résidus (RSD) entre les données brutes et lissées est à l'extérieur des limites d'acceptabilité, le score final des données brutes et lissées correspondantes à ce RSD équivaut à mauvais. Sinon, le score équivaut à bon.
- Si le score de toutes les méthodes équivaut à bon pour les données brutes et lissées, le score final équivaut à bon pour ces données.

Il est à noter que même si une donnée est évaluée comme étant mauvaise par les ADQATs, elle peut être de bonne qualité en réalité. Par exemple, les limites d'acceptabilité des caractéristiques des données (gamme locale réaliste, pente, etc.) ont été déterminées selon certains percentiles des caractéristiques des données historiques. Il se peut donc que les caractéristiques des données soient en dehors des limites d'acceptabilité sans que les données mesurées soient de mauvaise qualité. L'intervention humaine dans l'évaluation de la qualité des données sera toujours nécessaire, cependant elle peut être minimisée.

Suite à l'application de cette séquence de méthodes permettant la génération d'un score final de validation, les données brutes et lissées sont classifiées soit bonnes, incertaines ou mauvaises. Même si certaines données brutes sont classifiées comme étant mauvaises, les données lissées à partir de ces données brutes peuvent être de bonne qualité d'où l'importance de générer un score final de validation pour les deux types de données.

### **7.3 Diagnostic des capteurs**

Comme abordé dans la section 5.6 Résultats et analyses du suivi en temps réel, un aspect qui pourrait être aussi implanté aux ADQATs est le diagnostic des capteurs basé sur les caractéristiques des données mesurées et des fautes détectées. Brièvement, cela permettrait d'économiser du temps et de l'énergie en identifiant le problème influençant la qualité des données et l'action à poser pour remédier à ce problème. Par exemple, un bruit anormalement élevé peut être causé par l'affaiblissement de la batterie, tandis qu'un signal ne possédant aucune variation indique probablement un mauvais fonctionnement du matériel (Ni et al., 2009). D'ailleurs, la méthode employant l'écart-type des résidus présentée au chapitre 6 ADQATs serait en mesure de diagnostiquer ces types de mauvais fonctionnement. Grâce à l'exploitation de plusieurs méthodes d'évaluation de la qualité des données, les ADQATs développés dans le cadre de ce projet permettraient un diagnostic plus efficace des capteurs.

## 8. Conclusion

La collecte des données pour l'évaluation des ADQATs a constitué une partie importante du travail de ce mémoire. Tout d'abord, avant la mise en place d'une station de mesure en temps réel, il est nécessaire de connaître les gammes et les dynamiques générales de la qualité de l'eau dans laquelle les capteurs seront installés. Ceci était l'objectif de l'étude préliminaire du ruisseau Notre-Dame. Les données récoltées durant cette étude ont permis, entre autres, de déterminer si la gamme de mesure des capteurs était cohérente avec la gamme de la qualité de l'eau et de guider le choix de certaines solutions standards à utiliser pour la calibration des capteurs.

Le suivi en temps réel du ruisseau Notre-Dame consistait à mesurer les dynamiques de la qualité de l'eau lors d'évènements pluviaux et a permis de récolter des données pour évaluer les ADQATs qui ont été développés dans le cadre de ce mémoire. Au départ, il était prévu d'installer deux stations de mesure soit en amont et en aval du sous-bassin versant urbain du ruisseau Notre-Dame. Cependant, seulement une station a été installée en aval de ce ruisseau en raison de la charge de travail que nécessitait l'installation initiale et l'entretien d'une station de mesure prototype. Ce suivi a démontré l'importance dans le choix de l'emplacement et l'installation des capteurs afin de minimiser l'entretien requis et leur impact sur la qualité des données. L'égout pluvial situé à quelques mètres en amont des capteurs et qui avait un impact sur la représentativité des données mesurées, et l'obstruction des débris par la cage de protection sont deux exemples démontrant l'importance de l'emplacement et de l'installation des capteurs.

Le suivi en temps réel du ruisseau Notre-Dame a permis de bien mesurer les changements rapides de la qualité de l'eau et de mesurer les dynamiques qui n'étaient pas mesurables lors de l'étude préliminaire. Ceci a été réalisé grâce à une haute fréquence de mesure qui est possible qu'avec des capteurs. Par exemple, les capteurs mesurant la température ont confirmé qu'à cause de l'imperméabilisation de la surface de ruissèlement, la température de l'eau augmente lors d'un évènement pluvial estival. De plus, la grande quantité de données récoltées durant le suivi en temps réel comparativement à l'étude préliminaire a permis d'augmenter la fiabilité des modèles hydrologique et de la qualité de l'eau construits par le groupe de recherche autour du projet rivEAU. Cependant, un suivi en temps réel doit être

accompagné d'ADQATs afin que la qualité des données mesurées soit connue et accrue. Comme démontré précédemment, les conditions difficiles dans lesquelles les capteurs sont installés influencent la qualité des données mesurées. Si cette dernière n'est pas évaluée, les données pourraient avoir un effet néfaste sur le résultat de leur utilisation telle que la modélisation.

Les ADQATs présentés dans ce mémoire ont démontré qu'ils sont des outils efficaces pour l'évaluation de la qualité des données grâce à l'exploitation de plusieurs méthodes d'évaluation basées sur différentes caractéristiques des données et du système. Même s'ils ont été évalués sur seulement un cas réel, ils exploitent des méthodes permettant leur application à différents types d'eau tels que les eaux de surface, les eaux unitaires d'un réseau d'assainissement, les eaux usées d'une usine d'épuration, etc. De plus, ils sont basés sur des méthodes univariées nécessitant peu de connaissance du processus mesuré et s'implantant facilement. En effet, ces méthodes sont principalement implantées à partir de données historiques.

Les méthodes hors ligne permettent de détecter des fautes causées par l'encrassement ou par un mauvais fonctionnement général du capteur. Par exemple, le « standard solution control chart » et le « manual cleaning control chart » construits pour les MeS ont permis de détecter des fautes causées par des effets de dérive et de saut (« drift and shift »). Ces fautes sont probablement liées à un mauvais fonctionnement du système de nettoyage automatique du Solitax sc et à la détérioration des fenêtres optiques.

Les méthodes en ligne suivent une séquence d'exécution (élimination des données aberrantes, lissage et détection de fautes) et cette dernière est appropriée pour la détection de fautes dans ce type de séries temporelles. Le fait de remplacer les données aberrantes par les données prédites assure qu'elles n'ont pas d'effet sur le lissage des données et l'utilisation des données lissées élimine la corruption du bruit dans le calcul des caractéristiques des données. De plus, l'exploitation de plusieurs caractéristiques des données (la gamme locale réaliste, la pente et l'écart-type des résidus) et de l'information des modèles et du lisseur (données aberrantes, « loss-of-control », fraction de données prédites et le « runs test ») a l'avantage d'informer sur le comportement de la série et les performances des modèles et du lisseur.

L'utilisation d'une haute fréquence de mesure dans un suivi en temps réel comporte plusieurs avantages dans l'évaluation de la qualité des données tant au niveau des méthodes hors ligne que des méthodes en ligne. Tel que discuté à la section 7.1 Importance de la fréquence de mesure, cela diminue l'effet de la variation de la qualité de l'eau sur la détection de fautes à l'aide du « manual cleaning control chart » et augmente les performances des méthodes en ligne pour plusieurs raisons.

L'étape suivante au développement des ADQATs serait l'intégration d'un algorithme permettant la génération et l'interprétation des scores pour attribuer un score final de validation aux données. Ainsi, la qualité des données mesurées pourrait être déterminée selon leur score final. Un algorithme qui pourrait également être intégré aux ADQATs est le diagnostic des capteurs basé sur les caractéristiques des données mesurées et des fautes détectées. Ceci permettrait d'identifier les causes possibles influençant la qualité des données (ex. affaiblissement de la batterie) et donc de déterminer l'action à poser pour y remédier.



## Bibliographie

- Aldegheri M. (1979). *Mesure des débits à partir des vitesses*. ORSTOM : Institut français de recherche scientifique pour le développement en coopération, Paris, France.
- Andersen K.L. and Wagner P. (1990). Testing on-line monitors for ammonia, orthophosphate and suspended solids in activated sludge treatment plants. *Instrumentation, Control and Automation of Water and Wastewater Treatment and Transport Systems (Advances in Water Pollution Control 10)*, Ed. Briggs R., Pergamon Press, London, 667–675.
- APHA (1992). *Standard Methods for the Examination of Water and Wastewater*. 18e Edition, Washington DC, USA.
- Aspegren H., Nyberg U. and Andersson B. (1993). Integration of on-line instruments in the practical operation of the Klagshamn wastewater treatment plant. *Med. Fac. Landbouww. Univ. Gent*, 58, 2019–2028.
- AWWA, WEF and APHA (2005). *Standard Methods for the Examination of Water & Wastewater*. 21e Edition, Washington DC, USA.
- Beaupré M. (2010). Characterization of on-line sensors for water quality monitoring and process control. Mémoire, Université Laval, Québec, Qc, Canada.
- Beck M.B., Watts J.B. and Winkler S. (1998). An environmental process control laboratory: At the interface between instrumentation and model development. *Water Sci. Technol.*, 37(12), 353-362.
- Bertrand-Krajewski J.-L., Laplace D., Joannis C. and Chebbo G. (2000). *Mesures en hydrologie urbaine et assainissement*. Tec&Doc, Paris, France.
- Bertrand-Krajewski J.-L., Barraud S., Lipeme Kouyi G., Torres A. and Lepot M. (2007). Event and annual TSS and COD loads in combined sewer overflows estimated by continuous in situ turbidity measurements. *Proceedings of the 11th International Conference on Diffuse Pollution*, Belo Horizonte, Brazil, 26 au 31 août 2007.
- Bertouex P.M. (1989). Constructing control charts for wastewater treatment plant operation. *Res J Water Pollut C*, 61(9), 1534-1551.
- Bolton R.J. and Hand D.J. (2001). Unsupervised profiling methods for fraud detection. *Proceedings of Credit Scoring and Credit Control VII*, Edinburgh, UK, 5 au 7 septembre 2001.
- Branisavljevic N., Prodanovic D. and Pavlovic D. (2010). Automatic, semi-automatic and manual validation of urban drainage data. *Water Sci. Technol.*, 62(5), 1013-1021.
- Brunet J., Labarrère M., Jaume D., Rault A. and Vergé M. (1990). Détection et diagnostic de pannes. Hermès, Paris, France.
- Bulut A., Singh A.K., Shin P., Fountain T., Jasso H., Yan L. and Elgamal A. (2005). Realtime nondestructive structural health monitoring using support vector machines and wavelets. *Proceedings of the SPIE Advanced Sensor Technologies for Nondestructive Evaluation and Structural Health Monitoring*, Vol. 5770, 180–189, San Diego, CA, USA, 9 mai 2005.
- Clement L., Thas O., Ottoy J.P. and Vanrolleghem P.A. (2007) Data management of river water quality data: A semi-automatic procedure for data validation. *Water Resour. Res.*, 43, W08429, doi:10.1029/2006WR005187.
- Dochain D. and Vanrolleghem P.A. (2001). *Dynamical Modelling and Estimation in Wastewater Treatment Processes*. IWA Publishing, London, UK.

- Du Preez L.A., Husselmann A.J., Acton N.R. and Lange L. (1998). Establishing a network of on-line monitors at the purification works and in the distribution network of Rand Water. *Water Sci. Technol.*, 37(9), 65-71.
- Edwards H.O. (1998). An instrument for the measurement of colour and turbidity in natural waters. *Water Sci. Technol.*, 37(12), 263-267.
- enva. Spectro-Lyser Information Sheet, [En ligne]. <http://watertreatment.enva.com/Technologies/upload/File/Spectro%20-%20Lyser.pdf>. Consulté le 25 juillet 2011.
- EPA. (2003). Enhancing Storm Water Infiltration to Reduce Water Temperature Downstream, [En ligne]. <http://www.epa.gov/owow/NPS/natlstormwater03/08Dorava.pdf>. Consulté le 25 juillet 2011.
- Hach (2004). sc100™ Inductive Conductivity Analysis System: Operation Manual, 1<sup>st</sup> edn.
- Hach (2004). SOLITAX sc: Operation Manual, 1<sup>st</sup> edn.
- Hach (2006). LDO™ Dissolved Oxygen Sensor: USER MANUAL, 6<sup>th</sup> edn.
- Hach (2006). pHD sc Digital Differential pH/ORP Sensors: USER MANUAL, 4<sup>th</sup> edn.
- Hach (2008). Nitrate: DOC316.53.01070, 5<sup>th</sup>edn.
- Hach (2008). Nitrogen, Ammonia: DOC316.53.01081, 5<sup>th</sup>edn.
- Hach (2008). Nitrogen, Total: DOC316.53.01087, 5<sup>th</sup>edn.
- Hach (2008). Phosphorus, Reactive (Orthophosphate) and Total: DOC316.53.01125, 5<sup>th</sup>edn.
- Hach (2008). Potassium: DOC316.53.01127, 5<sup>th</sup>edn.
- Hach (2008). Sigma 950 Flow Meter: INSTRUMENT MANUAL, 5<sup>th</sup> edn.
- Hach. Manuel d'utilisateur: SIGMA 950 débitmètre - Préleveur d'échantillons - Notice d'utilisation-HACH, [Enligne]. [http://www.hachlange.fr/shop/action\\_q/documents\\_list/no\\_docs/TRUE/lkz/FR/spkEmeaEnJO64/M/QhdO8A/z/fr/TOKEN/i6A68bbe3mscE6UII](http://www.hachlange.fr/shop/action_q/documents_list/no_docs/TRUE/lkz/FR/spkEmeaEnJO64/M/QhdO8A/z/fr/TOKEN/i6A68bbe3mscE6UII). Consulté le 25 juillet 2011.
- Hill D.J. and Minsker B.S. (2010). Anomaly detection in streaming environmental sensor data: A data-driven modeling approach. *Environ Modell Softw*, 25(9), 1014-1022.
- Horsburgh J.S., Jones A.S., Stevens D.K., Tarboton D.G. and Mesner N.O. (2010). A sensor network for high frequency estimation of water quality constituent fluxes using surrogates. *Environ Modell Softw*, 25(9), 1031-1044.
- Jaccon G. (1986). *Tracé de la courbe de tarage et calcul des débits*. ORSTOM : Institut français de recherche scientifique pour le développement en coopération, Paris, France.
- Jeppsson U., Alex J., Pons M.N., Spanjers H. and Vanrolleghem P.A. (2002). Status and future trends of ICA in wastewater treatment - a European perspective. *Water Sci. Technol.*, 45(4-5), 485-494.
- John G.H. (1995). Robust decision trees: removing outliers from databases. *Proceedings of the 1st International Conference on Knowledge Discovery and Data Mining*, 174-179, Montréal, Canada, 20 au 21 août 1995.
- Johnson M. (1998). Remote turbidity measurement with a laser reflectometer. *Water Sci. Technol.*, 37(12), 255-261.

- Kirchner J.W., Feng X.H., Neal C. and Robson A.J. (2004). The fine structure of water-quality dynamics: the (high-frequency) wave of the future. *Hydrol. Process.*, 18(7), 1353-1359.
- Kozuma R., Kitamura M., Sakuma M. and Yokoyama Y. (1994). Anomaly detection by neural network models and statistical time series analysis. *Neural Networks 1994. IEEE World Congress on Computer Intelligence*, Orlando, FL, USA, 27 juin au 2 juillet 1994.
- Krajewski W.F. and Krajewski K.L. (1989). Real-time quality-control of streamflow data - a simulation study. *Water Resour Bull*, 25(2), 391-399.
- Lombard V., Toloméo S., Bertrand-Krajewski J.-L., Debray R., Comte C. and de Bénédictis J. (2010). Conception et mise en place de stations de mesure des flux polluants dédiées à la gestion intégrée d'un système d'assainissement. *Proceedings NOVATECH 2010*, Lyon, France, 27 juin au 1 juillet 2010.
- Menéndez Martínez A., Biscarri Trivino F., Sanchez Gomez A.B., Gomez Gutierrez A.A. and Castaño Rubiano M. (2001). Uncertainty and redundancy in flow metrology. *Proceedings A&E2001 Automatique et Environnement*, Saint-Etienne, France, 4 au 6 juillet 2001.
- Micrology Laboratories. Coliscan Easygel. [En ligne] [http://micrologylabs.mennonite.net/Home/Our\\_Methods/Coliscan\\_Media/Coliscan\\_Easygel](http://micrologylabs.mennonite.net/Home/Our_Methods/Coliscan_Media/Coliscan_Easygel). Consulté le 1<sup>er</sup> octobre 2009.
- Micrology Laboratories. Question list.[En ligne]. <http://ww.micrologylabs.com/Home/FAQ> Consulté le 1<sup>er</sup> octobre 2009.
- Micrology Laboratories. 2010. [En ligne]. Dilutions. [http://micrologylabs.mennonite.net/Home/Our\\_Methods/Dilutions](http://micrologylabs.mennonite.net/Home/Our_Methods/Dilutions). Consulté le 10 janvier 2011.
- Montgomery D.C. (1990). *Forecasting and Time Series Analysis*. 2<sup>nd</sup> edn, McGraw-Hill, New York, USA.
- Montgomery D.C. (2001). *Introduction to Statistical Quality Control*. 4<sup>th</sup> edn., Wiley, New York, USA.
- Mourad M. and Bertrand-Krajewski J.L. (2002). A method for automatic validation of long time series of data in urban hydrology. *Water Sci. Technol.*, 45(4-5), 263-270.
- Nelson, Lloyd S. (1984). The Shewhart control chart tests for special causes. *Qual. Tech.*, 16, 237.
- NMKL (1990). *Quality Assurance Principles for Chemical Food Laboratories*. Nordic Council of Ministers, Copenhagen, Denmark.
- Ni K., Ramanathan N., Nabil M., Chehade H., Balzano L., Nair S., Zahedi S., Kohler E., Pottie G., Hansen M. and Srivastava M. (2009). Sensor network data fault types. *ACM T Sensor Network*, 5(3), 25.
- Pfannkuche J. and Schimdt A. (2003). Determination of suspended particulate matter concentration from turbidity measurements: particle size effects and calibration procedures. *Hydrol. Proc.* 17(10), 1951-1963.
- Patry G.G. and Takács I. (1995). Modelling, simulation and control of large-scale wastewater treatment plants: An integrated approach. *Med. Fac. Landbouww. Univ.Gent*, 60, 2335-2343.
- Pressl A., Winkler S. and Gruber G. (2004). In-line river monitoring - new challenges and opportunities. *Water Sci. Technol.*, 50(11), 67-72.
- Olsson G. and Newell B. (1999). *Wastewater Treatment Systems: Modelling, Diagnosis and Control*, IWA Publishing, London, UK.

- Ragot J., Darouach M., Maquin D. and Bloch G. (1990). Validation de données et diagnostic, Hermès, Paris, France.
- Ramaswamy S., Rastogi R., Shim K. (2000). Efficient algorithms of mining outliers from large data sets. *Proceedings of the ACM SIGMOD Conference on Management of Data*, 427–438, Dallas TX, USA, 2 juin 2000.
- Rieger L., Siegrist H., Winkler S., Saracevic E., Votava R. and Nadler J. (2002). In-situ measurement of ammonium and nitrate in the activated sludge process. *Water Sci. Technol.*, 45(4–5), 93–100.
- Rieger L. and Vanrolleghem P.A. (2008). monEAU: A platform for water quality monitoring networks. *Water Sci. Technol.*, 57(7), 1079-1086.
- Rousseeuw P., Leroy A. (1996). *Robust Regression and Outlier Detection*, 3<sup>rd</sup> edn, John Wiley & Sons, New York, USA.
- s::can (2007). Manual s::can spectrometer probe, version 1.0, Vienna.
- s::can (2007). Manual ammo::lyser, version 1.0, Vienna.
- Schimek M. G. (2000). *Smoothing and Regression: Approaches, Computation, and Application*, John Wiley & Sons, New York, USA.
- Sikow M. and Pursiainen J. (1995). Use and maintenance of on-line measurements at Suomenoja wastewater plant. *Preprints of IAWQ Specialized Conference on Sensors in Wastewater Technology*, Copenhagen, Denmark, 25 au 27 octobre 1995.
- Strobl R.O., Robillard P.D., Shannon R.D., Day R.L. and McDonnell A.J. (2006). A water quality monitoring network design methodology for the selection of critical sampling points: Part I. *Environ Monit Assess*, 112(1-3), 137-158.
- Tang J., Chen Z., Fu A., Cheung D. (2002). A robust outlier detection scheme in large data sets. *Proceedings of the 6th Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Taipei, Taiwan, 6 au 8 mai 2002.
- Thomann M., Rieger L., Frommhold S., Siegrist H. and Gujer W. (2002). An efficient monitoring concept with control charts for on-line sensors. *Water Sci. Technol.*, 46(4-5), 107-116.
- Tomlinson M.S. and De Carlo E.H. (2003). The need for high resolution time series data to characterize Hawaiian streams. *J. Am. Water Resour. Assoc.*, 39(1), 113-123.
- Vandenbergh V., van Griensven A. and Bauwens W. (2002). Detection of the most optimal measuring points for water quality variables: application to the river water quality model of the River Dender in ESWAT. *Water Sci. Technol.*, 46(3), 1-7.
- Vandenbergh V., Goethals P.L.M., Van Griensven A., Meirlaen J., De Pauw N., Vanrolleghem P. and Bauwens W. (2005). Application of automated measurement stations for continuous water quality monitoring of the Dender River in Flanders, Belgium. *Environ Monit Assess*, 108(1-3), 85-98.
- Vanrolleghem P.A. and Lee D.S. (2003). On-line monitoring equipment for wastewater treatment processes: state of the art. *Water Sci. Technol.*, 47(2), 1-34.
- Wacheux H., Da Silva S. and Lesavre J. (1993). Inventory and assessment of automatic nitrate analyzers for urban sewage works. *Water Sci. Technol.*, 28(11–12), 489–498.
- Wacheux H., Million J.-L., Guillo C. and Alves E. (1996). NH<sub>4</sub><sup>+</sup> automatic analysers for wastewater treatment plant: evaluation test at laboratory and field level. *Water Sci. Technol.*, 33(1), 193–201.
- Wang J. and Thongngamdee S. (2003). On-line electrochemical monitoring of (TNT) 2,4,6-trinitrotoluene in natural waters. *Anal. Chim. Acta*, 485(2), 139-144.

- Winkler S., Rieger L., Saracevic E., Pressl A., Gruber G. (2004). Application of ion sensitive sensors in water quality monitoring. *Water Sci. Technol.*, 50(11), 105-114.
- Winkler S., Zessner M., Saracevic E. and Fleischmann N. (2008). Intelligent monitoring networks - transformation of data into information for water management. *Water Sci. Technol.*, 58(2), 317-322.
- YSI (2002). YSI Environmental Operations Manual.



# ANNEXE 1. Fiche de jaugeage des vitesses d'écoulement avec un micro-moulinet

<b>Date:</b>
<b>Endroit:</b>
<b>Mesuré par:</b>

	Début	Fin
<b>Heure</b>		
<b>Hauteur d'eau</b>		

Point de mesure	# moulinet	Hauteur par rapport au lit	nb tours	temps (sec)

<b>Commentaire:</b>
---------------------



## ANNEXE 2. Fiche d'échantillonnage de l'étude préliminaire

### Sampling

Date	Time	Location	Measured by
------	------	----------	-------------

### General comments

Weather conditions (T, rain, humidity, pressure,..)	Flow conditions
Water conditions (visible)	Comments

### Onsite measurements

Measurement	Comment
Temperature [°C]	
Dissolved Oxygen [mg/l]	
pH [-log(H <sup>+</sup> )]	
Conductivity [mS/cm]	
Turbidity [NTU]	
Last calibration of probe	
Water height [cm]	
Measurement started at	Finished at

### Storage

Sample ID	Stored at 4 °C when	Stored at 4 °C where
Comments		



## **ANNEXE 3. Matériels et méthodologie d'une journée typique de l'étude préliminaire**

### **Au ruisseau Notre-Dame**

#### **Matériels :**

- Règle graduée
- Sonde YSI 6600EDS
- Bouteilles d'échantillonnage
- Marqueur permanent
- Crayon
- Montre
- Fiches d'échantillonnage
- Lampe de poche, en cas de noirceur
- Bottes d'eau

#### **Procédure:**

- Identifier les bouteilles.
- Remplir les fiches d'échantillonnage.
- Au point d'échantillonnage amont, prendre la hauteur d'eau au point de mesure désigné à cet effet à l'aide de la règle graduée. Noter l'heure, la date et la hauteur d'eau sur la fiche d'échantillonnage.
- Rincer la bouteille d'échantillonnage avec l'eau du ruisseau. Prendre un échantillon et noter l'heure et la date sur la fiche d'échantillonnage. L'échantillon doit être pris à la mi-hauteur, si possible.
- Mettre la sonde YSI 6600EDS dans l'eau, sans l'accoter sur le lit du ruisseau et attendre 3 minutes (3 lectures, 1 lecture/minute).
- Noter l'heure du début et de la fin des lectures avec la sonde YSI 6600EDS sur la fiche d'échantillonnage.
- Compléter la fiche d'échantillonnage.
- Répéter cette procédure pour le point d'échantillonnage aval.

Note : Ne jamais marcher en amont du point de mesure, afin d'éviter la suspension des sédiments.

Ne pas laisser la sonde YSI 6600EDS dans la voiture en cas de gel.

## Au laboratoire

- Mettre les échantillons dans la chambre froide et mettre les fiches d'échantillonnage à l'endroit prévu. Remplir le tableau d'échantillonnage dans le cahier de laboratoire.
- Rincer la sonde YSI 6600EDS avec de l'eau distillée et mettre le capteur pH dans une solution tampon à pH 4.

## Procédure au laboratoire

- Compter le nombre de UFC sur les boîtes de Petri de la journée précédente qui sont à l'incubateur depuis 24 h.
- Prendre les échantillons de 18 heures de la journée précédente et de 10 heures de la journée même qui sont dans la chambre froide et faire les analyses nécessaires.
- Filtrer les échantillons pour les mesures des MeS et mettre les filtres à l'étuve (105°C) 2 heures.
- Préparer les boîtes de Petri servant aux mesures des pathogènes. Faire 3 boîtes de Petri par échantillon soit une boîte de Petri avec un volume d'échantillon de 5 mL non dilué, une boîte de Petri avec un volume d'échantillon de 1 mL non dilué et une boîte de Petri avec un volume d'échantillon de 1mL dilué 10 fois. Ainsi, il est possible de mesurer de façon optimale les E.coli et coliformes totaux qui sont présents à des concentrations différentes.
- Mesurer le pH à l'aide du pH/ion meter.
- Pour les métaux lourds, mettre un volume d'échantillon dans un contenant hermétique et l'acidifier à un pH 2 avec de l'acide nitrique. Ensuite, le mettre à la chambre froide.
- Effectuer les analyses du phosphore total, de l'azote total, de l'orthophosphate, du nitrate et de l'ammonium avec les tests Hach TNT.
- Commencer les tests Hach TNT avec les analyses du phosphore total et de l'azote total (1 h de digestion). Ensuite, pendant que le phosphore total et l'azote total sont en digestion, effectuer les analyses de l'orthophosphate, du nitrate et de l'ammonium.
- Analyser l'orthophosphate, le nitrate et l'ammonium sur un échantillon témoin (« blank sample »), lorsque l'échantillon est très turbide.
- Terminer les analyses du phosphore total et de l'azote total.
- S'il n'y a pas de test Hach TNT de disponible pour l'analyse d'un paramètre, acidifier à pH 2 un certain volume de l'échantillon (Ammonium → Acide Hydrochlorique - Nitrate, phosphore total et azote total → Acide Sulfurique).
- Noter le volume d'acide ajouté, afin de connaître le volume de base à ajouter pour neutraliser l'échantillon lors de l'analyse avec les tests Hach TNT et le facteur de correction.
- Mettre les filtres des MeS qui sont à l'étuve depuis 2 heures au dessiccateur pendant 30 minutes.
- Peser les filtres des MeS.

- Transférer les données mesurées par la sonde YSI 6600EDS sur l'ordinateur au laboratoire. L'espace mémoire de la sonde est limité (9 jours environ).
- Vérifier la capacité restante des batteries de la sonde YSI 6600EDS.
- Vérifier la calibration de la sonde YSI (à faire de façon hebdomadaire, selon YSI incorporated).

Note : L'orthophosphate ne se préserve pas au-delà de 48h, mais dans ce cas-ci des échantillons ont quand même été congelés pour la préservation.  
Lors des analyses en laboratoire, toujours s'assurer que l'échantillon soit bien mélangé afin d'avoir une homogénéisation de ce dernier.



## ANNEXE 4. Résultats de l'étude préliminaire (Suite)

Cette annexe présente certains résultats qui n'ont pas été présentés dans la section 4.4 Résultats et analyses de l'étude préliminaire, ainsi qu'une brève discussion sur les dynamiques observées.

### Hach TNT

#### Azote total

La figure 76 présente les concentrations d'azote total mesurées aux points d'échantillonnage amont et aval durant le mois d'octobre. Il y a quelques valeurs absentes, car certains échantillons n'ont pas été analysés. Selon la figure 76, les concentrations en azote total se comportent de façon similaire aux concentrations d'ammonium, c'est-à-dire qu'elles diminuent lorsque le débit augmente et elles augmentent graduellement vers la fin du mois d'octobre. Il y a quelques valeurs qui semblent aberrantes.

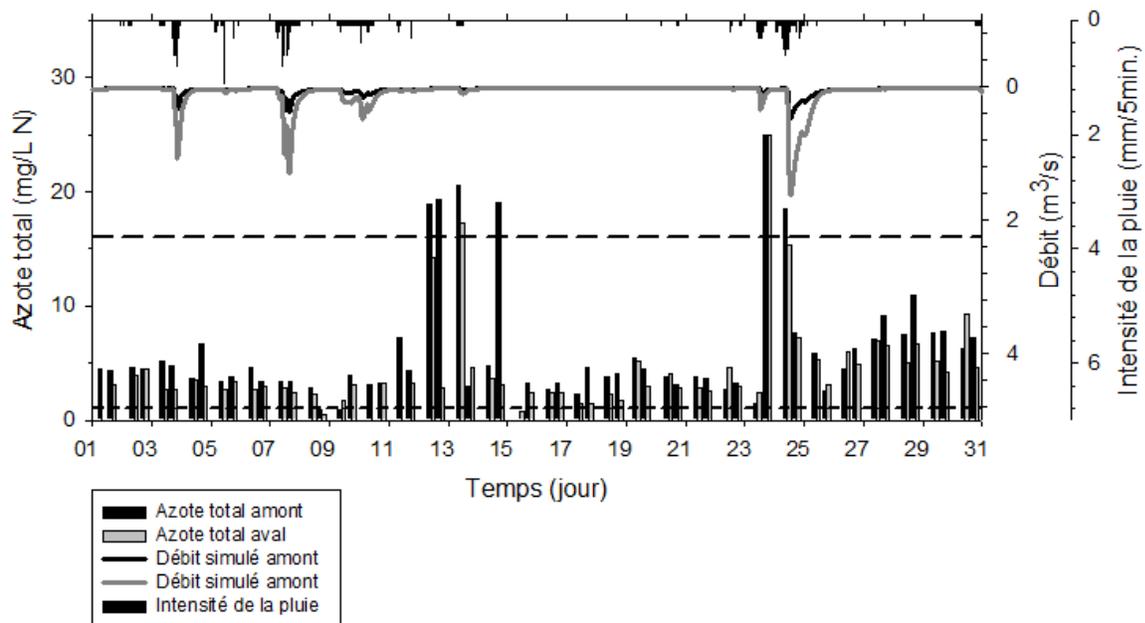


Figure 76: Concentrations d'azote total en amont et en aval au mois d'octobre 2009 (ligne à court pointillée : limite inférieure, ligne à long pointillé : limite supérieure)

La figure 77 et la figure 78 présentent respectivement les concentrations d'azote total mesurées en amont et en aval pour les journées du 1<sup>er</sup> et du 19 octobre. Ces figures montrent que les concentrations en azote total en amont et en aval et l'écart entre ces concentrations varient durant ces deux journées même si le débit est relativement constant. Comme pour les concentrations d'ammonium et de nitrate, les concentrations d'azote total sont généralement plus faibles en aval qu'en amont et ceci pourrait s'expliquer par la dilution causée par l'augmentation du débit vers l'aval.

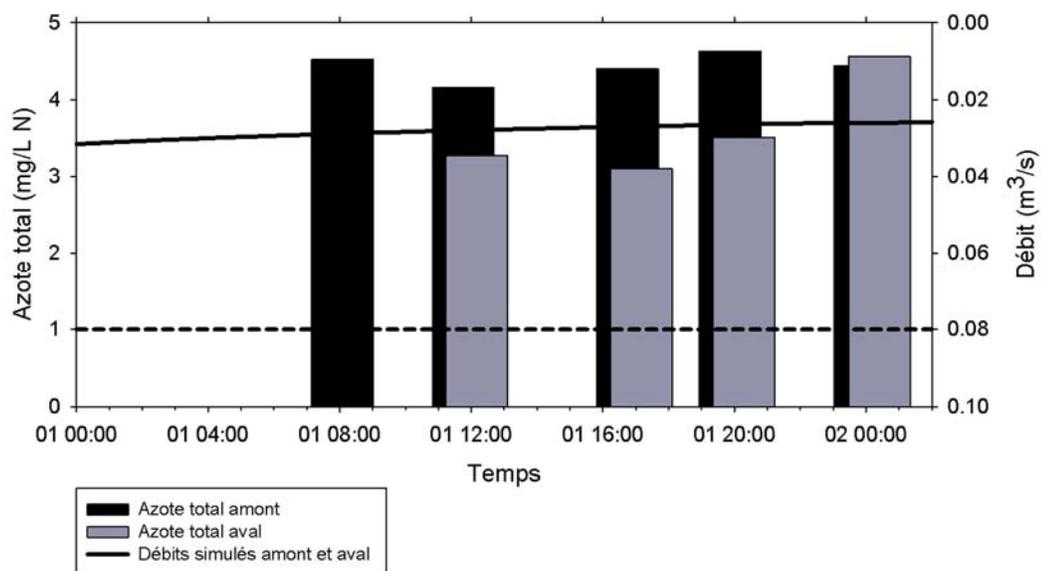


Figure 77: Concentrations d'azote total en amont et en aval le 1<sup>er</sup> octobre 2009 (ligne à court pointillée : limite inférieur)

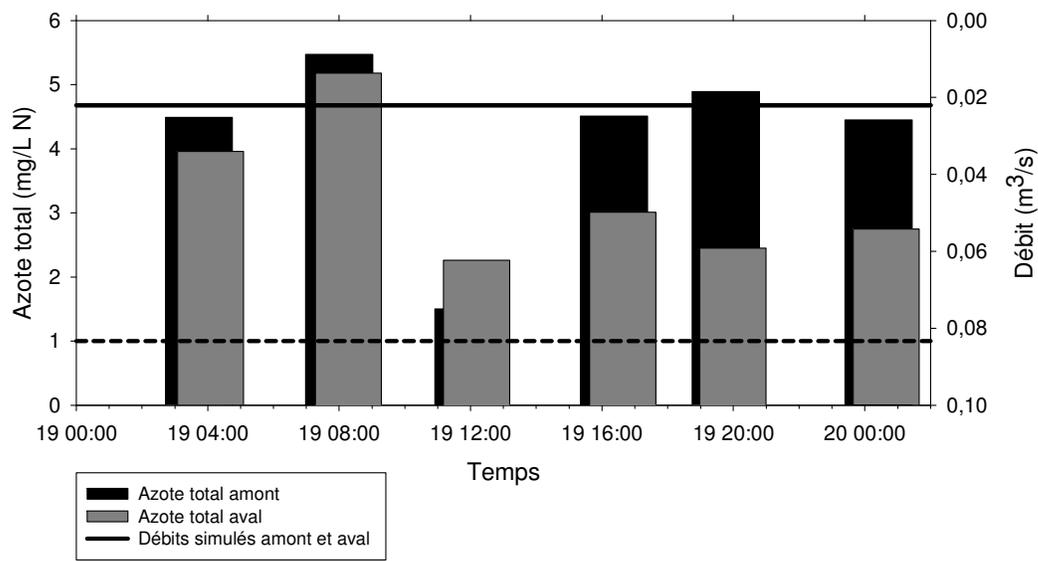


Figure 78: Concentrations d'azote total en amont et en aval le 19 octobre 2009  
(ligne à court pointillée : limite inférieur)

### Orthophosphate et phosphore total

Les figures concernant les concentrations d'orthophosphate et de phosphore total ne sont pas présentées, car il y avait trop de mesures qui se situaient en dehors de la gamme de mesure des tests Hach TNT.

## **Pathogènes**

Les pathogènes ont été mesurés durant tout le mois d'octobre. Cependant, après plusieurs échantillons, il a été observé qu'une boîte de Petri n'était pas suffisante pour obtenir un nombre optimal d'E.coli et de coliformes totaux, soit entre 30 et 300 (Micrology Laboratories, 2010). Trois boîtes de Petri par échantillon ont donc été utilisées à partir des échantillons AM-2-14 et AV-2-14. Les volumes pipetés et les dilutions des échantillons pour les trois boîtes de Petri étaient un volume de 5 ml non dilué, un volume de 1 ml non dilué et un volume de 1 ml dilué 10 fois.

Lorsqu'il n'a aucun pathogène de compté sur une boîte de Petri, la valeur de la concentration de pathogènes (UFC/100 mL) attribuée à l'échantillon est plus petite que le facteur de multiplication pour obtenir une concentration sur 100 mL d'échantillon. Par exemple, pour un volume d'échantillon de 1 mL dilué 10 fois, il faut multiplier le nombre de pathogènes compté par 1000 pour obtenir la concentration de pathogènes en UFC/100 mL d'échantillon. Dans ce cas, s'il n'y a aucun pathogène de compté, la valeur attribuée est donc plus petite que 1000 UFC/100 mL (<1000 UFC/100 mL). Ce type de valeur est indiqué par un espace vide sur les figures ci-après. De plus, lorsque le nombre de pathogènes compté sur une boîte de Petri est plus petit que 20, ceci est considéré comme discutable statistiquement quant à la précision de cette valeur (Micrology Laboratories, 2009).

### **E.coli**

La figure 79 présente les concentrations d'E.coli mesurées durant le mois d'octobre. Selon cette figure, les concentrations d'E.coli augmentent lorsque le débit augmente et la concentration d'E.coli en aval est généralement plus élevée que celle en amont. Par ailleurs, l'impact de l'utilisation de trois boîtes de Petri avec une dilution et un volume d'échantillon différents s'observe à la figure 79. En effet, avant le 14 octobre, beaucoup de boîtes ne comptaient aucun E.coli à cause d'une trop grande dilution ou d'un trop petit volume d'échantillon pipeté.

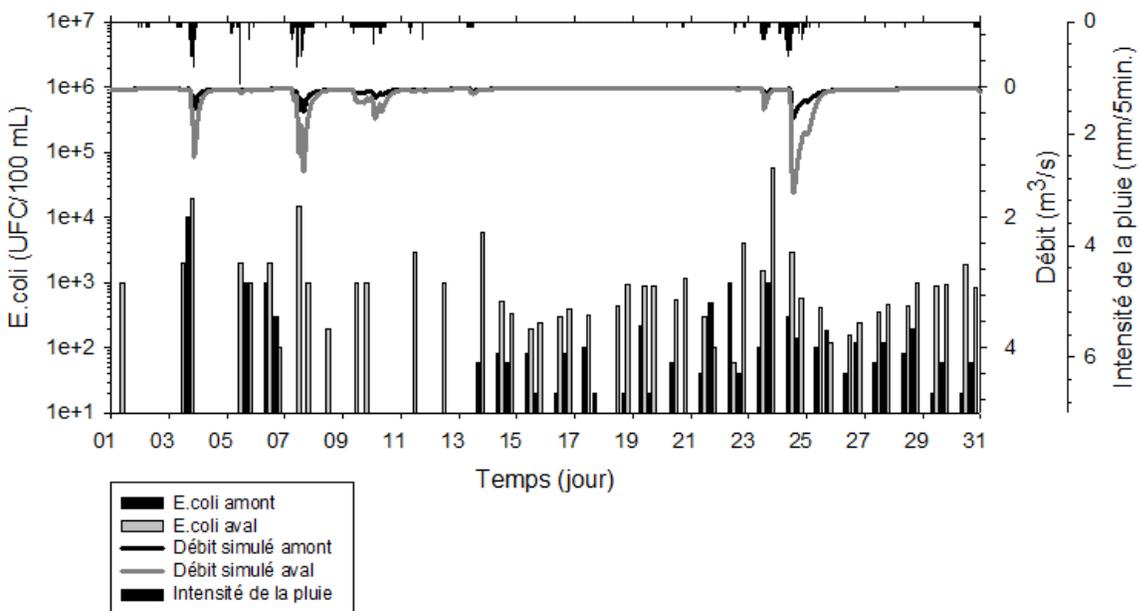


Figure 79: Concentrations d'E.coli en amont et en aval au mois d'octobre 2009

La figure 80 présente les concentrations d'E.coli mesurées en amont et en aval pour la journée du 19 octobre. Sur cette figure, on remarque que la concentration d'E.coli est généralement plus élevée en aval qu'en amont et elle varie beaucoup en amont comparativement en aval où elle demeure relativement constante.

Les concentrations d'E.coli pour le 1er octobre ne sont pas présentées, car les dilutions et les volumes d'échantillon utilisés n'étaient pas optimaux et, par conséquent, aucun E.coli n'a été compté sur la majorité des boîtes de Petri.

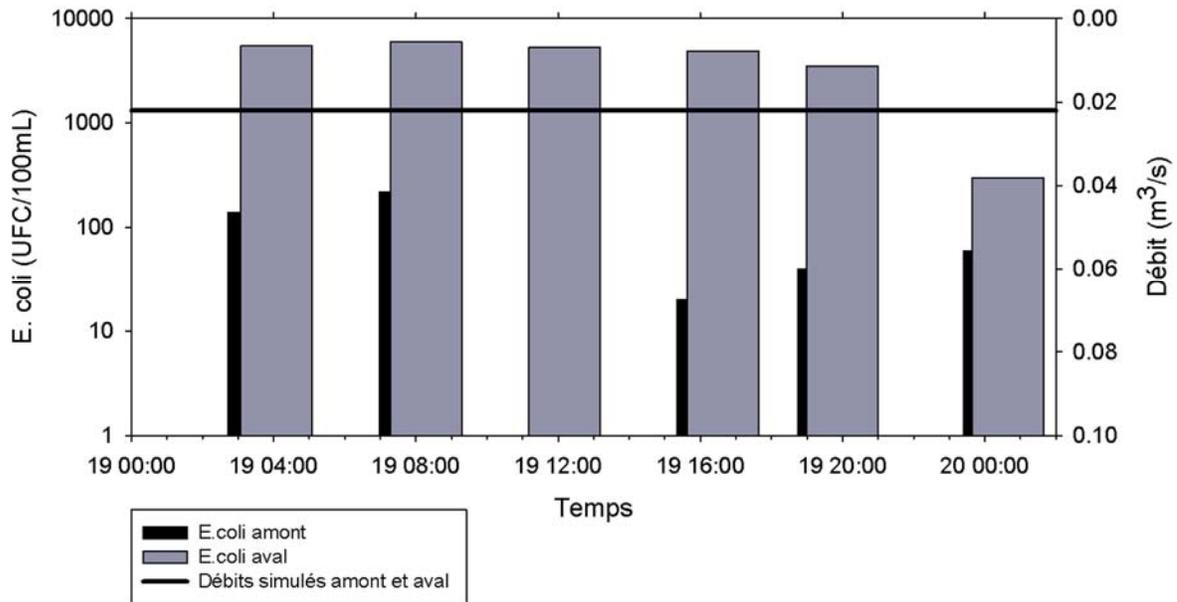


Figure 80: Concentrations d'E.coli en amont et en aval le 19 octobre 2009

### Coliformes totaux

La figure 81 présente les concentrations de coliformes totaux mesurées durant le mois d'octobre. Cette figure montre que les concentrations de coliformes totaux ont la même tendance que les concentrations d'E.coli, c'est-à-dire qu'elles augmentent lorsque le débit augmente. Contrairement aux concentrations d'E.coli, il n'y a pas de différence majeure entre les concentrations de coliformes totaux en amont et en aval et, en temps sec, les concentrations de coliformes totaux sont relativement constantes.

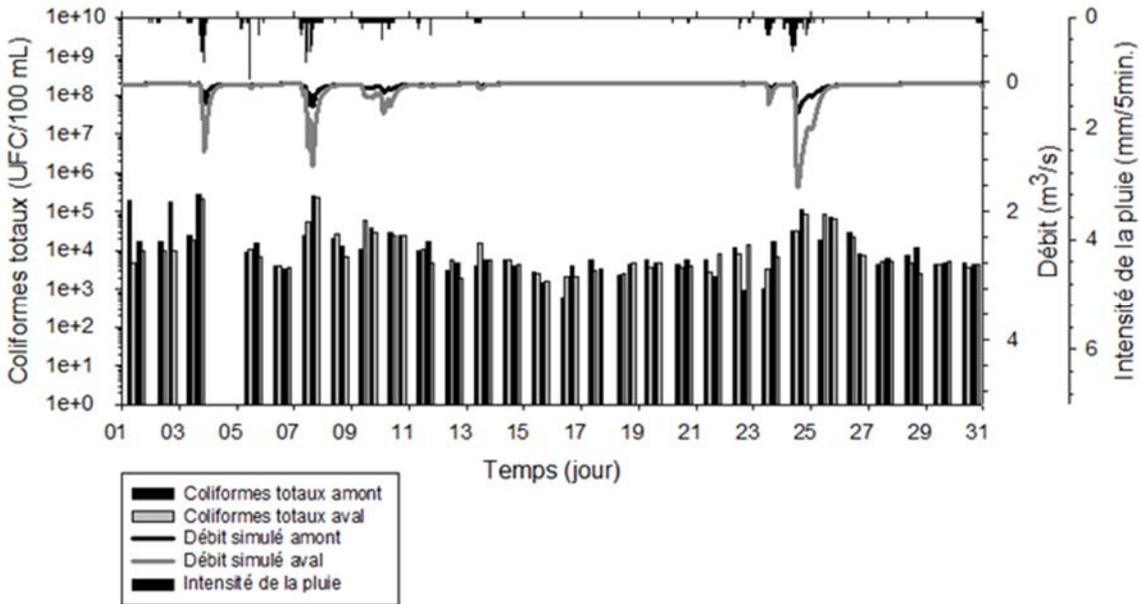


Figure 81: Concentrations de coliformes totaux en amont et en aval au mois d'octobre 2009

La figure 82 et la figure 83 présentent respectivement les concentrations de coliformes totaux en amont et en aval pour les journées du 1er et 19 octobre. La figure 82 montre que les concentrations de coliformes totaux varient légèrement au cours de la journée du 1er octobre, tandis que la figure 83 démontre qu'elles sont relativement constantes pour la journée du 19 même si le débit est constant pour ces deux journées.

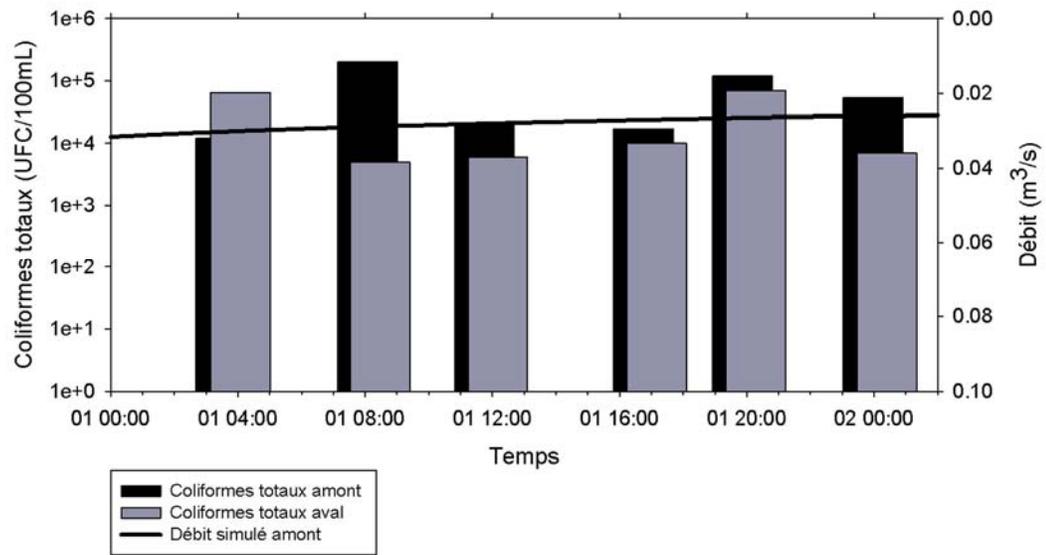


Figure 82: Coliformes totaux en amont et en aval le 1 octobre 2009

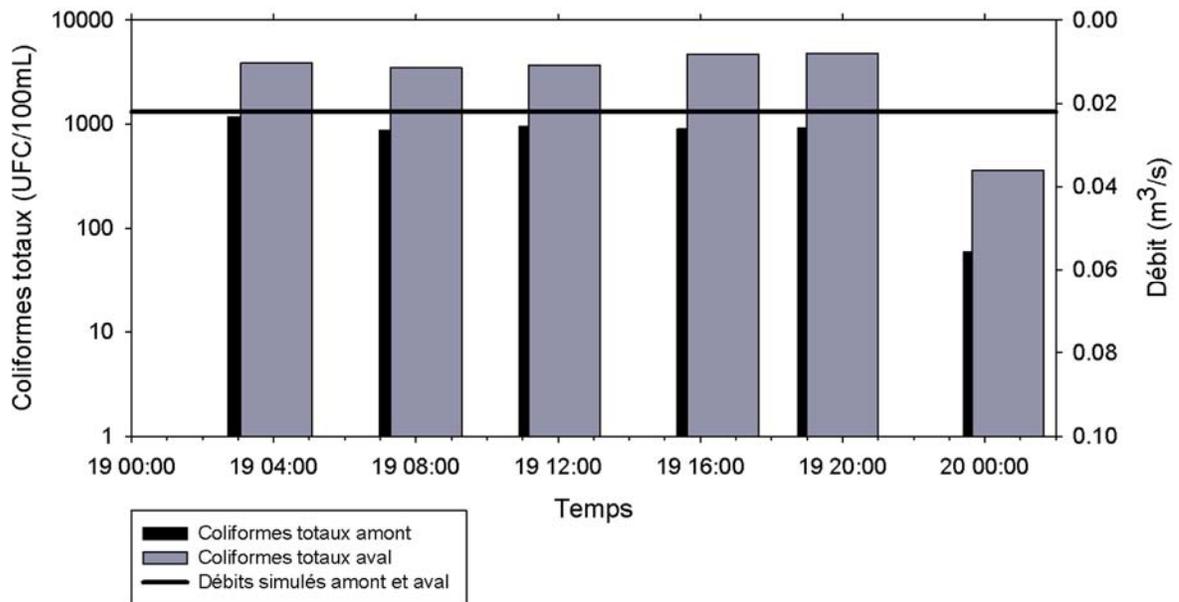


Figure 83: Concentrations de coliformes totaux en amont et en aval le 19 octobre 2009

## Métaux lourds

La figure 84 et la figure 85 présentent respectivement les concentrations de manganèse total et de zinc total mesurées durant le mois d'octobre. Lorsque les valeurs ne sont pas indiquées sur ces figures, cela signifie que les concentrations sont soit sous la limite de détection ou qu'elles n'ont pas été mesurées. La figure 84 montre que lors d'évènements pluviaux les concentrations de manganèse diminuent tant en amont qu'en aval, à l'exception du 23 et 24 octobre où les concentrations en aval ont augmenté. En temps sec, la concentration de manganèse semble augmenter graduellement en amont, tandis qu'elle demeure relativement constante en aval. En observant la figure 85, on remarque que, contrairement aux concentrations de manganèse total, les concentrations de zinc total sont constantes durant le mois d'octobre et elles augmentent lorsque le débit augmente.

Les concentrations de cuivre total ont été mesurées, mais elles ne sont pas présentées, car les valeurs mesurées étaient souvent sous la limite de détection. Aussi, les figures présentant les profils journaliers des métaux lourds du 1 et 19 octobre ne sont pas présentées, car les concentrations étaient relativement stables au cours des deux journées.

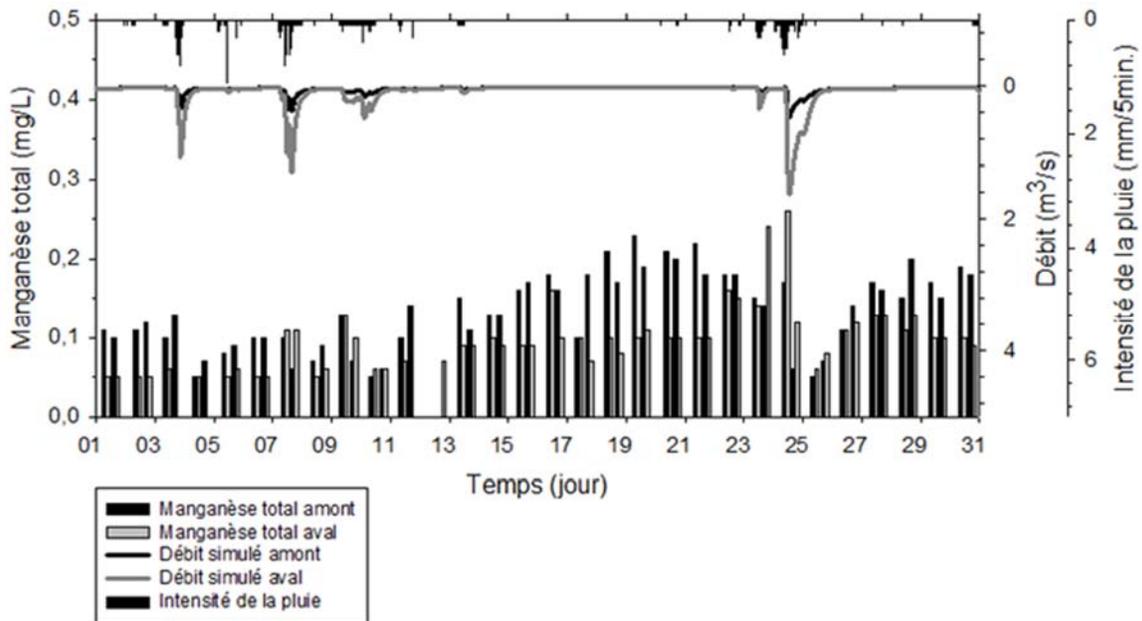


Figure 84: Concentrations de manganèse total en amont et en aval au mois d'octobre 2009

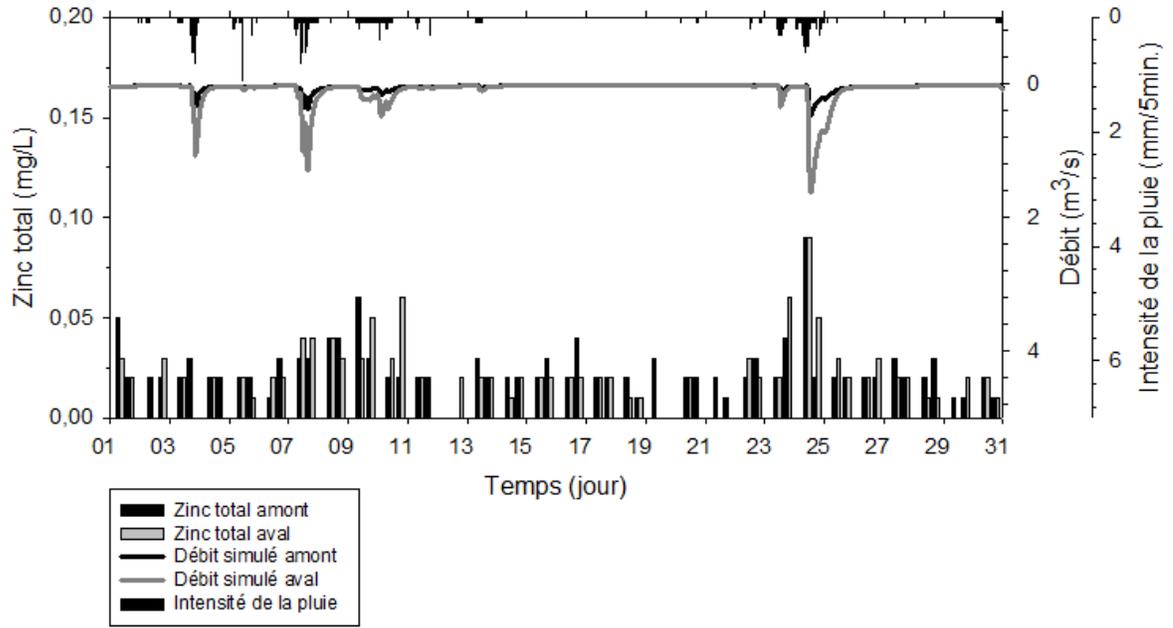


Figure 85: Concentrations de zinc total en amont et en aval au mois d'octobre 2009

## **ANNEXE 5. Jaugeage du débit durant l'étude préliminaire**

Le tableau 7 et le tableau 8 présentent la bathymétrie des sections transversales aux points d'échantillonnage amont et aval relevée selon la méthode expliquée dans la section 3.4 Jaugeage du débit.

Comme expliqué dans la section 3.4 Jaugeage du débit, l'étalonnage des micro-moulinets (172-112-5, 111987-6 et 111134) s'est effectué en déterminant la droite décrivant la relation entre la vitesse d'écoulement mesurée avec un ADV et la vitesse de rotation mesurée par le micro-moulinet. Cependant, la variation des vitesses mesurées par l'ADV pour une même vitesse d'écoulement était très grande. La vitesse d'écoulement a donc été obtenue en calculant une moyenne sur plusieurs vitesses d'écoulement mesurées. Le tableau 9 présente les vitesses de rotation et d'écoulement mesurées lors de l'étalonnage des trois micro-moulinets. Ce tableau montre qu'il y a une variation entre les vitesses moyennes d'écoulement mesurées pour un même débit. Certaines vitesses qui semblaient incohérentes ont donc été supprimées. Cette variation peut être attribuable au fait que lors d'un changement de débit, ce dernier prenait un certain temps avant de se stabiliser.

Les trois micro-moulinets cessaient de tourner à une vitesse moyenne de 2,70 cm/s. Ces points n'ont pas été inclus pour la détermination des droites d'étalonnage des micro-moulinets. Le tableau 10 présente les points qui ont été conservés pour la détermination des droites d'étalonnage des trois micro-moulinets.

La figure 86, la figure 87 et la figure 88 présentent respectivement les droites d'étalonnage des micro-moulinets 172-112-5, 111987-6 et 111134. Les droites présentées sur ces figures sont celles qui possèdent le meilleur coefficient de détermination ( $R^2$ ) et leur équation est présentée sur chaque figure. Les équations des droites d'étalonnage des micro-moulinets ont été appliquées aux mesures des vitesses de rotation dans le ruisseau pour obtenir les vitesses d'écoulement.

Tableau 7: Bathymétrie de la section transversale au point d'échantillonnage amont

Point	Distance (cm)	Profondeur (cm)	Commentaires
	0	0	rive gauche dans le sens de l'écoulement, point d'origine
	50	-25	
	70	-35	
	90	-54	
7	111	-67	point de mesure de vitesse
6	146	-67	point de mesure de vitesse
5	200	-68	point de mesure de vitesse et de la hauteur
4	250	-63	point de mesure de vitesse
3	300	-65	point de mesure de vitesse
2	350	-61	point de mesure de vitesse
	370	-52	
1	400	-36	point de mesure de vitesse

Tableau 8: Bathymétrie de la section transversale au point d'échantillonnage aval

Point	Distance (cm)	Profondeur (cm)	Commentaires
	0	0	rive gauche dans le sens de l'écoulement, point d'origine
	30	-9	
	45	-20	
	50	-35	
7	70	-43	point de mesure de vitesse
6	100	-48	point de mesure de vitesse
5	150	-54	point de mesure de vitesse
4	200	-59	point de mesure de vitesse
3	250	-60	point de mesure de vitesse
2	300	-61	point de mesure de vitesse et de la hauteur
1	350	-53	point de mesure de vitesse
	385	-37	
	395	-19	
	400	-13	

Tableau 9: Étalonnage des micro-moulinets

# Débit	# moulinet	nb tours	Temps (s)	Vitesse de rotation (nb tours/s)	Vitesse moyenne ADV (cm/s)
1	172-112-5	184	60	3,07	11,66
	111987-6	108	60	1,80	13,73
	111134	52	60	0,87	13,56
2	172-112-5	197	60	3,28	14,99
	111987-6	210	100	2,10	14,14
	111134	99	100	0,99	13,03
3	172-112-5	428	100	4,28	19,39
	111987-6	201	100	2,01	21,21
	111134	116	100	1,16	18,07
4	172-112-5	250	100	2,50	5,71
	111987-6	150	100	1,50	9,63
	111134	72	100	0,72	11,66
5	172-112-5	380	100	3,80	21,41
	111987-6	207	100	2,07	22,49
	111134	100	100	1,00	22,32
6	172-112-5	83	100	0,83	6,51
	111987-6	59	100	0,59	1,98
7	172-112-5	0	100	0,00	2,70
	111987-6	0	100	0,00	2,70
	111134	0	100	0,00	2,70
8	172-112-5	170	100	1,70	8,25
	111987-6	100	100	1,00	8,24
	111134	52	100	0,52	9,16
9	172-112-5	228	100	2,28	12,69
	111987-6	135	100	1,35	11,15
	111134	65	100	0,65	13,81
10	172-112-5	250	100	2,50	14,84
	111987-6	142	100	1,42	13,13
	111134	71	100	0,71	14,28

Tableau 10: Points conservés pour la détermination des droites d'étalonnage

172-112-5		111987-6		111134	
Vitesse de rotation (nb tours/s)	Vitesse moyenne ADV (cm/s)	Vitesse de rotation (nb tours/s)	Vitesse moyenne ADV (cm/s)	Vitesse de rotation (nb tours/s)	Vitesse moyenne ADV (cm/s)
3,07	11,66	1,80	13,73	0,87	13,56
3,28	14,99	2,10	14,14	0,99	13,03
4,28	19,39	2,01	21,21	1,16	18,07
3,80	21,41	1,50	9,63	0,72	11,66
0,83	6,51	2,07	22,49	1,00	22,32
1,70	8,25	1,00	8,24	0,52	9,16
2,28	12,69	1,35	11,15	0,65	13,81
2,50	14,84	1,42	13,13	0,71	14,28

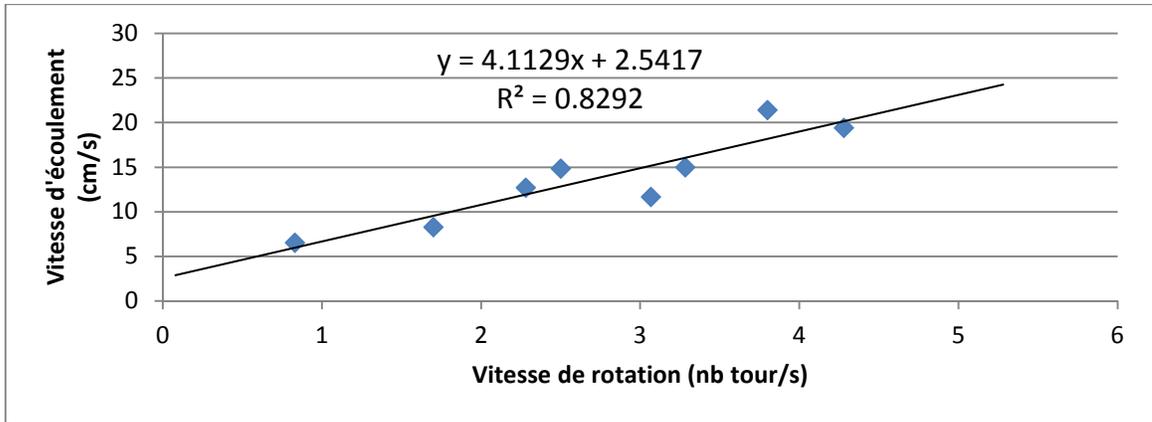


Figure 86: Droite d'étalonnage du micro-moulinet 172-112-5

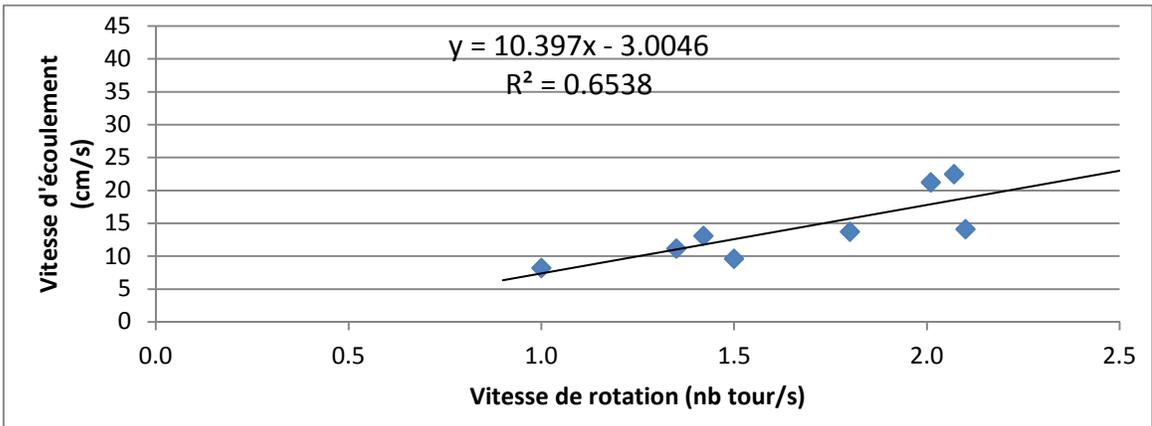


Figure 87: Droite d'étalonnage du micro-moulinet 111987-6

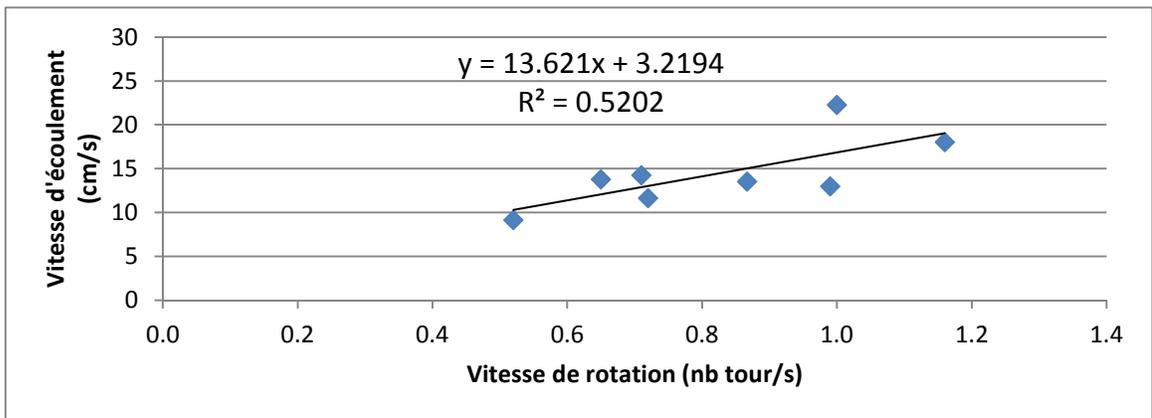


Figure 88: Droite d'étalonnage du micro-moulinet 111134

Le tableau 11 représente un exemple d'une fiche de mesure des vitesses d'écoulement avec un micro-moulinet remplie lors d'un jaugeage du débit. Cette fiche présente aussi les vitesses d'écoulement calculées aux points de mesure des vitesses de rotation.

Tableau 11: Exemple de fiche de mesure des vitesses d'écoulement avec un micro-moulinet

<b>Date:</b>	2010-10-22
<b>Endroit:</b>	Amont
<b>Mesuré par:</b>	Étienne Gaborit

	<b>Début</b>	<b>Fin</b>
<b>Heure</b>	10h00	10h30
<b>Hauteur d'eau</b>	11.5	12

Point	# moulinet	H. p/r au lit	nb tours	temps (sec)	nb tours/sec	Vitesse (cm/sec)
3	172-112-5	3	63	60	1.1	6.9
4	172-112-5	3	54	60	0.9	6.2
5	172-112-5	3	137	60	2.3	11.9
6	172-112-5	3	110	60	1.8	10.1
7	172-112-5	3	166	60	2.8	13.9
4	172-112-5	6	126	60	2.1	11.2
5	172-112-5	6	198	60	3.3	16.1
6	172-112-5	6	170	60	2.8	14.2
7	172-112-5	6	139	60	2.3	12.1
5	172-112-5	10	237	60	4.0	18.8

À la suite du calcul des vitesses d'écoulement, il faut les intégrer à la surface d'écoulement. L'intégration a été faite à l'aide du logiciel Golden Software Surfer 8. Le « gridding method » utilisé pour le calcul du débit est le « natural neighbor » qui se compare à la méthode des polygones de Thiessens. Une anisotropie avec un ratio de 8 a été imposée. Les valeurs horizontales sont donc 8 fois plus importantes que les valeurs verticales dans l'interpolation. Ceci se justifie par le fait que la variation des vitesses d'écoulement est plus importante

verticalement qu'horizontalement. Le tableau 12 et le tableau 13 présentent les débits jaugés en amont et en aval.

Tableau 12 : Débits jaugés au point d'échantillonnage amont

Hauteur (cm)	Débit (l/s)
11,75	19,7
13,50	23,7
14,00	28,2
16,50	59,5
28,25	172,6
29,50	113,2

Tableau 13 : Débits jaugés au point d'échantillonnage aval

Hauteur (cm)	Débit (l/s)
20,0	32,2
22,0	43,7
23,0	51,6
24,5	75,6
35,0	201,1
48,0	529,7

La figure 89 et la figure 90 présentent respectivement les courbes de tarage aux points d'échantillonnage amont et aval développées à partir des débits présentés au tableau 12 et au tableau 13. Les débits correspondant aux hauteurs de 16,5 et 28,25 cm en amont ont été supprimés, car ils semblaient aberrants. Ceci est probablement attribuable au fait que les vitesses mesurées sont hautement en dehors de la gamme d'étalonnage du micro-moulinet utilisé. L'équation de chaque courbe de tarage décrivant la relation hauteur-débit est présentée sur ces figures.

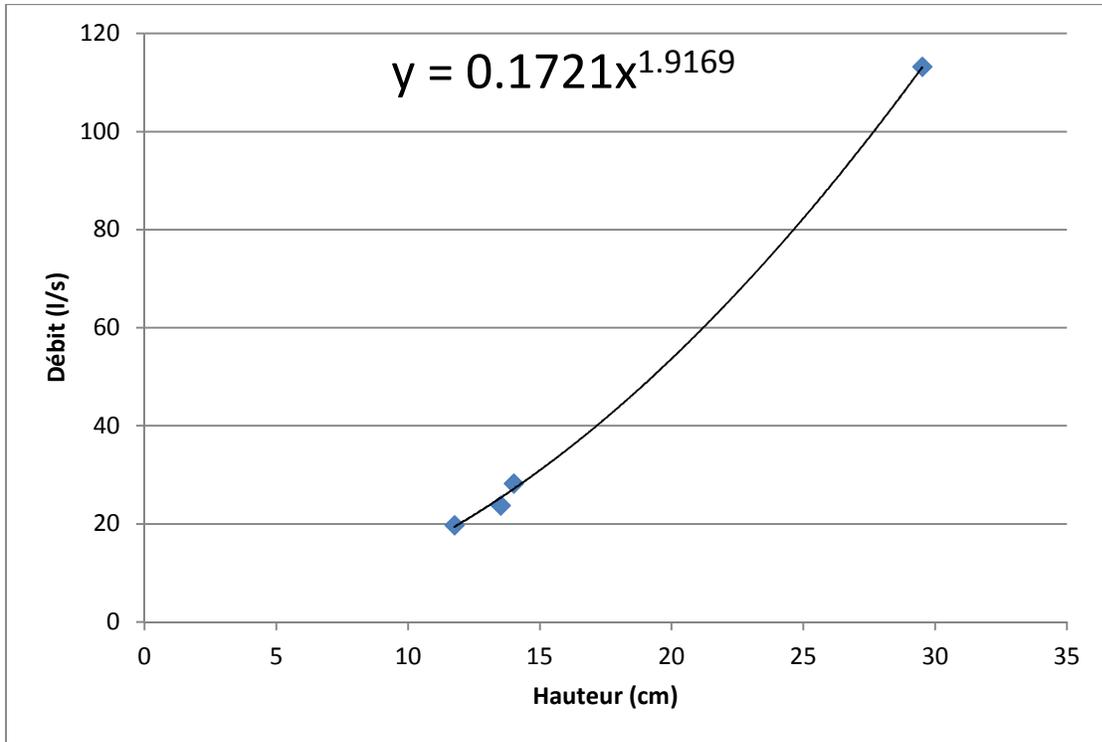


Figure 89: Courbe de tarage au point d'échantillonnage amont

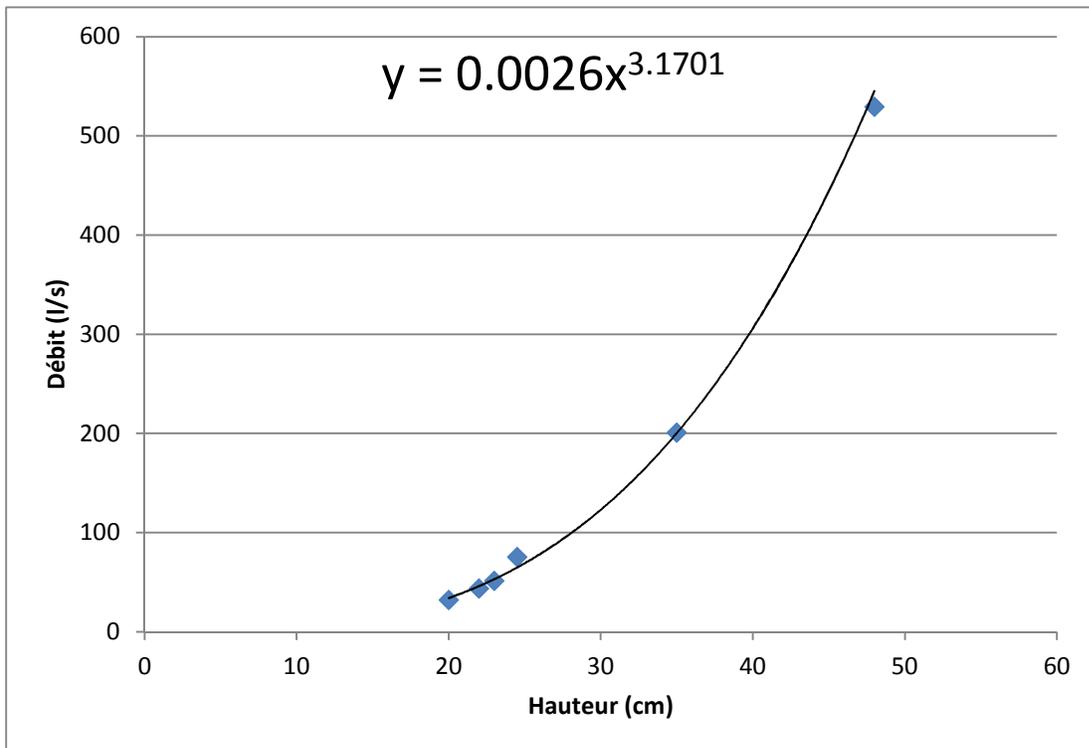


Figure 90: Courbe de tarage au point d'échantillonnage aval



## **ANNEXE 6. Jaugeage du débit durant le suivi en temps réel**

Cette annexe présente les résultats concernant le jaugeage du débit durant le suivi en temps réel. Comme mentionné dans la section 3.4 Jaugeage du débit, la méthode de jaugeage point par point a été remplacée par la méthode verticale par verticale à la suite d'un bris du premier montage utilisé pour le jaugeage du débit. Les résultats de ces deux méthodes sont présentés respectivement selon cet ordre. Cette annexe ne répète pas l'information qui a déjà été discutée à l'annexe 5.

Le tableau 14 présente la bathymétrie de la section transversale au point d'installation des capteurs bulle à bulle et de vitesse Doppler relevée selon la méthode expliquée dans la section 3.4 Jaugeage du débit.

Durant le suivi en temps réel, seulement le micro-moulinet le plus sensible a été utilisé pour la méthode de jaugeage du débit point par point. En effet, après l'étude préliminaire, il a été déterminé que ce micro-moulinet permettait de mesurer la gamme de vitesses d'écoulement rencontrée dans le ruisseau Notre-Dame. Le tableau 15 montre les vitesses de rotation et d'écoulement mesurées lors de l'étalonnage du micro-moulinet 172-112-5. Selon ce tableau, le micro-moulinet ne tourne plus à une vitesse moyenne de 2,78 cm/s. Ce point n'a pas été inclus dans la détermination de la courbe d'étalonnage.

La figure 91 présente la droite d'étalonnage du micro-moulinet 172-112-5 ainsi que son équation. La droite présentée sur cette figure est celle qui possède le meilleur coefficient de détermination ( $R^2$ ).

Tableau 14 : Bathymétrie de la section transversale au point de mesure aval

Point	Distance (cm)	Profondeur (cm)	Commentaires
0	0	5	Rive gauche dans le sens de l'écoulement, point d'origine
1	150	64	
2	200	68	
3	250	73	
4	300	123	
5	350	142	
6	400	143	
7	450	144	Point d'installation des capteurs
8	500	145	
9	550	146	
10	600	140	
11	650	102	
12	700	87	
13	891	0	

Tableau 15 : Étalonnage du micro-moulinet 172-112-5

# Débit	# moulinet	nb tours	Temps (s)	Vitesse de rotation (nb tours/s)	Vitesse moyenne ADV (cm/s)
1	172-112-5	172	100	1,72	8,51
2	172-112-5	196	100	1,96	9,77
3	172-112-5	243	100	2,43	12,82
4	172-112-5	220	100	2,2	11,33
5	172-112-5	267	100	2,67	14,45
6	172-112-5	302	100	3,02	11,99
7	172-112-5	331	100	3,31	15,50
8	172-112-5	394	100	3,94	19,22
9	172-112-5	440	100	4,4	20,73
10	172-112-5	0	100	0	2,78
11	172-112-5	105	100	1,05	7,57
12	172-112-5	72	100	0,72	5,53
13	172-112-5	116	100	1,16	7,86
14	172-112-5	155	100	1,55	10,52

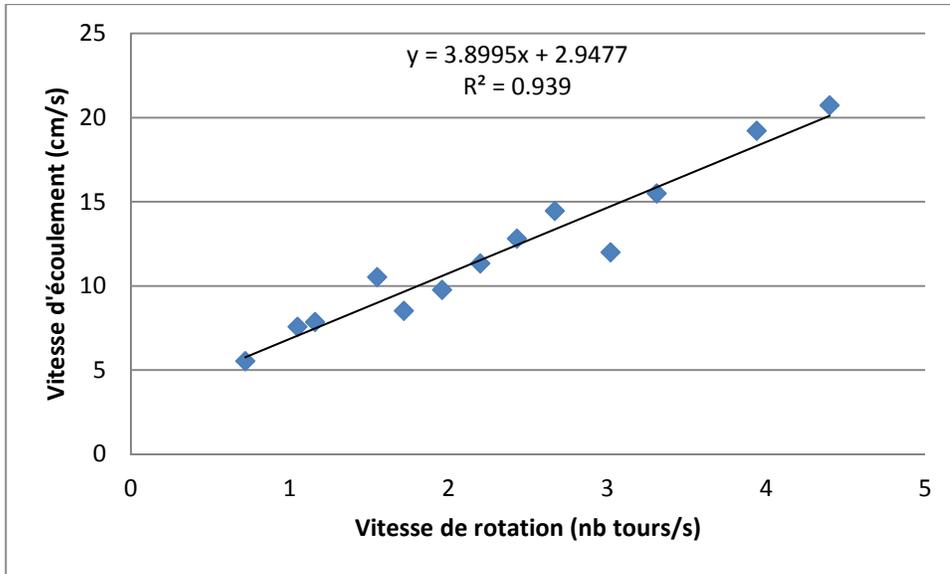


Figure 91: Étalonnage du micro-moulinet 172-112-5

Le micro-moulinet utilisé pour la méthode verticale par verticale est déjà étalonné par la compagnie fournisseur et affiche des mesures de vitesse en m/s. Cela a donc simplifié grandement le jaugeage du débit. Le calcul du débit s'est fait selon la méthode de la section médiane.

Le tableau 16 présente les débits jaugés au point de mesure aval selon les deux méthodes de jaugeage du débit.

Tableau 16 : Débits jaugés au point de mesure aval

Méthode	Hauteur (cm)	Débit (l/s)
Point par point	44,5	302,5
	32,3	74,3
	30,0	59,7
	38,8	192,1
	21,5	14,0
Verticale par verticale	36,8	234,9
	38,3	259,5
	41,5	340,1
	38,9	273,5
	47,0	417,2
	46,0	406,2
	46,8	446,5
	48,7	535,6
	49,7	598,1
	37,1	237,7
	37,4	224,2
	37,3	188,9

La figure 92 montre la courbe de tarage développée au point de mesure aval à partir des débits présentés dans le tableau 16.

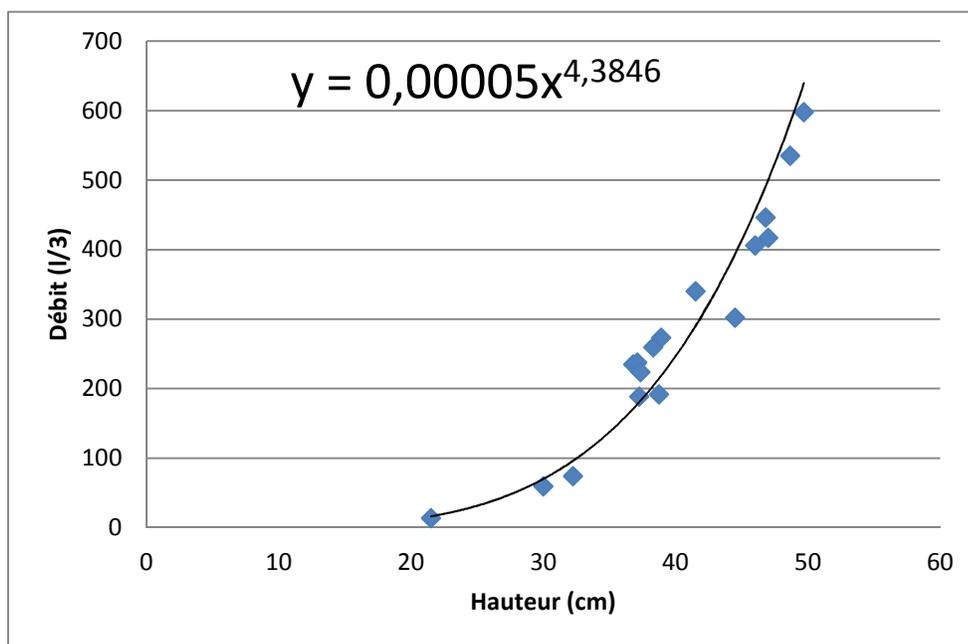


Figure 92 : Courbe de tarage au point de mesure aval

## ANNEXE 7. Chartes de contrôle

Les figures suivantes présentent les « standard solution control charts » obtenus durant le suivi en temps réel. Il est important de mentionner que les tests préliminaires pour déterminer les limites de contrôle des « standard solution control charts » n'ont pas été effectués. L'auteur de ce mémoire jugeait donc par lui-même si la différence entre la valeur mesurée et la valeur de la solution standard ou de la méthode redondante était significative pour qu'un entretien soit nécessaire. Le « standard solution control chart » du Spectro::lyser n'est pas montré, car ce capteur est complexe et il est difficile de lui appliquer cette méthode d'évaluation de la qualité des données. Aussi, il n'y a pas de « standard solution control chart » pour le capteur de vitesse Doppler, car aucune méthode redondante n'était disponible.

Les « cleaning control charts » ne sont pas présentés, car, à la fin du suivi en temps réel, il a été observé qu'un intervalle de mesure de 2 minutes était trop grand pour leur utilisation. En effet, pour les paramètres mesurés à un intervalle de 2 minutes, une différence significative était souvent détectée entre les valeurs mesurées avant et après un nettoyage. Cette différence était causée par un changement naturel de la qualité de l'eau. Le « cleaning control chart » a donc seulement été utilisé pour la mesure des MeS par le Solitax sc dont l'intervalle de mesure était de 5 secondes. Il est présenté dans la section ADQATs.

La figure 93 présente le « standard solution control chart » du capteur bulle à bulle obtenu durant le suivi en temps réel. On remarque qu'il y a un point qui s'éloigne particulièrement des autres. Cet écart est causé par le fait que le capteur bulle à bulle s'était décroché de son emplacement. La variation des autres points est principalement reliée à l'incertitude de mesure causée par des vaguelettes présentes à la surface de l'eau.

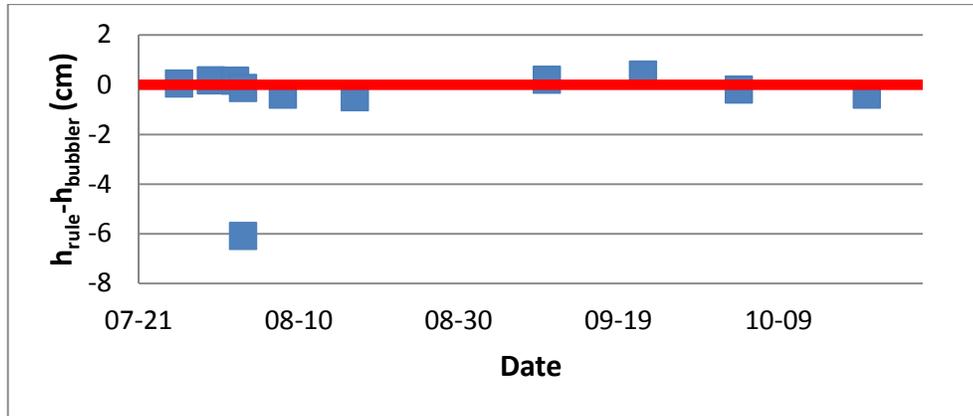


Figure 93: « Standard solution control chart » du capteur bulle à bulle

La figure 94 présente le « standard solution control chart » du capteur 3700 sc Conductimètre inductif obtenu durant le suivi en temps réel. On remarque qu'il y a généralement une différence entre la valeur mesurée et celle de la solution standard. Cela indique donc qu'un entretien est nécessaire. Cependant, il a été observé que la mesure de ce capteur est influencée par les parois du récipient dans lequel est mesurée la solution standard. Il est donc possible que le capteur mesurait des données de bonnes qualités malgré les fautes détectées.

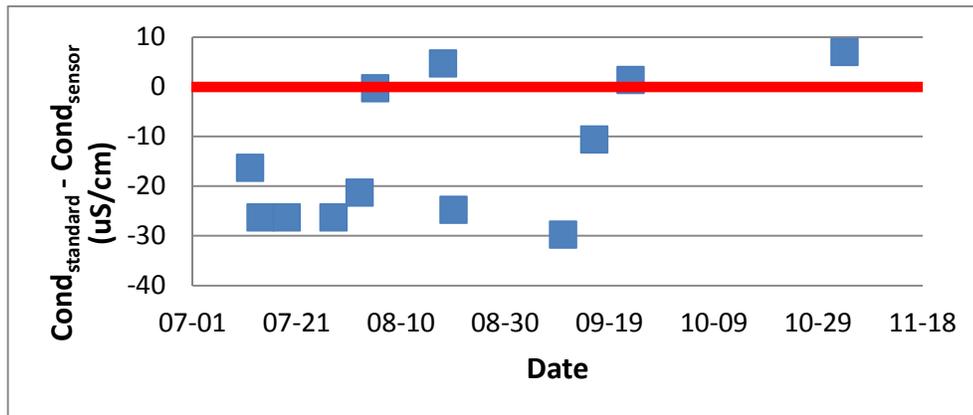


Figure 94 : « Standard solution control chart» du capteur 3700 sc Conductimètre inductif

La figure 95 présente le « standard solution control chart » du capteur pHD sc différentiel numérique obtenu durant le suivi en temps réel. Cette figure présente les valeurs mesurées dans des solutions standards de pH 7 et 10. On remarque qu'il y a souvent une différence

entre la valeur mesurée et celle de la solution standard. Cette différence est principalement causée par une dérive du signal du capteur. Il était donc nécessaire de le vérifier et de le calibrer régulièrement pour assurer une bonne qualité des données mesurées.

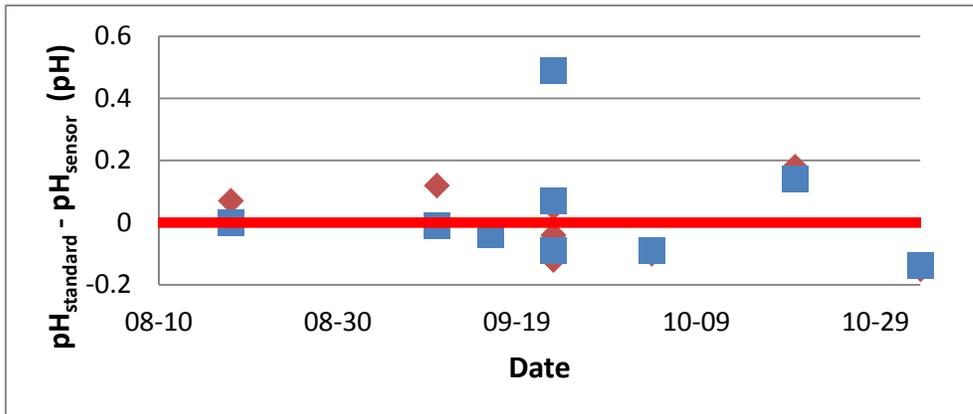


Figure 95 : « Standard solution control chart » du capteur pHD sc différentiel numérique (pH7 : marqueur en losange, pH 10 : marqueur carré)

La figure 96 présente le « standard solution control chart » du LDO obtenu durant le suivi en temps réel. Cette figure montre que ce capteur est généralement influencé par une dérive du signal. Il doit donc être vérifié et calibré régulièrement pour garantir une bonne qualité des données mesurées.

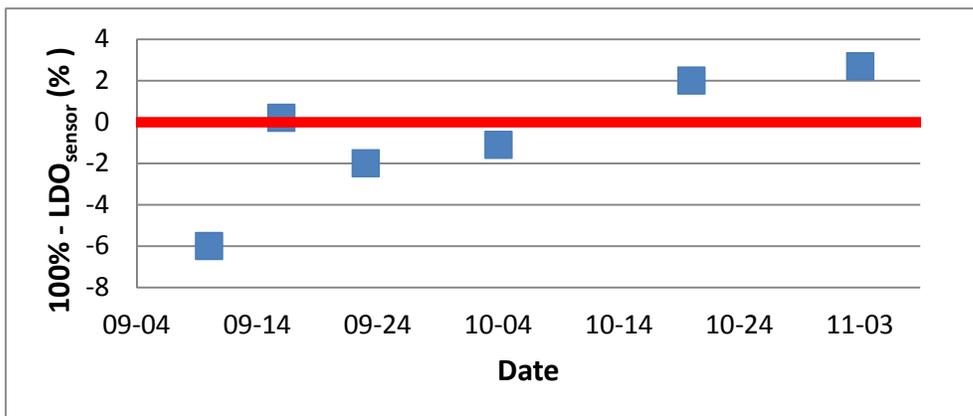


Figure 96 : « Standard solution control chart » du capteur LDO

La figure 97 présente les « standard solution control charts » obtenus durant le suivi en temps réel des capteurs mesurant la température soit le capteur 3700 sc Conductimètre inductif, le capteur pHD sc différentiel numérique et le capteur LDO. Généralement, les différences entre les valeurs mesurées par les capteurs et la méthode redondante sont faibles à l'exception de la dernière vérification. Cet écart est causé par le fait que la calibration de la température s'effectue avec un seul point (« one point calibration ») et qu'elle ne couvre plus la gamme de mesure lorsque la température du cours d'eau varie grandement. Dans ce cas-ci, les capteurs ont été calibrés en été lorsque l'eau était relativement chaude et l'écart est mesuré en automne lorsque l'eau s'est refroidie. Aussi, il n'était pas possible de calibrer la température du capteur LDO, c'est pourquoi la température mesurée par ce capteur était plus élevée que la température mesurée par les deux autres capteurs.

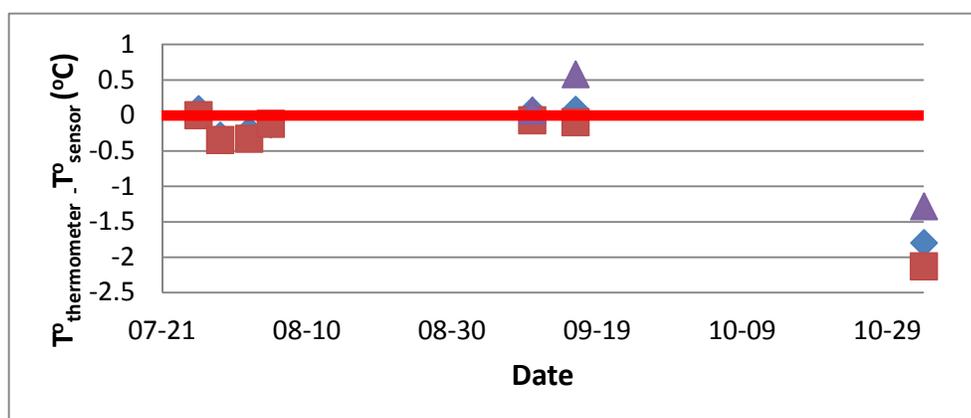


Figure 97 : « Standard solution control charts » pour la mesure de la température des capteurs 3700 sc Conductimètre inductif (marque en losange), pHD sc différentiel numérique (marqueur carré) et LDO (marqueur en triangle)

## ANNEXE 8. Méthodes en ligne de détection de fautes

Cette annexe présente l'application des méthodes en ligne de détection de fautes (« on-line methods ») sur deux différents paramètres de la qualité de l'eau soit les MeS non calibrées (SS\*) et mesurées par un capteur de turbidité et la conductivité. De plus, les méthodes ont été testées sur différents intervalles de mesure en sous-échantillonnant la série de SS\* possédant un intervalle de mesure de 5 secondes. La conductivité a été mesurée à un intervalle de 2 minutes, c'est pourquoi les méthodes ont été testées seulement sur cet intervalle de mesure.

Le tableau 17 présente les valeurs des constantes des modèles de prédiction de la valeur ( $\alpha_x$ ) et de l'erreur ( $\alpha_e$ ), et du lisseur ( $h$ ) déterminées pour les différents intervalles de mesure. Les valeurs des constantes des modèles et du lisseur entre les parenthèses ont été déterminées respectivement avec le « RMSE » et le « cross-validation criteria » comme décrit dans le chapitre 6 ADQATs. Cependant, ces valeurs n'étaient pas jugées optimales et elles ont donc été ajustées. Par exemple, l'erreur entre les données mesurées et les valeurs prédites était parfois corrélée attribuable à une valeur de la constante  $\alpha_x$  trop petite. De plus, l'intervalle de prédiction était parfois très variable à cause d'une valeur de la constante  $\alpha_e$  trop grande. La figure 98 et la figure 99 montrent l'effet des constantes  $\alpha_x$  et  $\alpha_e$  sur la méthode de détection des données aberrantes appliquée à une série de SS\* dont l'intervalle de mesure est de 5 secondes. Les valeurs des constantes  $\alpha_x$  et  $\alpha_e$  à la figure 98 et à la figure 99 sont respectivement celles qui ont été déterminées à l'aide du « RMSE » et celles qui ont été ajustées.

L'erreur entre les données lissées et mesurées était parfois corrélée à cause d'une trop grande valeur de la constante  $h$  déterminée avec le « cross-validation criteria ». La figure 100 et la figure 101 montrent l'effet de la constante  $h$  sur la méthode de lissage appliquée à une série SS\* dont l'intervalle de mesure est de 30 secondes. Les valeurs de la constante  $h$  à la figure 100 et à la figure 101 sont respectivement celle qui a été déterminée à l'aide du « cross-validation criteria » et celle qui a été ajustée.

Les constantes des modèles ont été ajustées subjectivement en observant sur un graphique des séries de données jugées de bonne qualité, ainsi que les valeurs prédites et les intervalles

de prédiction. La constante  $\alpha_e$  a été fixée à 0,1, car l'auteur de ce mémoire a observé qu'elle n'est pas prioritaire dans la méthode de détection des données aberrantes et qu'il est préférable d'avoir un intervalle de prédiction relativement stable. La constante du lisseur a été ajustée en se basant sur les résultats du « runs test » indiquant s'il y a une corrélation dans l'erreur entre les données mesurées et lissées. La constante K multipliant l'écart-type de l'erreur de prédiction pour déterminer l'intervalle de prédiction a été fixée à 5 qui est la valeur proposée par Montgomery (1990).

Tableau 17 : Valeurs des constantes des modèles et du lisseur pour différents intervalles de mesure et pour les paramètres SS\* et conductivité

Paramètre	Intervalle de mesure	$\alpha_x$	$\alpha_e$	h
SS*	5 secondes	0,1 (0,04)	0,1 (0,27)	13 (13)
SS*	15 secondes	0,2 (0,06)	0,1 (0,64)	6 (6)
SS*	30 secondes	0,25 (0,2)	0,1 (0,11)	3 (5)
SS*	1 minute	0,35 (0,3)	0,1 (0,1)	2 (3)
SS*	2 minutes	0,55 (0,42)	0,1 (0,23)	1 (2)
Conductivité	2 minutes	0,36 (0,6)	0,69 (0,1)	1 (4)

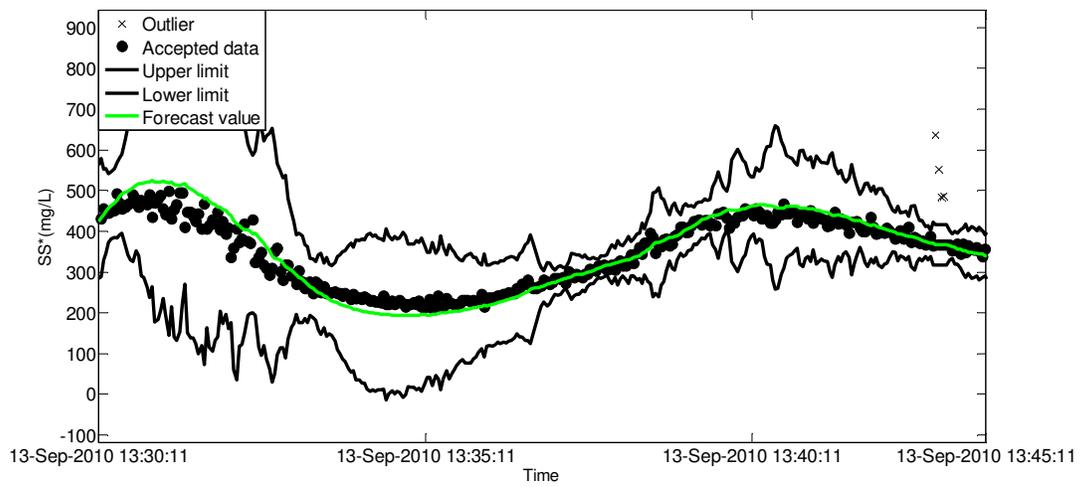


Figure 98: Méthode de détection des données aberrantes sur les SS\*  
 ( $\alpha_x=0,04$ ;  $\alpha_e=0,27$ ;  $K=5$ ;  $\Delta t=5$  sec.)

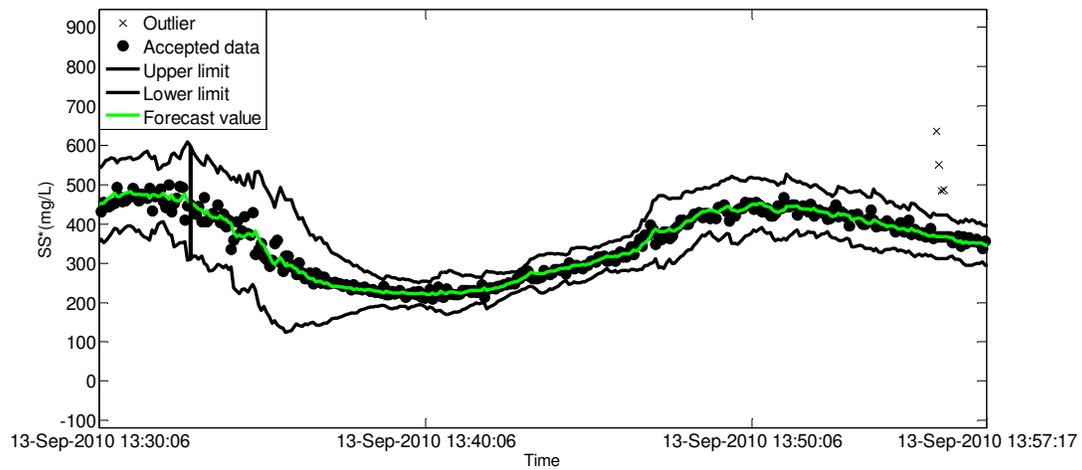


Figure 99 : Méthode de détection des données aberrantes sur les SS\*  
 ( $\alpha_x=0,1$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $\Delta t=5$  sec.)

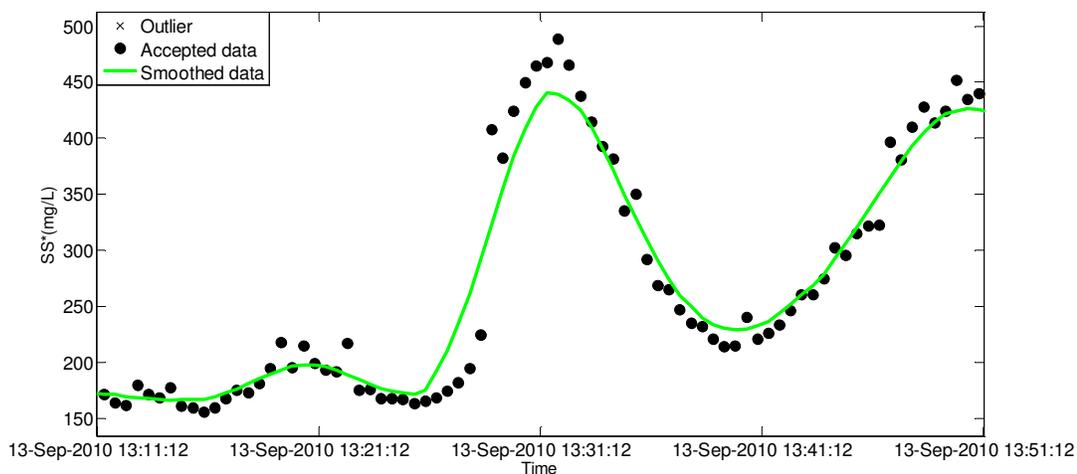


Figure 100 : Méthode de lissage des données sur les SS\*  
( $h=5$ ;  $\Delta t=30$  sec.)

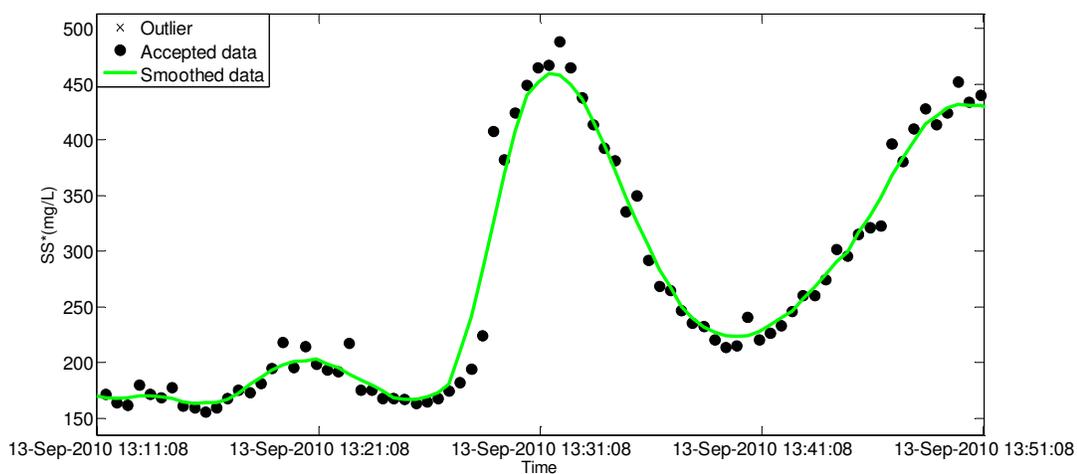


Figure 101 : Méthode de lissage des données sur les SS\*  
( $h=3$ ;  $\Delta t=30$  sec.)

Les figures ci-après démontrent l'application des méthodes en ligne de détection de fautes. Pour chaque intervalle de mesure, des figures illustrent, premièrement, l'application de la méthode de détection des données aberrantes et, ensuite, l'application des méthodes de détection de fautes. Le « runs test » et le RSD sont toujours calculés sur 30 données, même si l'intervalle de mesure varie.

## SS\* (5 sec.)

La figure 102 montre l'application de la méthode de détection des données aberrantes sur une série de SS\* dont l'intervalle de mesure est de 5 secondes. On observe que la méthode détecte plusieurs données aberrantes et qu'il y a un « loss-of-control ».

La figure 103 est un agrandissement de la figure 102 montrant le « loss-of-control ». On observe un changement brut dans la série de données, indiqué par la flèche, ce qui crée le « loss-of-control ». Le modèle n'est donc pas en mesure de prédire ce changement. Cependant, la réinitialisation de la méthode de détection des données aberrantes permet de récupérer la plupart des données détectées comme état aberrantes à cause du « loss-of-control ». Les données prédites représentent généralement bien les données mesurées, en ce sens que l'erreur entre les données mesurées et prédites ne semble pas être corrélée. Aussi, l'intervalle de prédiction est relativement restrictif.

La figure 104 présente l'application des méthodes de détection de fautes sur la série de SS\* dont l'intervalle de mesure est de 5 secondes. Selon le « runs test », on remarque que généralement aucune corrélation significative n'est détectée dans l'erreur entre les données lissées et mesurées, à l'exception d'un endroit. Les données lissées représentent donc relativement bien les données mesurées. Par contre, des pentes croissantes et décroissantes, et des RSD anormalement élevés sont détectés.

La figure 105 est un agrandissement de la figure 104 où le « loss-of-control » se produit. Le « runs test » détecte une corrélation significative dans l'erreur qui s'explique par le fait que les données lissées ne représentent pas bien les données mesurées. Conséquemment, le RSD calculé est élevé pour ces données. De plus, une autre faute est détectée avec le RSD qui est anormalement élevé dans la région encadrée. Cette faute indique un bruit anormalement élevé dans le signal.

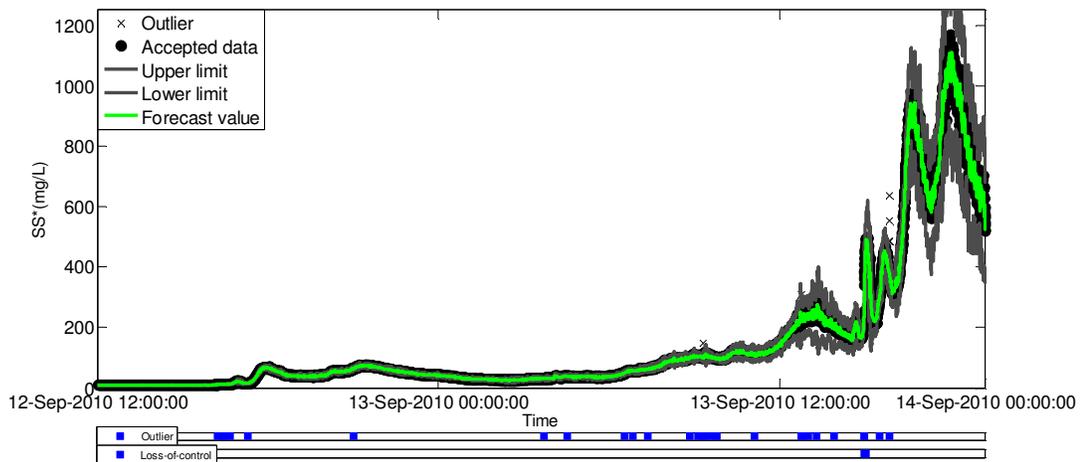


Figure 102 : Méthode de détection des données aberrantes sur les SS\*  
 ( $\alpha_x=0,1$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $\Delta t=5$  sec.)

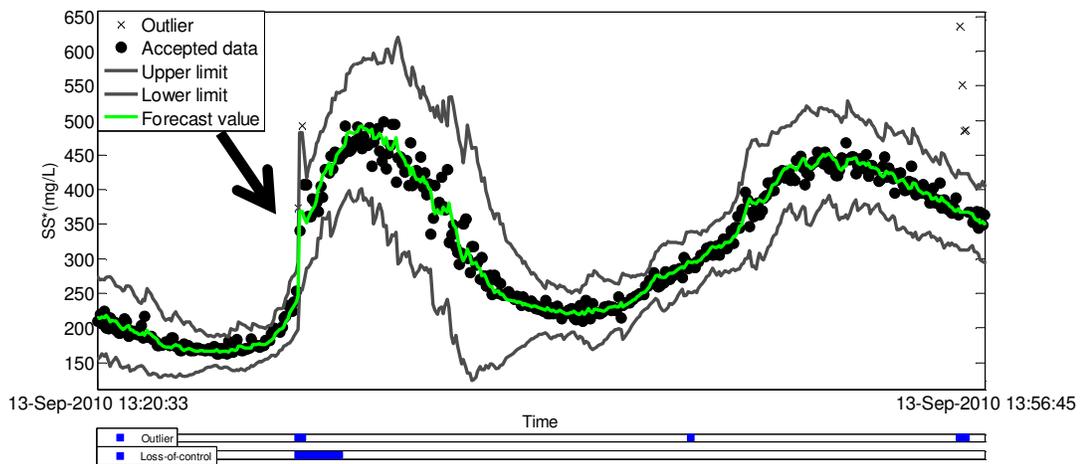


Figure 103 : Méthode de détection des données aberrantes sur les SS\*, agrandissement de  
 la figure 102 ( $\alpha_x=0,1$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $\Delta t=5$  sec.)

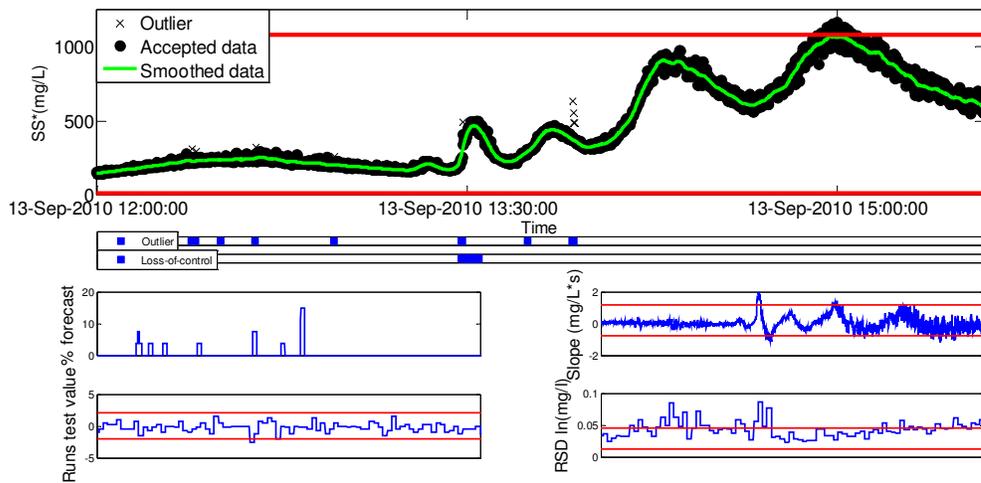


Figure 104 : Méthodes de détection de fautes sur les SS\*  
 $(\alpha x=0,1; \alpha e=0,1; K=5; h=13; \Delta t=5 \text{ sec.})$

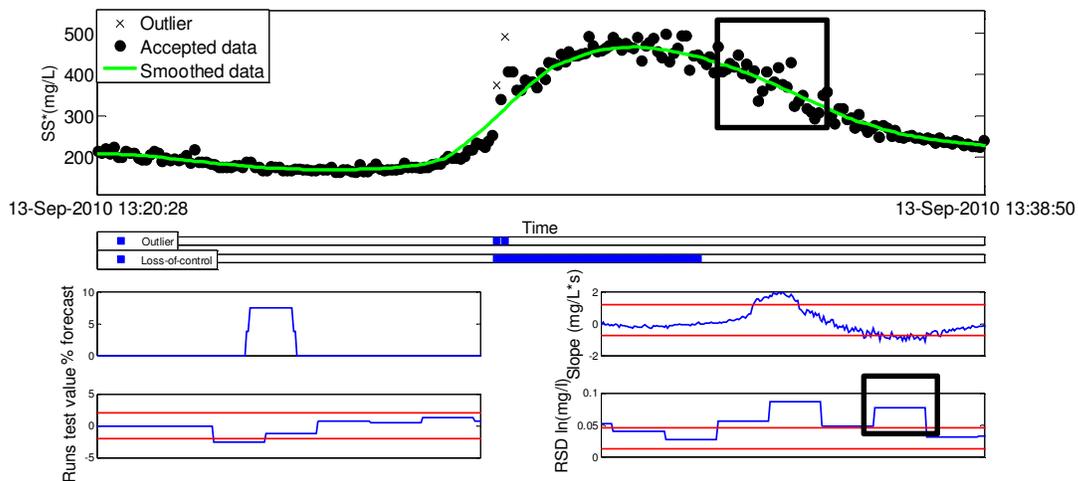


Figure 105 : Méthodes de détection de fautes sur les SS\*, agrandissement de la figure 104  
 $(\alpha x=0,1; \alpha e=0,1; K=5; h=13; \Delta t=5 \text{ sec.})$

### **SS\* (15 sec.)**

La figure 106 montre l'application de la méthode de détection des données aberrantes sur une série de SS\* dont l'intervalle de mesure est de 15 secondes. L'augmentation de l'intervalle de mesure de 5 secondes à 15 secondes ne semble pas influencer la performance de la méthode de détection des données aberrantes. Il y a moins de données aberrantes détectées, cependant, le nombre de données mesurées est moindre. De plus, un « loss-of-control » se produit au même endroit que pour l'essai avec un intervalle de mesure de 5 secondes.

La figure 107 est un agrandissement de la figure 106 montrant le « loss-of-control ». Comme pour l'essai avec un intervalle de mesure de 5 secondes, la réinitialisation de la méthode de détection des données aberrantes permet de récupérer la majorité des données détectées comme étant aberrantes à cause du « loss-of-control ». L'intervalle de prédiction est un peu moins restrictif attribuable à une légère augmentation de l'erreur de prédiction. Toutefois, les données prédites par le modèle représentent généralement bien les données mesurées.

La figure 108 présente l'application des méthodes de détection de fautes sur la série de SS\* dont l'intervalle de mesure est de 15 secondes. Selon le « runs test », aucune corrélation significative de l'erreur n'est détectée entre les données lissées et mesurées. Cependant, comme pour l'essai avec un intervalle de 5 secondes, des pentes croissantes et décroissantes, et des RSD anormalement élevés sont mesurés.

La figure 109 est un agrandissement de la figure 108 où le « loss-of-control » se produit. Le « runs test » ne détecte aucune corrélation significative de l'erreur, même si on en observe une visuellement à l'endroit indiqué par une flèche. Ceci est probablement causé par un nombre de données trop élevé utilisé dans le calcul du « runs test ». Des fautes sont aussi détectées avec le RSD qui est anormalement élevé. Cependant, étant donné que l'intervalle de mesure est plus grand et que le nombre de données utilisé pour calculer le RSD demeure le même, les fautes sont moins localisées que pour l'essai avec un intervalle de mesure de 5 secondes.

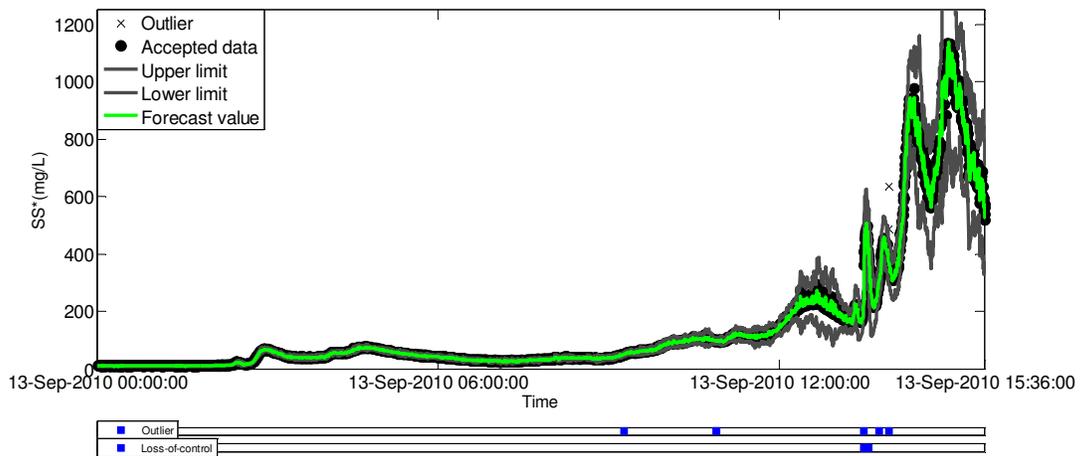


Figure 106 : Méthode de détection des données aberrantes sur les SS\*  
 ( $\alpha x=0,2$ ;  $\alpha e=0,1$ ;  $K=5$ ;  $\Delta t=15$  sec.)

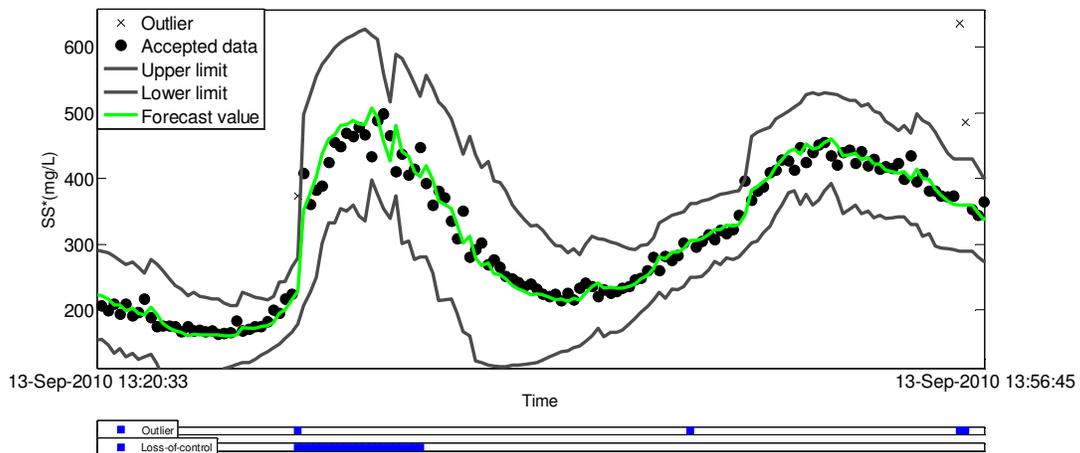


Figure 107 : Méthode de détection des données aberrantes sur les SS\*, agrandissement de  
 la figure 106 ( $\alpha x=0,2$ ;  $\alpha e=0,1$ ;  $K=5$ ;  $\Delta t=15$  sec.)

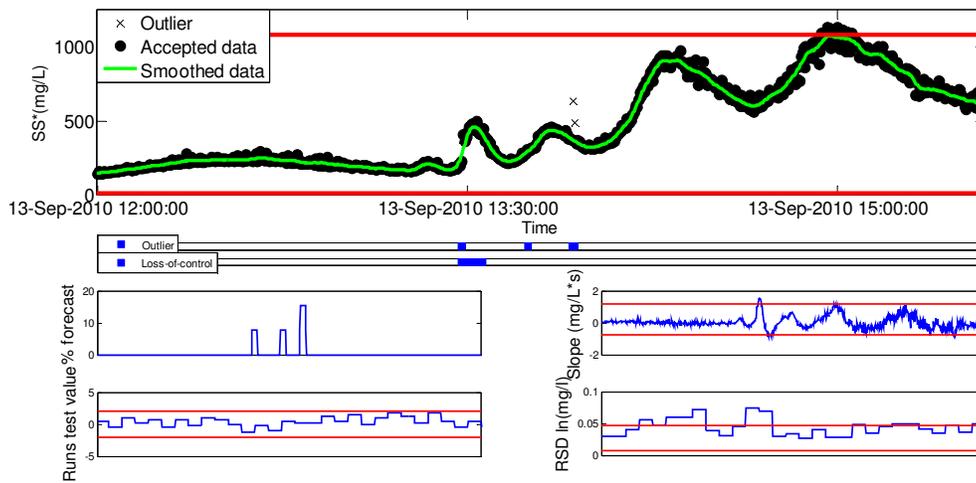


Figure 108 : Méthodes de détection de fautes sur les SS\*  
 ( $\alpha_x=0,2$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $h=6$ ;  $\Delta t=15$  sec.)

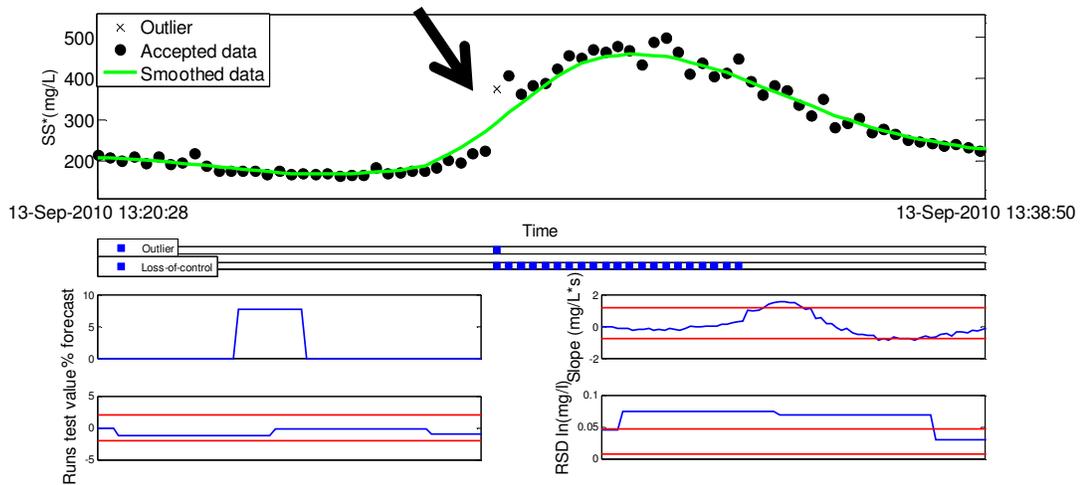


Figure 109 : Méthodes de détection de fautes sur les SS\*, agrandissement de la figure 108  
 ( $\alpha_x=0,2$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $h=6$ ;  $\Delta t=15$  sec.)

## SS\* (30 sec.)

La figure 110 illustre l'application de la méthode de détection des données aberrantes sur une série de SS\* dont l'intervalle de mesure est de 30 secondes. Contrairement aux essais précédents, deux « loss-of-controls » se produisent avec cet intervalle de mesure.

La figure 111 est un agrandissement de la figure 110 montrant le premier « loss-of-control ». On remarque que l'intervalle de prédiction est très restrictif lorsque la variation dans les données est relativement faible. Alors, un « loss-of-control » se produit lorsqu'il y a un changement dans la tendance, car les données consécutives au changement se situent à l'extérieur de l'intervalle de prédiction. Ce changement est plus graduel avec un intervalle de mesure plus petit et, c'est pourquoi, pour des intervalles de 5 et de 15 secondes, le premier « loss-of-control » ne se produit pas.

La figure 112 est un agrandissement de la figure 110 montrant le deuxième « loss-of-control ». Comme pour les essais précédents, la réinitialisation de la méthode de détection des données aberrantes permet de récupérer la plupart des données détectées comme étant aberrantes à cause du « loss-of-control ». Cependant, l'intervalle de prédiction est moins restrictif, car l'erreur de prédiction est généralement plus élevée. En effet, même si la valeur du paramètre  $\alpha_x$  a été ajustée, les données prédites par le modèle ne représentent pas bien les données mesurées lorsqu'il y a une variation rapide.

La figure 113 présente l'application des méthodes de détection de fautes sur la série de SS\* dont l'intervalle de mesure est de 30 secondes. Les résultats de l'application des méthodes pour un intervalle de mesure de 30 secondes sont similaires à ceux obtenus pour un intervalle de mesure de 15 secondes. Cependant, les fautes détectées avec le RSD sont moins localisées à cause que l'intervalle de mesure est plus grand et que le nombre de données pour calculer le RSD demeure le même.

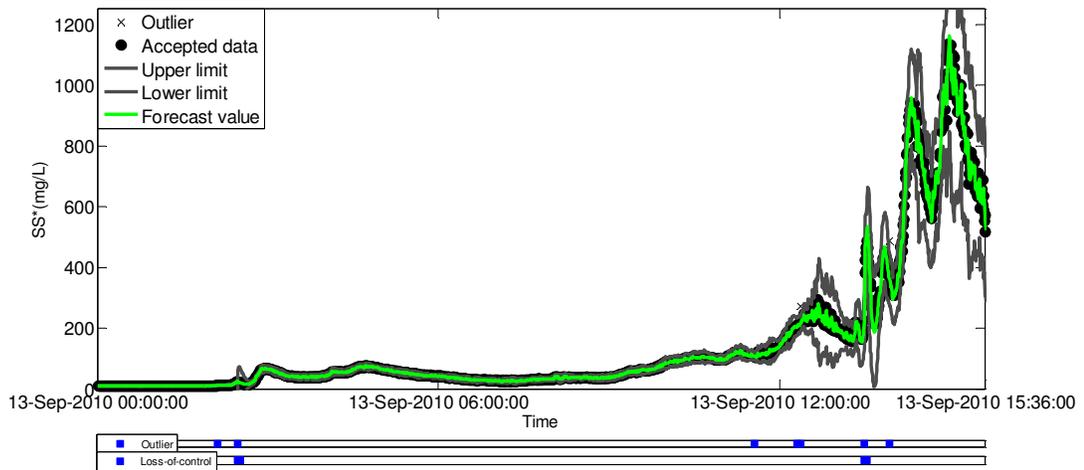


Figure 110 : Méthode de détection des données aberrantes sur les SS\*  
 ( $\alpha=0,25$ ;  $\alpha e=0,1$ ;  $K=5$ ;  $\Delta t=30$  sec.)

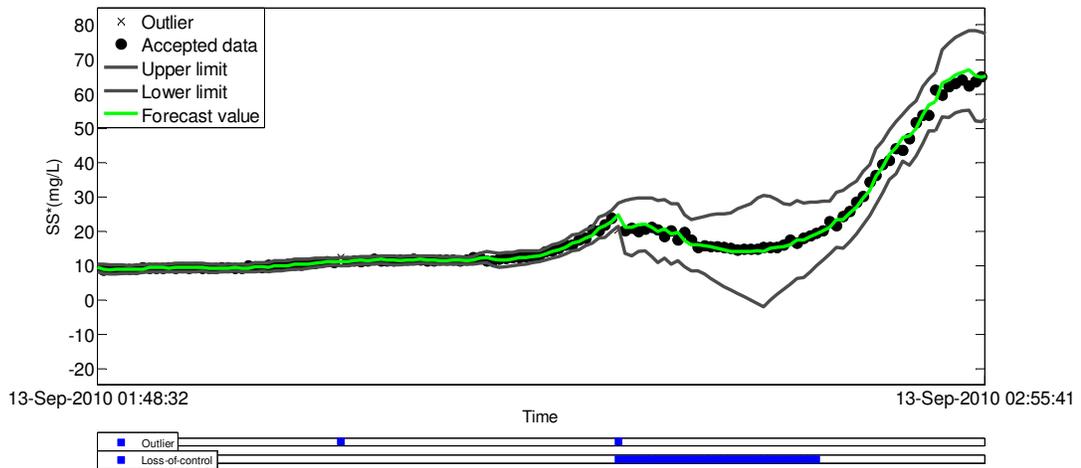


Figure 111 : Méthode de détection des données aberrantes sur les SS\*, agrandissement du premier « loss-of-control » de la figure 110 ( $\alpha=0,25$ ;  $\alpha e=0,1$ ;  $K=5$ ;  $\Delta t=30$  sec.)

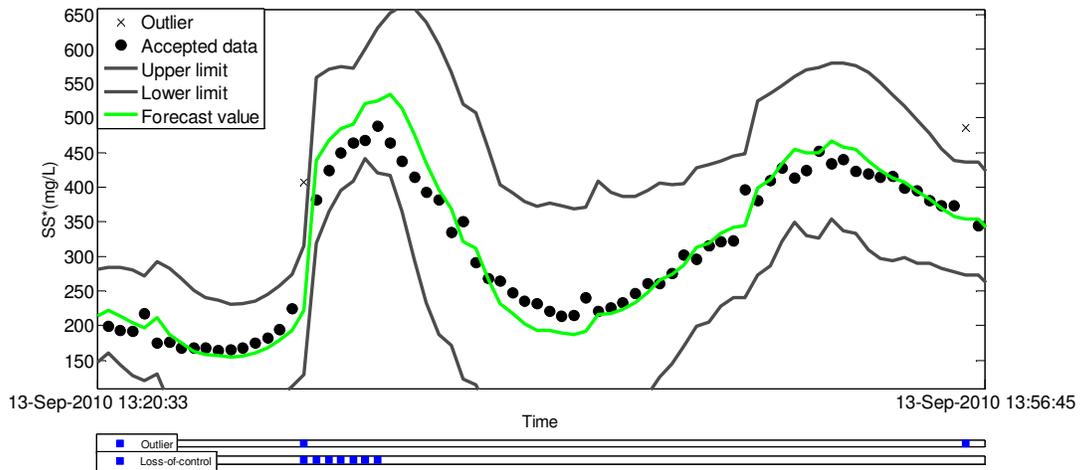


Figure 112 : Méthode de détection des données aberrantes sur les SS\*, agrandissement du deuxième « loss-of-control » de la figure 110 ( $\alpha_x=0,25$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $\Delta t=30$  sec.)

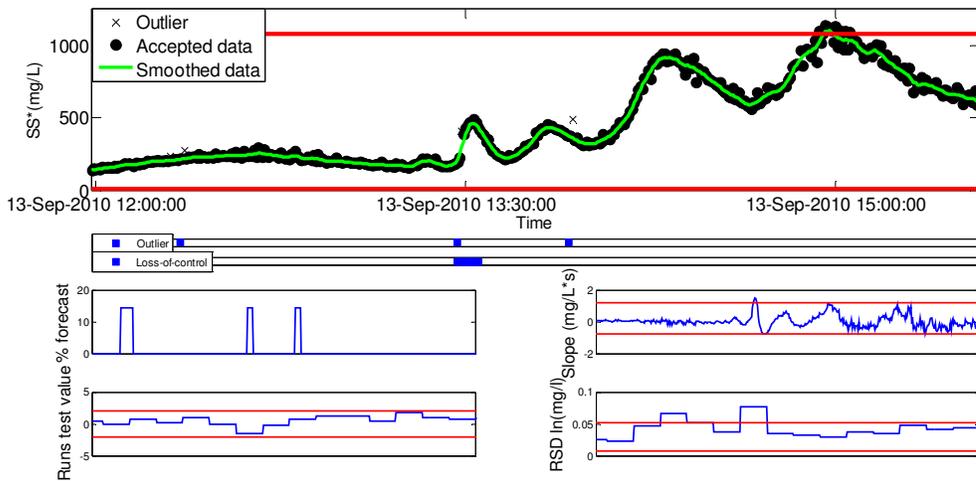


Figure 113 : Méthodes de détection de fautes sur les SS\* ( $\alpha_x=0,25$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $h=3$ ;  $\Delta t=30$  sec.)

## SS\* (60 sec.)

La figure 114 démontre l'application de la méthode de détection des données aberrantes sur une série de SS\* dont l'intervalle de mesure est de 1 minute. Contrairement à l'essai précédent, seulement un « loss-of-control » se produit avec cet intervalle de mesure. Ceci s'explique par le fait que l'erreur de prédiction est généralement plus grande attribuable et, donc, l'intervalle de prédiction est moins restrictif.

La figure 115 est un agrandissement de la figure 114 montrant le « loss-of-control ». Contrairement aux essais précédents avec des intervalles de mesure plus petits, la réinitialisation de la méthode de détection des données aberrantes est moins efficace et elle ne permet pas la récupération de la majeure partie des données détectées comme étant aberrantes. Comme pour l'essai dont l'intervalle de mesure est de 30 secondes, même si la valeur du paramètre  $\alpha_x$  a été ajustée, les données prédites par le modèle ne représentent pas bien les données mesurées lorsqu'il y a une variation dans ces dernières. L'erreur de prédiction est donc élevée et corrélée et, par conséquent, l'intervalle de prédiction est moins restrictif.

La figure 116 présente l'application des méthodes de détection de fautes sur la série de SS\* dont l'intervalle de mesure est de 1 minute. Les résultats de l'application des méthodes pour un intervalle de mesure de 1 minute sont similaires à ceux obtenus pour les essais possédant des intervalles de mesure de 15 et de 30 secondes. Cependant, moins de fautes sont détectées avec le RSD. Aussi, étant donné que la réinitialisation de la méthode de détection des données aberrantes ne récupère pas la majeure partie des données détectées comme étant aberrantes, les données lissées ne représentent pas bien les données mesurées correspondantes au « loss-of-control ». En effet, lorsqu'une donnée est détectée comme étant aberrante, celle-ci est remplacée par la donnée prédite et c'est la donnée prédite qui est ensuite utilisée par le lisseur Kernel. La valeur du « % forecast » où se produit le « loss-of-control » est donc très élevée, ce qui indique la possibilité que les données lissées ne soient pas représentatives des données mesurées. Par conséquent, les caractéristiques des données (pente, « runs test » et RSD) correspondantes aux valeurs de « % forecast » élevées sont douteuses.

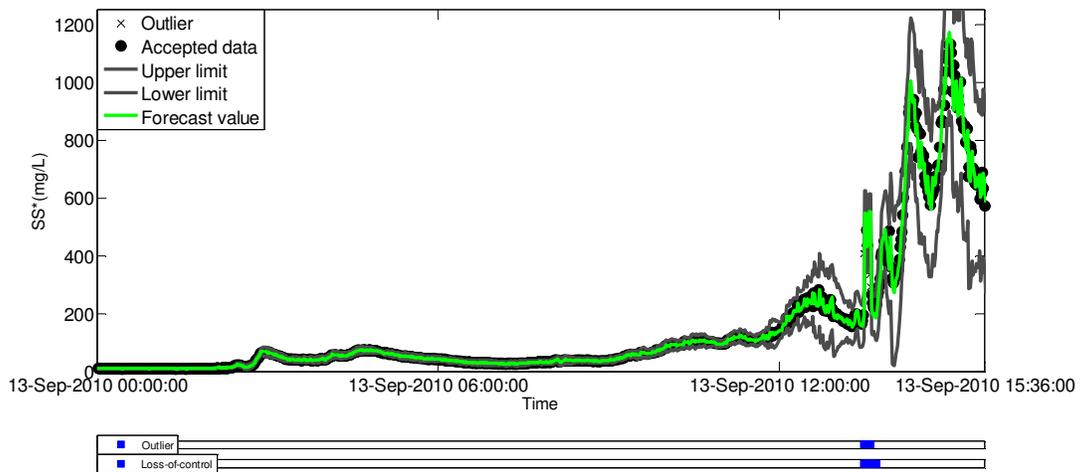


Figure 114 : Méthode de détection des données aberrantes sur les SS\*  
 ( $\alpha_x=0,35$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $\Delta t=1$  min.)

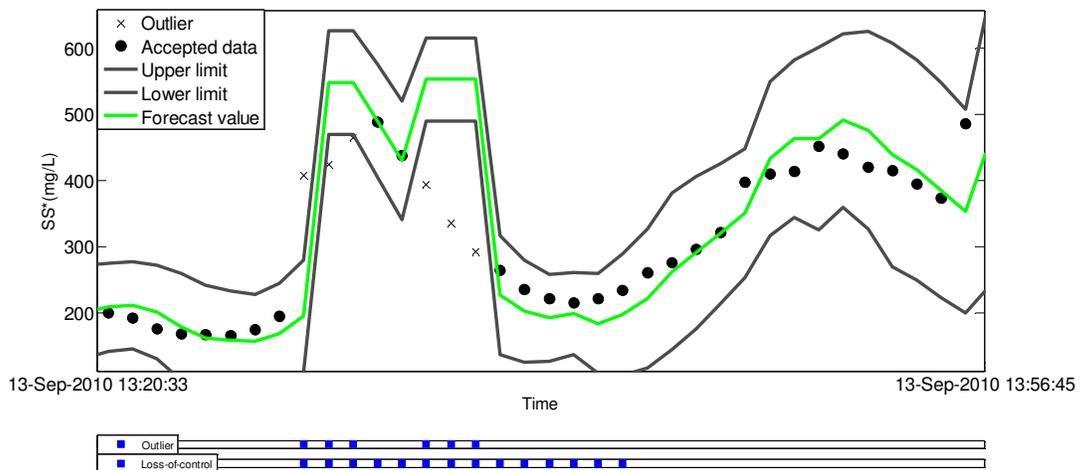


Figure 115 : Méthode de détection des données aberrantes sur les SS\*, agrandissement du  
 « loss-of-control » de la figure 114 ( $\alpha_x=0,35$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $\Delta t=1$  min.)

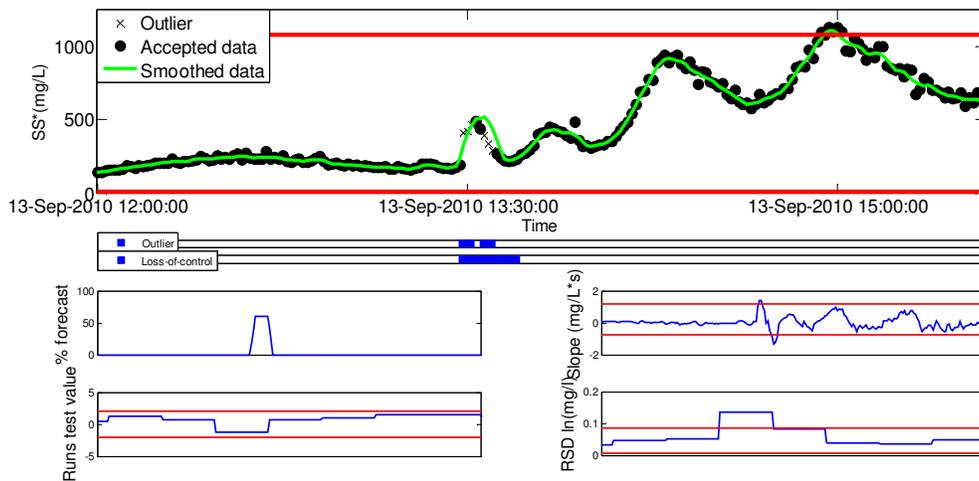


Figure 116 : Méthodes de détection de fautes sur les SS\*  
 $(\alpha=0,35; \alpha e=0,1; K=5; h=2; \Delta t=1 \text{ min.})$

### SS\* (120 sec.)

La figure 117 présente l'application de la méthode de détection des données aberrantes à une série de SS\* dont l'intervalle de mesure est de 2 minutes. Contrairement aux résultats obtenus pour l'essai avec un intervalle de mesure de 30 secondes, il y a seulement un « loss-of-control ». Le deuxième « loss-of-control » ne se produit pas, car seulement quatre données sont détectées comme aberrantes étant donné que l'intervalle de mesure est plus grand.

La figure 118 est un agrandissement de la figure 117 montrant le « loss-of-control ». Comme pour l'essai dont l'intervalle de mesure est de 30 secondes, le « loss-of-control » est attribuable au fait que la variation des données est relativement faible et, donc, l'intervalle de prédiction est très restrictif. Aussitôt qu'il y a une variation dans les données, il y a un « loss-of-control », car les données consécutives au changement sont à l'extérieur de l'intervalle de prédiction.

La figure 119 présente l'application des méthodes de détection de fautes sur la série de SS\* dont l'intervalle de mesure est de 2 minutes. On observe un « % forecast » élevé à l'endroit où des données aberrantes consécutives sont détectées, ce qui signifie que les caractéristiques calculées à partir des données lissées correspondantes à ce « % forecast » élevé sont douteuses. En effet, on observe qu'à cet endroit les données lissées ne représentent pas bien les données brutes. Il n'y a pas eu de réinitialisation de la méthode de détection des données

aberrantes, car le nombre de données consécutives détectées comme étant aberrantes est de quatre et la réinitialisation a été configurée pour qu'elle s'effectue après cinq données. Aussi, le « runs test » détecte une corrélation significative de l'erreur entre les données lissées et mesurées attribuable à une fréquence de changement de signe de l'erreur trop élevée. Les données lissées ne représentent donc pas bien les données mesurées. Par contre, aucune faute n'est détectée avec le calcul de la pente ou du RSD.

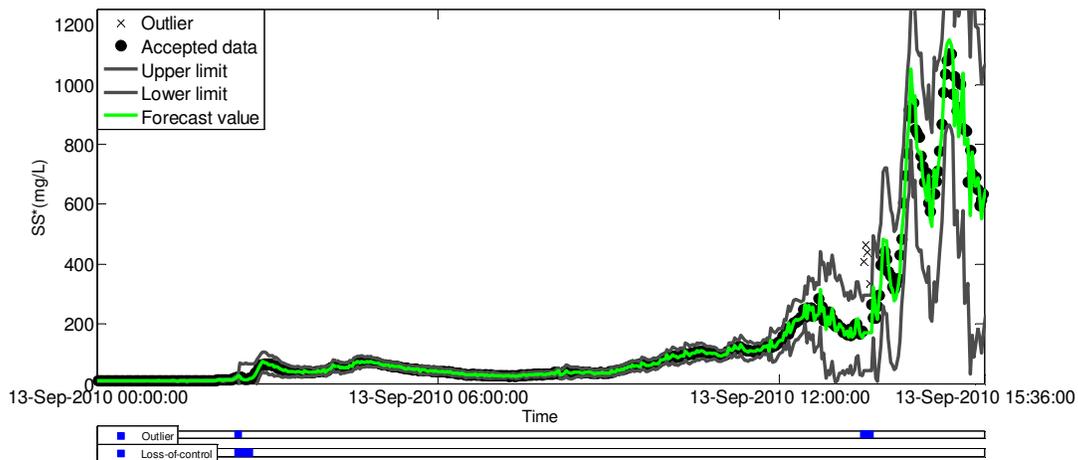


Figure 117 : Méthode de détection des données aberrantes sur les SS\* ( $\alpha=0,55$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $\Delta t=2$  min.)

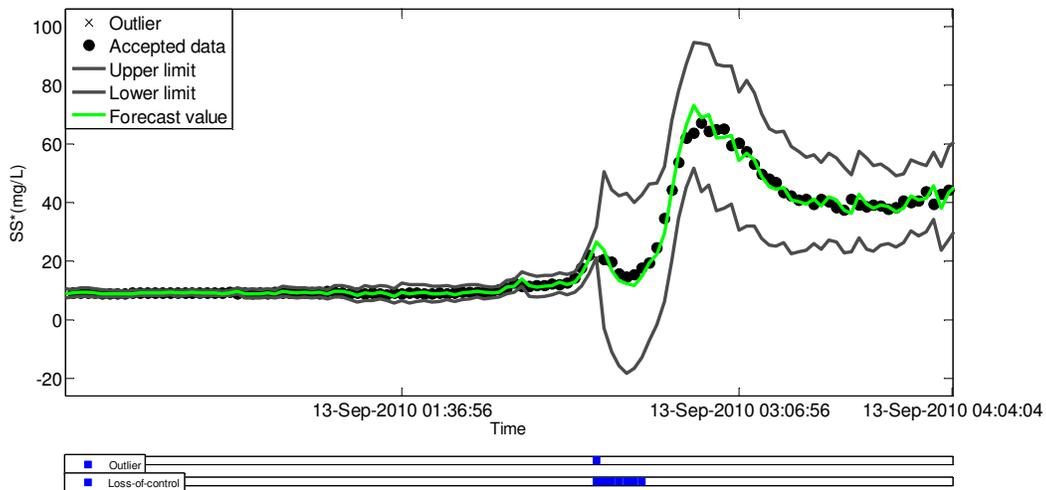


Figure 118 : Méthode de détection des données aberrantes sur les SS\*, agrandissement du « loss-of-control » de la figure 117 ( $\alpha=0,55$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $\Delta t=2$  min.)

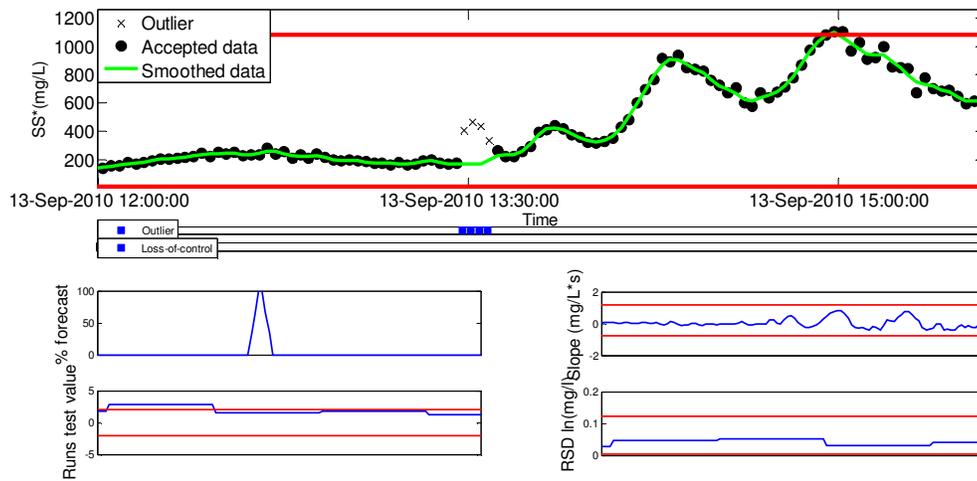


Figure 119 : Méthodes de détection de fautes sur les SS\*  
 ( $\alpha=0,55$ ;  $\alpha e=0,1$ ;  $K=5$ ;  $h=1$ ;  $\Delta t=2$  min.)

### Conductivité (120 sec.)

Les prochaines figures illustrent l'application des méthodes de détection de fautes sur une série de conductivité. Contrairement aux séries de SS\*, ce type de série n'est généralement pas influencé par la présence de données aberrantes si le capteur de conductivité demeure en place dans le milieu. La figure 120 illustre l'application de la méthode de détection des données aberrantes sur une série de conductivité dont l'intervalle de mesure est de 2 minutes. On remarque qu'il y a plusieurs « loss-of-controls ».

La figure 121 est un agrandissement de la figure 120 montrant un des « loss-of-controls ». Comme pour l'essai avec le paramètre SS\* dont l'intervalle de mesure est de 2 minutes, les « loss-of-controls » sont causés par un changement brut dans les données mesurées. De plus, l'erreur de prédiction est grande et corrélée lorsque la variation de la pente des données est rapide.

La figure 122 est un agrandissement de l'encadré de la figure 120. Cette section de la série possède un bruit extrêmement élevé. On remarque qu'une seule donnée est détectée comme étant aberrante et qu'aucun « loss-of-control » ne se produit. Ceci s'explique par le fait que le bruit dans la série augmente graduellement. L'erreur de prédiction et, donc, l'intervalle de prédiction augmentent graduellement en s'adaptant au bruit.

La figure 123 présente l'application des méthodes de détection de fautes sur la série de conductivité dont l'intervalle de mesure est de 2 minutes. On remarque que le « runs test » détecte souvent une corrélation significative de l'erreur causée par une fréquence de changement de signe de l'erreur trop élevée. Aussi, le « % forecast » est élevé pour certaines données lissées, ce qui rend les caractéristiques calculées à partir de ces données douteuses. Certaines fautes sont détectées avec des pentes croissantes et décroissantes anormalement élevées, et avec le RSD à l'endroit où le bruit est extrêmement élevé.

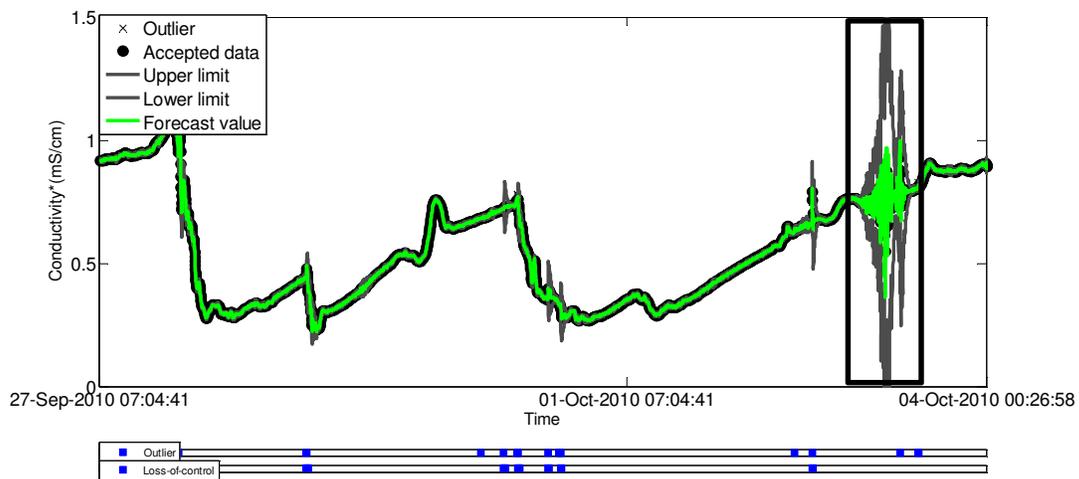


Figure 120 : Méthode de détection des données aberrantes sur la conductivité  
( $\alpha=0,6$ ;  $\alpha e=0,1$ ;  $K=5$ ;  $h=1$ ;  $\Delta t=2$  min.)

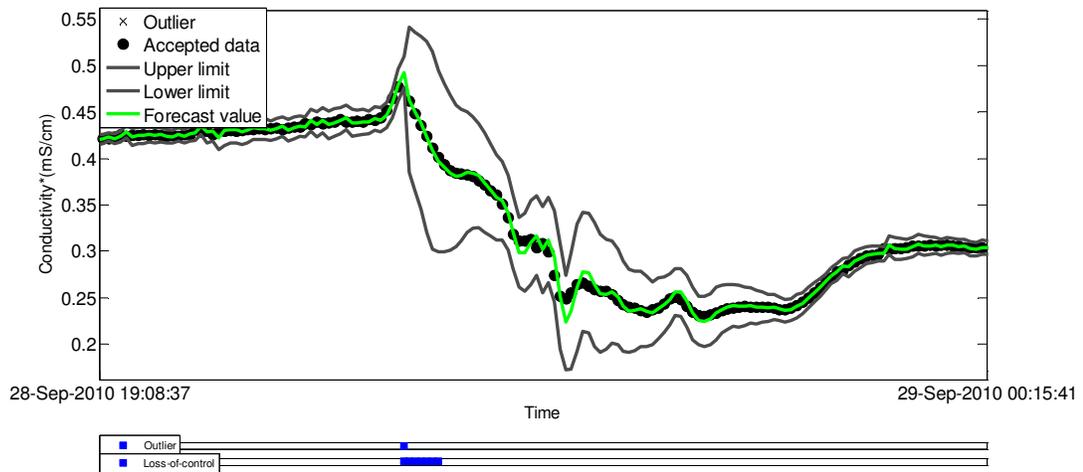


Figure 121 : Méthode de détection des données aberrantes sur la conductivité, agrandissement d'un « loss-of-control » de la figure 120 ( $\alpha_x=0,6$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $h=1$ ;  $\Delta t=2$  min.)

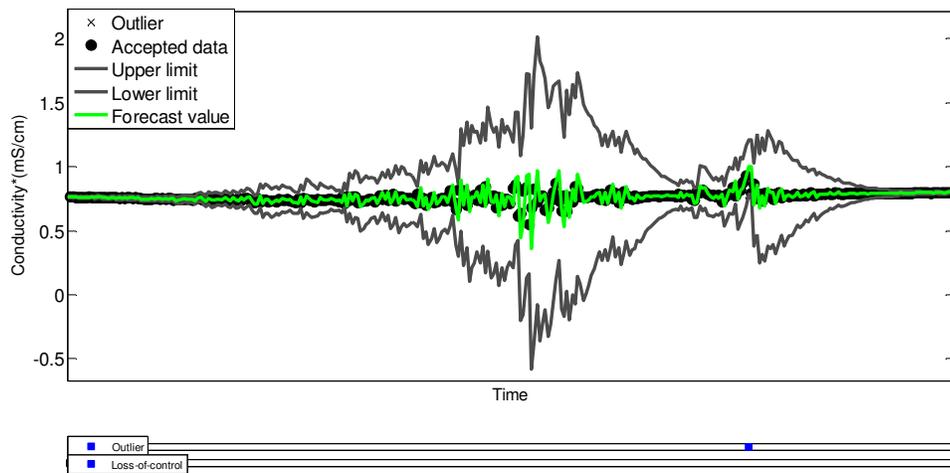


Figure 122 : Méthode de détection des données aberrantes sur la conductivité, agrandissement de l'encadré de la figure 120 ( $\alpha_x=0,6$ ;  $\alpha_e=0,1$ ;  $K=5$ ;  $h=1$ ;  $\Delta t=2$  min.)

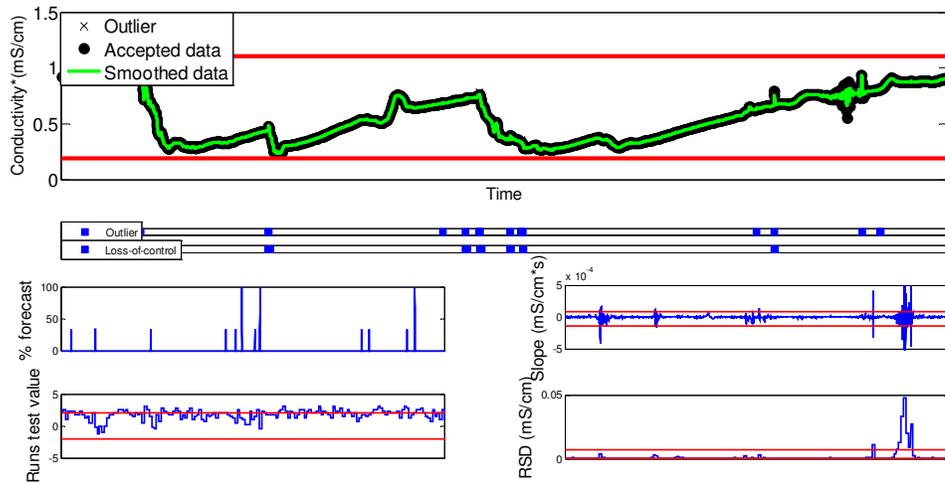


Figure 123 : Méthode de détection des fautes sur une série de conductivité  
 ( $\alpha x=0,6$ ;  $\alpha e=0,1$ ;  $K=5$ ;  $h=1$ ;  $\Delta t=2$  min.)

Les essais précédents ont démontré que les méthodes de détection de fautes fonctionnent relativement bien pour les séries de SS\* et de conductivité possédant un intervalle de mesure de 2 minutes. Cependant, tel que discuté à la section 7.1 Importance de la fréquence de mesure, leurs performances sont grandement améliorées si la fréquence de mesure est plus élevée.