

J. Bernier<sup>1</sup>, C. Cornut<sup>1</sup>, A. Hicks<sup>1</sup>, P.A. Vanrolleghem<sup>2</sup>, V. Rocher<sup>1</sup>

1 SIAAP – Direction Innovation et Environnement – 82 avenue Kléber – 92700 Colombes  
2 Groupe modelEAU, Université Laval – Département de génie civil et de génie des eaux, 1065, av. de la Médecine, Québec, QC, Canada, G1V 0A6

### Contexte scientifique et technique

#### Multiplication des capteurs/analyseurs in-situ en station d'épuration, dans le réseau d'assainissement et dans le milieu naturel

- Connaissance de l'état du système en tout temps
- Prise de décision sur l'exploitation des systèmes accélérée
- Meilleure caractérisation / compréhension des dynamiques rapides
- Augmentation des possibilités de contrôle de procédé
- Alimentation de modèles d'aide à la décision en temps réel
- Pourtant...



#### Utilisation encore relativement faible des données récoltées

Les capteurs sont soumis à plusieurs facteurs pouvant affecter la qualité de l'information remontée :

- Encrassement temporaire ou durable des organes de mesure par la matière organique, les particules ou les bactéries présentes dans le milieu de mesure
- Problèmes logiciels ou d'alimentation électrique
- Vieillesse des organes de mesure



Ces capteurs génèrent un volume important de données qu'il est très difficile de traiter manuellement, d'où le besoin d'outils pour en faciliter la qualification

### Algorithmes de qualification évalués

#### 1. Construction d'un modèle prédictif simple (exemple : lissage exponentiel double, DES)

$$S_t = \alpha y_t + (1 - \alpha)(S_{t-1} + b_{t-1})$$

$$b_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)b_{t-1}$$

$$F_{t+1} = S_t + b_t$$

S = variable lissée  
y = valeur de la variable mesurée  
b = tendance linéaire lissée  
F = prédiction du modèle  
 $\alpha, \gamma$  = coefficients de lissage

#### 2. Calcul des intervalles de confiance du modèle

$$e_T = F_t - y_t$$

$$\hat{\Delta}_{t+1} = \eta |e_T| + (1 - \eta)\hat{\Delta}_t$$

$$F_{t+1} \pm 1.25L\hat{\Delta}_{t+1}$$

e = erreur de prédiction  
 $\Delta$  = erreur de prédiction lissée  
L = coefficient multiplicateur de l'erreur  
 $\eta$  = coefficient de lissage

#### 3. Comparaison de la nouvelle mesure et qualification

$$y_{t+1} \in F_{t+1} \pm 1.25L\hat{\Delta}_{t+1} \rightarrow OK$$

*Sinon, donnée douteuse*

#### 4. Traitement optionnel

- Lissage des données
- Remise à zéro du modèle si besoin
- Passage supplémentaire de l'algorithme en ordre chronologique inverse
- Application de seuils supplémentaires

### Exemple de résultats d'application en réseau d'assainissement

#### La station de mesure pilote du poste de Clichy (Hauts-de-Seine)

##### Situation géographique



##### Méthode d'application utilisée

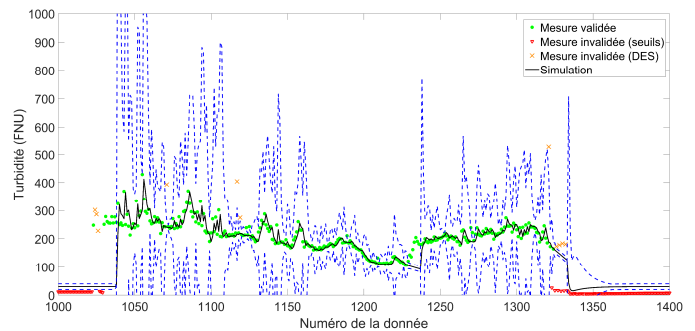
- Identification d'une période « propre » pour les données de chaque sonde, hors pH
- Construction de quatre types de modèles prédictifs simples : lissage double exponentiel, lissage triple exponentiel (2 variantes) et réseau de neurone linéaire simple couche
- Calibration des paramètres de lissage de l'erreur de prédiction et du calcul des intervalles de confiance sur une période d'un mois comprenant relativement peu d'erreurs
- Validation successive sur deux périodes d'un mois comprenant progressivement plus d'erreurs :
  - a) Validation quantitative (% de détection, faux négatifs et faux positifs)
  - b) Validation qualitative (inspection visuelle des données invalidées)

##### Équipements installés

Paramètres	Appareils (Hach Lange)	Fréquence	Gammaes
NH <sub>4</sub> <sup>+</sup>	AMTAX	15 min	5 à 100 mg N/L
Conductivité	CDC401	5 min	0,01 µS/cm à 200,0 mS/cm
pH	pH SC	5 min	-
Turbidité	Solitax SC	5 min	0,001-4000 FNU
UV	UVAS	5 min	0-1 500 m <sup>-1</sup> (2 mm)

#### Exemple de résultats typiques de qualification des données

##### Exemple d'application sur la turbidité (modèle DES)



- Utilisation d'un seuil minimal de 30 FNU (étude statistique des séries de données) et application de l'algorithme en sens chronologique et inversé (affichage des intervalles en sens chronologique uniquement)
- Quelques données aberrantes bien détectées
- L'application dans les deux sens permet d'éviter des faux négatifs (voir données ~1230)
- Présence de quelques faux négatifs en début et fin de période, due à l'alternance entre donnée d'apparence valide et donnée située sous le seuil minimal (valeur nulle, rouge)
- Sévérité de l'algorithme à calibrer et adapter en fonction de la sévérité des utilisateurs

### Travaux R&D prévus (Mocopee – Phase II)

- Application plus large sur les séries de données de capteurs au SIAAP
- Ajouts de tests de comparaison des propriétés des séries de données dans la chaîne de validation
- Travail d'application de méthodes plus avancées de détection des dérives de capteurs
- Travail de développement de méthodes d'auto-calibration des paramètres de sévérité des algorithmes