

# Avancées en assimilation de données : développements méthodologiques et comparaisons de méthodes

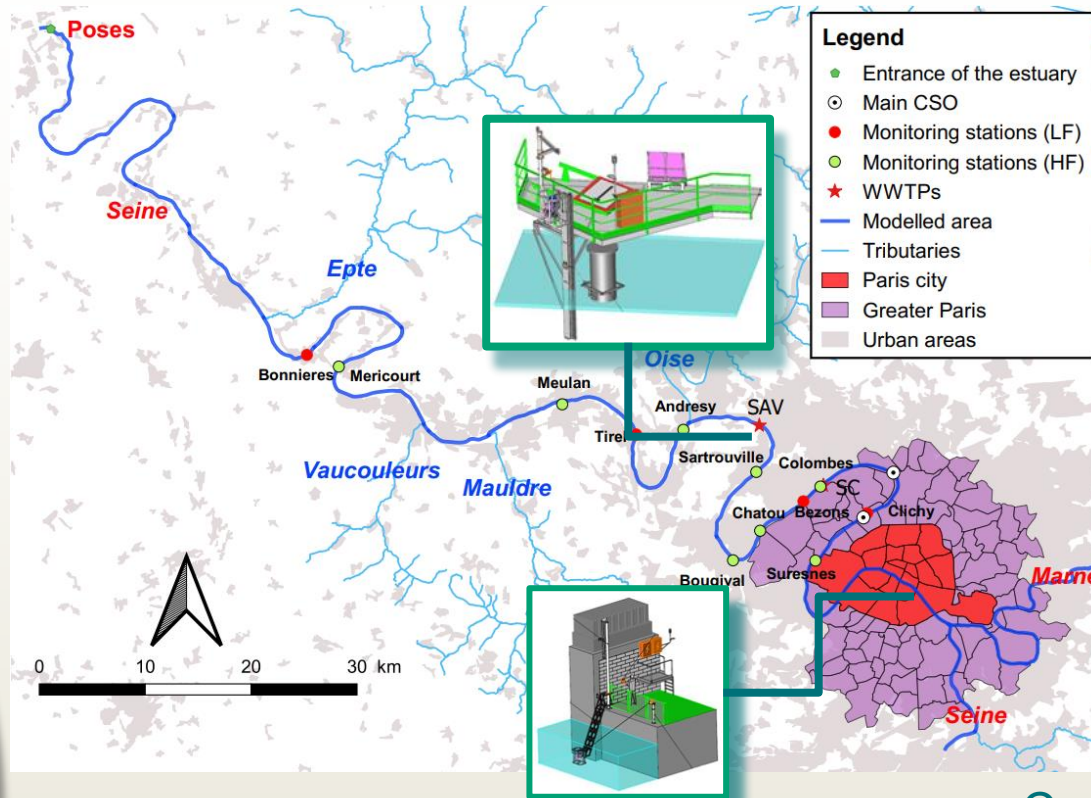
*Masihullah Hasanyar<sup>1</sup>, Shuaitao Wang<sup>2</sup>, Thomas Romary<sup>1</sup>,  
Nicolas Flipo<sup>1</sup>*

1

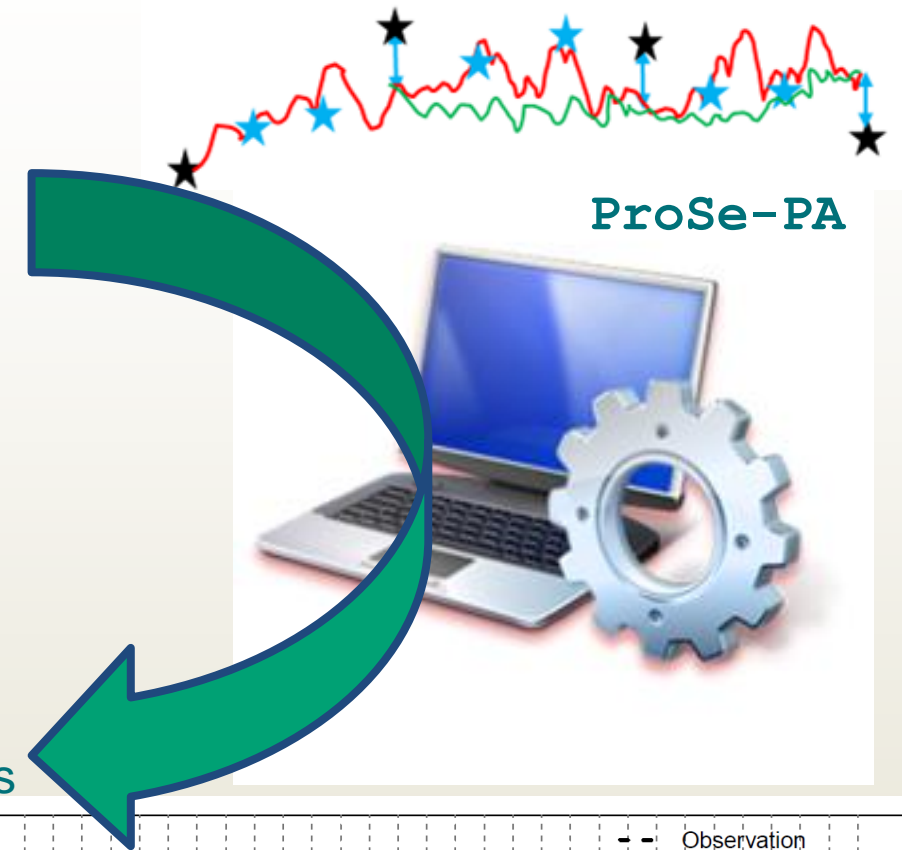
2

# Synergie mesures haute fréquence - modèle

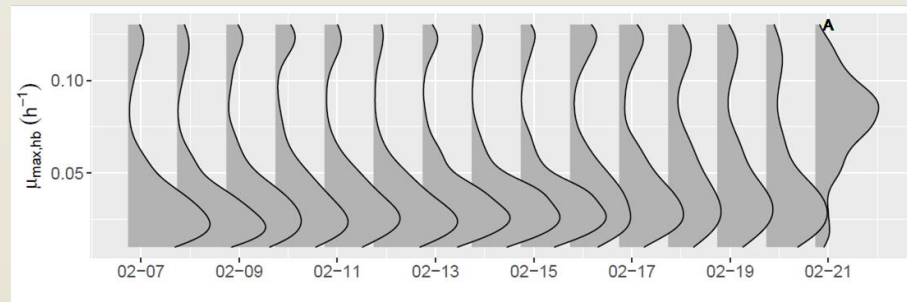
## Suivi du milieu



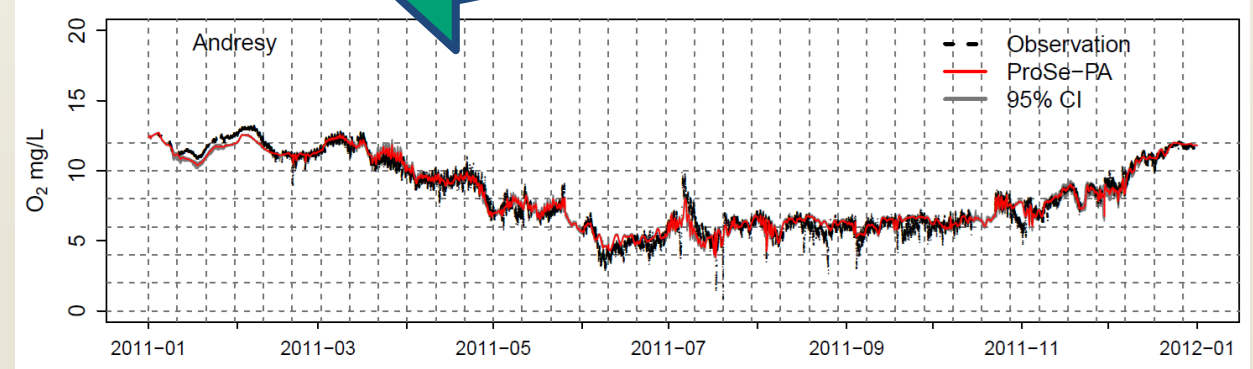
## Assimilation de données



## Paramètres physiques et physiologiques



## Concentrations



# Etat des lieux de la modélisation hydro-biogéochimique au PIREN Seine

- + Première implémentation d'un filtre particulaire (PF) dans le logiciel ProSe-PA, (Wang et al., 2019)
  - + Originalité par rapport à communauté internationale qui privilégie aujourd'hui le filtre EnKF (Ensemble Kalman Filter), (Cho et al., 2020 ; Wang et al., 2022)

→ Quel filtre utiliser en modélisation de la qualité de l'eau ?

- + Simulation de l'oxygène dissous très précise (Wang et al., 2022)
  - + en hiver à basse température (réoxygénation à l'interface air-eau domine, problème plutôt linéaire),
  - + pendant les efflorescences algales (dynamique des communautés de phytoplancton, problème non linéaire)
- + Des problèmes demeurent lors des étiages (sur-estimations pouvant atteindre 1 mgO<sub>2</sub>/L) (Wang et al., 2022)
  - + Nécessité de mieux caractériser la matière organique rapidement biodégradable
    - Peut-on quantifier la biodégradabilité de la matière organique à l'aide de l'assimilation de données ?

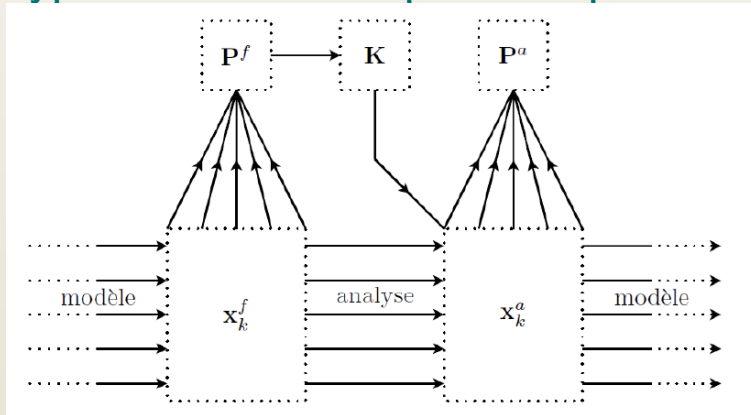
# Comparaison des filtres : Filtre de Kalman d'ensemble (EnKF) vs. Filtre Particulaire (PF)

Corriger à chaque pas d'assimilation les distributions statistiques des paramètres →  
Théorème de Bayes

$$f_{\mathbf{Z}}(\mathbf{z}|\mathbf{y}^*) \propto f_{\mathbf{Y}^*}(\mathbf{y}^*|\mathbf{z})f_{\mathbf{Z}}(\mathbf{z})$$

## EnKF (Evensen et al., 1994)

- Basé sur le filtre de Kalman (Kalman, 1960)
- Hypothèse gaussienne pour les paramètres
- Hypothèse linéaire pour les prédictions

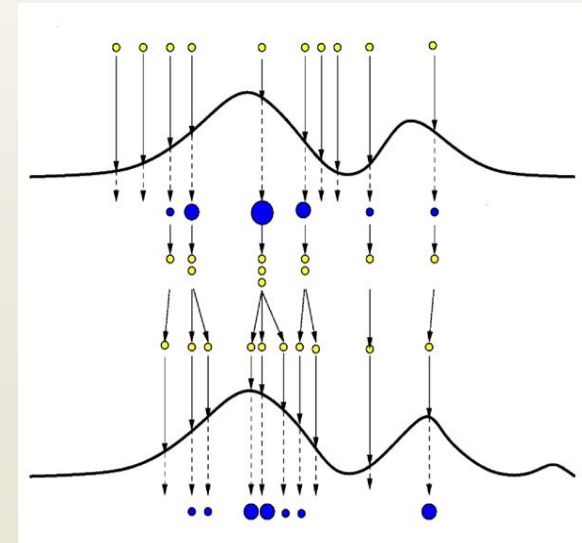


Daget, (2007)

Distribution a posteriori gaussienne

## PF

- Pas d'hypothèse gaussienne et non linéarité possible
- Particules pondérées par des poids



Doucet et al., (2001)

Distribution a posteriori multi-modal possible

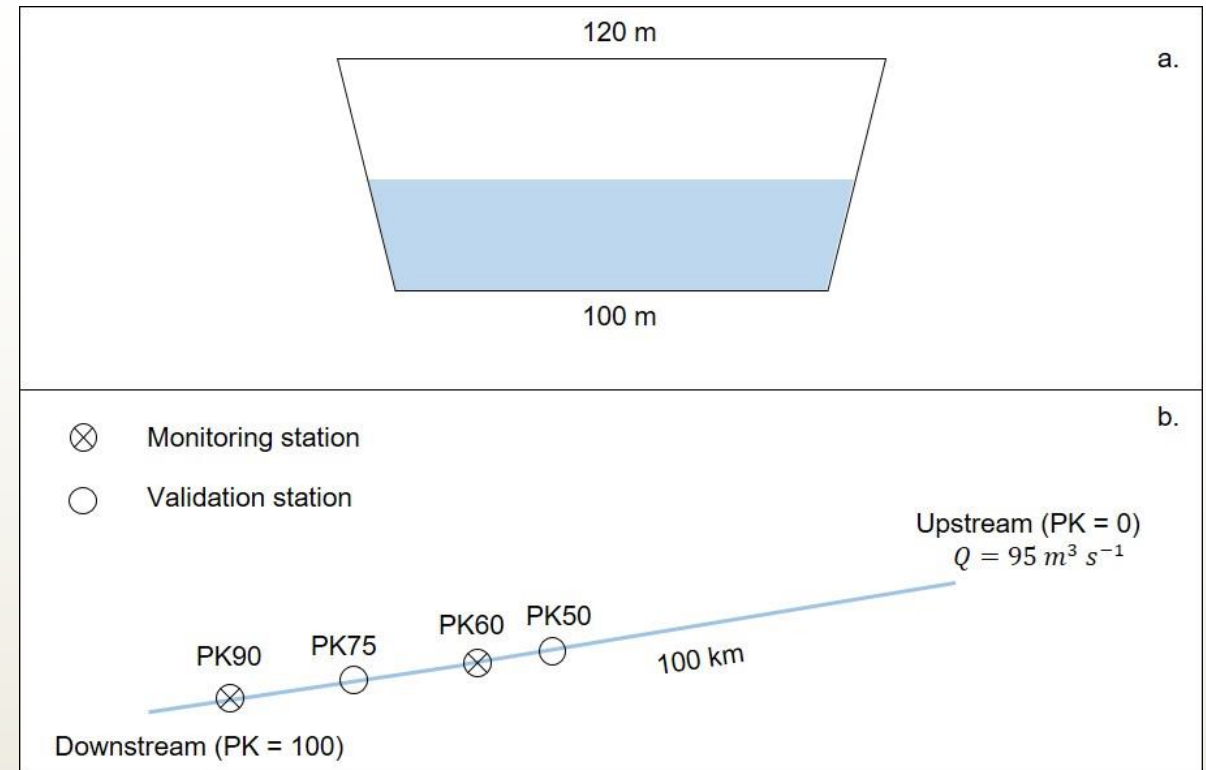
# Quel filtre utiliser ? Description des cas tests

## Géométrie et stations

- Un tronçon de 100 km
- 2 stations d'assimilation (PK60 et PK90)
- 2 stations de validation (PK50 et PK75)

## Deux scénarios

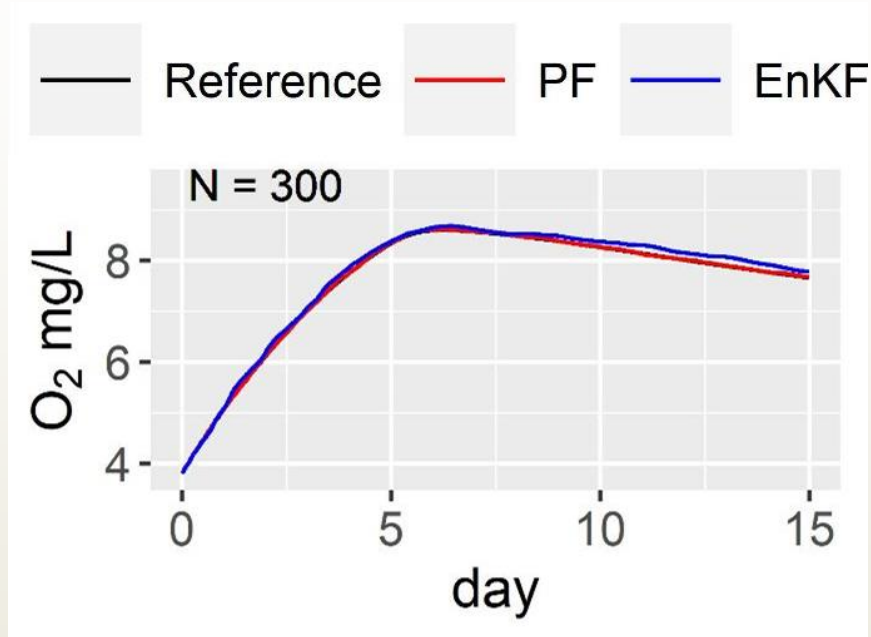
- Problème linéaire (réoxygénation)  
 $K_{\text{navig}}$  (coefficient de réaération liée à la navigation)
- Problème non linéaire (dégradation matière organique)  
 $\mu_{\text{max}}$  (taux de croissance) et  $Y$  (rendement de croissance)



Wang et al., (soumis)

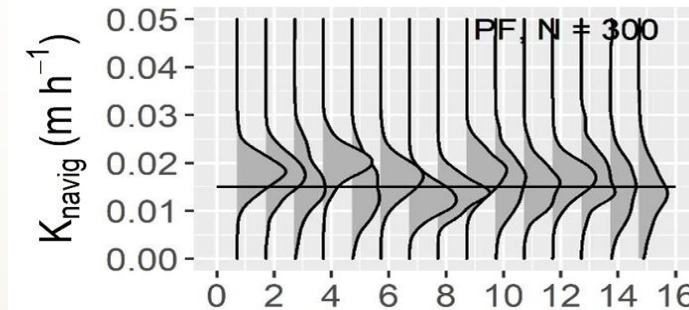
# Problème linéaire (réoxygénation à l'interface air-eau)

Concentrations à la station de validation

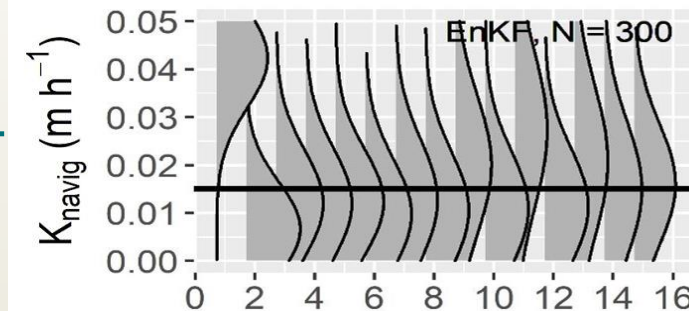


Distributions a posteriori journalières,  $K_{\text{navig}}$

PF



EnKF



PF et EnKF permettent de bien simuler l'oxygène avec une précision un peu meilleure pour le PF

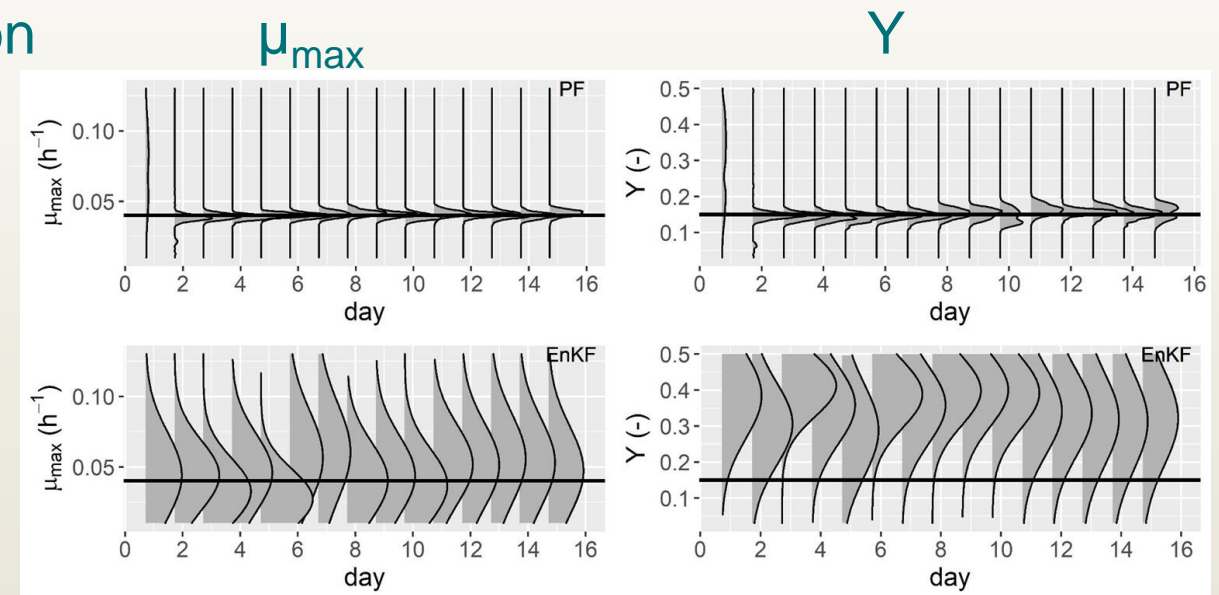
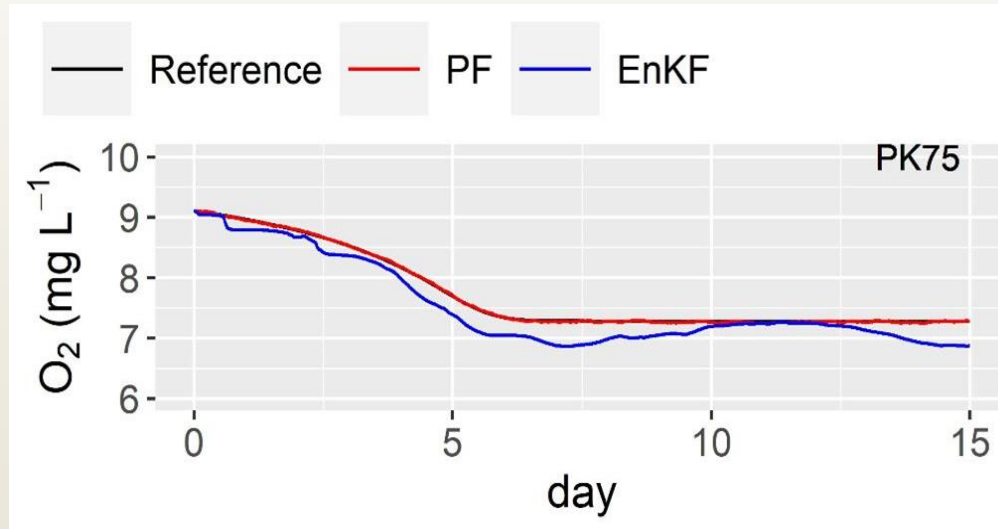
Wang et al., (soumis)

# Problème non linéaire (dégradation de la matière organique)

++ PF : Très bonnes performances (oxygène dissous et propriétés des micro-organismes)

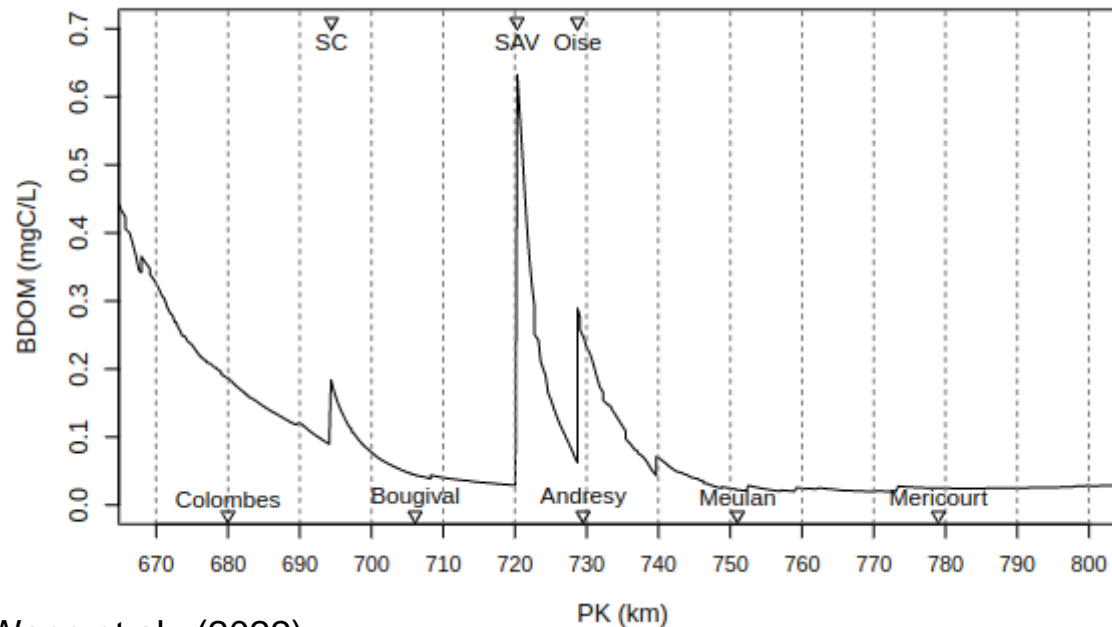
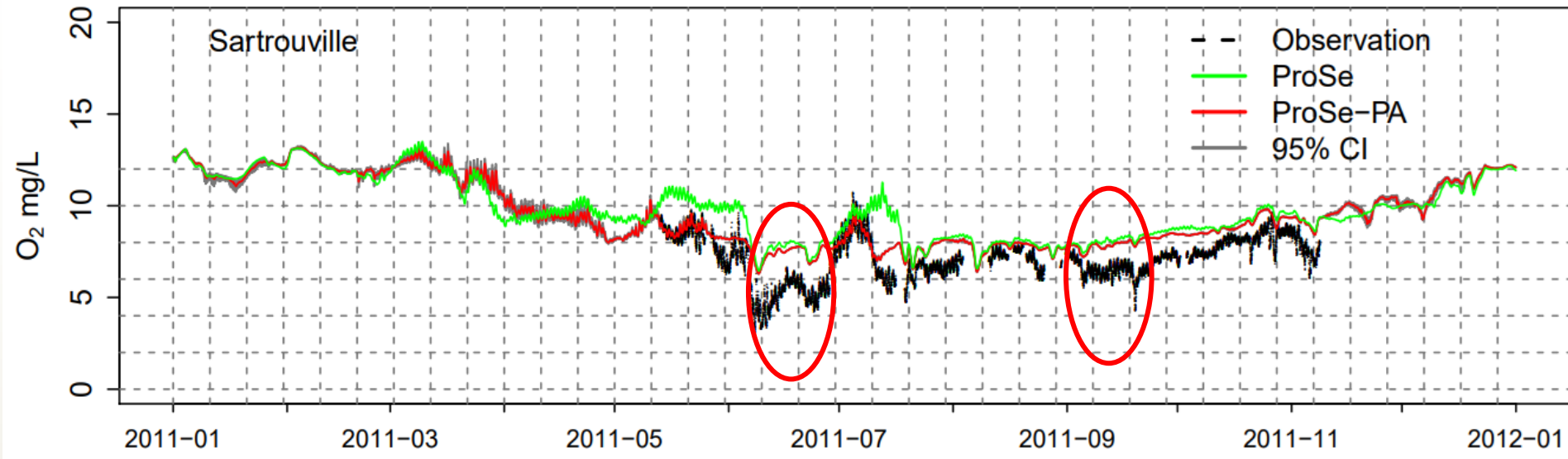
- EnKF : Des erreurs non négligeables sur l'oxygène (jusqu'à 0.5 mgO<sub>2</sub>/L) associées à une mauvaise caractérisation physiologique

## Concentrations à la station de validation



Wang et al., (soumis)

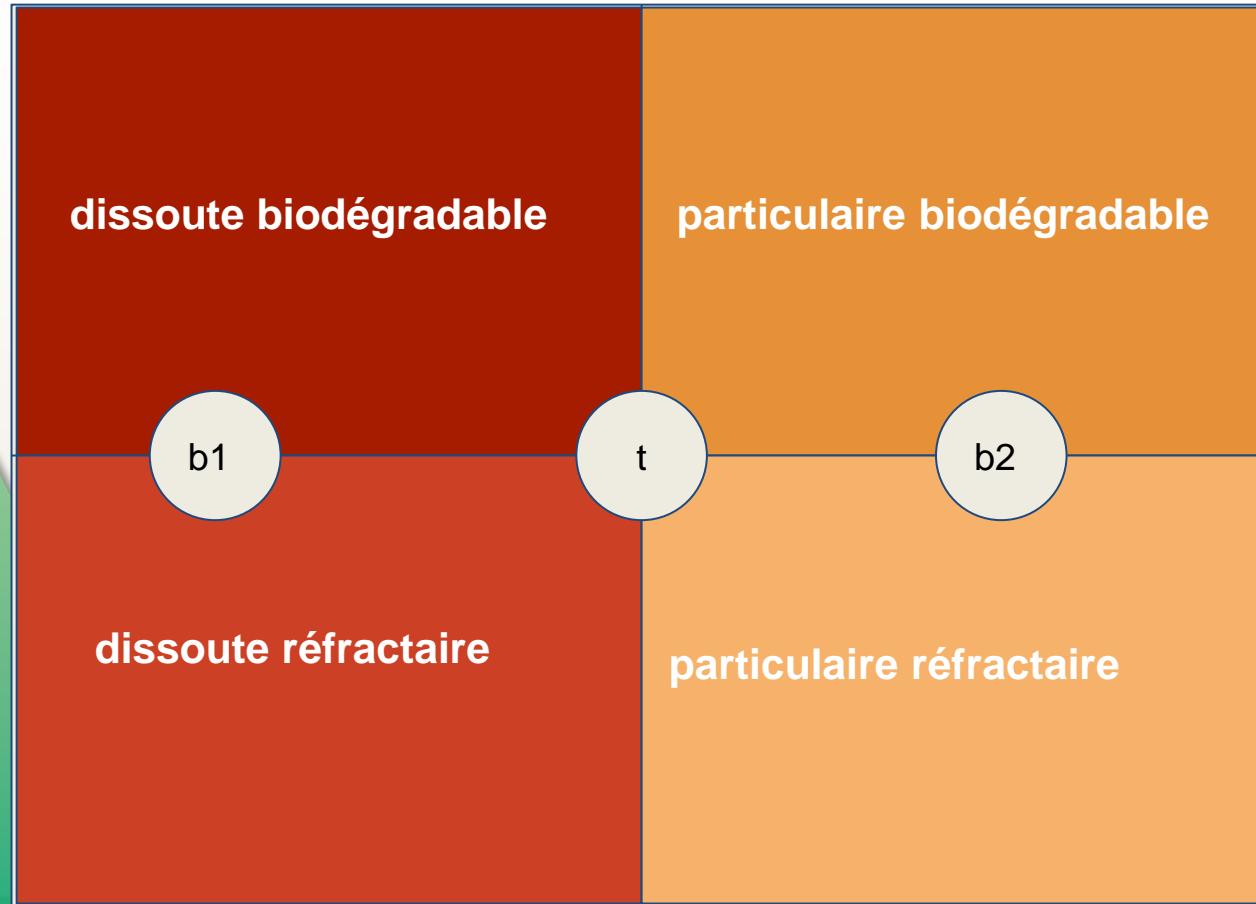
# Écart entre O<sub>2</sub> simulée et observée à l'été



Comment représenter explicitement le MO dans ProSe-PA?

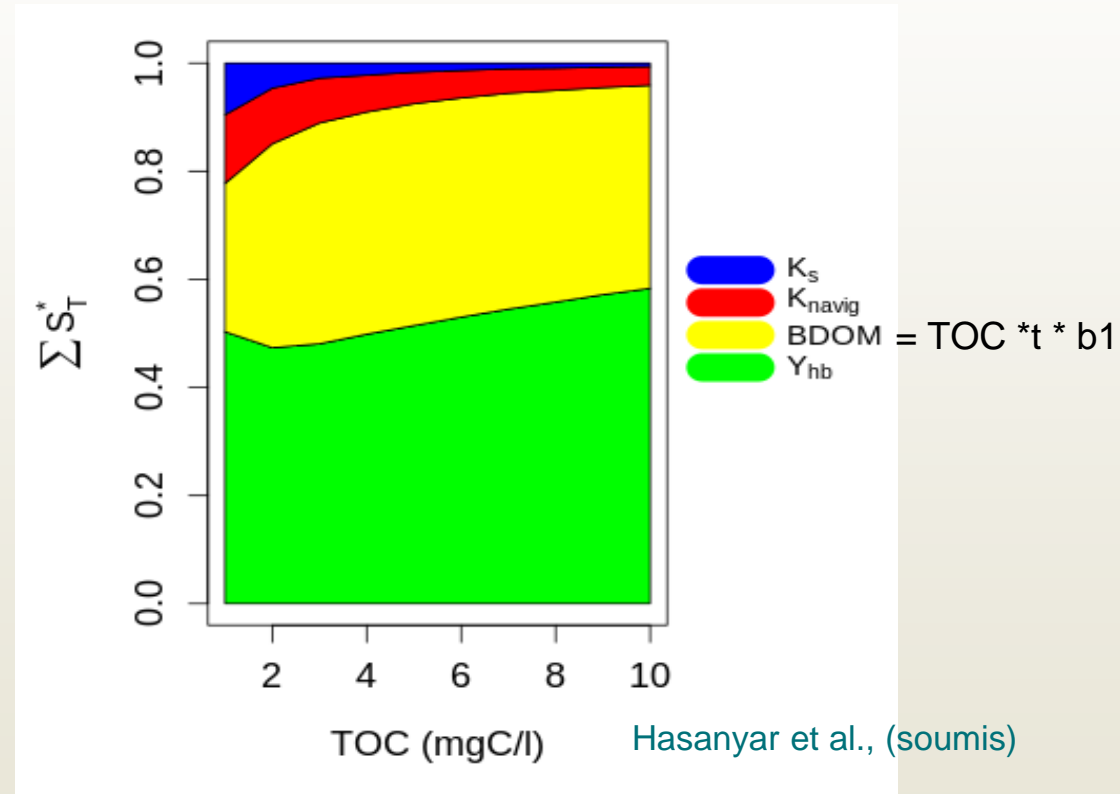
- Nécessité de paramétrer la répartition de la MO

# Les paramètres représentant la répartition de la MO et analyse de sensibilité



Modèle de répartition de la matière organique

- la croissance bactérienne, la mortalité bactérienne, et le rendement bactérien sont les paramètres les plus influents



- pour les fortes croissances nettes bactériennes : la fraction de BDOM est également influente

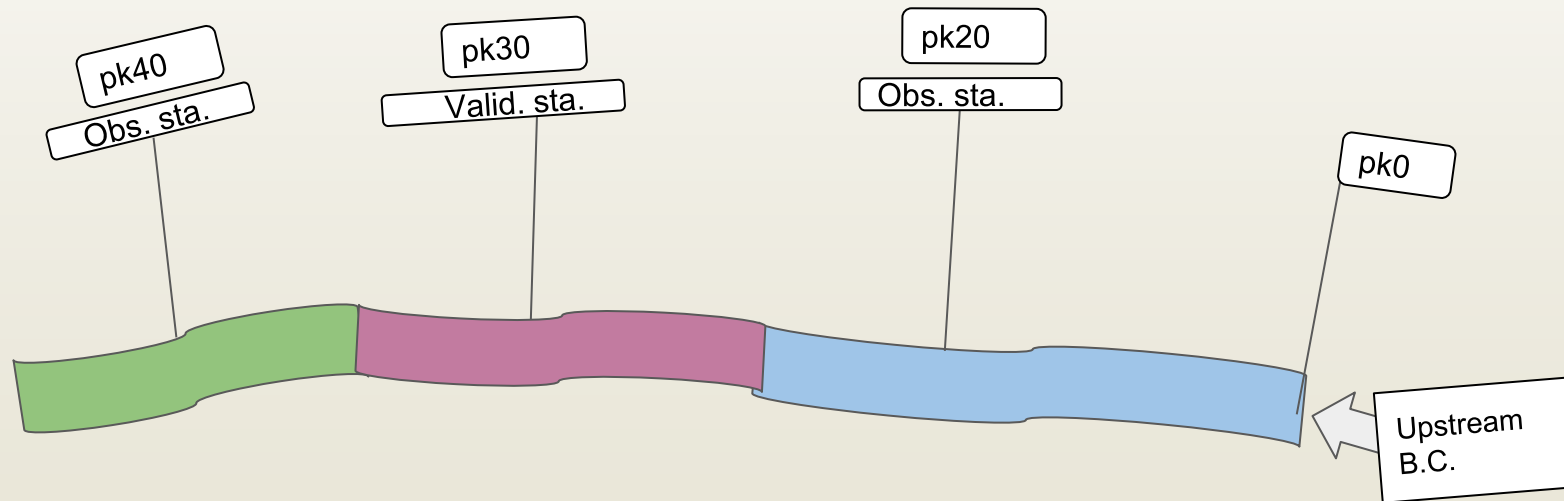
# Intégration d'un modèle de matière organique dans ProSe-PA

1. Modèle de répartition de la MO pour les conditions limites

1. Intégration de la biodégradabilité de la MO (b1) dans les paramètres ajustés par ProSe-PA

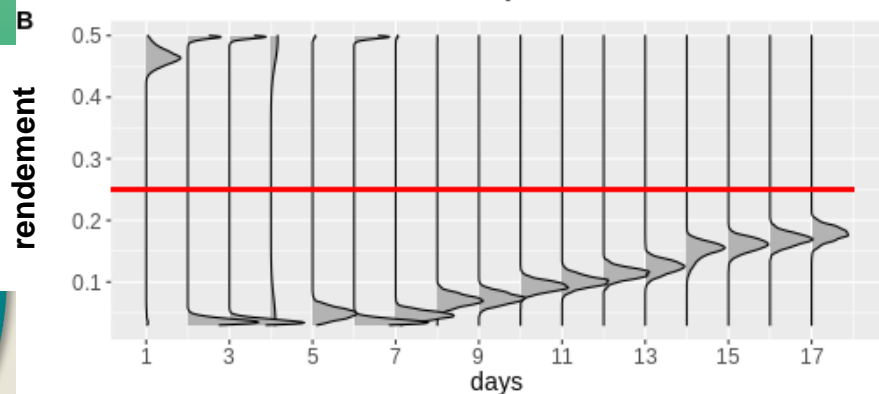
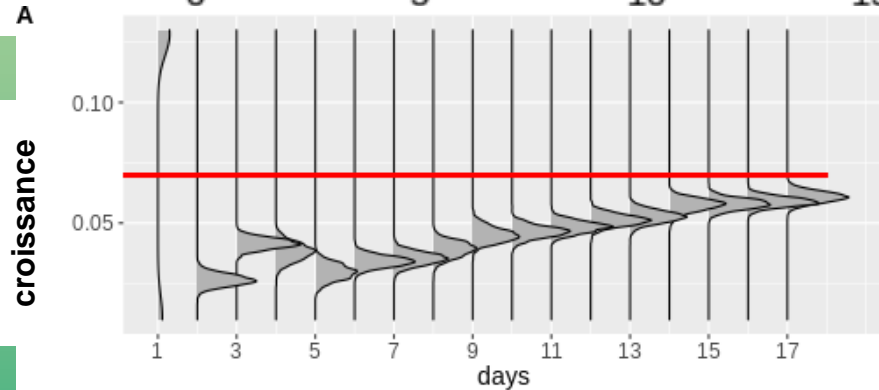
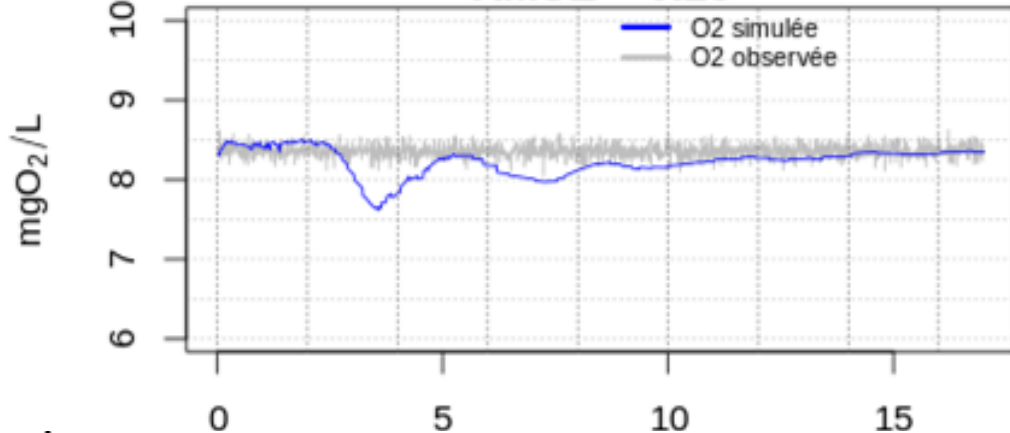
$$\text{BDOM} = b1 * \text{DOM}$$

1. Création d'un cas d'étude



# La biodégradabilité de la MO peut biaiser l'estimation

RMSE = 0.23



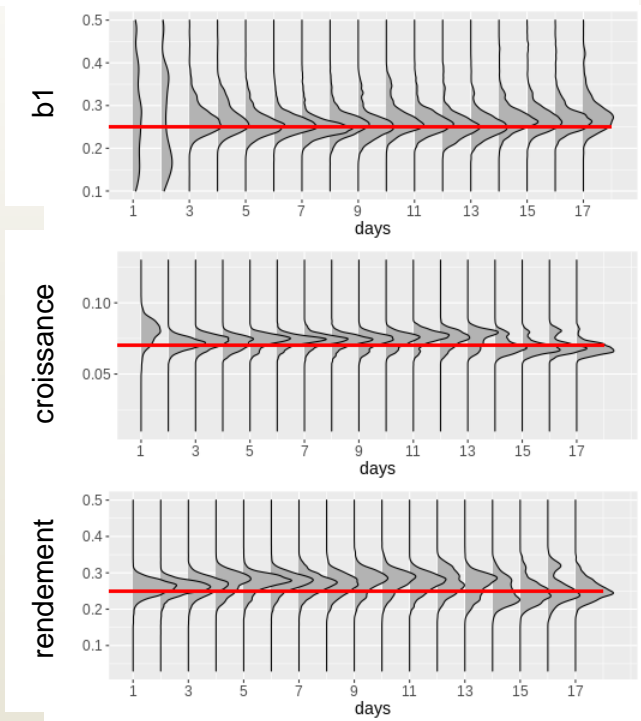
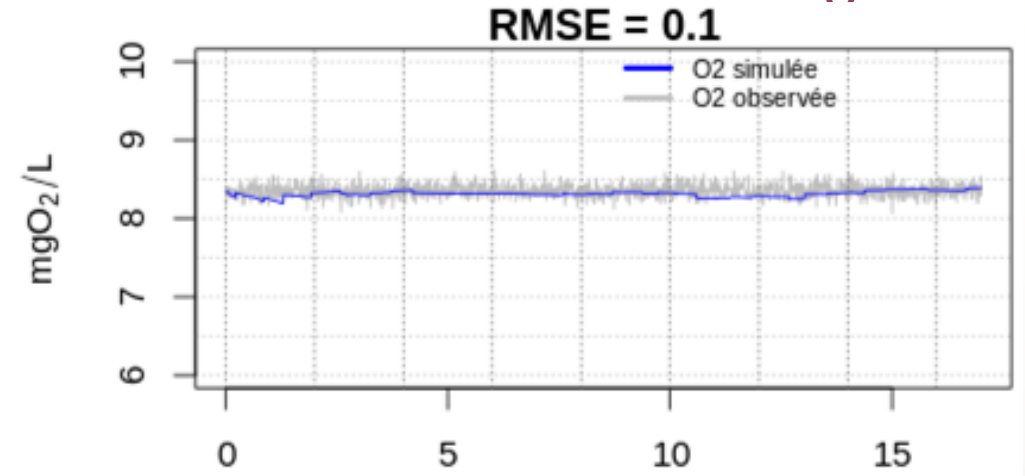
## Estimation des paramètres croissance et rendement :

1. simulation direct avec  $b1 = 0.35$  et valeurs ciblées de la croissance et du rendement
2. perturbation des observation avec erreur gaussien 1%
3. on fixe  $b1=0.25$  et on lance ProSe-PA

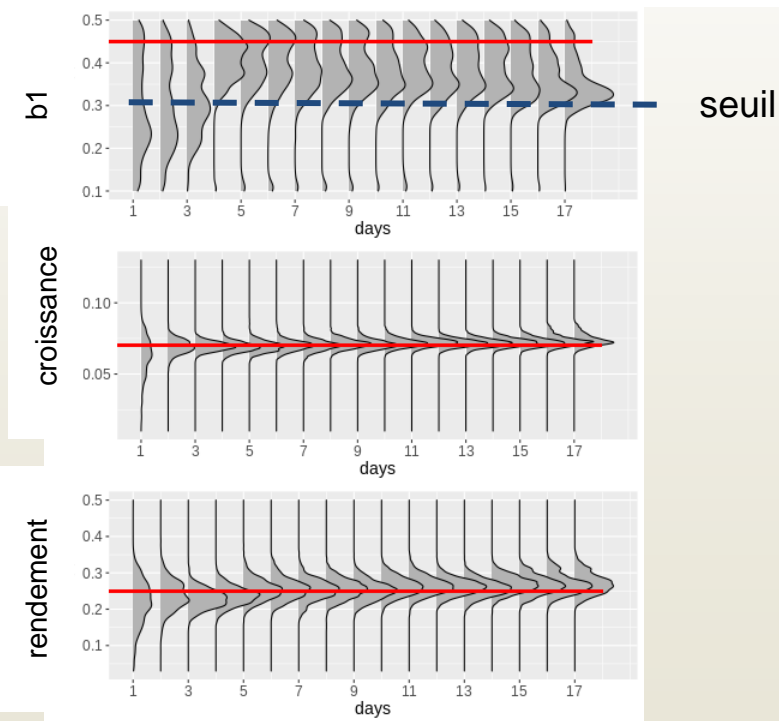
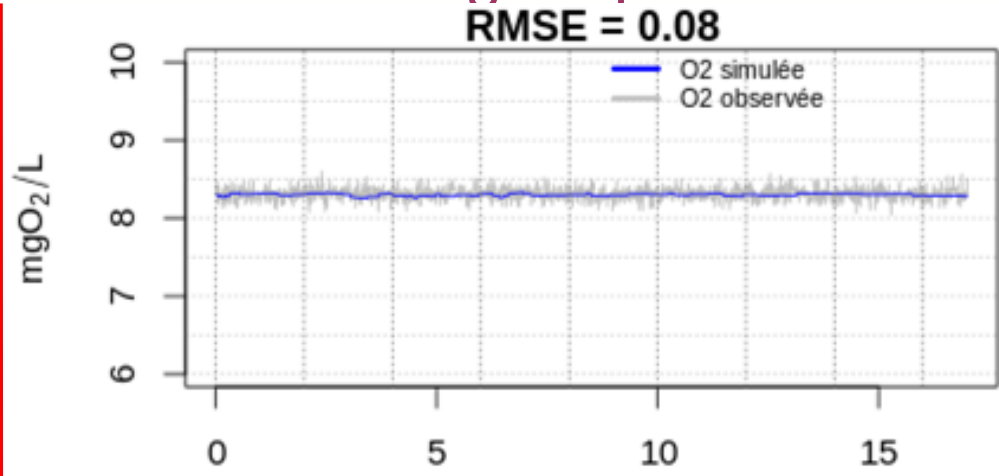
Résultats : à  $b1=0.25$  la MO est limitante, et le PF ne peut pas identifier les bonnes valeurs des paramètres de croissance

→ Nécessaire d'inclure  $b1$  dans le filtre

# Caractérisation de la biodégradabilité de matière organique dissoute



**BDOM limitante**



**BDOM non-limitante**

**BDOM**

# Conclusion

- ❖ PF est le bon choix méthodologique pour mettre en synergie mesures et modèles de qualité de l'eau
- ❖ ProSe-PA est capable d'identifier les paramètres physiologiques des bactéries hétérotrophes et de caractériser la biodégradabilité du carbone organique dissous dans les grandes rivières
  - exactement pour les milieux appauvris en BDOM
  - la valeur du seuil de limitation pour les milieux riches en BDOM
  - Perspective: Où positionner une station d'assimilation ?

Merci de votre attention

