Modélisation des eaux souterraines dans la plaine alluviale de la Bassée : un outil pour comprendre la dynamique d'un socio-hydrosystème complexe

Clémence Saias[1], Adrien Renaud[1], Nicolas Flipo[2], Anne Jost[1*]

[1] Sorbonne Université, CNRS, EPHE, Metis, F-75005 Paris

[2] Mines Paris-PSL, Centre de Géosciences, F-77305 Fontainebleau

[*] anne.jost@sorbonne-universite.fr

Résumé

Situé au cœur du bassin versant de la Seine et au sein du réseau Zone Atelier Seine, l'Observatoire de la Bassée est une plate-forme de recherche essentielle pour comprendre les processus hydrologiques liés au défi stratégique d'une gestion durable des ressources en eau. L'observatoire se concentre sur le comportement de la plaine alluviale en tant qu'hydrosystème complexe et influencé par l'homme, en tenant compte de son évolution géohistorique de long terme. Grâce à son vaste réseau de stations de surveillance des eaux de surface et souterraines, l'observatoire met en évidence le rôle central des eaux souterraines et de leurs interactions avec les eaux de surface dans la dynamique actuelle de cette région. Nous présentons les développements récents du modèle hydrogéologique de la Bassée, qui combine la plate-forme de modélisation distribuée CaWaQS avec les utilitaires d'estimation des paramètres proposés dans la suite PEST. Cette intégration améliore la représentation de l'hétérogénéité de la plaine alluviale et fournira une base solide pour une prise de décision éclairée. Le modèle est conçu pour aider les parties prenantes à relever les défis liés à l'exploitation et à la conservation de la plaine alluviale dans un environnement en mutation.

Points clés

- *Modélisation hydrogéologique avancée et intégrée (CaWaQS + PEST)*
- *Estimation non linéaire des paramètres pour une meilleure représentation de l'hétérogénéité du système alluvial*
- *Aide à la décision pour la gestion de l'eau et les acteurs locaux*

Abstract

Located in the heart of the Seine catchment and part of the Zone Atelier Seine network,the Bassée Observatory is an essential research platform for understanding the hydrological processes associated with the strategic challenge of sustainable water resource management. The observatory focuses on the behaviour of the alluvial plain as a complex and human-influenced hydrosystem, considering its long-term geohistorical evolution. Through its extensive network of surface water and groundwater monitoring stations, the observatory highlights the central role of groundwater and its interactions with surface water in the current dynamics of this region. We present recent developments in the Bassée groundwater model, which combine the CaWaQS hydrogeological platform with the groundwater utilities of the PEST parameter estimation approach. This integration enhances the representation of the alluvial plain's heterogeneity and will provide a solid basis for informed decision making. The model is designed to assist stakeholders in addressing the challenges of operating and conserving the alluvial plain in a changing environment.

Key points

- Advanced groundwater modelling integration (CaWaQS + PEST)
- Nonlinear parameter estimation for an improved representation of alluvial system heterogeneity
- Decision support for water management and local stakeholders

Introduction

L'exposition croissante des territoires aux impacts directs et indirects des perturbations environnementales, en particulier climatiques, constitue aujourd'hui un enjeu majeur dans la planification des politiques publiques. Aux effets directs, par exemple la multiplication des épisodes de sécheresse, s'ajoutent des effets indirects, souvent plus difficiles à anticiper, tels que le renforcement des tensions sur les ressources en eau, impliquant des décisions d'arbitrage entre les différents usages. Le PIREN-Seine s'est ainsi donné pour tâche, au sein de l'axe 1 de son nouveau programme de phase, de documenter la vulnérabilité du bassin de la Seine et de ses territoires face aux événements exceptionnels, d'une part en caractérisant l'aléa, et d'autre part en s'interrogeant sur les modalités d'action en réponse à ces manifestations du changement climatique. Dans ce cadre, et pour ce qui relève de la disponibilité des ressources en eau, les modèles s'avèrent particulièrement utiles pour éclairer les débats et aider à la prise de décisions. En simulant le comportement des nappes et cours d'eau en réaction aux sollicitations environnementales et anthropiques, ils permettent de tester différentes hypothèses de fonctionnement, d'en évaluer les conséquences concrètes et ainsi, d'apporter des arguments en faveur de solutions acceptables par tous.

Les modèles hydrogéologiques combinent des équations mathématiques paramétrées pour représenter de façon simplifiée le comportement des systèmes aquifères. Lorsqu'ils sont à base physique, les paramètres en jeu ont une signification hydrodynamique et peuvent être mesurés sur le terrain à l'aide d'essais de pompage. Ces essais étant cependant coûteux à mettre en place, les données exploitables sont rares. Les niveaux piézométriques sont en revanche plus faciles à observer et servent alors utilement à ajuster les valeurs des paramètres, en inversant le problème posé. Lorsque le modèle reproduit fidèlement ce qui a été observé sur le

terrain, il peut alors être utilisé pour apporter des réponses quantitatives aux questions formulées en matière de gestion de l'eau, en y associant un degré de certitude conditionnant l'usage opérationnel du modèle.

L'objectif concret de l'étude ici présentée est de faire progresser à de telles fins l'ajustement des paramètres hydrodynamiques du modèle hydrogéologique de la plaine de la Bassée, dans le val de Seine amont, l'un des territoires phares du PIREN-Seine. Ce modèle est composé de deux aquifères, chacun situé dans l'une des deux principales formations géologiques, le substrat calcaire et les alluvions de la Seine, et couvre la totalité de la plaine alluviale à l'affleurement entre les deux confluences de la Seine. Il prend en compte l'essentiel des forçages anthropiques actuels, gravières et prélèvements notamment, dont les impacts ont pu être évalués dans un exercice de modélisation précédent (Jost et al., 2024). Ce modèle a été développé à l'aide de la plate-forme de modélisation CaWaQS (Flipo et al., 2023), outil de base du PIREN-Seine.

Pour un modèle complexe comme ce dernier, l'utilisation d'algorithmes d'optimisation s'avère nécessaire dès lors qu'il s'agit d'étalonner ses paramètres hydrodynamiques. La suite PEST (*Model-Independent Parameter Estimation and Uncertainty Analysis*, Doherty, 2015) fournit de tels outils. Comme son nom complet l'indique, PEST est indépendante du modèle : avec une interface appropriée, elle peut être couplée à n'importe quel modèle hydrogéologique. L'interface avec la plate-forme de modélisation MODFLOW (Langevin et al., 2022) est par exemple déjà bien documentée, de même qu'avec la plate-forme MARTHE (Aissat, 2021). Suivant ces exemples, l'objectif de ce travail a été prioritairement de construire une telle interface entre CaWaQS et PEST, dénommée pyWaQS. Après une présentation rappelant les fondamentaux des outils qui ont été combinés, soit le modèle CaWaQS et les principes de l'algorithme utilisé dans PEST, ce rapport décrit l'interface développée entre PEST et CaWaQS (partie 1). Cette interface a été testée et analysée sur un cas test (partie 2), avant d'être appliquée au modèle de la plaine alluviale de la Bassée (partie 3).

1 Développement de l'interface pyWaQS

1.1 Les outils de modélisation

1.1.1 La plate-forme de modélisation des hydrosystèmes CaWaQS

CaWaQS (*CAtchment WAter Quality Simulator*) (Flipo et al., 2023) est la plate-forme de modélisation des hydrosystèmes du PIREN-Seine (Fig. 1). La structure de CaWaQS permet à l'utilisateur de choisir entre différentes bibliothèques, en fonction du type d'hydrosystème qu'il souhaite modéliser. Pour obtenir un modèle complet d'un hydrosystème comprenant la surface, les rivières, la zone non saturée et la zone saturée, quatre bibliothèques principales sont utilisées : *libfp* (calcul du bilan hydrologique à l'aide d'un modèle de réservoirs), *libhyd* (simulation du débit des rivières à l'aide du modèle de Muskingum et de la formule de Manning-Strickler), *libnsat* (modélisation de l'infiltration verticale dans la zone vadose) et *libaq* (résolution de l'équation de la diffusivité dans la zone saturée pour obtenir les charges hydrauliques). Dans cette étude, les deux bibliothèques mobilisées sont *libaq* (Labarthe, 2016) et *libwet* (Jost et al., 2023), cette dernière gérant les échanges entre les eaux souterraines et les lacs. Dans cette configuration, seule l'équation de la diffusivité pour la nappe captive est disponible. L'hypothèse est donc émise que les charges calculées avec cette équation sont suffisamment cohérentes pour ne point nécessiter de complexité additionnelle.

Dans *libaq*, le calcul de l'écoulement horizontal est fondé sur l'équation d'écoulement des eaux souterraines dans la zone saturée, en utilisant l'hypothèse de Dupuit qui stipule que le gradient hydraulique est nul dans la direction z et en supposant que le tenseur de transmissivité $\overline{\overline{T}}$ est diagonal (équation 1). L'équation est résolue à l'aide de la méthode semi-implicite sur un maillage tridimensionnel.

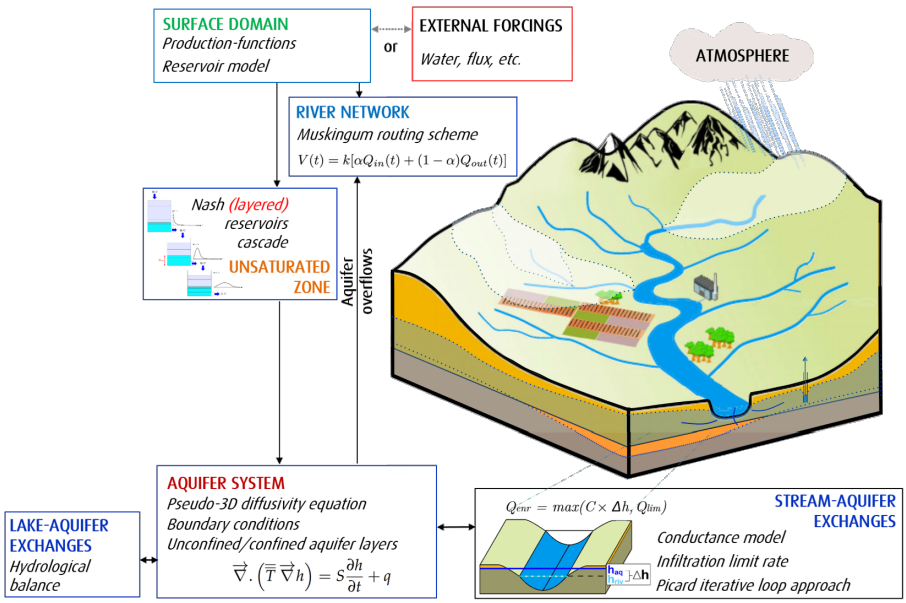


Figure 1. Illustration simplifiée de la structure de la plate-forme CaWaQS3.x (Gallois et al., 2025). Cette étude fait appel aux modules relatifs au système aquifère et aux échanges entre celui-ci et les lacs de gravière.

$$\frac{\partial(T_{xx} \cdot \frac{\partial h}{\partial x})}{\partial x} + \frac{\partial(T_{yy} \cdot \frac{\partial h}{\partial y})}{\partial y} = S\frac{\partial h}{\partial t} + q_s \tag{1}$$

où :

- T_{xx} (resp. T_{yy}) en [L^2/T] : transmissivité dans la direction x (resp. y)
- h [L] : charge hydraulique
- S [-]: coefficient d'emmagasinement
- q_s [L/T]: terme source.

La bibliothèque *libwet* calcule la charge hydraulique à l'instant $n+1$ dans un lac à l'aide du bilan hydrologique suivant :

$$h_l^{n+1} = h_l^n + \Delta t \cdot \frac{P - E + R + S_p}{A_s} \tag{2}$$

avec :

- h_l^{n+1} et h_l^n [L] : la charge hydraulique dans le lac aux pas de temps $n+1$ et n
- Δt [T] : la différence de temps entre les pas de temps $n+1$ et n
- P [L^3/T] : les précipitations
- E [L^3/T] : l'évaporation
- R [L^3/T] : le ruissellement
- S_p [L^3/T] : les échanges entre l'aquifère et le lac
- A_s [L^2] : la surface du lac.

S_p est calculé comme la somme de chaque échange à travers une interface entre une maille du lac et une maille de l'aquifère. S'il y a M interfaces, nous avons donc :

$$S_p = \sum_{m=1}^{M} Q_m \tag{3}$$

avec Q_m le débit échangé à travers l'interface m. La loi de Darcy donne alors :

$$S_p = \sum_{m=1}^{M} C_m \cdot (h_{a_m}^{n+1} - h_l^n) \tag{4}$$

avec C_m [L^2/T] la conductance de l'interface m et $h_{a_m}^{n+1}$ la charge hydraulique dans l'aquifère à l'interface m au pas de temps $n+1$. Enfin, pour résoudre l'équation 2, avec S_p défini comme indiqué dans l'équation 4, un paramètre implicite (θ_w) est défini et h_l^n dans l'équation 4 est remplacée par $(1-\theta_w) \cdot h_l^n + \theta_w \cdot h_l^{n+1}$.

1.1.2 Mode opératoire de la suite PEST

La résolution de l'équation de la diffusivité (équation 1) permet de calculer les charges hydrauliques en tout point du maillage choisi. Elle nécessite de connaître la distribution spatiale des propriétés hydrodynamiques du système aquifère étudié, la transmissivité T et le coefficient d'emmagasinement S, paramètres du modèle auxquels s'ajoute la conductance C_m dans le cas des échanges nappe-lac (équation 4). Or, les connaissances *a priori*, fondées sur les mesures sur site, les valeurs rapportées dans la littérature et l'expertise du modélisateur, sont indicatives mais insuffisantes. La nécessaire détermination des valeurs des paramètres du modèle en tout point du maillage est alors réalisée par modélisation inverse, en s'appuyant sur les observations de charge disponibles. Les valeurs des paramètres sont ajustées de telle sorte que les résultats du modèle reproduisent fidèlement les observations de terrain. Les programmes tels que PEST automatisent ce processus de calibration en minimisant une fonction objectif relative à la différence entre charges simulées par le modèle et charges observées. Notons que cette opération d'ajustement ne concerne cependant que les paramètres. De nombreux éléments du modèle, portant par exemple sur sa géométrie ou la nature de ses conditions aux limites, sont à modifier manuellement par l'utilisateur.

La suite PEST (*Model-Independent Parameter Estimation and Uncertainty Analysis*) dispose d'outils avancés pour calibrer tout modèle numérique. Les méthodes fondamentales appliquées par PEST sont ici très brièvement décrites et sont détaillées dans la documentation (Doherty, 2015; Fienen et al., 2025). PEST utilise l'algorithme de Gauss-Levenberg-Marquardt (GLM) qui optimise de façon itérative les paramètres du modèle jusqu'à ce qu'une valeur minimale de la fonction objectif Φ choisie soit trouvée.

Le processus commence par une estimation initiale des paramètres p_0, d'après les informations *a priori*. L'équation approximant la relation entre les paramètres et les observations par une fonction linéaire est la suivante :

$$h = h_0 + J(p - p_0) + \varepsilon \tag{5}$$

où :

- h [$\in \mathbb{R}^n$] est le vecteur des n observations,
- $h_0 = f(p_0)$ sont les sorties du modèle pour le jeu de paramètres initial,
- p [$\in \mathbb{R}^m$] est le vecteur des m paramètres des points pilotes, les inconnues du processus de calibration,
- J [$\in \mathcal{M}_{n,m}(\mathbb{R})$] est la matrice jacobienne de la fonction f évaluée en p_0, calculée en m itérations pour chaque dérivée partielle à l'aide d'une approximation par différences finies,
- ε [$\in \mathbb{R}^n$] est la différence entre les observations et les résultats du modèle, souvent appelée « erreur » ou « bruit », du fait du modèle ou d'une erreur dans la mesure.

Pour calibrer un modèle, PEST ajuste p jusqu'à ce que ε soit aussi petit que possible. Toutes les composantes d'erreur ne fournissant pas des informations d'égale importance sur la précision du modèle, des pondérations sont appliquées au vecteur ε. L'objectif du processus de calibration est ainsi de trouver $\bar{p}$ tel que :

$$\bar{p} = \min_p \Phi(p) \tag{6}$$

$$\text{avec} \quad \Phi(p) = (h - h_0 - J(p - p_0))^T Q (h - h_0 - J(p - p_0)). \tag{7}$$

Φ est la fonction objectif et Q est la matrice diagonale contenant le carré des facteurs de pondération des observations. S'il existe, $\bar{p}$ est nécessairement la solution de l'équation suivante :

$$\nabla\Phi|_{p=\bar{p}} = 0. \tag{8}$$

A chaque itération, PEST calcule un nouvel ensemble de paramètres p à partir d'un ensemble précédent p_0, des résultats du modèle calculés avec cet ensemble de paramètres précédent h_0 et de la matrice jacobienne de la relation observation/paramètre évaluée en p_0 :

$$p = p_0 + (J^T Q J)^{-1} J^T Q (h - h_0), \tag{9}$$

en supposant que $J^T Q J$ est inversible.

Ce processus itératif, appelé méthode de Gauss-Newton, est modifié par l'introduction d'un paramètre d'échelle $\lambda \in \mathbb{R}$ introduit par Levenberg pour les cas de non-linéarité importante et ajusté par Marquardt pour donner jour à l'équation de l'algorithme GLM utilisée par PEST :

$$p = p_0 + (J^T Q J + \lambda Diag(J^T J))^{-1} J^T Q (h - h_0). \tag{10}$$

L'un des défis lors de l'ajustement d'un modèle est la non-unicité inhérente associée à la solution inverse. Comme il y a en pratique plus d'inconnues (les paramètres) que de variables connues (les observations), ce problème est mal posé et il existe une infinité de solutions, c'est-à-dire une infinité d'ensembles de paramètres donnant le même ensemble de résultats du modèle. La régularisation désigne le processus nécessaire consistant à transformer ce problème mal posé en un problème bien posé, c'est-à-dire avec une solution unique, que PEST entreprend avant d'appliquer l'algorithme GLM. La solution choisie est celle qui minimise le potentiel d'erreur des prédictions, ce qui équivaut à choisir le champ de paramètres présentant la variance d'erreur minimale. La décomposition en valeurs propres (SVD) et la régularisation de Tikhonov sont les deux méthodes de régularisation employées. La méthode de décomposition en valeurs propres réduit le problème inverse en ne prenant en compte que les combinaisons linéaires des paramètres qui ont une influence suffisante sur le modèle. La régularisation de Tikhonov aborde le problème dans l'autre sens : elle enrichit l'ensemble d'observations avec des connaissances expertes sur les paramètres, par exemple des indications sur leur distribution spatiale. La régularisation de Tikhonov peut ainsi aider le processus de calibration à prendre en compte des relations non linéaires complexes entre les paramètres, telles que l'anisotropie.

1.2 Principes de pyWaQS

Pour coupler la suite d'algorithmes PEST et le modèle CaWaQS, une interface nommée pyWaQS a été développée. Cette interface est commandée par un script *shell* qui exécute une succession de commandes décrites ci-dessous et représentées sur la figure 2 pendant une itération du processus d'identification des paramètres. Dans cet exemple, il y a 5 000 mailles, 100 paramètres (représentés par les indices p_1, ... , p_{100}) et 15 observations de la charge hydraulique (représentées par les indices o_1, ... , o_{15}). Le choix des paramètres ajustables par PEST est l'une des premières étapes de sa configuration. Si à chaque maille du modèle distribué est associé un paramètre à estimer, le temps d'exécution de l'algorithme peut devenir prohibitif. Pour réduire le nombre de paramètres, on utilise ici la méthode géostatistique des points pilotes, soit une sélection d'emplacements particuliers où estimer la valeur des paramètres avant leur interpolation à tout le maillage. Les observations sont comparées par PEST à chaque itération aux sorties du modèle. Bien que ces observations puissent être disponibles à de courts pas de temps, pour la même raison d'efficacité numérique, PEST ne peut pas prendre en compte toutes leurs variations : des moments spécifiques doivent être choisis. Dans cet exemple de pyWaQS fonctionnant en régime transitoire, deux moments échantillonnés (t_1 et t_2) ont été choisis.

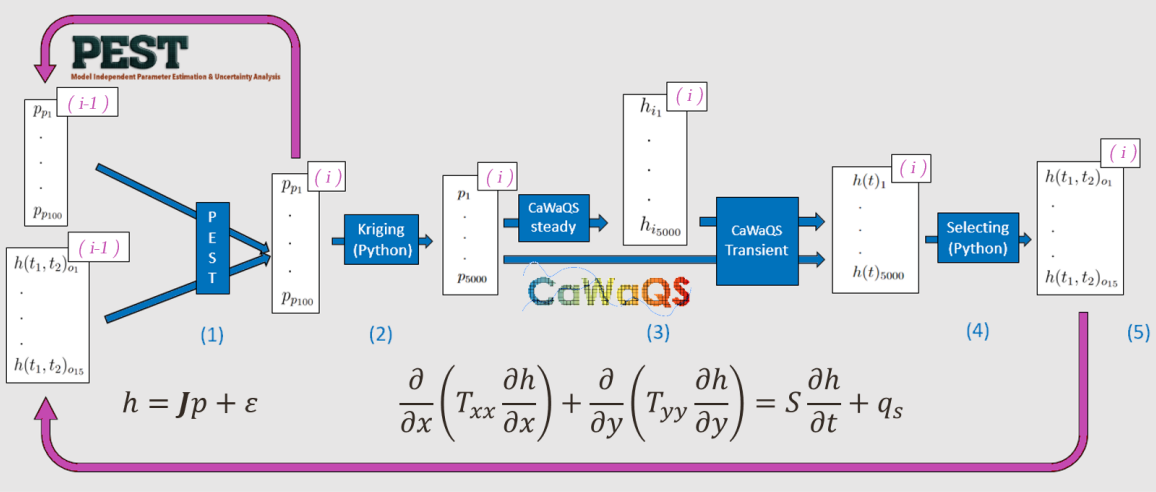


Figure 2. Principe de fonctionnement de l'interface pyWaQS (Saias, 2025) au cours d'une itération du processus d'étalonnage. Exemple choisi pour 5 000 mailles, 100 points pilotes et 15 piézomètres.

Voici comment fonctionne l'interface :

1. A partir de l'ensemble des paramètres des points pilotes de l'itération précédente (ou de l'ensemble initial de paramètres s'il s'agit de la première itération), PEST effectue l'une des deux opérations suivantes : il modifie la valeur d'un paramètre afin de progresser dans le calcul de la matrice jacobienne ; ou, si la matrice jacobienne est déjà complète (après 100 itérations dans cet exemple), PEST produit un nouvel ensemble de paramètres à l'aide de l'équation 10. Dans les deux cas, l'ensemble de paramètres a été mis à jour.

2. PEST ne prend en compte que les paramètres des points pilotes. Cependant, CaWaQS a besoin d'une valeur par maille. Les valeurs des paramètres des points pilotes sont ainsi interpolées géostatistiquement à l'ensemble du modèle par krigeage à l'aide d'un script python, en utilisant un modèle de semi-variogramme déjà calculé. Le même script python renseigne les données d'entrée nécessaires au modèle CaWaQS.

3. CaWaQS est d'abord exécuté en régime permanent afin d'obtenir les valeurs initiales de charge pour chaque maille, puis en régime transitoire afin de simuler les variations de charge. Si l'on souhaite calibrer un modèle en régime permanent, seule une exécution de CaWaQS est nécessaire à cette étape.

4. Le fichier *shell* exécute ensuite un autre script python pour sélectionner les valeurs des charges simulées au niveau des points d'observation, et aux moments d'échantillonnage choisis si le régime est transitoire. Ces valeurs sont ensuite écrites dans le format approprié pour que PEST puisse les lire.

5. Si le jacobien n'est pas encore entièrement calculé, une autre itération est exécutée. Si c'est le cas, PEST calcule la nouvelle valeur de la fonction d'objectif Φ. S'il a atteint sa valeur cible, PEST s'arrête. Sinon, il lance une autre série de 100 itérations pour calculer une nouvelle matrice jacobienne.

Le fichier *shell* qui exécute pyWaQS est intégré dans PEST via une ligne dans le fichier de commande. Ce fichier de commande contient toutes les données dont PEST a besoin pour effectuer son processus de calibration : données paramétriques, observations, informations *a priori* pour les méthodes de régularisation. Toutes ces informations fournies à PEST sont détaillées par la suite, d'abord pour un cas d'étude synthétique puis pour l'ajustement du modèle hydrogéologique de la plaine alluviale de la Bassée.

Table 1. Configuration paramétrique du cas d'étude synthétique.

Paramètres et valeurs associées		Unités
Transmissivité	$2 \cdot 10^{-2}$ pour les graviers, $6 \cdot 10^{-3}$ pour le sable et $4 \cdot 10^{-5}$ pour les sables argileux	[$m^2.s^{-1}$]
Coefficient d'emmagasinement	$0,21$ pour les graviers, $0,33$ pour les sables et $0,06$ pour les sables argileux	[–]
Epaisseur	10	[m]
Charge hydraulique de la rivière	138	[m]
Flux latéral	$2 \cdot 10^{-3}$	[$m^3.s^{-1}.km^{-1}$]
Recharge (régime permanent)	150	[$mm.an^{-1}$]
Recharge (régime transitoire)	Fig. 3c	[$mm.an^{-1}$]
Charges initiales	Issues du régime permanent pour une recharge de 1 500 $mm.an^{-1}$	[m]

2 Cas d'étude

Un cas synthétique a été construit pour développer et tester pyWaQS dans une configuration simple avant son application au cas d'étude plus complexe de la Bassée.

2.1 Modèle CaWaQS

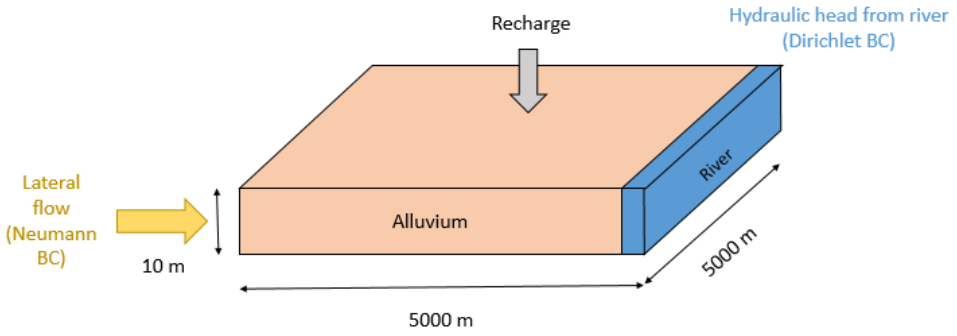
Le cas synthétique (Fig. 3) est celui d'un aquifère alluvial hétérogène couvrant une zone de 5 km^2, rechargé par un aquifère crayeux à l'amont et alimentant une rivière à l'est. Le modèle comprend une couche d'alluvions discrétisée en 2 519 mailles de dimensions respectives en x, y et z en mètres $100 \cdot 100 \cdot B$, où B est la profondeur de la couche considérée, ici égale à 10 m. Le flux latéral entrant à l'amont est représenté par une condition aux limites de type Neumann et la rivière à l'aval, par une condition aux limites de Dirichlet (Fig. 3a). Un terme source de recharge est imposé sur toutes les mailles alluviales. Les paramètres du modèle sont récapitulés dans le tableau 1. Les figures 3d et 3e présentent la distribution spatiale des paramètres hydrodynamiques, fonction de la lithologie des alluvions (Fig. 3b).

Une simulation en régime permanent et une simulation en régime transitoire ont été conduites avec CaWaQS, dont les résultats sont présentés en figure 4 pour la première. L'influence des lentilles de sable argileux est particulièrement perceptible, avec des valeurs de charge plus élevées, alors que les isopièzes sont plus espacées dans la partie graveleuse plus perméable à l'aval. A l'exception de ces irrégularités dues à la lithologie hétérogène choisie, les isopièzes sont parallèles et diminuent d'ouest en est jusqu'à atteindre la charge hydraulique fixe dans la rivière à 138 mètres.

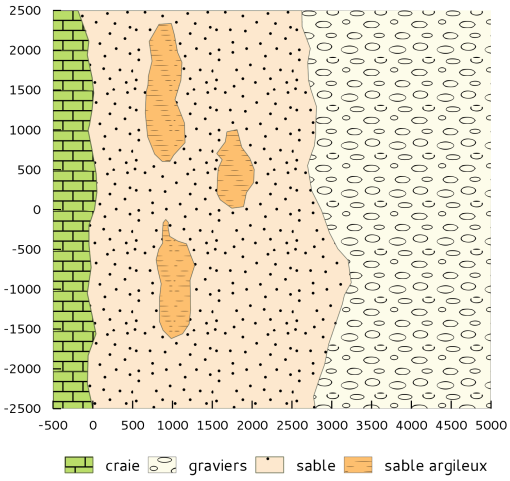
Pour la simulation transitoire, seule la recharge est modifiée et varie en fonction du temps de manière sinusoïdale, avec une période d'un an et selon une amplitude quelque peu exagérée (Fig. 3c) afin d'obtenir des variations de charge conséquentes dans l'aquifère, aux fins de l'exercice de calibration.

2.2 PEST et pyWaQS

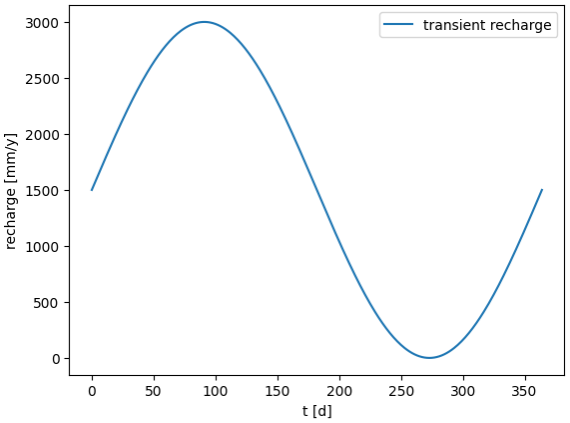
Le cas d'étude qui vient d'être décrit a servi de base au test de pyWaQS en conditions contrôlées. L'objectif est de trouver une solution au problème inverse de l'équation de la diffusivité, c'est-à-dire un ensemble de paramètres correspondant de façon satisfaisante à un ensemble de charges observées. Le cas étant fictif, cet ensemble observé a été choisi comme le résultat de simulations CaWaQS.



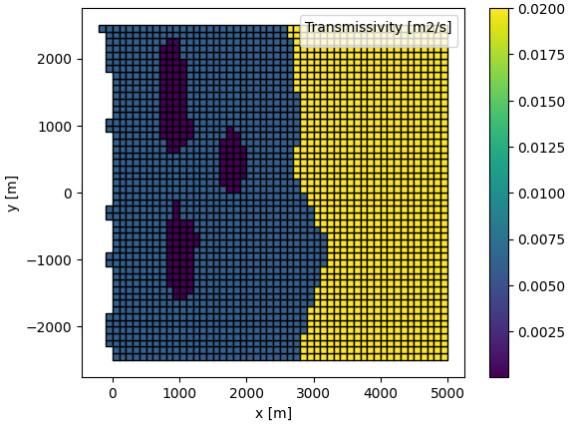
(a) *Conditions aux limites*



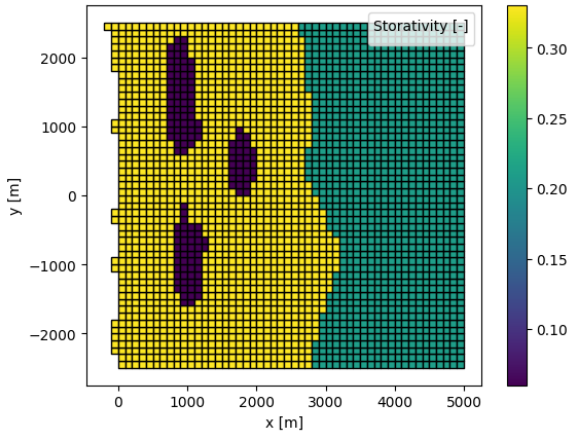
(b) *Lithologie*



(c) *Recharge transitoire*



(d) *Transmissivité*



(e) *Coefficient d'emmagasinement*

Figure 3. Configuration du cas d'étude synthétique. Les distances sont exprimées en mètres. Seules les alluvions sont représentées dans le modèle.

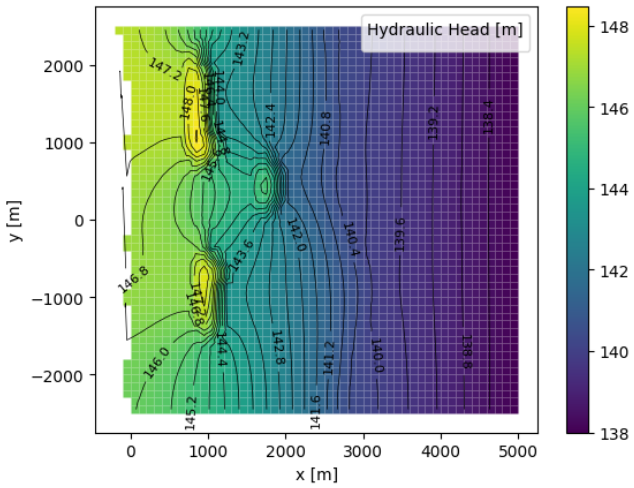


Figure 4. Charges hydrauliques (m) dans la couche alluviale en fonction de la position, calculées par CaWaQS pour la simulation en régime permanent.

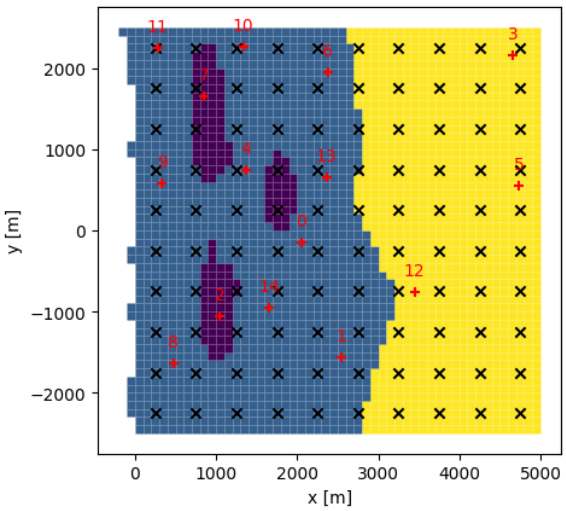


Figure 5. Points pilotes (noirs) et points d'observation (rouges) sur la carte lithologique du cas test.

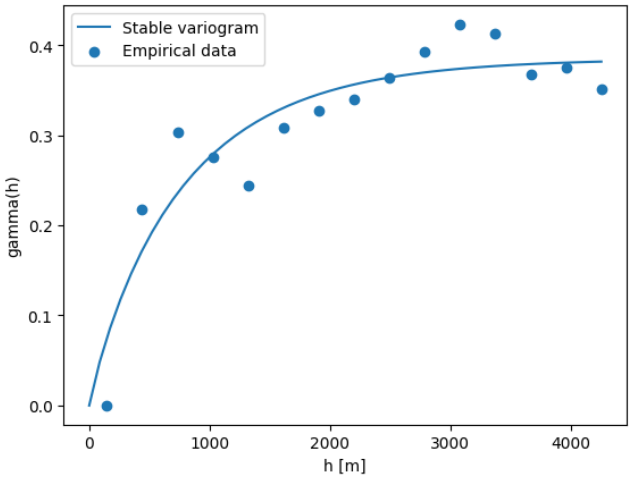


Figure 6. Semi-variogramme approché à partir de l'ensemble des données paramétriques.

Les points d'observation (h dans l'équation 5) sont au nombre de 15 et leur emplacement est présenté en figure 5, de même que les 100 points pilotes (p dans l'équation 5) répartis régulièrement sur la grille. Les observations correspondent aux résultats de simulations CaWaQS conduites selon la configuration décrite précédemment (section 2.1), et extraites au niveau des piézomètres choisis. Les valeurs initiales des paramètres des points pilotes fournies à PEST pour lancer la première itération du processus, provenant en principe de connaissances *a priori* du terrain étudié, sont ici issues de l'ensemble de paramètres qui a été utilisé pour modéliser les charges « observées » avec CaWaQS.

La régularisation de Tikhonov nécessite des informations supplémentaires sur les paramètres, à ajouter au vecteur h, afin que la matrice jacobienne calculée comporte plus de lignes que de colonnes. Cet ensemble d'informations *a priori* est ici considéré comme le logarithme des éléments de l'ensemble de données des paramètres. La régularisation de Tikhonov nécessite également la matrice de covariance de la fonction d'information *a priori*, qui dans notre cas est égale au logarithme de la fonction paramétrique.

$$\begin{aligned} C(h) &= Cov(p(x+h), p(x)) \\ C(h) &= C(0) - \gamma(h) \\ \text{avec} \quad \gamma(h) &= \frac{E((p(x+h) - p(x))^2)}{2} \end{aligned} \tag{11}$$

La quantité $\gamma(h)$ est appelée semi-variogramme. La plage, ou distance à partir de laquelle deux paramètres deviennent plus ou moins indépendants, a été fixée à 4 400 m après essais et erreurs et la forme générale du semi-variogramme, à un modèle stable. La figure 6 montre le semi-variogramme calculé. Il convient de noter que le semi-variogramme est utilisé par pyWaQS à deux fins : pour calculer la matrice de covariance de la fonction paramétrique et pour effectuer le krigeage des paramètres (voir figure 2).

En régime transitoire, un jeu de charges initiales est nécessaire en entrée de CaWaQS. Celles-ci dépendent naturellement du champ de transmissivité, qui est modifié à chaque itération PEST. Par conséquent, pour une simulation transitoire, CaWaQS doit être exécuté deux fois à chaque itération PEST : une fois en régime permanent pour calculer les charges initiales puis en régime transitoire. Afin d'estimer le jeu de paramètres du modèle, PEST doit avoir accès aux variations de charge au niveau des points d'observation. Un intervalle d'échantillonnage a donc été choisi ($t_s = 36$ jours), qui permet un très bon ajustement des variations de charge observées (Fig. 7). Avec 15 piézomètres d'observation et 11 échantillonnages par puits, les données d'observation contiennent 165 éléments.

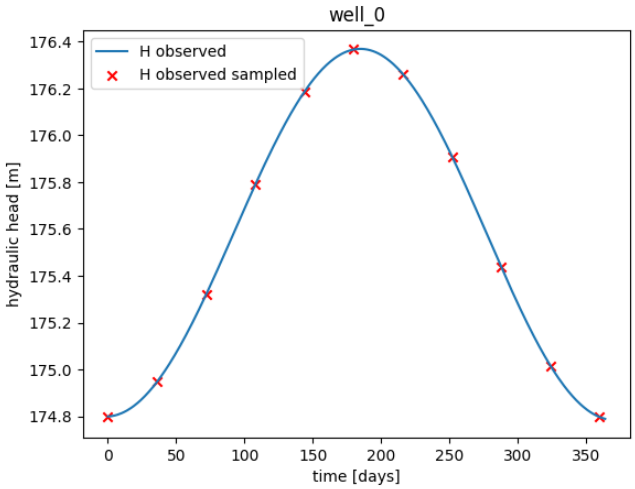


Figure 7. Echantillonage des charges observées au piézomètre 0.

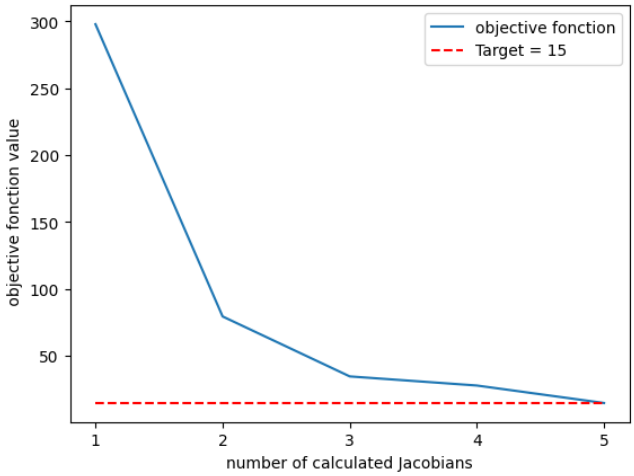
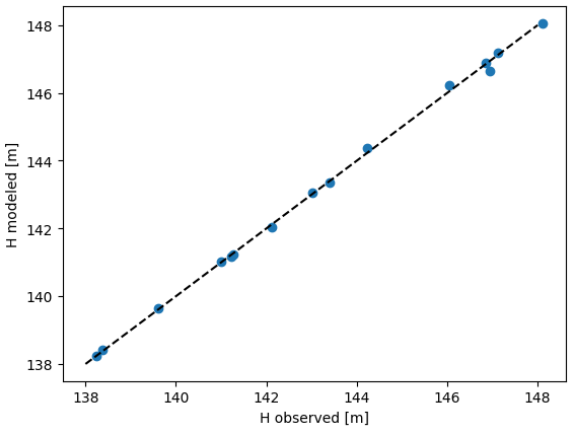


Figure 8. Evolution de la fonction objectif en régime permanent.

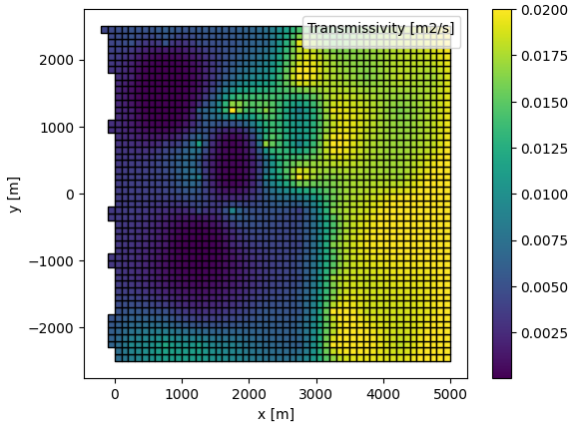
2.3 Résultats

Comme le montre la figure 8, le processus piloté par pyWaQS a convergé en régime permanent. Le modèle a atteint la valeur cible de la fonction objectif (égale à 15, car il y a 15 points d'observation) après cinq calculs de la matrice jacobienne. A l'aide d'une machine à 4 cœurs, le processus a nécessité environ 50 minutes pour converger.

A l'issue du processus, les charges hydrauliques simulées se rapprochent de manière très satisfaisante des charges « observées » (Fig. 9a). Le coefficient de corrélation des deux ensembles de charges hydrauliques est supérieur à 99,9 % et l'écart quadratique moyen normalisé est égal à 1,0 %. Rappelons que les charges dites « observées » étant en réalité simulées, il n'y a pas de bruit dû au processus de mesure. Enfin, les transmissivités résultant du processus de calibration PEST via l'interface pyWaQS sont représentées sur la figure 9b. Bien qu'elles ne soient pas identiques à celles qui ont été utilisées pour simuler les charges « observées » (voir Fig. 3d), la répartition spatiale des différentes lithologies des alluvions est visible.



(a) *Comparaison des charges hydrauliques simulées avec celles « observées ».*



(b) *Transmissivités finales issues du processus de calibration de PEST.*

Figure 9. Résultats en régime permanent.

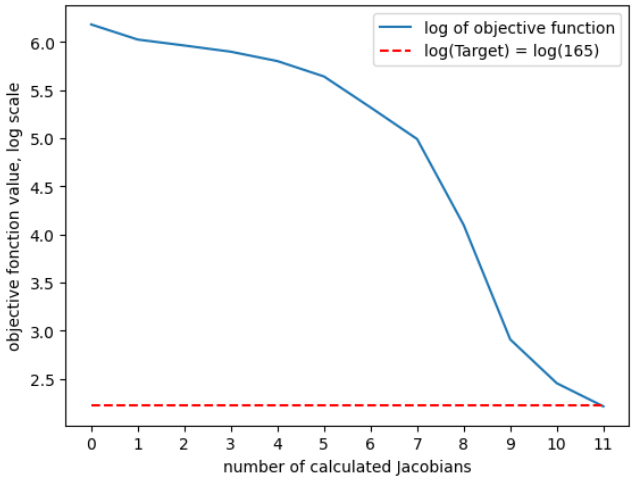


Figure 10. Evolution de la fonction objectif en régime transitoire.

En régime transitoire, le processus a convergé après 11 calculs de la matrice jacobienne, comme le montre la figure 10. A l'aide d'une machine à 4 cœurs, il a fallu environ 4 heures pour atteindre la cible de la fonction objectif, égale à 165.

L'étalonnage du modèle a été couronné de succès (Fig. 11), les charges modélisées correspondant aux observations avec un coefficient de corrélation supérieur à 99,9 % et un écart quadratique moyen normalisé inférieur à 0,4 %. La figure 12 présente les charges « observées » et simulées en fonction du temps sur un même graphique pour deux points d'observation. Des différences mineures sont observées, par exemple sur le piézomètre 14, mais dans l'ensemble, les variations de charge sont bien représentées. Trois indicateurs de performance ont été utilisés, le coefficient de Nash-Sutcliffe (NSE), le coefficient de Nash-Sutcliffe calculé sur le logarithme des données (NSElog) et l'efficacité de Kling-Gupta (KGE). A l'exception du KGE du piézomètre 14, ils sont tous supérieurs à 90 %.

Les transmissivités et les coefficients d'emmagasinement obtenus à l'issue du processus de calibration sont illustrés en figure 13. Ils ne correspondent pas parfaitement à ceux utilisés pour modéliser les « observations » mais de nouveau, les différentes lithologies sont apparentes. Les deux champs sont distincts l'un de l'autre, la distribution spatiale du coefficient d'emmagasinement étant plus nette que celle de la transmissivité,

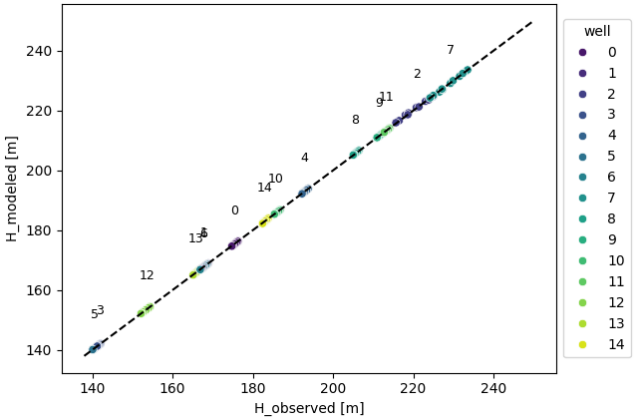
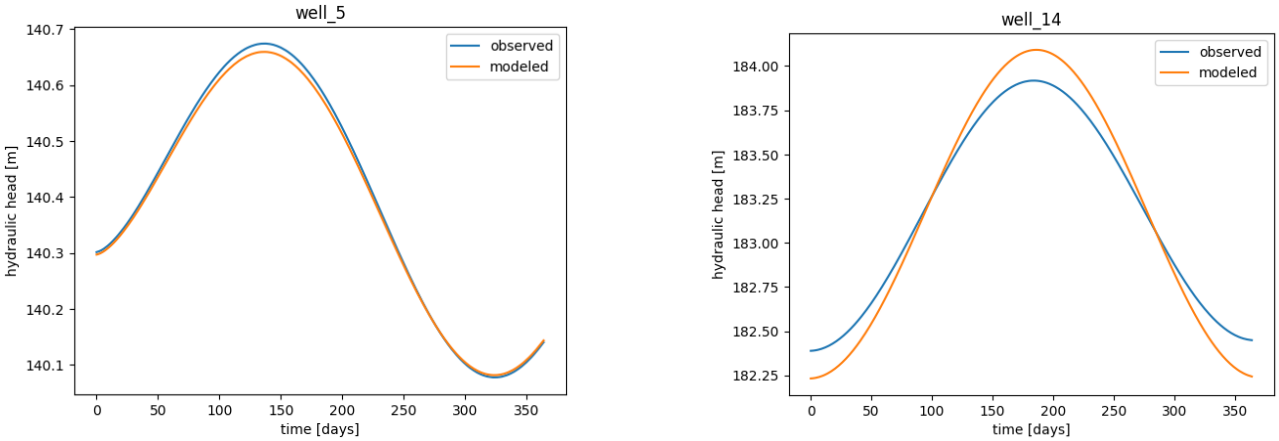
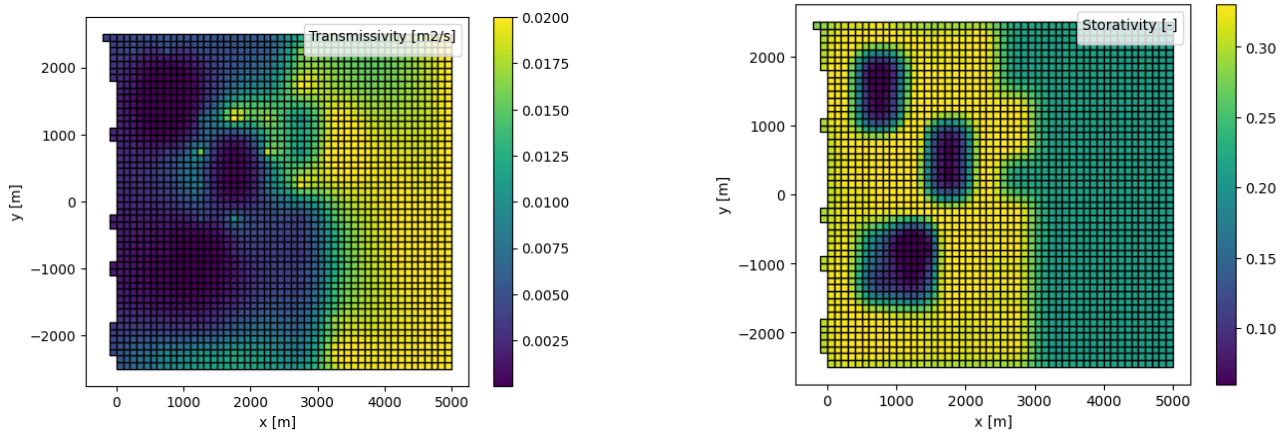


Figure 11. Comparaison des charges hydrauliques simulées avec celles « observées » en régime transitoire.



(a) *Piézomètre 5* (b) *Piézomètre 14*

Figure 12. Charges simulées et « observées » en deux piézomètres d'observation.



(a) *Transmissivité* (b) *Coefficient d'emmagasinement*

Figure 13. Champs de transmissivité et du coefficient d'emmgasinement résultant de la calibration PEST en régime transitoire.

probablement en lien avec des plages acceptables de paramètres bien plus restreintes dans le cas du coefficient d'emmagasinement (entre 10^{-3} et 1) par rapport à la transmissivité (entre 10^{-6} et 10^{-1} m.s^{-1}).

3 Application au système alluvial de la Bassée

3.1 Modèle hydrogéologique de la plaine alluviale

Le modèle hydrogéologique de la plaine de la Bassée a été développé au cours de plusieurs étapes. La première version de ce modèle, qui ne comportait alors qu'une seule couche alluviale, a été réalisée par Labarthe (2016) afin de décrire les échanges entre la Seine et le système aquifère, à l'aide de la bibliothèque *libaq*. Le modèle a été ensuite complexifié (Jost et al., 2017; Picourlat et al., 2018) avec l'ajout d'une couche sous-jacente représentant l'aquifère de la craie et la prise en compte des gravières à l'aide de la bibliothèque *libwet*. Dans la présente version, le maillage a été simplifié afin d'en réduire le nombre d'éléments, avec des mailles de 50 à 400 mètres de côté pour les alluvions et de 100 à 800 mètres pour la craie.

Le modèle local de la Bassée est intégré dans un modèle régional couvrant l'ensemble du bassin versant de la Seine (Flipo et al., 2023), à partir duquel sont déduites la plupart de ses conditions aux limites : recharge issue des précipitations, telle que calculée par la bibliothèque *libfp* et flux imposé aux limites latérales des deux couches. Les charges calculées par la bibliothèque *libhyd* qui simule le débit des rivières étant trop éloignées de la réalité (Fig. 14), elles ont été remplacées le long du linéaire de la Seine et de ses principaux affluents par les valeurs issues du modèle numérique de terrain (MNT) au pas de 1 mètre (RGE ALTI®, IGN). Le modèle prend également en compte les prélèvements d'après la Banque nationale des prélèvements quantitatifs en eau (BNPE). Les montants annuels ont été convertis en montants quotidiens, en partant du principe que l'irrigation n'est opérationnelle que de mai à septembre. Les conditions aux limites pour la couche alluviale sont résumées sur la figure 15, qui rappelle également la localisation des gravières.

3.2 Mise en œuvre de pyWaQS

Des points pilotes ont été régulièrement répartis sur le maillage de l'aquifère des alluvions, tous les 800 mètres dans les directions x et y. Ce schéma de répartition a abouti au nombre de 547 points pilotes (Fig. 16). Il a été décidé d'exclure les gravières lors du choix de la localisation des points pilotes, la conduc-

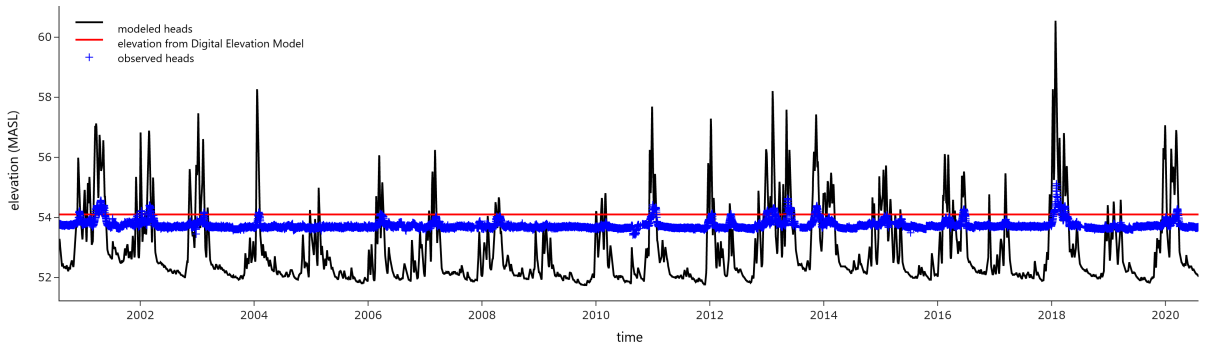


Figure 14. Comparaison des charges hydrauliques de la Seine telles qu'observées et calculées par CaWaQS à Bray-sur-Seine.

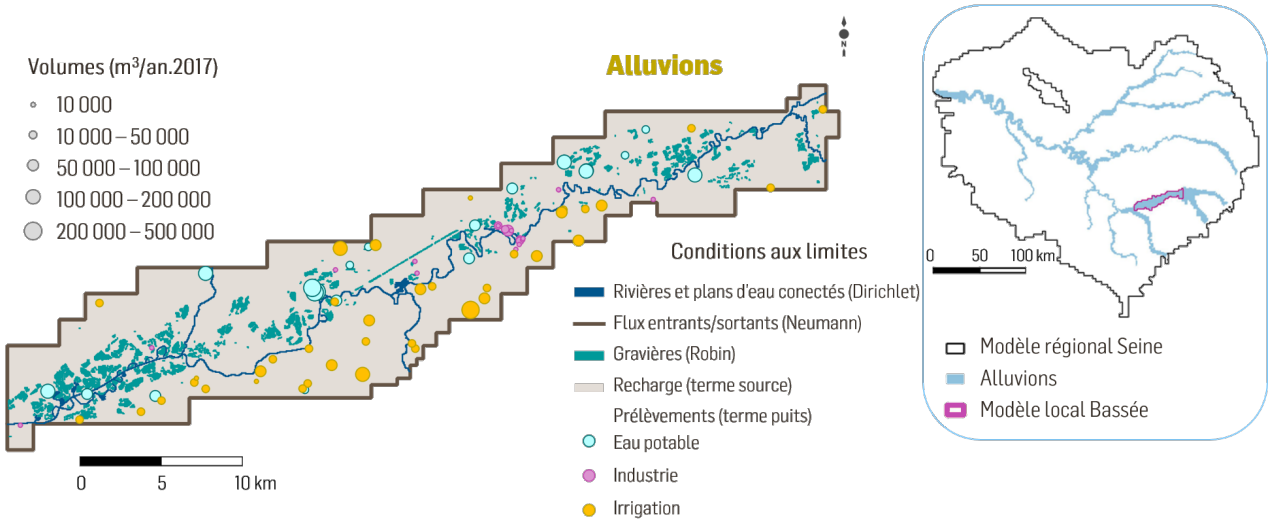


Figure 15. Modèle hydrogéologique de la plaine de la Bassée : couche des alluvions.

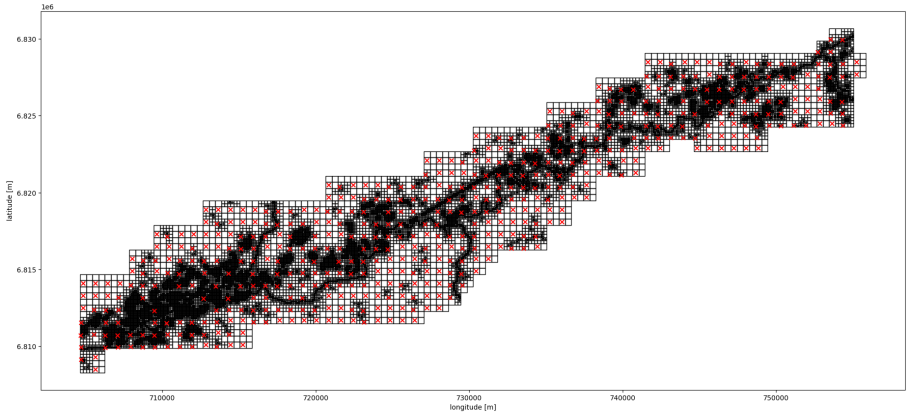


Figure 16. Localisation des points pilotes au sein du maillage de la couche des alluvions.

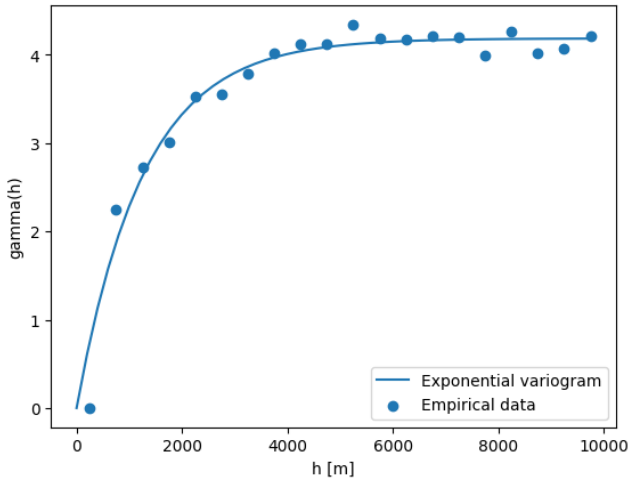


Figure 17. Semi-variogramme du champ de transmissivité initial.

tance C_m qui les caractérise n'ayant pas été retenue à ce stade comme paramètre à calibrer et restant homogène à l'échelle de la plaine.

Les valeurs initiales associées à chaque point pilote sont issues d'estimations antérieures du champ de transmissivité dans les alluvions (Labarthe, 2016), établies par estimation successive des flux (Pasquier et Marcotte, 2006) et de données plus anciennes, remontant à septembre 1965 (Mégnien, 1965) et septembre 1994 (Greiner, 1997). Le semi-variogramme, essentiel pour le krigeage des paramètres et pour la régularisation de Tikhonov, a été calculé à partir de cette carte initiale, sur la base du logarithme de la transmissivité et d'un modèle exponentiel, en supposant que la corrélation spatiale du champ initial est pertinente (Fig. 17). Aucune information spatialisée préalable n'est en revanche disponible pour ce qui concerne le coefficient d'emmagasinement. Nous ne présenterons de fait ici que la mise en œuvre de pyWaQS en régime permanent, pour la détermination du champ de transmissivité de la plaine alluviale. L'application au régime transitoire pour l'estimation conjointe du coefficient d'emmagasinement, d'autant plus coûteuse en temps de calcul, n'a pas encore abouti de façon satisfaisante.

La limite inférieure pour la conductivité hydraulique associée à la transmissivité a été fixée à 10^{-9} $m^2.s^{-1}$ et la limite supérieure, à 0,1 $m.s^{-1}$. Pour chaque point pilote, la limite supérieure de transmissivité a donc été fixée à 0,1 fois l'épaisseur de la maille.

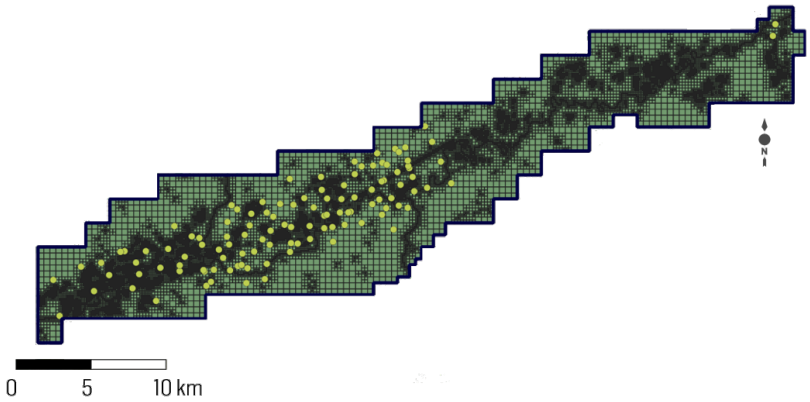


Figure 18. Localisation des points d'observation au sein du maillage de la couche des alluvions.

Les données d'observation proviennent de piézomètres crépinés dans l'aquifère alluvial de la plaine de la Bassée, piézomètres mis en place pour la plupart au cours des dix dernières années, en lien avec les travaux d'aménagement et de surveillance de la plaine, en particulier dans le cadre du programme PIREN-Seine (*e.g.*, Baratelli et al., 2021). Les années hydrologiques disposant du plus grand nombre de données sont 2016-2017 et 2017-2018. Afin d'utiliser un nombre maximum de points d'observation, l'état de référence a été fixé au 4 novembre 2017, en période de basses eaux. Les 116 piézomètres sélectionnés sont quasiment tous situés dans la partie aval de la plaine (Fig. 18). La simulation transitoire couvrant l'année hydrologique 2016-2017 est toujours en cours d'exploitation.

3.3 Résultats

PyWaQS a été lancée en configuration parallèle sur le supercalculateur MCMeSU de Sorbonne Université. La figure 19 présente la carte de la distribution spatiale de la conductivité hydraulique générée par le couplage CaWaQS-PEST. Sa large plage est caractéristique de l'hétérogénéité des dépôts alluviaux. Comparée à la distribution précédemment estimée (Labarthe, 2016), on note une distribution décalée vers les valeurs les plus élevées de conductivité hydraulique, en particulier une plus grande occurrence des valeurs de conductivité supérieure à 10^{-2} m.s^{-1}.

La figure 20 compare les charges hydrauliques simulées avec celles observées aux piézomètres. Elles ne correspondent pas parfaitement, mais la corrélation est évidente, confirmée par un coefficient de corrélation de 95 %.

En échantillonnant la fonction de distribution de probabilité *a posteriori* du champ de conductivité hydraulique, un grand nombre de réalisations a pu être généré et servir de base pour évaluer la capacité prédictive du modèle. Les figures 21a et 21b rendent ainsi respectivement compte de la justesse et de la précision du modèle. Elles sont prometteuses en matière d'analyse de l'incertitude des prédictions du modèle, même si l'effort d'étalonnage peut encore être poursuivi, compte tenu de la large palette d'outils d'optimisation disponibles dans la suite PEST.

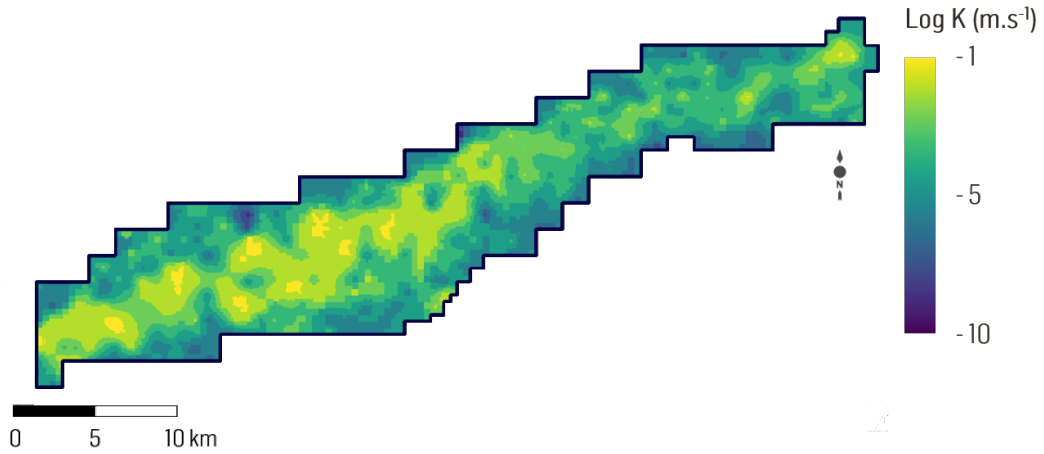


Figure 19. Champ de conductivité hydraulique en sortie de calibration PEST.

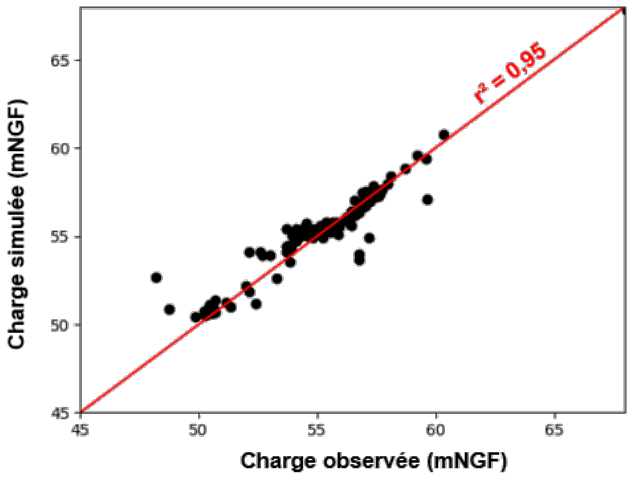
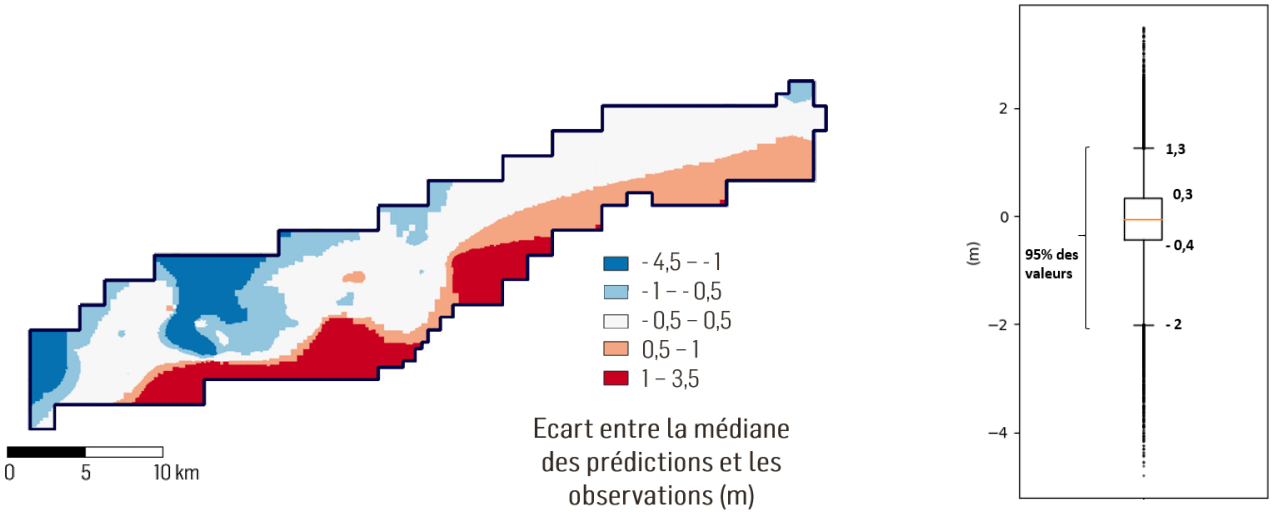
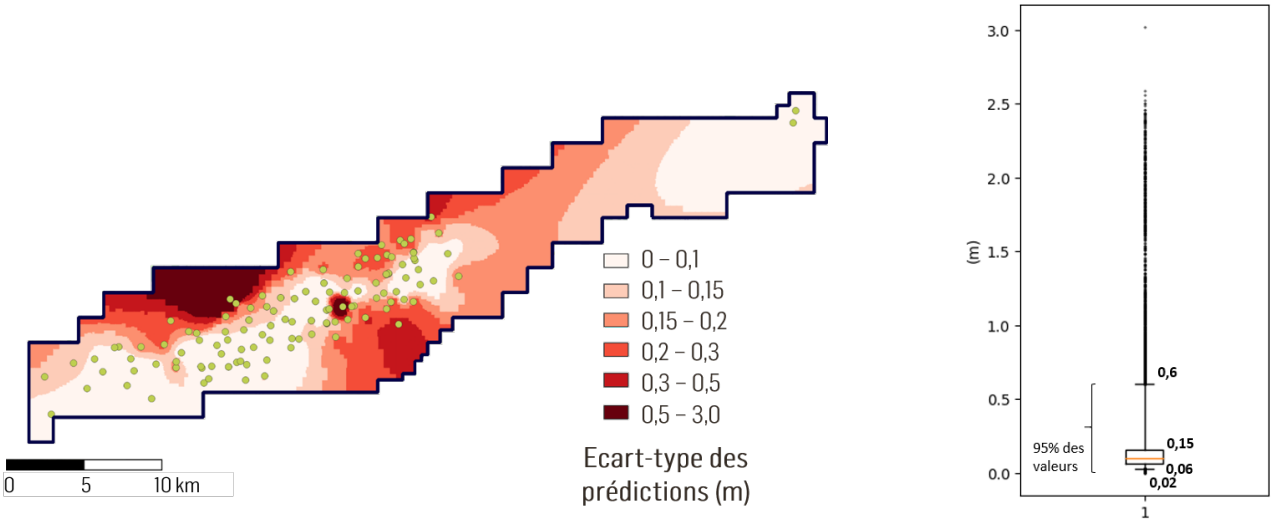


Figure 20. Comparaison des charges hydrauliques simulées aux charges observées après calibration.



(a) *Ecart entre la médiane des prédictions et les observations.*



(b) *Ecart-type des prédictions.*

Figure 21. Analyse d'incertitudes.

Conclusion et perspectives

Après avoir examiné en détail les processus mathématiques sur lesquels reposent la plate-forme de modélisation des hydrosystèmes CaWaQS et la suite d'algorithmes PEST, une interface a été créée entre les deux logiciels afin de permettre la calibration de modèles hydrogéologiques CaWaQS à l'aide des outils d'optimisation fournis par PEST. Cette interface a été testée pour l'estimation des paramètres hydrodynamiques (transmissivité et coefficient d'emmagasinement) d'un cas test simplifié, avec des résultats tout à fait satisfaisants. Bien que les champs de paramètres résultant du processus d'étalonnage soient différents de ceux considérés comme réels, les charges simulées correspondent parfaitement à celles observées. Après avoir compilé les données disponibles dans la plaine alluviale de la Bassée, l'interface pyWaQS a été utilisée pour tenter d'estimer les paramètres hydrodynamiques du modèle CaWaQS développé dans cette plaine. A ce jour, le champ de transmissivité calculé à l'issue du processus a permis d'obtenir une corrélation étroite entre les charges hydrauliques simulées et observées.

Les premières tentatives pour estimer la distribution spatiale du coefficient d'emmagasinement permettant, conjointement au champ de transmissivité, de simuler la dynamique des charges observées doivent être poursuivies. A plus long terme, de nouvelles données pourront être ajoutées dans le modèle et le processus de calibration. Par exemple, l'intégration des hauteurs d'eau en rivière et sur les gravières fournies par la mission satellitaire SWOT devrait apporter des contraintes supplémentaires utiles pour renforcer la robustesse de l'évaluation des flux échangés par le compartiment souterrain avec les eaux de surface.

C'est l'un des termes du bilan hydrologique qui pourrait évoluer si la recharge venait à diminuer ou les prélèvements dans la plaine, augmenter. Dans un contexte de pression accrue sur la ressource en eau, le modèle est alors un outil facilitant la scénarisation des aménagements possibles et leur spatialisation et pouvant orienter à terme les décisions à prendre pour anticiper des situations critiques. Ce travail est à réaliser en concertation avec les échelles décisionnaires pertinentes, en particulier la Commission Locale de l'Eau (CLE) du SAGE Bassée-Voulzie, concernée par la définition de quantités d'eau mobilisables de façon durable dans le système craie-alluvions. La large palette des outils PEST (White et al., 2020) offre un cadre approprié pour optimiser des débits de pompage, tant sur le plan spatial que quantitatif, tout en intégrant des contraintes sociales et environnementales, notamment celles liées aux sécheresses futures, et l'ensemble des incertitudes associées (*e.g.*, Coulon et al., 2024). Seule une calibration rigoureuse des paramètres en amont garantira que l'optimisation s'appuie sur un modèle fiable. Des réflexions méthodologiques seront à mener en lien avec les acteurs sur le plan de la définition des contraintes à respecter (*e.g.*, maintien du débit minimal d'un cours d'eau, seuil de niveau d'eau, qualité).

Pour retrouver les outils et données

La plupart des données utilisées pour la validation du modèle sont issues du Hub'eau.

CaWaQS (Flipo et al., 2023) est distribué sous licence EPL v2.0. Sa version la plus à jour est téléchargeable sur son dépôt gitlab de développement.

La suite PEST: *Model-Independent Parameter Estimation and Uncertainty Analysis* (Fienen et al., 2025) est téléchargeable sur son dépôt github de développement.

Bibliographie

Aissat, R. (2021). *Mise en œuvre d'algorithmes d'optimisation pour l'estimation des paramètres distribués des modèles hydrogéologiques régionaux – Application au Modèle Nord-Aquitain (MONA)*. Thèse de Doctorat, Université Bordeaux 3.

Baratelli, F., E. Parlanti, N. Flipo, S. Guillon, A. Rivière, A. Jost, M. Sourzac, B. Mercier, A. Martinez, R. Nespoulet, et al. (2021). Campagnes de prélèvements multi-traceurs pour la caractérisation hydro-biogéochimique de la plaine alluviale de la Bassée. Rapport, PIREN-Seine.

Coulon, C., J. T. White, A. Pryet, L. Gatel, et J.-M. Lemieux (2024). An ensemble-based approach for pumping optimization in an island aquifer considering parameter, observation and climate uncertainty. *Hydrology and Earth System Sciences 28*, 303–319.

Doherty, J. (2015). *Calibration and Uncertainty Analysis for Complex Environmental Models*. Brisbane, Australia: Watermark Numerical Computing. ISBN : 978-0-9943786-0-6.

Fienen, M., J. White, et M. Hayek (2025). Parameter ESTimation With the Gauss–Levenberg–Marquardt Algorithm: An Intuitive Guide. *Groundwater 63*, 93–104.

Flipo, N., N. Gallois, et J. Schuite (2023). Regional coupled surface–subsurface hydrological model fitting based on a spatially distributed minimalist reduction of frequency domain discharge data. *Geosci. Model Dev. 16*, 353–381.

Gallois, N., N. Flipo, B. Labarthe, A. Rivière, F. Baratelli, M. Maillot, S. Wang, et A. Jost (2025). *CaWaQS 3.52 CAtchment WAter Quality Simulator: User guide, concepts and equations*. Mines Paris/ARMINES - Université PSL, Fontainebleau.

Greiner, I. (1997). *Analyse de la variabilité fonctionnelle de la plaine alluviale de la Seine et modélisation des transferts de nitrates*. Thèse de Doctorat, Université Pierre et Marie Curie Paris VI.

Jost, A., G. Dass, F. Picourlat, S. Wang, L. Lestel, D. Eschbach, N. Flipo, et A. Ducharne (2024). Exploring historical anthropogenic influences on groundwater in the alluvial plain of the Upper Seine River. EGU General Assembly, Vienna, Austria, 14–19 Apr 2024, EGU24-17913.

Jost, A., S. Wang, B. Labarthe, et N. Flipo (2017). Impacts hydrodynamiques des gravières dans la plaine alluviale de la Bassée. Rapport, PIREN-Seine.

Jost, A., S. Wang, T. Verbeke, F. Colleoni, et N. Flipo (2023). Hydrodynamic relationships between gravel pit lakes and aquifers: Brief review and insights from numerical investigations. *Comptes Rendus Géoscience 355*, 1–25.

Labarthe, B. (2016). *Quantification des échanges nappe-rivière au sein de l'hydrosystème Seine par modélisation multi-échelle*. Thèse de Doctorat, Université PSL.

Langevin, C. D., A. M. Provost, S. Panday, et J. D. Hughes (2022). Documentation for the MODFLOW 6 Groundwater Transport Model. Rapport 6-A61, U.S. Geological Survey, Reston, VA.

Mégnien, C. (1965). Possibilités aquifères des alluvions du val de Seine entre Nogent-sur-Seine et Montereau, incluant la carte géologique et géomorphologique de la Bassée. Rapport 6-A61, BRGM DSGR.

Pasquier, P. et D. Marcotte (2006). Steady- and transient-state inversion in hydrogeology by successive flux estimation. *Advances in Water Resources 29*, 1934–1952.

Picourlat, F., A. Jost, S. Teillaud, P. Passy, F. Baratelli, et N. Flipo (2018). Impact des aménagements anthropiques dans la plaine de la Bassée : le cas des gravières. Rapport, PIREN-Seine.

Saias, C. (2025). Estimation of the hydrodynamic parameters in the aquifer system model of Bassee alluvial plain, France. Mémoire de Master, Sorbonne Université.

White, J. T., R. J. Hunt, M. N. Fienen, et J. E. Doherty (2020). Approaches to highly parameterized inversion: PEST++ Version 5, a software suite for parameter estimation, uncertainty analysis, management optimization and sensitivity analysis. Rapport 7-C26, U.S. Geological Survey, Reston, VA.