



Utilisation du calcul parallèle pour étudier numériquement la macro dispersion dans les aquifères

Anthony Beaudoin, Jean-Raynald de Dreuzy, Jocelyne Erhel

► To cite this version:

Anthony Beaudoin, Jean-Raynald de Dreuzy, Jocelyne Erhel. Utilisation du calcul parallèle pour étudier numériquement la macro dispersion dans les aquifères. CFM 2007 - 18ème Congrès Français de Mécanique, Aug 2007, Grenoble, France. hal-03361147

HAL Id: hal-03361147

<https://hal.science/hal-03361147>

Submitted on 1 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Utilisation du calcul parallèle pour étudier numériquement la macro dispersion dans les aquifères

Anthony Beaudoin¹, Jean Raynald de Dreuzy² & Jocelyne Erhel³

¹ LMPG - Université du Havre - anthony.beaudoin@univ-lehavre.fr

² Géosciences - Université de Rennes 1 - aupepin@univ-rennes1.fr

³ IRISA - INRIA de Rennes - joceylne.erhel@irisa.fr

Abstract :

The heterogeneity of natural geological formations has a major impact in the contamination of groundwater by migration of pollutants. In order to get an asymptotic behavior of the solute dispersion, numerical simulations require large scale computations. We have developed a fully parallel software : parallel flow model and parallel transport model. Our performance results on a distributed memory parallel architecture show the efficiency of our two parallel models, for the whole range of geological parameters studied.

Résumé :

L'hétérogénéité des formations géologiques a un fort impact sur la contamination des nappes phréatiques par des polluants. Pour déterminer les lois asymptotiques de la dispersion des polluants, les logiciels doivent réaliser des grandes simulations numériques. Nous avons développé un logiciel complètement parallélisé : modèles parallèles Ecoulement et Transport. L'étude des performances sur une architecture à mémoire partagée montre l'efficacité de nos deux modèles parallèles, pour une large gamme de valeurs des paramètres géologiques étudiés.

Key-words :

macro dispersion, heterogeneous saturated porous media, Monte Carlo simulations

1 Introduction

Dans un sol saturé, la migration d'un soluté inerte est gouvernée par deux mécanismes. La convection est un processus d'entraînement du soluté par l'écoulement souterrain. La dispersion est un processus d'étalement du soluté. A l'échelle des pores, l'hétérogénéité de la distribution de la taille des grains provoque des variations de l'écoulement, engendrant un étalement du soluté, appelé dispersion cinématique. A l'échelle de l'aquifère, l'hétérogénéité des formations géologiques provoque des variations de la perméabilité générant un écoulement hétérogène. Cela induit un étalement du soluté, appelé macro dispersion (Salandin *et al.* (1998)). La difficulté d'accéder au milieu naturel rend impossible l'étude in-situ de ces mécanismes. Les simulations numériques de type Monte Carlo sont alors une solution de substitution intéressante. Le principe du Monte Carlo est de répéter plusieurs fois quatre étapes de calcul : générer une perméabilité aléatoire sur une grille, résoudre le problème d'écoulement sur cette grille, évaluer la trajectoire de particules injectées dans cette grille et déterminer la macro dispersion à partir des trajectoires (Gelhar (1993)). La quantification de la macro dispersion nécessite de résoudre les problèmes de flux et de transport sur des grilles de très grandes tailles aboutissant à des systèmes linéaires de plusieurs millions d'inconnues. Cela pose deux problèmes : espace mémoire limité sur un ordinateur parallèle et temps de résolution du système linéaire important. Des progrès récents sur les solveurs parallèles et l'architecture des ordinateurs parallèles permettent de résoudre ces difficultés. Nous utilisons ces progrès pour développer un logiciel de calcul

parallèle. Les simulations numériques réalisées à partir de ce logiciel permettent effectivement de quantifier la macro dispersion dans des milieux hétérogènes.

2 Modèle parallèle

2.1 Modèle physique

2.1.1 Modèle Ecoulement

Le domaine de calcul est une grille rectangulaire 2D, de dimensions L_x et L_y (en m). L_x et L_y sont respectivement parallèle et perpendiculaire à la direction de l'écoulement. Le milieu poreux est supposé isotrope et caractérisé par une perméabilité aléatoire K (en m/s). Cette perméabilité K suit une distribution logarithmique $Y = \ln(K)$, définie par une moyenne m_y (en m/s) et une fonction de covariance C_y :

$$C_y(\mathbf{r}) = \sigma_y^2 \exp(-|\mathbf{r}|^2/\lambda_y) \quad (1)$$

σ_y^2 et λ_y désignent la variance et la longueur de corrélation. La longueur de corrélation λ_y est comprise entre 0.1 m et 100 m (Tompson *et al.* (1998)); et la variance σ_y^2 entre 0 et 7 (Neuman *et al.* (1993)). $\mathbf{r}$ est un vecteur de séparation (en m) entre deux vecteurs positions $\mathbf{X}$ et $\mathbf{X}'$. Etant permanent et sans sources ni puits, l'écoulement est décrit en couplant l'équation de conservation de masse avec la loi de Darcy :

$$\nabla \cdot \mathbf{V} = 0 \quad \text{avec} \quad \mathbf{V} = -K \nabla h \quad (2)$$

$\mathbf{V}$ et h désignent la vitesse de Darcy (en m/s) et la charge hydraulique (en m). Pour les conditions aux limites, on impose une condition de Neumann $\partial h / \partial n = 0$ sur les frontières supérieure et inférieure. Sur les autres frontières, on impose deux conditions de Dirichlet $h = 0$ et 1 m.

2.1.2 Modèle Transport

Le soluté étant inerte, sa migration est gouvernée par la convection, la diffusion moléculaire et la dispersion cinématique. Pour simplifier l'étude, la dispersion cinématique est négligée. A l'échelle de Darcy, cette migration de soluté est décrite par une équation moyenne ou homogénéisée, appelée équation de convection diffusion :

$$\frac{\partial C}{\partial t} + \nabla \cdot (C\mathbf{V}) - \nabla \cdot (D \nabla C) = 0 \quad (3)$$

C (en g/m^3) et D (en m^2/s) sont la concentration du soluté et le tenseur de diffusion. Dans le cas d'une diffusion homogène et isotrope, le tenseur de diffusion D est donné par $D = D_m I$ avec D_m un coefficient de diffusion effective et I le tenseur unité. Le coefficient D_m est proportionnel au coefficient de diffusion moléculaire et aussi à des paramètres microstructuraux comme la tortuosité. Pour les conditions aux limites, la condition initiale est l'injection du soluté. Le ratio entre la diffusion moléculaire et la convection est mesuré par le nombre de Péclet $Pe = \lambda U / D_m$ dont les valeurs sont comprises entre 100 et l'infini. Infini désigne un régime convectif pure.

2.2 Modèle numérique

2.2.1 Modèle Ecoulement

La perméabilité K est générée avec une simulation spectrale FIM (Fourier Integral Method) (Yao (2004)). La transformée de Fourier est évaluée à l'aide de la librairie FFTW (Frigo *et*

al. (2005)). Cette librairie calcule la transformée de Fourier sur un cluster de processeurs. La grille est alors distribuée sur les processeurs selon les colonnes. Cela résoud le problème de stockage mémoire des grilles de calcul. Les équations d'écoulement sont discrétisées sur une grille régulière en utilisant une méthode de volumes finis, permettant de conserver localement la masse et les flux sur les éléments. Dans les deux directions, la maille de la grille est fixée à $\Delta x = \Delta y = 1$ m. Le système linéaire $Ax = B$ est caractérisé par une matrice creuse A , définie positive et symétrique. La solution x est la charge hydraulique et le second membre B représente les conditions de Dirichlet. Nous résolvons ce système linéaire avec la librairie HYPRE contenant un algorithme parallèle multigrille V-cycle, appelé SMG (Structure Multi-Grid) (Falgout *et al.* (2005)).

2.2.2 Modèle Transport

L'équation de transport est résolue avec une méthode particulière dont les deux avantages sont absence de diffusion numérique et bonne performance pour une précision correcte. Dans cette représentation lagrangienne, l'équation de transport est discrétisée avec un schéma explicite d'ordre 1 (Pokrajac (2002)) :

$$\mathbf{X}(t + dt) = \mathbf{X}(t) + dt \cdot \mathbf{V} + \sqrt{2D_m dt} Z \mathbf{w} \quad (4)$$

$\mathbf{X}$ est la position d'une particule. Z et w désignent respectivement un nombre aléatoire obtenu à partir d'une distribution gaussienne et un vecteur unitaire. Le pas de temps dt (en secondes) est proportionnel soit au temps de diffusion nécessaire pour sortir d'une maille ou soit au temps convectif qui est égale à la taille d'une maille divisée par le minimum des vitesses calculées sur une maille. A l'intérieur d'une maille, la vitesse des particules est obtenue par une simple interpolation linéaire respectant la conservation de la masse (Pollock (1988)). Ce modèle Transport est facilement parallélisable puisque les particules inertes n'interagissent pas entre elles. La difficulté réside dans l'accès à des données distribuées sur les processeurs. Gérant un sous domaine de la grille, chaque processeur calcule la trajectoire de la particule présente dans son sous domaine. Il communique aux processeurs voisins les particules sortantes de son sous domaine par des opérations, send et recv, non bloquantes et assurées par la librairie MPI.

2.3 Analyse de la performance

2.3.1 Modèle Ecoulement

Pour analyser la performance de notre modèle parallèle Ecoulement, nous comparons le solveur parallèle HYPRE à un autre solveur parallèle PSPASES. PSPASES utilise une méthode directe basée sur la factorisation de Cholesky $A = LL^T$. Pour les simulations numériques, nous fixons la variance σ_y^2 à 1 et les dimensions de la grille $L_x = L_y = 512$ à 2048 m. Le nombre de processeurs P varie de 2 à 64. Les temps CPU (en secondes) obtenus sont représentés sur la fig. 1. Pour des petites grilles, PSPASES est plus efficace que HYPRE. Pour des grilles moyennes, les deux solveurs donnent des temps CPU similaires. Cependant PSPASES reste plus efficace pour un grand nombre de processeurs. Pour des grandes grilles, HYPRE est plus efficace et PSPASES ne tourne que sur un grand nombre de processeurs à cause d'un problème d'espace mémoire.

2.3.2 Modèle Transport

Pour analyser la performance de notre modèle parallèle Transport, nous injectons 1000 particules dans une grille de dimensions $L_x = L_y = 1024$ et 4096 m. Le nombre de Péclet Pe est

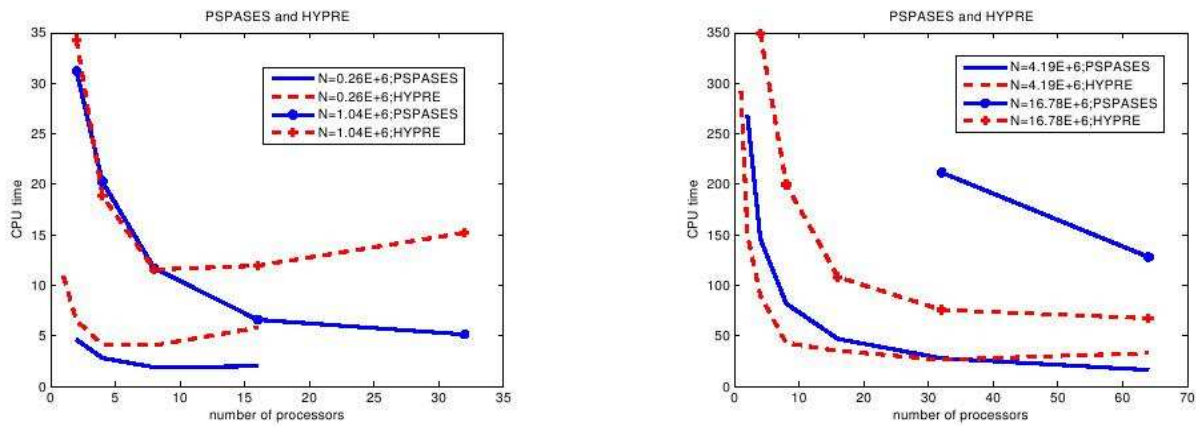


Figure 1: Temps CPU (en secondes) de PSPASES (lignes bleues) et de HYPRE (lignes rouges) pour un nombre variable de processeurs P avec une variance $\sigma_y^2 = 1$ et une taille variable de système N .

fixé à 100 et l'infini. La variance σ_y^2 varie de 0,25 à 9. Le nombre de processeurs P varie de 2 à 32. Les temps CPU (en secondes) obtenus sont représentés sur la fig. 2. Dans le cas $Pe = \infty$, la performance est bonne pour toutes les simulations. L'augmentation de la valeur de la variance provoque une légère augmentation du temps CPU et une légère détérioration de la performance. Dans le cas $Pe = 100$, on observe la même tendance. Les temps CPU sont plus importants car il faut générer des nombres aléatoires pour simuler la diffusion et les particules mettent plus de temps à sortir de la grille à cause de la diffusion.

2.4 Analyse de la macro dispersion

L'objectif de notre étude est d'évaluer les macro dispersions (longitudinale $D_L(t)$ et transversale $D_T(t)$) et de déterminer leur régime asymptotique. Ces macro dispersions (en m^2/s) sont évaluées à partir des trajectoires des particules (voir fig. 3) :

$$D_L(t) = \frac{1}{M} \sum_{j=1}^M \frac{d(\langle x^2(t) \rangle_j - \langle x(t) \rangle_j^2)}{dt} \quad \text{avec} \quad \langle x^k(t) \rangle_j = \frac{1}{Np} \sum_{i=1}^{Np} x_i(t)_j^k \quad (5)$$

M et Np sont le nombre de simulations réalisées et le nombre de particules injectées. j et i sont le numéro de la simulation réalisée et le numéro de la particule injectée. k désigne l'ordre du moment calculée (1 ou 2). $x_i(t)$ est l'abscisse de la particule i . Pour la macro dispersion transversale, on remplace L et x par T et y dans les expressions. Sur la fig. 4, nous avons représenté ces deux macro dispersions pour une variance σ_y^2 variant de 0,25 à 9. Le nombre de Péclet Pe est infini. Pour observer les régimes asymptotiques de ces macro dispersions, nous avons dû injecter 10 000 particules sur des grilles de dimensions $L_x = 16384$ m et $L_y = 8192$ m à chaque simulation et réaliser 100 simulations. Sur un cluster SUN de 2 noeuds de 32 machines (2,2 Ghz AMD Opteron bi-processeurs avec 2 Go de RAM), cela a nécessité un temps de calcul de 149 600 secondes soit environ 42 heures.

3 Conclusion

Nous avons développé un logiciel complètement parallèle pour simuler l'écoulement et le transport de soluté inerte dans des milieux poreux saturés hétérogène. Dans ce papier, nous présentons les deux modèles parallèles : Ecoulement et Transport. Pour une large gamme de valeurs

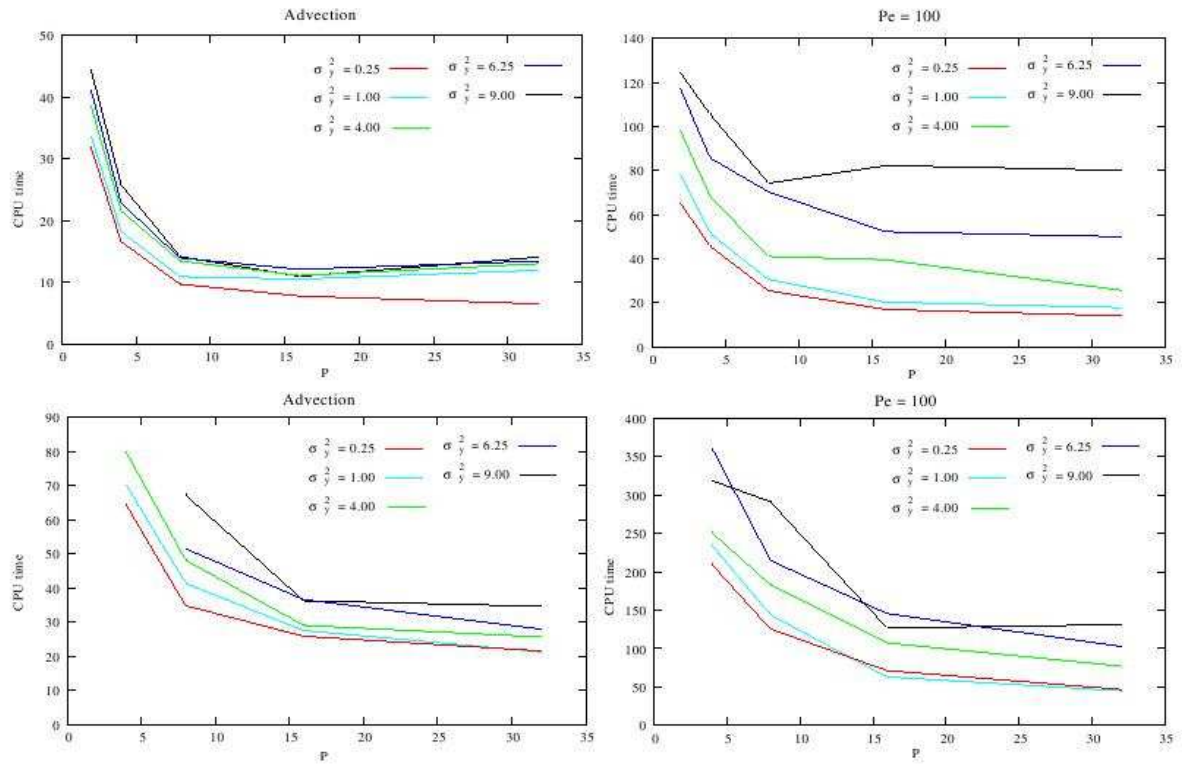


Figure 2: Temps CPU (en secondes) pour un nombre variable de processeurs P avec une variance σ_y^2 comprise entre 0,25 et 1, et un nombre de Péclet $Pe = 100$ ou infini (Figures haut : $L_x = L_y = 1024$ m - Figures bas : $L_x = L_y = 4096$ m).

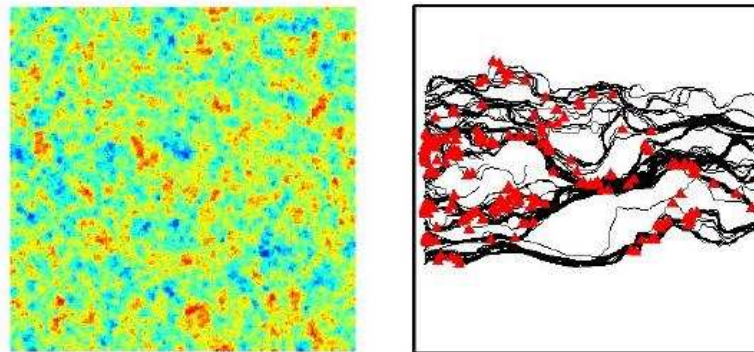


Figure 3: Exemple de simulation numérique (Figure gauche : champ de perméabilité aléatoire - Figure droite : nuage de particules injectées).

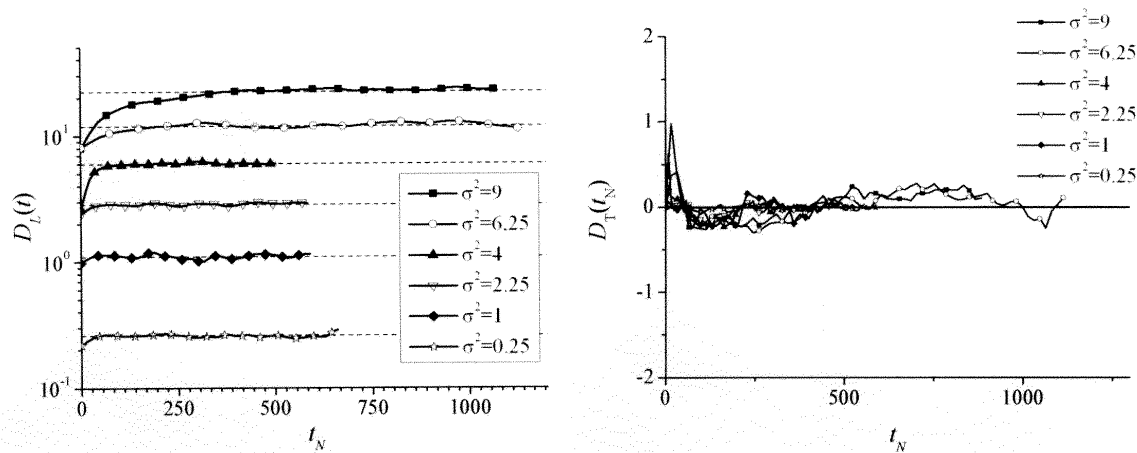


Figure 4: Evolution en fonction de $t_N = t/(\lambda_y/V_{moy})$ des macro dispersions (en m^2/s), longitudinale $D_L(t)$ et transversale $D_T(t)$, pour différentes valeurs de la variance σ_y^2 .

des paramètres géologiques étudiés, l'étude des performances montre que nos deux modèles sont efficaces et robustes. Cela nous permet de réaliser des simulations sur des grilles de calcul de grande taille et ainsi remonter sans ambiguïté aux valeurs asymptotiques des macro dispersions longitudinale et transversale.

References

- Falgout, R.D., Jones, J.E., Yang U.M. 2005 Pursuing scalability for hypre's conceptual interfaces. *ACM Trans. Math. Softw.* **31** 326-350
- Frigo, M., Johnson, S.G. 2005 The design and implementation of FFTW3. *Proceedings of the IEEE* **93** 216-231
- Gelhar, L. 1993 Stochastic Subsurface Hydrology. *Engelwood Cliffs, New Jersey*
- Neuman, S.P., Orr, S. 1993 Prediction of steady state flow in nonuniform geologic media by conditional moments : exact nonlocal formalism, effective conductivities and weak approximation. *Water Resources Research* **29** 341-364
- Pollock, D.W. 1998 Semianalytical computation of path lines for finite difference models. *Ground Water* **26** 743-750
- Pokrajac, D. 2002 An efficient algorithm for high accuracy particle tracking in finite elements *Advances in Water Resources* **25** 353-369
- Salandin, P., Fiorotto, V. 1998 Solute transport in highly heterogeneous aquifers *Water Resources Research* **39**
- Tompson, A., Falgout, R.D., Smith, S.G., Bosl, W.J., Ashby S.F. 1998 Analysis of subsurface contaminant migration and remediation using high performance computing. *Advances in Water Resources* **22** 203-221
- Yao, T. 2004 Reproduction of the mean, variance and variogram model in spectral simulation. *Mathematical Geology* **36** 487-505