Subsurface flow with uncertainty : applications and numerical analysis issues

Julia Charrier
I2M, Université Aix-Marseille
collaboration with A. Debussche, J. Erhel, J-R. De Dreuzy, A. Beaudoin, G. Pichot

CEMRACS 2017

1. Subsurface flow with uncertainty : a simple model
   - Heterogeneities in hydrogeology
   - The general model
   - A classical model for the permeability law
   - Examples of applications
2. Basic and important mathematical properties
   - Spatial regularity issues
   - Unboundedness isuues
   - High dimensionnality of the randomness
3. A detailed application to hydrogeology : study of the migration of pollutants in an aquifer
   - Presentation of the problem
   - Description of the numerical method
   - Numerical analysis of the method

1 Subsurface flow with uncertainty : a simple model
- Heterogeneities in hydrogeology
- The general model
- A classical model for the permeability law
- Examples of applications

2 Basic and important mathematical properties

3 A detailed application to hydrogeology : study of the migration of pollutants in an aquifer

What is hydrogeology ?

"The word "hydrogeology" can be understood as a combination of "hydraulics" and "geology"...
"Hydrogeology" is thus the science where the two are combined : finding the solution of the flow (and transport) equations in a complex, only partly identified, geological system".

Dealing with spatial heterogeneity,
Marsily *et al.*,
Hydrogeology Journal, 2005

What is hydrogeology ?

- Hydraulics : well-known PDEs (Darcy, transport,...)
- Geological properties : lack of data (few well tests/pumping tests) and more importantly multiple scales of Heterogeneity (frome pore to aquifer)

How to deal with heterogeneity in hydrogeology ?

Early approach : average to get equivalent homogeneous properties

- The experiments themselves provide an average (Darcy 1856 until $\simeq$ 1960)
- Quantitative hydrogeology and numerical modelling (from $\simeq$ 1960) :
  - A few pumping test values (wells) provide local averaged permeability values which are interpolated in the aquifer.
  - Matheron 1967 : aquifer properties are described as random variables (for the purpose of averaging and not to describe heterogeneity). Lognormal laws are used to describe the permability.
- This approach works quite well to compute pressure
- It is limited for predictions of oil recovery, groundwater contamination problems
  $\rightsquigarrow$ such problems are very sensitive to high-permeability channels, faults, law permeability barriers
- pressure variations due to heterogeneity are small, whereas those of velocities and travel times are large

Stochastic modeling to describe heterogeneity

- Geostatistics (Delhomme 1976...) and stochastic hydrogeology (Gelhar 1976, Dagan 1985...) use random variables (lognormal laws) taking into account spatial covariance to descibe heterogeneity.
- Heterogeneity can be described by a "structure" defined by the spatial covariance (Marsily 1986, Chiles 1999...)
- Calibration of the parameters (inverse problems) and provide Monte-Carlo simulations of aquifer models to estimate uncertainty on the flow and transport (Delhomme 1979,Ramaro et al. 1995, Zimmermann et al. 1998...)
- "Geostatistics make better use of the data without askig for more" : few degrees of freedom (variance range and type of covariance)

Groundwater steady flow

- We consider an isotropic porous medium with constant porosity and permability **a**.
- We denote by **u** the hydraulic head and by **v** the Darcy velocity.
- **Darcy law** $v = -a\nabla p$ together with **mass conservation** $\mathrm{div}\ v = 0$ yield the steady flow equation:

$$\mathrm{div}(a(x)\nabla u(x)) = 0 \quad \forall x \in D \subset \mathbb{R}^d$$
$$+\text{boundary conditions.}$$

$\rightsquigarrow$ in practice : typically we take a box with mixed homogeneous Neumann conditions (up/down) and non-homogeneous Dirichlet conditions (left/right).

Uncertainty in groundwater steady flow

- Because of :
  - the heterogeneity of the permeablity a
  - the lack of data

  $\rightsquigarrow$ we use a stochastic model : the permeability **a** is not known exactly instead we suppose that we know its law : the permeability is then a random field $a(\omega, x)$
- More precisely a is a function $a : \Omega \times D \to \mathbb{R}$. We can see a as a random variable taking values into $\mathcal{C}^0(\bar{D})$
- the steady flow equation with uncertainty is then:

$$\mathrm{div}_x(a(\omega, x)\nabla_x u(\omega, x)) = 0 \quad \forall x \in D \subset \mathbb{R}^d, \ \omega \ a.e.$$

$+$boundary conditions.

Equation and first assumptions

- Let D be a bounded open $\mathcal{C}^2$ domain of $\mathbb{R}^d$, $(\Omega, \mathcal{F}, P)$ a probability space and $f \in L^2(D)$.
- We look for $u : \Omega \times D \to \mathbb{R}$ such that for almost every ω

$$\left\{ \begin{array}{rcll} -\mathrm{div}_x(a(\omega,x)\nabla u_x(\omega,x)) & = & f(x) & x \in D \\ u(\omega,.) & = & 0 & \text{on } \partial D. \end{array} \right.$$

- Remark : to ensure that the equation is well posed we will require that we have for almost all ω : $a(\omega,\cdot) \in L^\infty(D)$

$$0 < a_{min}(\omega) \leq a(\omega,x) < a_{max}(\omega) < +\infty,$$

for almost every x.

A model for the law of the permeability : lognormal law

- A widely used model in hydrogeology : $a : \Omega \times \bar{D} \to \mathbb{R}$ is a lognormal homogeneous random field, which means that : $a(\omega, x) = e^{g(\omega,x)}$, where g is a gaussian random field,
  $\leadsto$ i.e. any linear combination $\lambda_1 g(x_1, \omega) + ... + \lambda_n g(x_n, \omega)$ is a gaussian random variable.
- The law of g is determined by its expected value $x \mapsto \mathbb{E}[g(\omega, x)]$ and its covariance function :
  $\mathrm{cov}[g](x, y) = \mathbb{E}[(g(\omega, x) - \mathbb{E}[g(\omega, x)])(g(\omega, y) - \mathbb{E}[g(\omega, y)]]$
- we suppose g to be homogeneous,
  i.e. $\mathrm{cov}[g]$ only depends on $\|x - y\|$.
  $\leadsto$ it means that the law of the permeability field is invariant by any affine isometry of the spatial domain.
- Such a choice enables to modelize such very heterogeneous fields

Typical examples of covariance functions (Mattern class)

- the widely use case of an exponential covariance (Hoeksema et al. 1985, Gelhar 1986...) : $\mathrm{cov}[g](x,y) = \sigma^2 e^{-\frac{\|x-y\|}{\lambda}}$.
- the case of a gaussian covariance $\mathrm{cov}[g](x,y) = \sigma^2 e^{-\left(\frac{\|x-y\|}{\lambda}\right)^2}$.

$\rightsquigarrow$ two examples leading to very different mathematical and numerical properties.

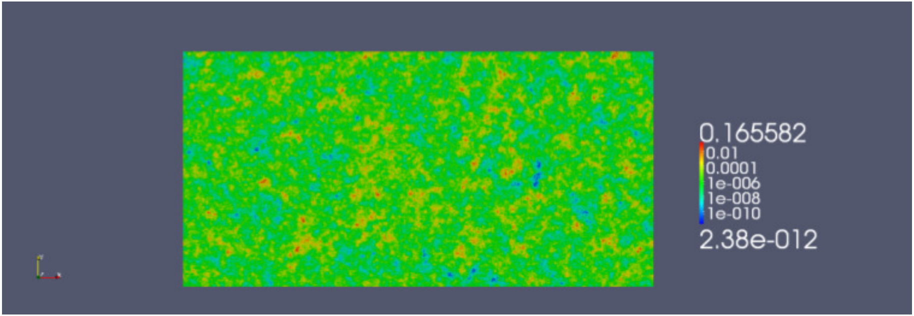

Figure : case of an exponential covariance with $\lambda = 10$, $\sigma = 3$, $\mathbb{E}[g] = -14$, realized by Géraldine Pichot with the software platform H2OLAB

Domains of application

- Study of groundwater pollution (managment of groundwater ressources)
- Oil and gas recovery
- Storage of nuclear waste
- Geological sequestration of carbon dioxide
- ...

Various quantities of interest

- The law of $u : \Omega \to H^1(D)$:
  the law of u is determined by the knowledge of all the values of the $\mathbb{E}[\varphi(u)]$ for a certain class of functions φ
- In practice it is not possible (and not really interesting) to know the law completely. We consider only particular test functions φ.
  - mean point values of the pressure $\mathbb{E}[u(x)]$
  - variance of point values of the pressure $\mathbb{E}[(u(x) - \mathbb{E}[u(x)])^2]$,
  - mean value of some norm of the pressure (typically $\mathbb{E}[\|u\|_{L^2(D)}]$,$\mathbb{E}[\|u\|_{L^\infty(D)}]$...),
  - mean value of outflow through a part Γ of the boundary $\mathbb{E}\left[\int_\Gamma -a(\omega,x)\nabla u(\omega,x).d\vec{\nu}\right]$
- exit times of transported particles (with or without diffusion).
- Failure probabilities and cdf.
  - pressure at a fault $\mathbb{P}(p(x) \geq c)$
  - exit time $\mathbb{P}(T_{exit} \leq t_{critical})$
- Densities of some functionnals of the solution.
- Rare events (failure probabilities, rare events).
- Inverse problems.

1 Subsurface flow with uncertainty : a simple model

2 Basic and important mathematical properties
- Spatial regularity issues
- Unboundedness isuues
- High dimensionnality of the randomness

3 A detailed application to hydrogeology : study of the migration of pollutants in an aquifer

Spatial regularity of lognormal fields

Since g is homogeneous we can write $\mathrm{cov}[g](x,y) = k(x-y)$.
We suppose in what follows that $k \in \mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R})$. We deduce the following spatial regularity result :

Proposition

There exists a version of a, still denoted a, such that for almost all ω, $a(\omega,.) \in \mathcal{C}^{0,\beta}(D)$ for any $\beta < \frac{1}{2}$.

Proof : Since k a L-lipschitz continuous function, we deduce that

$$\begin{aligned} \mathbb{E}[|g(x)-g(y)|^2] &= \mathbb{E}[g(x)^2] - 2\mathbb{E}[g(x)g(y)] + \mathbb{E}[g(y)^2] \\ &= 2(k(0) - k(x-y)) \leq 2L\|x-y\|. \end{aligned}$$

$g(x) - g(y)$ is a mean-free gaussian random variable, so for any positive integer p,

$$E[|g(x)-g(y)|^{2p}] \leq c_p(2L)^p\|x-y\|^p.$$

Proof of the spatial regularity result

We use the Kolmogorov's continuity theorem

Theorem (Kolmogorov)

Let $X(\omega, x) \; : \; \Omega \times D \subset \mathbb{R}^d \to \mathbb{R}^n$ be a stochastic process such that there exists constants C, $p > 1$ and $\varepsilon > 0$ such that for any $x, y \in D$ we have

$$\mathbb{E}[\|X(\omega, x) - X(\omega, y)\|^p] \leq C\|x - y\|^{d+\varepsilon},$$

then X admits a version $\tilde{X}$ such that for almost all ω, $\tilde{X}(\omega, .) \in \mathcal{C}^{0,\beta}(\bar{D})$ for any $\beta < \frac{\epsilon}{p}$.

and deduce that here exists a version of g which is a.s. Hölder-continuous with any exponent $\beta < \frac{p-d}{2p}$, it remains to let $p \to +\infty$.

More spatial regularity

Using the same arguments we can prove that if the covariance function is more regular (more precisely if the function k is more regular), the realizations of a will also have more spatial regularity.

Proposition

If $k \in \mathcal{C}^{n,\alpha}(\mathbb{R})$ (with $\alpha > 0$) then there exists a version of a, still denoted a such that for almost all ω, $a(\omega, .) \in \mathcal{C}^{[\beta],\beta-[\beta]}(D)$ for any $\beta < \frac{n+\alpha}{2}$.

Coming back to our two examples :

- exponential covariance : $\mathrm{cov}[g](x,y) = \sigma^2 e^{-\frac{\|x-y\|}{\ell}}$.
  $\mathrm{cov}[g]$ has only lipschitz regularity (as a function of $x - y$)
  $\Rightarrow$ the realizations of a are only Hölder continuous on $\bar{D}$ with any exponent $\beta < 1/2$.
- gaussian covariance : $\mathrm{cov}[g](x,y) = \sigma^2 e^{-\left(\frac{\|x-y\|}{\ell}\right)^2}$.
  $\mathrm{cov}[g]$ has $\mathcal{C}^\infty$ regularity (as a function of $x - y$)
  $\Rightarrow$ the realizations of a are $\mathcal{C}^\infty$ on $\bar{D}$.

Lognormal fields are not uniformly bounded from above or below

- a is neither uniformly bounded from above nor below with respect to ω. In other words, $a(\omega, x)$ can be arbitrary close to 0 and arbitrary close to $+\infty$.
- However, since we have seen that for almost every ω, $x \mapsto a(\omega, x)$ is continuous on $\bar{D}$, we can then define for almost all ω: $a_{min}(\omega) = \min_{x \in \bar{D}} a(\omega, x) > 0$ and $a_{max}(\omega) = \max_{x \in \bar{D}} a(\omega, x) < +\infty$.
- Conclusion : $\frac{1}{a_{min}(\omega)} \notin L^\infty(\Omega)$ and $a_{max}(\omega) \notin L^\infty(\Omega)$.

An integrability result

Proposition

$\frac{1}{a_{min}(\omega)} \in L^p(\Omega)$ and $a_{max}(\omega) \in L^p(\Omega)$ $\forall p > 0$. Moreover we have $a \in L^p(\Omega, \mathcal{C}_0^{0,\beta}(\bar{D}))$ $\forall p > 0$ and $\beta < 1/2$.

Proof :
$g : \Omega \to \mathcal{C}^0(\bar{D})/\mathcal{C}_0^{0,\beta}(\bar{D})$ is a gaussian random variable which takes values into a separable Banach space. We can apply Fernique's theorem

Theorem (Fernique)

If E is a separable Banach space and X a mean-free gaussian random variable with values in E, for any finite $p > 0$, we have $\mathbb{E}[e^{p\|X\|_E}] < \infty$.

An important consequence : an integrability property of the solution

Proposition

The equation

$$\begin{cases} -\mathrm{div}(a(\omega,.)\nabla u(\omega,.)) & = & f(x) & \text{on } D \\ u(\omega,.) & = & 0 & \text{on } \partial D. \end{cases}$$

admits a unique solution $u \in L^p(\Omega, H_0^1(D)), \forall p > 0.$

Proof

For a.e. ω, the equation admits a unique solution $u(\omega,.) \in H_0^1(D)$ with

$$\|u(\omega,x)\|_{H_0^1(D)} \leq \frac{C_D}{a_{min}(\omega)} \|f\|_{L^2(D)}.$$

Therefore

$$\|u\|_{L^p(\Omega,H_0^1(D))} \leq C_D \|f\|_{L^2(D)} \left\| \frac{1}{a_{min}} \right\|_{L^p(\Omega)} .$$

Approximation of a

We approximate the random field $a(\omega, x)$ by a function of x and of N random variables, i.e. in a finite dimensional stochastic space:

$$a(\omega, x) \rightsquigarrow \tilde{a}(Y_1(\omega), ..., Y_N(\omega), x).$$

Why approximate a?

- It can be used to simulate a: we need only to simulate N random variables to simulate $\tilde{a}$.
- It is the first and fundamental step of several numerical methods: in particular stochastic collocation methods.

Approximation of a using a Karhunen Loève expansion

- We consider the Hilbert-Schmidt operator:

$$f \in L^2(D) \quad \longmapsto \quad \left(x \mapsto \int_D cov[g](x,y) f(y) dy \right) \in L^2(D)$$

- It is a compact self-adjoint operator, hence there exists a sequence $(\lambda_n, b_n)_{n \in \mathbb{N}}$ of eigenpairs such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ and such that $(b_n)_{n \geq 0}$ is an hilbertian basis of $L^2(D)$.
- For $n \in \mathbb{N}$, the normalized coordinate of $g(\omega, \cdot)$ in this hilbertian basis with respect to b_n is $Y_n(\omega) = \frac{1}{\sqrt{\lambda_n}} \int_D g(\omega, x) b_n(x) dx.$
- Here, since g is gaussian, the $(Y_n)_{n \geq 1}$ are more precisely independent gaussian random variables.
- Then the Karhunen-Loève expansion of g is:

$$g(\omega, x) \overset{L^2(\Omega \times D)}{=} \sum_{n=1}^{+\infty} \sqrt{\lambda_n} b_n(x) Y_n(\omega).$$

Decay of the eigenvalues : exponential covariance

- We recall the definition of exponential covariance : $\mathrm{cov}[g](x,y) = \sigma^2 e^{-\frac{\|x-y\|}{\ell}}$. If we take the norm 1, we have analytic expressions and are able to deduce properties.
- We have for some constant c, $\lambda_n \leq \frac{c\sigma^2}{\ell n^2}$ and $\|b_n\|_\infty \leq C$.
- We can also observe numerically a plateau (of size about $1/\ell$) in the decrease of the λ_n

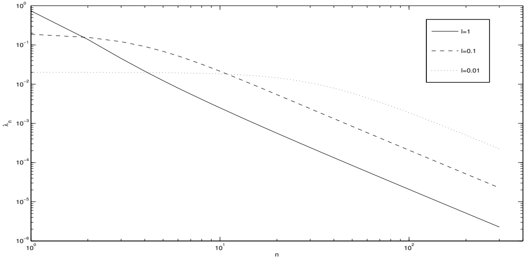

Figure : λ_n versus n, in logarithmic scale, for $\sigma = 1$ and different values of ℓ .

Decay of the eigenvalues : analytic covariance

We suppose that the covariance function $cov[g]$ is analytic on $D \times D$, then:

Theorem (Frauenfelder,Schwab,Todor, 2005)

There exists two constants $c_1, c_2 > 0$ such that for all $n \geq 1$

$$\lambda_n \leq c_1 e^{-c_2 n^{1/d}}.$$

For any $s > 0$ there exists a constant c_s such that for any $n \geq 1$,

$$\|b_n\|_\infty \leq c_s |\lambda_n|^{-s} \text{ and } \|\nabla b_n\|_\infty \leq c_s |\lambda_n|^{-s}.$$

Exemple of analytic covariance function : the gaussian covariance

$cov[g](x,y) = \sigma^2 e^{-\frac{\|x-y\|^2}{\ell^2}}$.

- More generally : the more regular the covariance function is, the most faster the eigenvalues decrease, and the most faster the KL expansion converges.
- The value of the covariance length has also influence on the speed of convergence of the KL expansion.

Approximation of a using a Karhunen Loève expansion

- We define the approximation a_N of a:

$$a_N(\omega, x) = e^{g_N(\omega,x)} = e^{\sum_{n=1}^{N} \sqrt{\lambda_n} b_n(x) Y_n(\omega)}.$$

- We define the approximation u_N of u as the solution of:

$$\begin{cases} -\mathrm{div}(a_N(\omega,.)\nabla u_N(\omega,.)) & = & f(x) & \text{on } D \\ u_N(\omega,.) & = & 0 & \text{on } \partial D. \end{cases}$$

- We have then $\tilde{a}_N(y, x) = e^{\sum_{n=1}^{N} \sqrt{\lambda_n} b_n(x) y_n}$.
- Hence $u_N(\omega, x) = \tilde{u}_N(Y(\omega, x))$ where $\tilde{u}_N$ is the solution of the deterministic parametrized in $\mathbb{R}^N$ PDE.
- The cost of stochastic collocation methods increases with N (exponentially for the basic collocation).

Strong and weak KL troncature error estimates : assumptions

Assumption

i) *The eigenfunctions* b_n *are Hölder continuous with exponent* α_0 *, where* $0 < \alpha_0 < 1/2$

ii) *The series* $\sum_{n\geq 1} \lambda_n \|b_n\|^2_{C^{0,\alpha_0}(\bar{D})}$ *is convergent*

iii) $f \in L^p(D)$ *for some* $p > d$

$\rightsquigarrow$ This is in particular the case for an exponential covariance with norm 1 and for a gaussian covariance .

Strong error bound

Under the previous assumption we define, for α such that $0 \leq \alpha \leq \alpha_0$ and $N \in \mathbb{N}$,

$$R_N^\alpha = \sum_{n>N} \lambda_n \|b_n\|^2_{C^{0,\alpha}(\bar{D})}.$$

Theorem (J.C, A.Debussche, 2014)

There exists for any α,β with $0 < \beta < \alpha < \alpha_0$ and $\beta < 1 - \frac{d}{p}$ a constant $C_s(p, q, \alpha, \beta)$ such that for any $N, q \in \mathbb{N}$ we have

$$\|u - u_N\|_{L^q(\Omega, \mathcal{C}^{1,\beta}(\bar{D}))} \leq C_s(p, q, \alpha, \beta)\sqrt{R_N^\alpha}\|f\|_{L^p(D)}.$$

Weak error bounds

Our goal is to compute the law of u, therefore we consider the weak error, i.e. the error commited by approximating the law of u by the law of u_N.

Theorem (J.C, A.Debussche, 2014)

Let α,β such that $0 < \beta < \alpha < \alpha_0$ and $\beta < 1 - \frac{d}{p}$, then

- *for any $\varphi \in \mathcal{C}^6(\mathbb{R}, \mathbb{R})$ whose derivatives have at most polynomial growth, there exists a constant $C_{w1}(\beta, f, p, \varphi)$ such that for all $N \in \mathbb{N}$*

$$\|\mathbb{E}_\omega[\varphi(u_N) - \varphi(u)]\|_{\mathcal{C}^{1,\beta}(D)} \leq C_{w1}(\beta, f, p, \varphi)R_N^\beta.$$

- *for any $\psi \in \mathcal{C}^4(\mathcal{C}^{0,\beta}(\bar{D}) \times \mathcal{C}^{1,\beta}(\bar{D}), \mathbb{R})$ whose differentials have at most polynomial growth, there exists a constant $C_{w2}(\beta, f, p, \varphi)$ such that for all $N \in \mathbb{N}$*

$$|\mathbb{E}[\psi(a, u) - \psi(a_N, u_N)]| \leq C_{w2}(\beta, f, p, \varphi)R_N^\beta.$$

Conclusion

Specificities of the random fields frequently used in hydogeology need to be taken into account :

- The bounds are not uniform wrt the random variable, it requires to be careful to get L^p bounds.
  $\rightsquigarrow$ you need to track all constants.
- The spatial regularity is often law and depend and the choice of the correlation function.
  $\rightsquigarrow$ you don't always have usual orders for the spatial discretization.
- The dimension of the randomness can be very high and depends strongly on the structure of the correlation function for the asymptotic behaviour but also strongly on the correlation length.
  $\rightsquigarrow$ It has to be considered in the choice of a numerical method.

1 Subsurface flow with uncertainty : a simple model

2 Basic and important mathematical properties

3 A detailed application to hydrogeology : study of the migration of pollutants in an aquifer
- Presentation of the problem
- Description of the numerical method
- Numerical analysis of the method

Transport-diffusion of a solute

- An inert solute is injected in the porous media at initial time : typically a pollutant.
- We consider only molecular diffusion, assumed to be homogeneous and isotropic.
- We denote by $c(\omega, x, t)$ the solute concentration
- The migration of the solute is then described by :

$$\frac{\partial c(\omega, x, t)}{\partial t} + v(\omega, x).\nabla_x c(\omega, x, t) - D\Delta_x c(\omega, x, t) = 0,$$

and boundary/initial conditions
$\rightsquigarrow$ we recall that v is the random Darcy velocity and we consider the advection-dominated case (Peclet number $>> 1$).

Quantities of interest

- First quantity of interest : the mean spread of the solute $S(t) = \mathbb{E}_\omega[S(\omega, t)]$, where

$$S_d(\omega, t) = \int_O c(\omega, x, t)(x_d - G_d(\omega, t))^2 dx$$

and

$$G_d(\omega, t) = \int_O c(\omega, x, t) x_d dx$$

$\rightsquigarrow$ it is the mean value of the spatial spread in each direction.

- Second quantity of interset : the mean macro-dispersion $\mathcal{D}(t)$ defined by

$$\mathcal{D}_d(t) = \mathbb{E}_\omega \left[\frac{dS_d(\omega, t)}{dt} \right].$$

$\rightsquigarrow$ it is the mean value of the speed of spreading in each direction.

- Goal : compute the mean values of spreading and macro-dispersion
$\rightsquigarrow$ asymptotic values of macrodispersion ?
- Question :how do molecular diffusion and high heterogeneity impact assymptotic macrodispersion ?

- A Monte-Carlo method to deal with uncertainty
  N independent realizations of the permeability field a : $a^1,\dots a^N$
- For each i, p_h^i is a finite element approximation (h is the mesh size) of the solution p^i of the flow equation

$$\mathrm{div}(a^i(\omega,x)\nabla p^i(x)) = 0$$

- $v_h^i = -a\nabla p_h^i$ is then the approximation of the Darcy velocity

Approximation of the mean spread

- We note that the solution $c(x,t)$ of

$$\frac{\partial c(x,t)}{\partial t} + v(x).\nabla_x c(x,t) - D\Delta_x c(x,t) = 0,$$

is the density of the solution of the SDE

$$dX(t) = v(X(t))dt + \sqrt{2D}dW(t)$$

- Therefore the spread $\mathcal{S}(t)$ can be expressed as

$$\mathcal{S}(t) = \mathbb{E}[(X(t) - \mathbb{E}[X(t)])^2]$$

- The solution of the SDE is appxoximated through an Euler scheme:

$$Y(t_{k+1}) = Y(t_k) + v_h(Y(t_k))\Delta t + \sqrt{2D}\Delta B_k,$$

- Approximation of the mean spread $\mathcal{S}(t)$:

$$\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(Y^{i,j}(t) - \frac{1}{M}\sum_{j=1}^{M} Y^{i,j}(t)\right)^2$$

Approximation of the macro-dispersion

- Thanks to Itô formula we have

$$\begin{aligned}\mathcal{D}_d(t) &= \mathcal{S}_d'(t)\\ &= \mathbb{E}\left[2(X_d(t)-\mathbb{E}[X_d(t)])(v^i(X_d(t))-\mathbb{E}[v^i(X_d(t)])+2D\right]\\ &= \mathbb{E}[f(X_d(t))],\end{aligned}$$

  where f depends on v.

- Therefore using an Euler scheme and a Monte Carlo method we compute the approximation of the mean macro-dispersion $\mathcal{S}(t)$:

$$\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} f_h^i(Y^{i,j}(t)),$$

  where

$$dY^{i,j}(t_{k+1}) = v_h^i(Y^{i,j}(t_k))\Delta t + \sqrt{2D}\Delta B_k^{i,j}.$$

We consider two assumptions :

- Assumption 1 : there exists $0 < \alpha < 1$, $r > d$ with $\alpha < 1 - \frac{d}{r}$ such that for any finite $q \geq 1$ $a \in L^q(\Omega, \mathcal{C}_b^{0,\alpha}(\mathbb{R}^d))$ and $\frac{1}{a_{min}} \in L^q(\Omega)$.
  $\rightsquigarrow$ it is fulfiled in the case of a lognormal field with exponential covariance
- Assumption 2 : there exists $0 < \alpha < 1$, $r > d$ with $\alpha < 1 - \frac{d}{r}$ such that for any finite $q \geq 1$ $a \in L^q(\Omega, \mathcal{C}_b^{1,\alpha}(\mathbb{R}^d))$ and $\frac{1}{a_{min}} \in L^q(\Omega)$.
  $\rightsquigarrow$ it requires more regularity, it is among others fulfiled in the case of a lognormal field with gaussian covariance.

Proposition

- *Under Assumption 1, we get that* $p \in L^q(\Omega, \mathcal{C}_b^{1,\alpha}(\mathbb{R}^d))$ *and* $v \in L^q(\Omega, \mathcal{C}_b^{0,\alpha}(\mathbb{R}^d))$ *for any finite* $q \geq 1$.
- *Under Assumption 2, we get that* $p \in L^q(\Omega, \mathcal{C}_b^{2,\alpha}(\mathbb{R}^d))$ *and* $v \in L^q(\Omega, \mathcal{C}_b^{1,\alpha}(\mathbb{R}^d))$ *for any finite* $q \geq 1$.

Total error on the generalized spread

We define the generalized spread :

$$\mathbb{E}_\omega[\psi(\mathbb{E}_\xi[\varphi(X(\omega,\xi,T))])].$$

Theorem (J.C. 2015)

Let $\varphi \in \mathcal{C}_b^{1,\alpha}(\mathbb{R}^d, \mathbb{R}^{d'})$ and $\psi \in \mathcal{C}_b^1(\mathbb{R}^{d'}, \mathbb{R}^{d''})$.

- *Under Assumption 1 we have the bound :*

$$\|er\|_{L^2(\Omega\times\Omega')} \leq C\left((\Delta t)^{\frac{\alpha}{2}} + h^\alpha |\ln h| + \frac{1}{\sqrt{M}} + \frac{1}{\sqrt{N}}\right).$$

- *Under Assumption 2 we have the bound :*

$$\|er\|_{L^2(\Omega\times\Omega')} \leq C\left((\Delta t)^{\frac{1+\alpha}{2}} + h|\ln h| + \frac{1}{\sqrt{M}} + \frac{1}{\sqrt{N}}\right).$$

Total error on the generalized macro-dispersion

- For some vector-valued functions φ and ψ we define the generalized spread :

$$\frac{d}{dt}\mathbb{E}_\omega[\psi(\mathbb{E}_\xi[\varphi(X(\omega,\xi,T))])],$$

which thanks to Itô formula is equal to

$$\mathbb{E}_\omega[D\psi(E_\xi[\varphi(X)]).\mathbb{E}_\xi[D\varphi(X).v(X)+D\Delta\varphi(X)]].$$

- The corresponding error is then

$$\bar{E}r = \mathbb{E}_\omega[D\psi(\mathbb{E}_\xi[\varphi(X)])\cdot\mathbb{E}_\xi[D\varphi(X).v(X)+D\Delta\varphi(X)]]$$
$$-\frac{1}{N}\sum_{i=1}^{N}\left[D\psi\left(\frac{1}{M}\sum_{j=1}^{M}\varphi(X_{n,h}^{i,j})\right)\cdot\frac{1}{M}\sum_{j=1}^{M}(D\varphi(X_{n,h}^{i,j})).v_h^i(X_{n,h}^{i,j})+D\Delta\varphi(X_{n,h}^{i,j}))\right]$$

$\rightsquigarrow$ we use similar techniques as for the spread, but we need more regularity.

Total error on the generalized macro-dispersion

Theorem (J.C 2015)

Let $\bar{\varphi} \in \mathcal{C}_b^{3,\alpha}(\mathbb{R}^d, \mathbb{R}^{d'})$ and $\bar{\psi} \in \mathcal{C}_b^2(\mathbb{R}^{d'}, \mathbb{R}^{d''})$, then under Assumption 2, we have

$$\|\bar{E}r(\omega,\xi)\|_{L^2_\omega} \leq C\left((\Delta t)^{\frac{1+\alpha}{2}} + h|ln(h)| + \frac{1}{\sqrt{M}} + \frac{1}{\sqrt{N}}\right)$$

Difficulties :

- v has not enough regularity to apply the classical results
- we need explicit and sharp bounds (because the constant will depend on ω and we need integrability with respect to ω)
- we need to deal with the spatial approximation at the same time as the time discretization in the weak error
- in the case of the macro-spreading, we have to deal with the presence of the random function v_h inside "the test function".

Numerical results and answer to the hydrogeology question

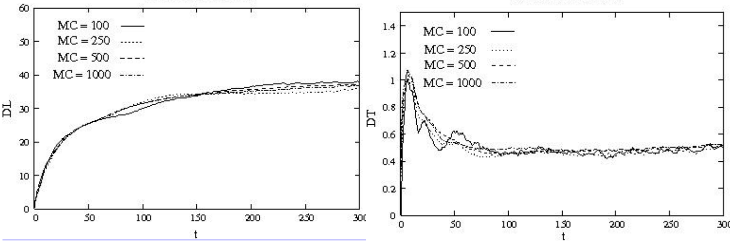

Figure : Case of an exponential covariance with $\lambda = 10$, $\sigma = 1,5$, $D = 0,1$, $d = 2$, 10000 particles, realized with the software platform H20LAB

"For large heterogeneities ($\sigma^2 > 1$), diffusion induces a significant longitudinal macro-dispersion decrease and a transerve macro-dispersion increase larger than expected.".

Asymptotic dispersion in 2D heterogeneous porous media determined by parallel numerical simulations, de Dreuzy *et al.*, Water Resources Research, 2007

Perspectives

- Use spatial ergodicity type properties to get better theoretical convergence rates in order to :
  - justify theoretically the use of stochastic particular methods
  - estimate the speed of convergence to the asymptotic value
  - get sharp a priori error estimates
  - fit the discretization parameters efficiently
- More complicated problems :
  - fractured media
  - chemical reactions
  - add cinematic diffusion
  - computation of exit times

Thank you for your attention !