



**HAL**  
open science

# Parameter Identification by High-Resolution Inverse Numerical Model Based on LBM/CMA-ES: Application to Chalk Aquifer (North of France)

Lahcen Zouhri, Sami Kaidi, Hassan Smaoui

► **To cite this version:**

Lahcen Zouhri, Sami Kaidi, Hassan Smaoui. Parameter Identification by High-Resolution Inverse Numerical Model Based on LBM/CMA-ES: Application to Chalk Aquifer (North of France). *Water*, 2021, 13, 10.3390/w13111574 . hal-03286264

**HAL Id: hal-03286264**

**<https://hal.science/hal-03286264>**

Submitted on 14 Jul 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.



Distributed under a Creative Commons Attribution 4.0 International License

## Article

# Parameter Identification by High-Resolution Inverse Numerical Model Based on LBM/CMA-ES: Application to Chalk Aquifer (North of France)

Lahcen Zouhri <sup>1,\*</sup> , Sami Kaidi <sup>2</sup> and Hassan Smaoui <sup>2</sup> 

<sup>1</sup> AGHYLE, Institut Polytechnique UniLaSalle Beauvais, SFR Condorcet FR CNRS 3417, 19 Rue Pierre Waguet, BP 30313, CEDEX, 60026 Beauvais, France

<sup>2</sup> CEREMA Risques Eaux et Mer (REM) EPR HA, 134 Rue de Beauvais, 60280 Margny-Les-Compiègne, France; sami.kaidi@cerema.fr (S.K.); hassan.smaoui@cerema.fr (H.S.)

\* Correspondence: lahcen.zouhri@unilasalle.fr

**Abstract:** The present paper proposes the numerical solution of an inverse problem in groundwater flow (Darcy's equation). This solution was achieved by combining a high-resolution new code HYSFLO-LBM (Hydrodynamic of Subsurface Flow by Lattice Boltzmann Method), based on LBM, to solve the direct problem, and the metaheuristic optimization algorithm CMA-ES ES (Covariance Matrix Adaptation-Evolution Strategy) to solve the optimization step. The integrated optimization algorithm which resulted from this combination, HYSFLO-LBM/CMA-ES, was applied to the hydrogeological experimental site of Beauvais (Northern France), instrumented by a set of sensors distributed over 20 hydrogeological wells. Hydrogeological parameters measured by the sensors are necessary to understand the aquifer functioning and to serve as input data for the identification of the transmissivity field by the HYSFLO-LBM/CMA-ES code. Results demonstrated an excellent concordance between the integrated optimization algorithm and hydrogeological applied methods (pumping test and magnetic resonance sounding). The spatial distribution of the transmissivity and hydraulic conductivity are related to the heterogeneous distribution of aquifer formations. The LBM and CMA-ES were chosen for their proven excellent performance and lesser cost, in terms of both money and time, unlike the geophysical survey and pumping test. The model can be used and developed as a decision support tool for integrated water resources management in the region.

**Keywords:** groundwater; modeling; Lattice Boltzmann method; chalk aquifer; transmissivity; heterogeneity



**Citation:** Zouhri, L.; Kaidi, S.; Smaoui, H. Parameter Identification by High-Resolution Inverse Numerical Model Based on LBM/CMA-ES: Application to Chalk Aquifer (North of France). *Water* **2021**, *13*, 1574. <https://doi.org/10.3390/w13111574>

Academic Editor: Aldo Fiori

Received: 2 May 2021

Accepted: 24 May 2021

Published: 2 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Chalk formations form the most important and considerable aquifer in the northern part of France as groundwater represents 97% of the water supply in the “Hauts-de-France” region—especially for drinking water, public consumption, agricultural and industrial activities, and supporting river flows. Hydraulic relations between groundwater chalk and rivers allow a favorable exploitation for industrial sectors by pumping the water surface resources.

Groundwater flow is governed by the estimation of hydrodynamic properties, namely the transmissivity, hydraulic conductivity, and water content values. The mapping of these parameters is very complex because: (i) the cost of experimental tests in hydrogeological wells is very high in terms of time and budget, especially in the hydrogeological context related to deeper wells; and (ii) the application of geophysical surveys, especially magnetic resonance soundings (MRS), is not obvious as the quality of results are influenced and perturbed by noise measurements in the field, which depend on urban areas and can makes measurements very difficult to acquire. On the other hand, identified transmissivity values (the important parameter in the pumping test interpretation) are rare in the North of Paris basin (Hauts-de-France region) and, if they exist, are in private scientific reports

and not available in the public domain. Indeed, methods and techniques to estimate this hydrogeological parameter, such as geotechnical approaches which are based on the grain size, the intrinsic permeability, or on the analyses or pumping tests, show several limitations in terms of economics, spatial scale, time, and the heterogeneity of the hydrogeological formations. On the other hand, the pumping test, which generally requires a longer duration of two, three, or more days to produce good-quality results, presents several practical and technical problems, especially in urban sectors (e.g., disturbances to the local population, flooding of the water supply networks, and drawdown in the wells of the city). Therefore, the numerical approach is considered to be the better solution of the transmissivity estimation. Hence, it is necessary to identify other approaches to deriving suitable values in order to understand and to define the principal parameters which govern the heterogeneity of the groundwater flow and, especially, to identify the transmissivity by adopting indirect processes.

Numerous works have focused their theses on inverse modeling, which is an important tool in the management and planning of water resources as it simulates and characterizes hydraulic parameters to erratic properties in hydrogeological models [1–5]. In hydrogeology, inverse modeling is widely used for prospective pumping tests [6,7], to identify the permeability coefficient of rock mass [8], to manage irrigation networks [9], to manage water in the urban context [10], and to analyze the groundwater quality monitoring network [11]. In addition, several numerical models in hydrogeology have used genetic algorithms in order to find the optimal solution of inverse problems in many fields of hydrogeology and in particular to identify the field of transmissivity of aquifers [12].

Solutions of inverse problems are determined in two stages. The first is the resolution of the direct model to estimate the calculated simulated values of state variables necessary for the construction of the objective function (the error between the calculated and observed state values). The second step is the minimization of the objective function whose minimum is the desired identified value. The resolution of the direct problem is generally accomplished by classical discretization methods such as the finite difference method, the finite element method, and the finite volume method. Moreover, for the minimization phase, most of the classical optimization algorithms can be implemented to solve the inverse problem, but have the disadvantage to convergence towards the local minimum. To avoid this drawback, several works have focused on the use of metaheuristic optimization algorithms such as genetic algorithms (GA) [13–15] or evolutionary strategy algorithms (ESA) [16–18].

For two decades, the Lattice Boltzmann method (LBM) has been considered a serious alternative to classical computational fluid dynamic (CFD) methods for solving Navier–Stokes equations in complex flow configurations. Since its introduction, LBM has been successfully used in all disciplines related to flows in general, including porous medium, multiphase flows, wave propagation, multiphysics flows, and many others. The foundations of LBM are rooted in statistical physics and, in particular, kinetic theory. Indeed, LBM reproduces the movement of particles virtually placed on a predefined grid (called a lattice) which collide with each other and then propagate on this grid. In addition, the bases of the kinetic theory, on which LBM is based, consist of solving the Boltzmann equation, which describes the spatio-temporal evolution of the distribution function with a source term which represents the collision operation. By its design, LBM is naturally parallelizable and particularly suited to implementations on GPU (Graphics Processing Unit) architecture. It is this advantage which results in a considerable saving of computing time, which continues to make LBM more and more attractive. There is an abundance of literature on all aspects (theoretical and practical) of LBM, but we will only cite a few books that present it in detail [19–22].

In this work, we present the solution of an inverse problem allowing the identification of the heterogeneous transmissivity field of an experimental site instrumented in 20 wells. To do this, we developed a new integrated optimization algorithm called HYSFLO-LBM/CMA-ES, in which the direct model is solved by the code HYSFLO-LBM

and the metaheuristic optimization algorithm CMA-ES due to [23]. Both the LBM and CMA-ES are presented in Sections 2.2 and 2.3 of this paper. Finally, we underline that the main novelty of the integrated optimization algorithm (HYSFLO-LBM/CMA-ES) lies in the fact that this digital tool allows identification of the transmissivity (or diffusivity) tensor and its spatial heterogeneity. With this numerical tool, we do not use the concept of zonation (the computational domain is represented by a limited number of areas with constant transmissivity), which is not desirable for study areas with high spatial variability.

This paper is organized as follows. Section 1 presents the introduction and the motivation behind the choices of LBM and the CMA-ES algorithm as the basis of the HYSFLO-LBM/CMA-ES code. The second section is devoted to the mathematical description of the different phases implemented to solve the inverse problem. Application of the new integrated optimization algorithm HYSFLO-LBM/CMA-ES to a realistic case is described in Section 3. Finally, Section 4 is dedicated to the discussion.

## 2. Mathematical Formulation of the Model

### 2.1. Governing Equations

The flow of a fluid in a porous medium is formulated by Darcy's law. For a porous medium of negligible deformations crossed by a slow flow fluid, this law expresses that the infiltration velocity  $\vec{q}$  is proportional to the gradient of the hydraulic head  $\nabla\varphi$ . It should be noted that Darcy [24] established this law for a flow of water in a vertical cylindrical column homogeneously filled with sand. Moreover, Darcy's law was also found to be applicable to heterogeneous porous media and non-uniform flow. If we denote by  $\varphi$  the hydraulic head and by  $\overline{T}$  the transmissivity tensor, Darcy's law is then expressed by:

$$\vec{q} = -\overline{T}\nabla\varphi \quad (1)$$

In  $d$ -dimension space, the transmissivity tensor is represented by a matrix  $(t_{ij})_{\substack{1 \leq i \leq d \\ 1 \leq j \leq d}}$ .

Note that the system of Equation (1) is not closed (2 unknowns ( $\vec{q}$  and  $\varphi$ ) and 1 equation). To complete this system, we will use the continuity equation which assumes that the fluid is incompressible, or in other words, the divergence of the velocity is zero ( $\nabla \cdot \vec{q} = 0$ ). Finally, the association of the continuity equation and Darcy's law will give an elliptical partial differential equation (PDE) with unknown  $\varphi$ . For a given transmissivity field  $(t_{ij})_{\substack{1 \leq i \leq d \\ 1 \leq j \leq d}}$ , this PDE makes it possible to calculate the hydraulic head  $\varphi$  and the

velocity of the flow to be deduced in post-processing using Darcy's law. Moreover, if the flow occupies the domain  $\Omega$  ( $\Omega \subset \mathbb{R}^d$ ) of the boundary  $\partial\Omega$  such as  $\partial\Omega = \Gamma_D \cup \Gamma_N$  and  $\Gamma_D \cap \Gamma_N = \emptyset$ , the mathematical model describing the stationary flow in a porous medium is given by:

$$(DP) \begin{cases} -\nabla \cdot (\overline{T}\nabla\varphi) = f \text{ in } \Omega \\ \varphi = \varphi_D \text{ on } \Gamma_D \\ -\vec{n} \cdot (\overline{T}\nabla\varphi) = \varphi_N \text{ on } \Gamma_N \end{cases} \quad (2)$$

where  $\varphi_D$  is a known function for imposing the Dirichlet condition  $\Gamma_D$ .  $\varphi_N$  is also a known function for imposing the Neuman condition on the domain boundary  $\Gamma_N$ .  $f$  is the source/sink term in the domain  $\Omega$ . The vector  $\vec{n}$  denotes the vector normal to  $\Gamma_N$  oriented towards the outside of the domain.

For a given transmissivity field, the mathematical model (2) describing the stationary flow in a porous medium is also called the direct problem, which we denote hereafter as (DP). Alternatively, if the hydraulic head  $\varphi$  is known at a few points in the  $\Omega$  (of the measured values for example), the model (2) allows us, in this case, to determine the tensor of transmissivity  $T$ . In this case, the solution of the inverse problem is denoted by (IP).

Under certain regularity conditions, Ciarlet [25] theoretically showed that the mathematical model (2) admits a unique solution. For more details on the mathematical analysis of the problem (DP), the reader may also consult the book of Dautray and Lions [26].

To solve the direct problem (DP), we developed a numerical model based on the Lattice Boltzmann method. The foundations of this method are presented in the following subsection.

## 2.2. Lattice Boltzmann Method

The Lattice Boltzmann method (LBM) is a relatively new numerical method when compared with the classical approaches used in numerical simulation. It appeared in the 1990's and was initially deduced from the methods of gases on the network and cellular automata [27]. Unlike classical approaches based on the discretization of the Navier–Stokes equations (NSE), the Lattice Boltzmann method is based on the formalism of statistical physics, which consists of the numerical resolution of the Boltzmann equation. This equation is concerned not only with macroscopic quantities (speed, pressure, density), but directly with the distribution of the various particles constituting a fluid. This is called a mesoscopic representation. The simplicity and locality of the LBM method algorithm allows its easy and efficient parallelization. Thus, very quickly, this method was used for unsteady and incompressible CFD calculations [28]. Therefore, as long as the Mach number of the flow remains sufficiently low, LBM allows us, by its nature, to simulate the behavior of a fluid governed by unsteady, weakly compressible Navier–Stokes equations.

The Boltzmann equation (BE) plays a main role in the kinetic theory of gases. In the absence of external force, it expresses the convective transport equation of the velocity distribution  $f(\vec{x}, \vec{c}, t)$ . This distribution is none other than the probability of a particle to be found at the position  $\vec{x}$ , at the instant  $t$ , and with the speed  $\vec{c}$ . The Boltzmann equation is written as:

$$\frac{\partial f}{\partial t} + \vec{c} \cdot \nabla f = \Omega \quad (3)$$

where  $\Omega$  denotes the rate of change of  $f$  due to the collision of particles (also called the collision operator).  $\Omega$  can be approached by several simple models [29], but the most popular in LBM is the Bhatnagar–Gross–Krook (BGK) model (1954). BGK assumes that  $\Omega$  is a function of the  $f$  distribution, the equilibrium distribution  $f^{(eq)}$ , and the relaxation time  $\tau$  (the time necessary to bring the system back to the state of equilibrium), as:

$$\Omega = \frac{(f^{(eq)} - f)}{\tau} \quad (4)$$

The Equations (3), (4), and (6) constitute BE which must be solved numerically.

From the second principle of thermodynamics (or the application of the H-theorem), we can deduce that  $f^{(eq)}$  is of Maxwellian type of the form:

$$f^{(eq)}(\vec{x}, \vec{c}, t) = \frac{\rho}{(2\pi RT)^{d/2}} \exp\left(-\frac{(\vec{c} - \vec{v})^2}{2RT}\right) \quad (5)$$

where  $d$  is the space dimension ( $d = 1, 2, 3$ ),  $\rho$  and  $\vec{v}$  are, respectively, the macroscopic fluid density and velocity,  $R = k_B/m$  with  $k_B$  is the Boltzmann constant,  $m$  is the molecular mass, and  $T$  is the temperature of the system. Note that the artificial sound speed  $c_s$  is defined as  $c_s^2 = RT$ .

Numerical evaluation on a computer is very costly in terms of CPU computing time. We therefore often look for an approximate expression, based on polynomials for example. If we assume that  $\|\vec{v}\| \ll c_s$  then the Taylor expansion of the exponential function allows us to approximate (5) by:

$$f^{(eq)}(\vec{x}, \vec{c}, t) \simeq \frac{\rho}{(2\pi c_s^2)^{d/2}} \exp\left(\frac{-\|\vec{c}\|^2}{2c_s^2}\right) \left(1 + \frac{\vec{c} \cdot \vec{v}}{c_s^2} + \frac{1}{2} \left(\frac{\vec{c} \cdot \vec{v}}{c_s^2}\right)^2 - \frac{1}{2} \frac{\|\vec{v}\|^2}{c_s^2}\right) \quad (6)$$

To solve the BE, we start by defining a computational grid (called a lattice) which will be used for both spatial discretization  $\vec{x}$  and speed discretization  $\vec{c}$ . To recover the NSE by the numerical resolution of BE, it is imperative that the discrete speeds must be chosen so that the following quadratic equation of the moment  $\mathcal{M}^{(k)}$  of order  $(k)$  is exact:

$$\mathcal{M}^{(k)} = \int \vec{c}^k f^{(eq)}(\vec{x}, \vec{c}) = \sum_i \omega_i \vec{c}_i^k f^{(eq)}(\vec{x}, \vec{c}_i) \quad 0 \leq k \leq 3 \tag{7}$$

where  $\omega_i$  and  $\vec{c}_i$  are the coefficients and points of the quadrature Equation (7). In 2D computations, LBM uses the Gauss–Hermite quadrature equation with at least five points for the numerical integration. The quadrature equation being chosen allows us to determine the coefficient  $\omega_i$  and the direction  $\vec{c}_i$ . With these two parameters, we can then discretize BE by setting  $f_i(\vec{x}, t) = \omega_i f(\vec{x}, \vec{c}_i, t)$ , which must verify the BE in turn as:

$$\frac{\partial f_i}{\partial t} + \vec{c}_i \cdot \nabla f_i = \frac{1}{\tau} (f_i^{(eq)} - f_i) \tag{8}$$

In 2D, there are two types of lattice: square and hexagonal. The most commonly used lattice, which leads to more accurate results, is the square lattice D2Q9 (calculation in 2D with a discretization of the velocity variable  $\vec{c}$  in 9 directions, see Figure 1).

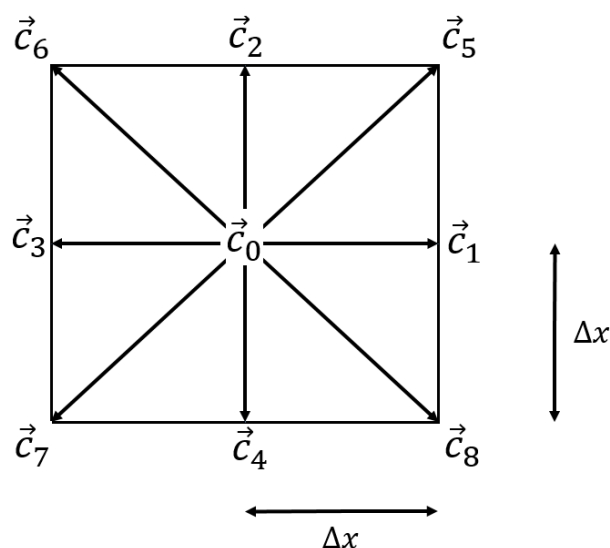


Figure 1. Discretization of the velocity space in 9-speed square lattice (D2Q9 lattice).

If the D2Q9 lattice is adopted  $\omega_i$ ,  $\vec{c}_i$ , and  $f_i^{(eq)}$  are given as:

$$\omega_0 = \frac{4}{9} \quad \omega_{1,2,3,4} = \frac{1}{9} \quad \omega_{5,6,7,8} = \frac{1}{36} \tag{9}$$

$$\vec{c}_i \begin{cases} (0, 0) & i = 0 \\ c \left( \cos \frac{(i-1)\pi}{2}, \sin \frac{(i-1)\pi}{2} \right) & i = 1, 2, 3, 4 \\ \sqrt{2}c \left( \cos \frac{(2i-9)\pi}{4}, \sin \frac{(2i-9)\pi}{4} \right) & i = 5, 6, 7, 8 \end{cases} \tag{10}$$

where  $c$  is lattice velocity  $c = \Delta x / \Delta t$ , with  $\Delta x$  being the lattice size and  $\Delta t$  being the time step.

$$f_i^{(eq)} = \omega_i \rho \left( 1 + \frac{\vec{c}_i \cdot \vec{v}}{c_s^2} + \frac{1}{2} \left( \frac{\vec{c}_i \cdot \vec{v}}{c_s^2} \right)^2 - \frac{1}{2} \frac{\|\vec{v}\|^2}{c_s^2} \right) \text{ for } i = 0, \dots, 8 \tag{11}$$

where  $c_s = c / \sqrt{3}$ . Note that, at this stage, we only give the discretization according to the velocity space. For the spatio-temporal discretization, LBM uses the characteristic method



but without interpolation, since the lattice should be chosen in such a way that  $\Delta \vec{x} = \vec{c} \Delta t$ . Thus, the final discretization of the BE is given by:

$$f_i(\vec{x} + \vec{c}_i \Delta t, t + \Delta t) - f_i(\vec{x}, t) = -\frac{\Delta t}{\tau} (f_i(\vec{x}, t) - f_i^{(eq)}(\vec{x}, t)) \text{ for } i = 0, \dots, 8 \quad (12)$$

Once this equation is solved, the macroscopic density and the velocity of the fluid can be recovered by:

$$\rho = \sum_{i=0}^8 f_i \text{ and } \rho \vec{v} = \sum_{i=0}^8 f_i \vec{c}_i \quad (13)$$

In this paragraph, we wanted to give a general description of LBM, and it is for this reason that we treated it as a fluid flow governed by the macroscopic model of NSE. As the aim of this paper is to solve the Darcy equation by LBM, we have to give only the new expression of  $f^{(eq)}$  (as mentioned above) which allows us to cover Darcy’s macroscopic equation. The Darcy equation is a special case of the NSE for slow flows (low Reynolds number) in which the inertia term is considered negligible. Therefore, the stationary Darcy’s equation is elliptical (flow dominated by diffusion processes) and resembles a diffusion equation. In a D2Q9 lattice, the equilibrium distribution function for a diffusion equation is given by:

$$f_i^{(eq)}(x, y) = \omega_i \varphi(x, y) \text{ for } i = 0, \dots, 8 \quad (14)$$

### 2.3. Solution of Darcy Equation by LBM

The solution of Darcy’s equation by LBM is done in two steps: collision and streaming. If we consider that the transmissivity tensor  $\bar{T}$  is of the form  $\bar{T} = T(x, y) \bar{I}$ . in the D2Q9 lattice, we have:

collision: 
$$f_i(x, y, t + \Delta t) = (1 - \lambda) f_i(x, y, t) + \lambda f_i^{(eq)}(x, y, t) \text{ } i = 0, \dots, 8 \quad (15)$$

streaming: 
$$f_i(x + \Delta x, y + \Delta y, t + \Delta t) = f_i(x, y, t + \Delta t) \text{ } i = 1, \dots, 8 \quad (16)$$

where  $\lambda$  is the dimensionless relaxation time ( $\lambda = \Delta t / \tau$ ). Developments in LBM show that the viscous effects of the Navier–Stokes equations are related to the relaxation time. It is the same for the transmissivity for the Darcy equation. Thus, the macroscopic transmissivity coefficient  $T(x, y)$  can be related to the relaxation factor  $\lambda$  as:

$$T = \frac{\Delta x^2}{3\Delta t} \left( \frac{1}{\lambda} - \frac{1}{2} \right) \quad (17)$$

Finally, the expression (15) is used to compute the equilibrium distribution function  $f_i^{(eq)}$  with the values of the weighting coefficients  $\omega_i$  given by expression (14). The hydraulic head  $\varphi$  can be calculated as:

$$\varphi(x, y) = \sum_{k=0}^8 f_k(x, y) \quad (18)$$

For the boundary conditions, LBM requires conditions on the different directions of the distribution function  $f_i$  while the boundary values are given on the primitive variable of the macroscopic model. It is, then, necessary to transform the boundary values from the macroscopic model to the mesoscopic model. To answer this question, several transformations have been proposed in the literature [30,31]. These transformations allow to impose various types of boundary conditions (Dirichlet, Neuman, Robin, etc.) on the pressure as well as on the velocity of the flow. In this paper, we use the Dirichlet-type boundary conditions on the four boundaries of the studied domain by applying the flux conservation principle which is expressed by:

$$f_i^{(eq)} - f_i = f_{opp(i)} - f_{opp(i)}^{(eq)} \quad (19)$$

where  $opp(i)$  denotes the direction opposite to the direction  $i$ . For instance, suppose we know  $f_7$  and  $\varphi$  on the boundary denoted  $\varphi_{wall}$ , then expression (19) allows us to transform the value of  $\varphi_{wall}$  into  $f_5$  by:

$$f_5^{(eq)} - f_5 = f_{opp(5)} - f_{opp(5)}^{(eq)} = f_7 - f_7^{(eq)} = f_7 - \omega_7 \varphi_{wall}$$

or

$$f_5 = \omega_7 \varphi_{wall} + \omega_5 \varphi_{wall} - f_7 \quad (20)$$

Finally, we give the Algorithm 1 implemented in the computer code HYSFLO-LBM to solve the Darcy equation, which was used for the direct problem.

---

#### Algorithm 1: LMB Steps

---

```

set numerical parameters ( $\omega_i, c_s, \bar{T}, \dots$ )
initialization
initialize  $\varphi_0$  by the measured values
compute  $f_k$  associated to  $\varphi_0$  from (Equation (14))
loop:  $t + \Delta t$ 
compute  $\lambda_{ij}$  from (Equation (17))
compute  $f_k^{(eq)}$  from (Equation (14))
collision from (Equation (15))
streaming from (Equation (16))
set boundary condition from (Equation (19))
compute  $\varphi$  from (Equation (18))
test  $\|\varphi - \varphi_0\| < tol$  exit
set  $\varphi_0 = \varphi$ 
go to the next time step

```

---

#### 2.4. CMAES Algorithm

CMAES belongs to the category of evolution strategy algorithms. It has become the primary algorithm for free-gradient optimization. It is well suited to solving continuous optimization problems where the objective function is not known explicitly and is not necessarily convex. Although it is stochastic in nature, the mutation stage of the CMA-ES algorithm is considered to be correlated and deterministic. This property makes this algorithm easy to implement and less computationally intensive. Moreover, like metaheuristic optimization algorithms, CMA-ES has the advantage of converging towards a global optimum. CMA-ES is gaining popularity and is becoming the benchmark algorithm in metaheuristic optimization. It has been successfully applied to several engineering disciplines including: environmental engineering [32], acoustics [33], electronics [34], hydrogeology [35], medicine [36], thermal and fluid flow [37], structural mechanics and failure [38], and many others. CMA-ES is particularly efficient for non-convex, poorly conditioned, multimodal optimization problems and with noisy evaluations of the objective function.

The CMA-ES algorithm is from the family of strategy evolution algorithms and, like genetic algorithms, is inspired by the Darwinian theory of evolution. CMA-ES is performed in four steps: initialization, selection, recombination, and mutation. These steps operate on a set of  $\mu$  parents to produce  $\lambda$  children, which we now denote by CMA-ES ( $\lambda, \mu$ ). In order to explore the search space, CMA-ES uses candidate populations according to a multivariate random distribution. The mutation step is considered to be the main operation in the CMA-ES algorithm. It allows us to produce a new population by adding a multivariate random vector to the parent population. The mutation also acts as the guide of CMA-ES for the different transformations (rotation and homothetic) of the adapted covariance matrix from the generated population. It should be noted that the mutation process is based on a set of parameters (called strategy parameters) which are updated automatically without user calibration, but by exploiting the various information from previous generations [16].



Like any iterative algorithm, CMA-ES begins with an initialization step for all of the algorithm’s control parameters. This step also consists to initialize a population of  $\lambda$  individuals randomly according to the multivariate normal distribution. Hansen [39] suggests considering a population of size  $\lambda = 4 + 3\ln(d)$ , where  $d$  is the size of the optimization variable. It also indicates that this value is reasonable for large optimization problems. In the rest of this paper, we denote by  $(\mathbf{X}_{i=1,\lambda}^{(g)} \in \mathbb{R}^d)$  the individual’s population of generation  $g$ , and by  $F$  the objective function.

After the initialization stage, CMA-ES enters the loop of generations until convergence towards the optimal solution. This loop begins by evaluating all of the individuals in the population by the objective function to obtain the fitness value  $F(\mathbf{X}_1^{(g)})$ ,  $F(\mathbf{X}_2^{(g)})$ , ...,  $F(\mathbf{X}_\lambda^{(g)})$ , CMA-ES then selects the “best”  $\mu$  ( $\mu \leq \lambda$ ) individual among the  $\lambda$  individuals based on their fitness value (“best” here designates the rank of the  $\mathbf{X}_i^{(g)}$  according to its fitness value: “smaller” if this is a minimization problem or “bigger” if it is a maximization problem). Then, we evaluate the weighted average vector on these  $\mu$  individuals noted  $\mathbf{m}^{(g)}$  and given by:

$$\mathbf{m}^{(g)} = \frac{\sum_{k=1}^{\mu} w_k \mathbf{X}_k^{(g)}}{\sum_{k=1}^{\mu} w_k} \tag{21}$$

where  $w_k$  are the weighting coefficients given by Hansen [39]. It is the expression (21) which will then contribute to the construction of the next generation of parents  $\mathbf{X}_{i=1,\lambda}^{(g+1)}$ .

The CMA-ES algorithm mutation step consists of adding to the population mean vector of the generation ( $g$ ), a noisy component according to a multivariate normal distribution as:

$$\mathbf{X}^{(g+1)} = \mathbf{m}^{(g)} + \sigma^{(g)} \mathcal{N}(\mathbf{0}, \mathbf{C}^{(g)}) \tag{22}$$

where  $\sigma$  is a positive parameter and  $\mathcal{N}(\mathbf{0}, \mathbf{C})$  denotes a vector of independent normal random numbers with zero mean and covariance matrix  $\mathbf{C}$  (symmetric and positive definite).

One can observe that for a good mutation which leads to a rapid convergence towards the optimum, it is necessary to choose judiciously the parameters  $\sigma$  and the matrix  $\mathbf{C}$ . However, in practice this choice turns out to be difficult. It is exactly at this point where the power of the CMA-ES algorithm appears, which automatically adapts these two parameters during successive generations and without user intervention. Thus, the CMA-ES algorithm can be summarized in the following steps named Algorithm 2:

---

**Algorithm 2: CMA-ES Steps**

---

- set numerical parameters ( $d$ : size of the optimization variable)
  - initialization  $\sigma, \mathbf{m}, \mathbf{C}$
  - generation:  $g + 1$
  - selection  $\rightarrow \mathbf{X}, \mathbf{m}, \mu$
  - mutation  $\rightarrow \mathbf{X} = \mathbf{m} + \sigma \mathcal{N}(\mathbf{0}, \mathbf{C})$
  - evaluation  $\rightarrow F(\mathbf{X})$
  - test  $\|F(\mathbf{X})\| < tol$  exit
  - adaptation  $\rightarrow \sigma, \mathbf{C}$
  - go to the next generation
- 

*2.5. The Integrated Optimization Algorithm (HYSFLO-LBM/CMA-ES)*

In hydrogeology, the transmissivity tensor  $\overline{\overline{T}}$  is one of the key parameters in the modeling of groundwater flows. Its estimation with precision is capital to deduce (by the Darcy equation) an accurate velocity field. A transmissivity field identified with good precision also guarantees precision of the diffusion coefficient of the studied porous medium, since the diffusion is deduced from the velocity field (and therefore from the transmissivity). Therefore, the identification with high accuracy of the tensor  $\overline{\overline{T}}$  is also necessary for the prediction of pollutant propagation in a porous medium. It is therefore

important to deploy the most sophisticated numerical tools to identify the transmissivity tensor  $\bar{\bar{T}}$ . This is explained in the following paragraphs.

If we have a series of observed values of the hydraulic potential ( $\varphi_{obs}$ ), we can formulate the identification of  $\bar{\bar{T}}$  by a mathematical optimization problem with constraints. In other words, the problem is equivalent to the search for the tensor  $\bar{\bar{T}}$  which minimizes the error between the observed values ( $\varphi_{obs}$ ) and those computed by the direct problem ( $\varphi_{comp}$ ). Recently, this problem was solved by a few researchers using manual regional adjustment of  $\bar{\bar{T}}$  until an error was obtained which they considered reasonable! With the spectacular advances in the field of mathematical optimization, and especially in metaheuristic algorithms, the identification of the tensor  $\bar{\bar{T}}$  can be treated as the resolution of an inverse problem (IP). In other words, identification of  $\bar{\bar{T}}$  becomes automatic without any manual adjustment. In this case, this problem can be formulated as the minimization of an objective function  $F$  (which is the error between  $\varphi_{obs}$  and  $\varphi_{comp}$ ) with direct model (Equation (2)) itself as a constraint. Thus, the inverse problem to solve is:

$$(IP) \left\{ \begin{array}{l} \text{Minimize } F(\bar{\bar{T}}), \bar{\bar{T}} \in \mathcal{A}_T \\ \text{subject to} \\ (DP) \left\{ \begin{array}{l} -\nabla \cdot (\bar{\bar{T}} \nabla \varphi) = f \text{ in } \Omega \\ \varphi = \varphi_D \text{ on } \Gamma_D \\ -\vec{n} \cdot (\bar{\bar{T}} \nabla \varphi) = \varphi_N \text{ on } \Gamma_N \end{array} \right. \end{array} \right. \quad (23)$$

where  $F$  is the objective function defined by  $F(\bar{\bar{T}}) = \|\varphi_{obs} - \varphi_{comp}(\bar{\bar{T}})\|$  and  $\mathcal{A}_T$  is the admissible set values of  $\bar{\bar{T}}$ .

The integrated optimization algorithm that we propose to the IP (23) is an iterative algorithm that converges to the desired solution  $\bar{\bar{T}}_{opt}$  such that:

$$\bar{\bar{T}}_{opt} = \underbrace{\min}_{\bar{\bar{T}} \in \mathcal{A}_T} \|\varphi_{obs} - \varphi_{comp}(\bar{\bar{T}})\| \quad (24)$$

Note that the problem (23) can be confronted with the problem of computational stability (sensitivity to measurement noise), but also with the problem of the uniqueness of the solution  $\bar{\bar{T}}_{opt}$ . This problem is called, the “ill-posed problem”. In fact, in the groundwater flow, different hydrogeological conditions can provide identical observations of the hydraulic potential or solute concentration. It is then impossible to determine uniquely the transmissivity (or diffusivity) tensor only from observations. Consequently, problem (IP) most often requires a well-chosen regularization strategy to guarantee a certain stability of the result.

Several solutions have been proposed in the literature to remedy the ill-posed problem in groundwater flow modeling. We cite for instance: (i) minimization of non-linearity and non-convexity during the execution of the optimization algorithm by transforming the optimization variable  $\bar{\bar{T}}$  (for example by considering  $Ln(\bar{\bar{T}})$  instead of  $\bar{\bar{T}}$ ); (ii) reduction of the number of unknowns of the transmissivity tensor  $\bar{\bar{T}}$  by adopting a zonation strategy; (iii) making realistic assumptions in order to restrict the range of variation of  $\bar{\bar{T}}$ ; and (iv) application of regularization procedures to reduce the fluctuations induced by the iterations during optimization.

Because it was formulated within a rigorous mathematical framework, the strategy of regularization made it possible to prove both the existence, the uniqueness, and the stability of the solutions of the inverse problems under certain regulatory assumptions. Several regularization strategies have been proposed in the literature. In his book, ref. [40] analyzes all of these strategies by indicating their conditions of applicability. In this paper, we use the Tikhonov regularization method which consists of modifying the objective function by introducing a regularization term based on the a priori information. In this case, the new objective function which will be considered in problem (IP) is:

$$F(\bar{\bar{T}}) = \|\varphi_{obs} - \varphi_{comp}(\bar{\bar{T}})\| + \alpha \|\bar{\bar{T}} - \bar{\bar{T}}_*\| \quad (25)$$

where  $\bar{\bar{T}}_*$  are some known values of transmissivity in the study domain (in absence of  $\bar{\bar{T}}_*$ , this value must be zero) and alpha is the regularization parameter. The choice of  $\alpha$  should be consistent with the inaccuracy of the input data. Ref. [41] proposed various practical choices which also prove to be useful for the efficiency (in the sense of convergence) of the algorithm implemented for solving the inverse problem (23).

Finally, the numerical solution of the problem (IP) by the integrated optimization algorithm HYSFLO-LBM/CMA-ES implemented for this study begins with an initial field of transmissivity  $\bar{\bar{T}}^{(0)}$ . CMA-ES uses this field as the solution at generation  $g = 0$ , then applies the selection and the mutation steps to obtain the new generation of  $\bar{\bar{T}}$ . With this new generation we can then solve the problem (DP) to obtain the hydraulic head field  $\varphi_{comp}(\bar{\bar{T}})$  necessary for the construction of the objective function (25). Finally, CMA-ES ends the iteration with the evaluation and convergence test steps before moving on to the adaptation step to perform a new iteration ( $g + 1$ ), if necessary, until the convergence of HYSFLO-LBM/CMA-ES towards  $\bar{\bar{T}}_{opt}$ , the optimal solution of the problem (IP). The CMA-ES code offers several convergence criteria. For all of the simulations presented in this paper, we adopted the convergence criterion relating to the evaluation of the objective function with a tolerance  $tol = 10^{-0.3}$  (see Algorithm 3). Algorithm 3 below summarizes the pseudocode of the proposed integrated optimization algorithm HYSFLO-LBM/CMA-ES.

---

#### Algorithm 3: HYSFLO-LBM/CMA-ES

---

```

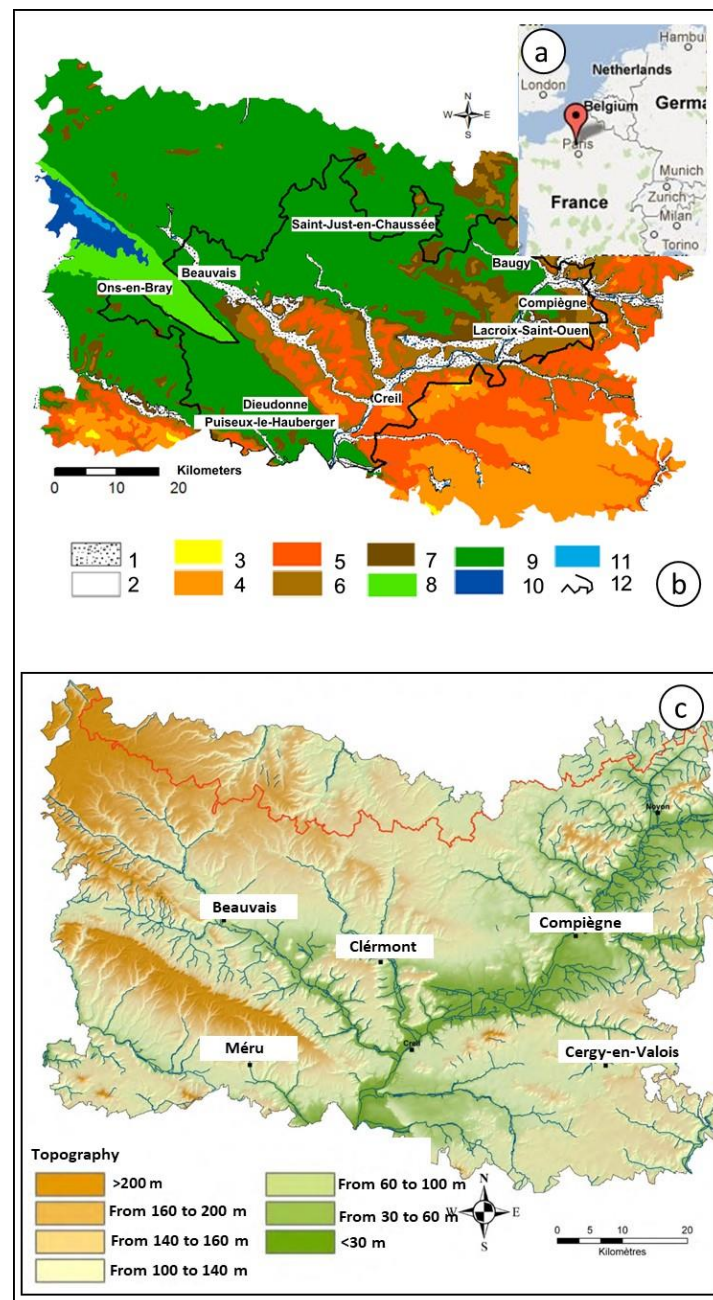
set the observations values vector  $\varphi_{obs}$ 
generation: g
initialization  $\bar{\bar{T}}^{(g)}$ 
generation: g + 1
selection from  $\bar{\bar{T}}^{(g)}$  + mutation  $\rightarrow \bar{\bar{T}}^{(g+1)}$ 
solve the (DP) by HYSFLO-LBM  $\rightarrow \varphi_{comp}(\bar{\bar{T}}^{(g+1)})$ 
construct the objective function according (25)
evaluation  $\rightarrow F(\bar{\bar{T}}^{(g+1)})$ 
test if  $\|F(\bar{\bar{T}}^{(g+1)})\| < tol$  then convergence to the  $\bar{\bar{T}}_{opt}$  and exit
adaptation  $\rightarrow \sigma^{(g+1)}, C^{(g+1)}$ 
go to the next generation

```

---

### 3. Realistic Case: Application to the Experimental Hydrogeological Site of Beauvais (Unconfined Aquifer)

This case is presented to demonstrate the successful application of the integrated LBM and CMAE-ES algorithm as the basis of the HYSFLO-LBM/CMA-ES code to solve the inverse problem in handling transport through fractured media, especially the chalk aquifer in the north of the Paris basin (Figure 2a).



**Figure 2.** (a) Location of the study area, (b) geological characteristics of the Beauvais [42], [1: Alluvium (Quaternary); 2: Anthropogenic embankments (Quaternary); 3: Upper and Lower Rupelian (Tertiary); 4: Upper and Lower Bartonian (Tertiary); 5: Lutetian = chalk, marl, and gravel (Tertiary); 6: Ypresian = ball clay and lignite of Soissons, sand of Cuise, and clay of Laon (Tertiary); 7: Thanetian = sand of Bracheux (Tertiary); 8: Lower Cretaceous from the Albian to the Neocomian (Secondary); 9: Upper Cretaceous from the Campanian to the Cenomanian (Secondary); 10: Portlandian Jurassic (Secondary); 11: Kimmeridgian-Sequanian Jurassic (Secondary); and 12: Limits of principal cities in Oise department], and (c) topography of the Oise department.

The experimental hydrogeological site of the Institut Polytechnique UniLaSalle Beauvais (SEHB) is located in the northern part of the Paris basin (in the Hauts-de-France region) and is characterized by a predominantly oceanic climate. The SEHB is equipped by the local meteorological station near the principal national meteorological monitoring system in the Beauvais-Tillé (Oise department). Precipitation, between 1985 and 2020, recorded an average value of 669.4 mm, with higher values toward the south and the west.

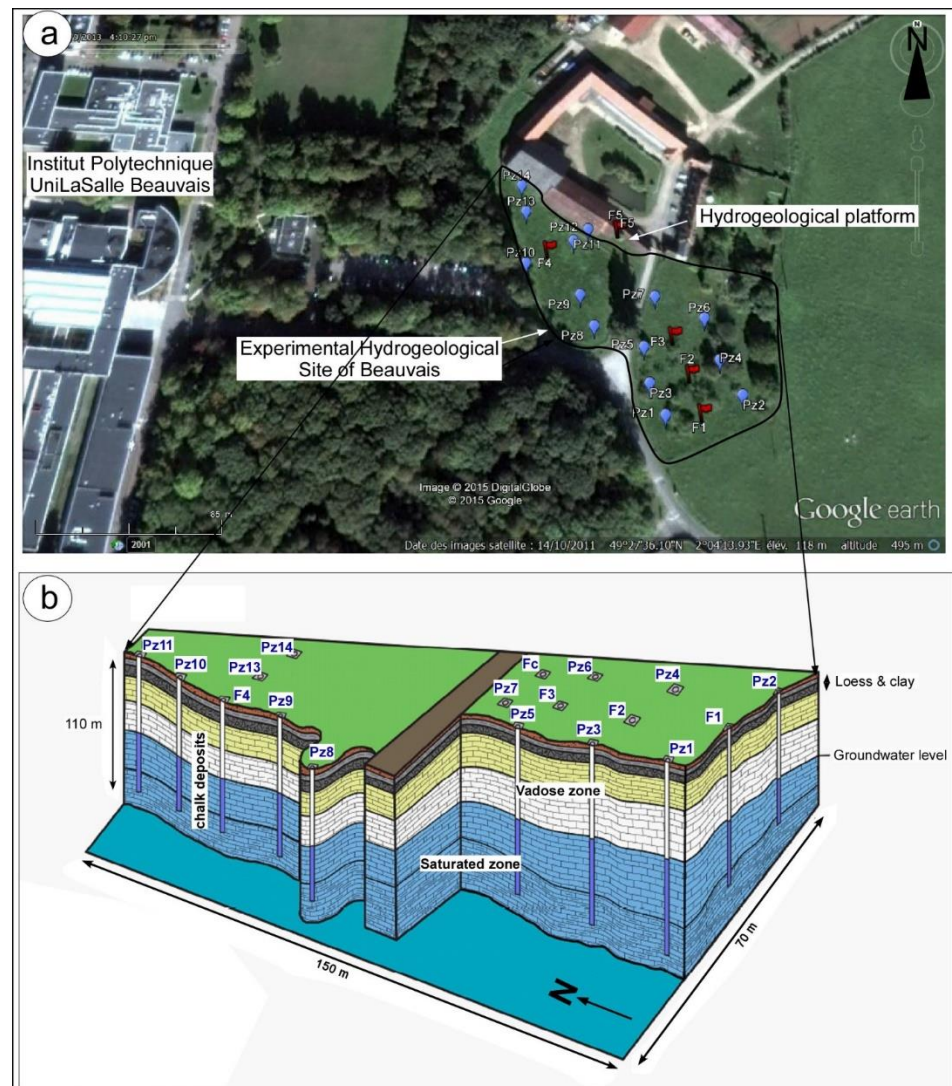
The geological information which was deduced from the field investigation (Figure 2b) in Oise [42] and from the analysis of wells revealed the following formations:

- the Jurassic formations which are composed of: (i) beige-grey lithographic limestone (the lower Portlandian, approximately 10 m thick); (ii) grey marls with limestone intercalations and dolomitic calcareous sandstone (the middle Portlandian, 150 m); and (iii) green-grey sand, sandy clay, or marls (the upper Portlandian, 15 m).
- principal formations of the Cretaceous Period which are, in this region: (i) the Cenomanian, which is comprised of “Gaise” (20 to 25 m): white-green siliceous limestone, from the base to a bluish gray sandy clay; white-green blood chalk, with or without chert, becoming glauconitic and clayish at the top; (ii) Turonian (110 m): marly chalk or grey chalk with chert; and (iii) Senonian (70 m): white chalk, yellow or gray with chert. The altitude of the Beauvais region varies between 57 m and 170 m (Figure 2c).

In order to follow the groundwater circulation and to improve knowledge of the chalk aquifer in the Oise department, the Hydrogeological Experimental Site of Beauvais (HESB) was built on 2 April, 2015. It is equipped with 20 hydrogeological drillings with a new hydrogeological platform (Figure 3a). The drilling diameter is approximately 125 mm for all the piezometers referenced by Pz (Figure 3b). But for all the wells referenced by F (Figure 3b) these diameters reach 160 mm. The depth of each well is about 110 m. Hydrogeological wells (Pz1 to Pz14 and F1 to F4) are georeferenced using a Differential Global Positioning System (DGPS) tool. This DGPS constitutes an improvement of GPS accuracy about 1 cm. It uses a network of fixed reference stations that transmit the difference between the positions indicated by the satellites and their known actual positions. On the other hand, water levels have been recorded by the piezometric manual probe SOLINST and by automatic datalogger (CTD DIVER, developed by Schlumberger water Service [43] and designed to measure, especially, water pressure), which contains electrical conductivity and temperature sensors, memory for storing measurements, and a battery. Recorded data are afterwards stored in the DIVER’s internal memory. Indeed, the DIVER consists of a pressure sensor, designed to measure water pressure, and a temperature sensor. The DIVER contains an autonomous datalogger and the duration of the measurements can be chosen by the user, ranging from seconds to hours. The various recorded measurements can be recovered either in the laboratory or at the measurement site using an optical communication system linking DIVER to a laptop computer.

Modeling of the Hydrogeological Experimental Site of Beauvais system requires expertise in terms of hydrodynamic characteristics, water table levels, the computed hydrological balance, transmissivity, and hydraulic conductivity. Generally, these parameters result from geostatistical concepts, laboratory experiments, pumping tests [44,45], and magnetic resonance sounding (MRS) for their hydrogeologic estimation [46]. Indeed, MRS investigations summarize the definition of the hydrodynamic parameters, especially in Beauvais city (in the North Paris Basin) and near the Hydrogeological Experimental Site, and will serve to ameliorate the database which should be used in the numerical processes of this paper and contribute to the knowledge of the heterogeneous permeable bodies [47] in the chalk aquifers.





**Figure 3.** (a) Location of the experimental hydrogeological site of Beauvais (UniLaSalle) and of the hydrogeological wells; (b) 3D conceptual model showing the hydrogeological wells and the chalk aquifer system.

#### 4. Modeling Results

The integrated optimization algorithm HYSFLO-LBM/CMA-ES was applied to the instrumented zone described in the previous paragraph. It is a rectangular area of 123 m wide and 133 m long. For high computational resolution, the size of the square lattice  $\Delta x$  was chosen to be 1 m. The study area has 20 measuring points of the hydraulic head  $\varphi$ . An interpolation procedure based on the Kriging techniques was performed to estimate the value of the hydraulic head on the entire computational grid ( $123 \times 133$  nodes) from the 20 measured points. These Kriged values were subsequently considered to be the field of observed values and denoted by  $\varphi_{obs}$ . The integrated optimization algorithm aimed to identify the field of the aquifer transmissivity during the year 2016. Table 1 presents the precipitation values of this year from which we have deduced the values of the source term  $f$  (see Equation (2)) necessary for the direct model HYSFLO-LBM. Finally, the direct model also required the boundary conditions for its forcing. The values of the Dirichlet boundary condition  $\varphi_D$  (see Equation (2)) at the four edges limiting the domain were calculated by evaluating  $\varphi_{obs}$  on these four limits as  $\varphi_D = \varphi_{obs}(\Gamma_D)$ . Moreover, the values of  $\bar{T}_*$  necessary for the regularization of the inverse problem (see Equation (25)) were



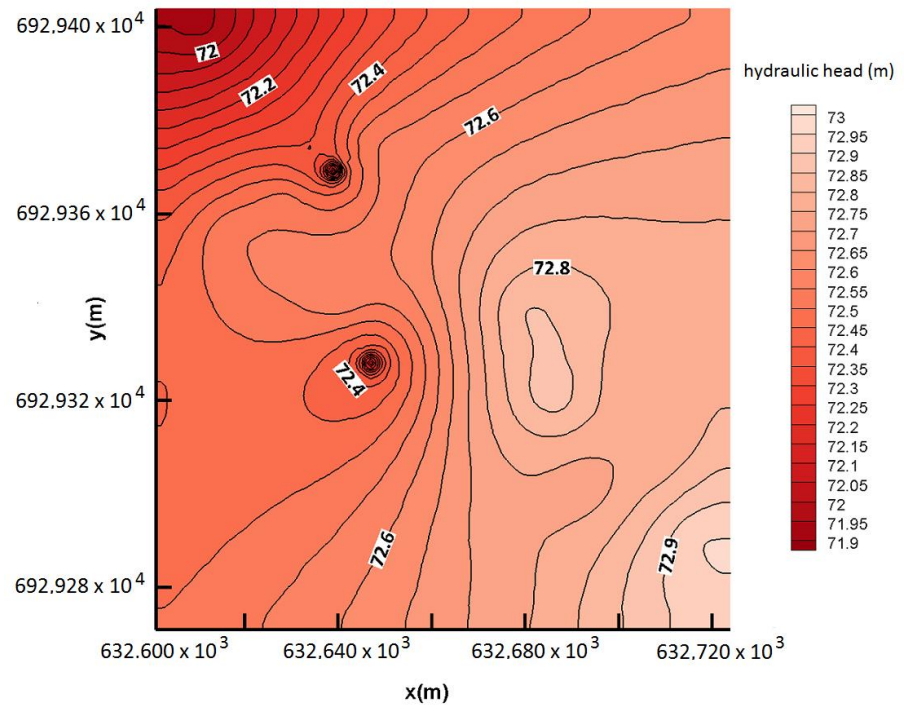
determined by using some values found in the literature from previous studies carried out on neighboring areas with the same hydrogeological characteristics.

The aquifer profits from principal refill, which is materialized by flow collected by the HESB. The lithological nature of the aquifer system is composed, especially, by the chalk deposits. It is very important to note that this formation is marked by dual porosity, with both interstices and cracks. The primary type of hydraulic conductivity linked to the interstitial porosity of the aquifer remains very low and generally does not exceed  $10^{-0.5}$  m/s [48]. The chalk fracturation at the surface permits higher hydraulic conductivities, between  $10^{-3}$  and  $10^{-2}$  m/s, which influence the runoff of the groundwater and create turbulent flows [49]. The comprehension of the infiltration and the identification of recharge and discharge periods is determined by the analysis and the interpretation of hydraulic head and rainfall evolution (Table 1). The hydraulic gradient is about 0.0047 between the wells F1 and F4 and 0.0066 between F2 and F4.

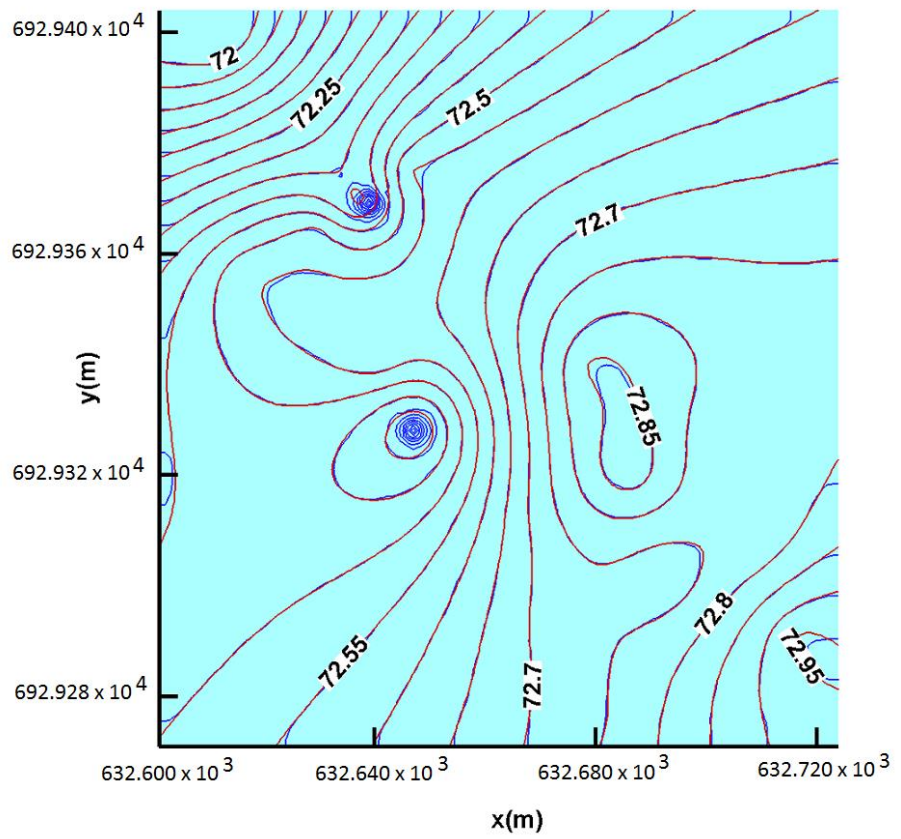
**Table 1.** Recharge periods in the Hydrogeological Experimental Site of Beauvais.

Recharge Period	Start Date (St)	End Date (Ed)	Rainfall: 1 Week before Recharge (mm)	Temperature (°C) before Recharge	Water Level (at the St) (m)	Water Level (at the Ed) (m)	Duration (Day)
2009/2010	2 November 2009	23 April 2010	39.5	7.9	73.872	75.262	171
2011	6 January 2011	19 April 2011	47	−4	73.335	73.715	95
2011/2012	30 November 2011	4 April 2012	33.5	10.0	73.33	74.54	102
2012/2013	4 October 2012	11 April 2013	46.6	13.0	71.53	73.51	186
2013/2014	11 October 2013	28 March 2014	29.5	9.8	69.84	71.29	173
2014/2015	5 January 2015	5 April 2015	31.3	2.5	70.42	71.43	89
2015/2016	3 February 2016	2 September 2016	55.4	8.7	70.92	72.47	210
2016/2017	8 February 2017	4 April 2017	40.2	6.9	71.78	72.21	54
Average					73.604	74.489	135

Based on hydro-geological knowledge, physical arguments, and piezometric analysis and interpretation, the hydrodynamic model is built in order to simulate the groundwater flow and to determine the distribution of the transmissivity values of the chalk aquifer in the HESB. The numerical process of achieving the objectives in this modeling expertise involves a series of procedures. They can be classified according to three principal steps: data collection; computer simulations, calibration, and analysis; and numerical results (Figure 4a–c). At the domain boundary and after hydrodynamic conditions, the piezometric values are imposed. Simulations are conducted in steady state with the reference values from 2016. The process of georeferenced nodes allows us to accelerate the modeling calibration (Figure 4b). This numerical method highlights an overlay between measured and computed hydraulic head, where the maximum recorded was 72.6 m and 71.92 m, respectively.

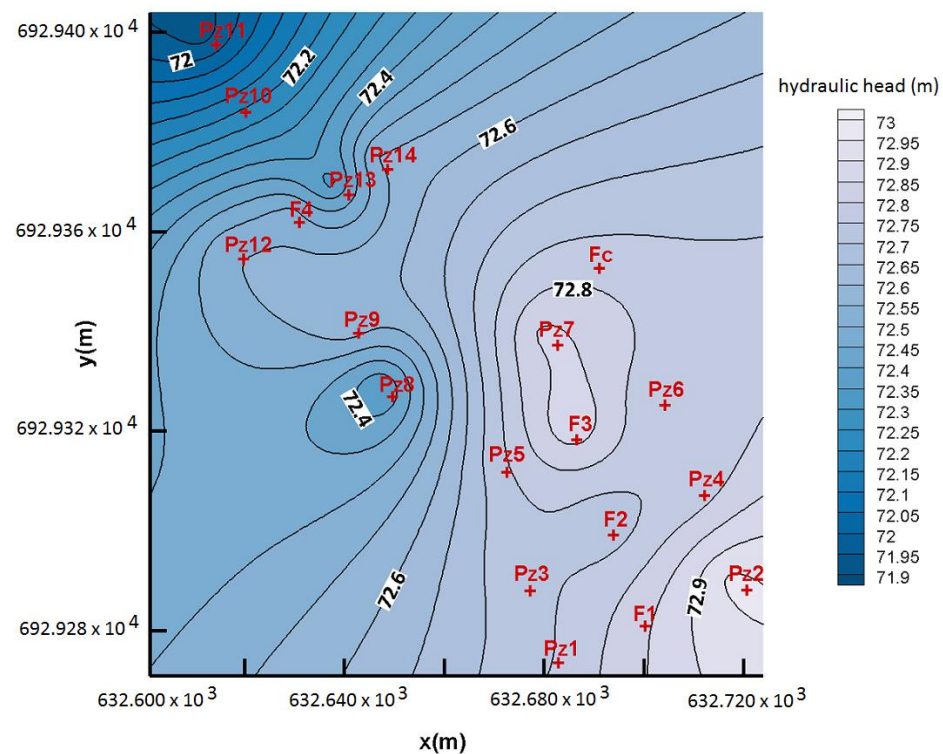


(a)



(b)

Figure 4. Cont.



(c)

**Figure 4.** Modeling results of the groundwater flow of the chalk aquifer (a) computed values of the piezometry, (b) calibration between calculated and measured values of the hydraulic head, (c) measured piezometry and groundwater flow of the chalk aquifer.

In order to stimulate the quality of the calibration, the absolute and the relative hydraulic head errors  $E_a$  and  $E_r$  are estimated using the following expression:

$$E_a = \|\varphi_{obs} - \varphi_{comp}\| \text{ and } E_r = \frac{\|\varphi_{obs} - \varphi_{comp}\|}{\|\varphi_{obs}\|} \quad (26)$$

Prior to interpreting the results, we would like to point out that we distinguish two areas: a diagonal band around the 20 measurement points and an area on either side of this diagonal. We recall that in the second zone it was necessary to complete the hydraulic head values by interpolation in order to have a reference values field  $\varphi_{obs}$  to formulate the objective function. It is quite obvious then that the results obtained in this zone will be more precise than that of the measurement zone, and occasionally with errors much lower than the model precision. Consequently, the interpretation and the discussion of the obtained results that we will present in the rest of the paper should concern only the diagonal band around the points of measurements.

The result, in the form of an errors map, displays map results displays relative and absolute errors in order to locate the points where there is a very large difference between measurement and simulates hydraulic heads. Generally, we note three principal zones which showed the existence of relatively small margins of absolute and relative error (Figure 5a,b):

- the first margin of absolute error from:  $8.0 \times 10^{-3}$  m to  $2.0 \times 10^{-2}$  m which corresponds to margin of the relative error between  $1.0 \times 10^{-2}\%$  and  $2.5 \times 10^{-2}\%$ ;
- the second margin of absolute error from:  $2.0 \times 10^{-2}$  m to  $2.8 \times 10^{-2}$  m which corresponds to margin of the relative error between  $2.5 \times 10^{-2}\%$  and  $3.5 \times 10^{-2}\%$ ;

- the third margin of absolute error from:  $2.8 \times 10^{-2}$  m to  $3.8 \times 10^{-2}$  m which corresponds to margin of the relative error between  $3.5 \times 10^{-2}\%$  and  $5.0 \times 10^{-2}\%$ .

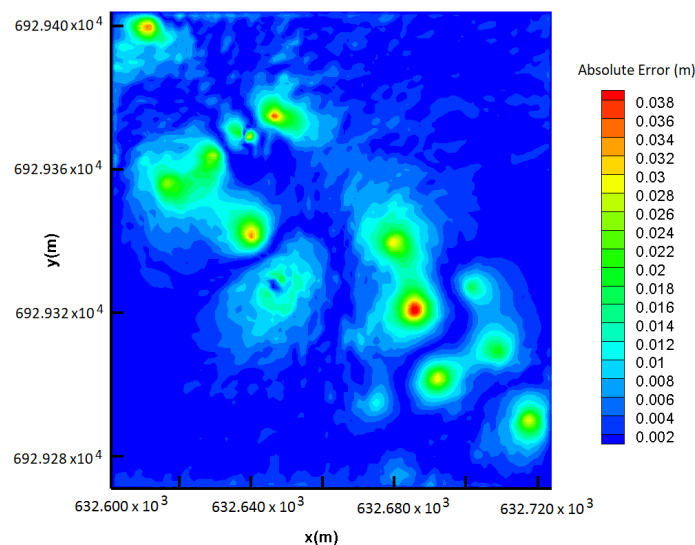
From these two figures we should retain that the absolute and relative errors do not exceed 4 cm and  $5.0 \times 10^{-2}\%$ , respectively. These two errors are located exactly at the hydrogeological well “F3” where the hydraulic mechanism is characterized by the groundwater dividing axis and a probable network of fractures, allowing control of the groundwater flow. Finally, in view of the relative errors (Figure 5a), and of the calibration figure (Figure 4b), the integrated optimization algorithm HYSFLO-LBM/CMA-ES allows us to obtain very good results both for the identification of the transmissivity and the simulation of the flow in the experimental site of Beauvais.

In the same way and for the same reasons cited above for the error fields (Figure 5a,b), we subsequently interpret the transmissivity and hydraulic conductivity fields only for the diagonal band, including the 20 measurement points.

The HYSFLO-LBM/CMA-ES code identifies the distribution of the hydraulic conductivity (or transmissivity) and its possible heterogeneity (Figure 6a,b). It is for this reason that the proposed integrated optimization algorithm punctually identifies (point by point) the tensor  $\bar{T}$ . This step was carried out by solving an optimization problem of the variable  $\bar{T}$  as a real vector of  $123 \times 133$  components which corresponds to the total number of the computation grid nodes. It should be noted that the CMA-ES algorithm is a stochastic search algorithm affected by certain errors induced by the generation of individuals randomly. In order to reduce these errors, we carried out about ten simulations to consider, as the final value  $\bar{T}_{opt}$ , only the average of these simulations. Consequently, Figure 6a,b shows, outside the diagonal band of the measurements, the heterogeneity of the hydraulic conductivity and of the tensor  $\bar{T}$ , but not of the random noises that could probably be interpreted by the reader.

Figure 6b shows lower, middle, and higher transmissivity values are about 0.0025, 0.0045, and  $0.007 \text{ m}^2/\text{s}$ , respectively, but globally, the magnitude order is about  $10^{-3} \text{ m}^2/\text{s}$ . This variation is very heterogeneous in the hydrogeological experimental site and is represented by the spatial distribution of transmissivity values (Figure 6a).

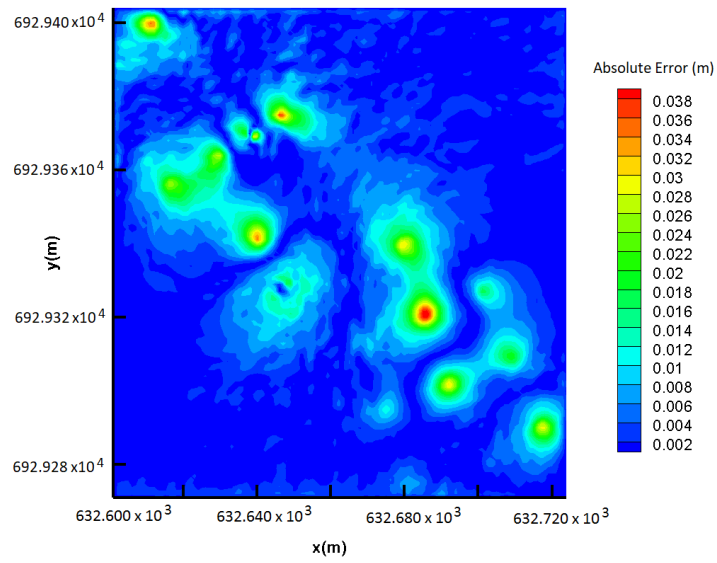
The same spatial distribution concerns the hydraulic conductivity. These values were obtained using simulated transmissivity values and the potential hydraulic head of the experimental site, and by considering the hydraulic character of the unconfined chalk aquifer. The hydraulic conductivity heterogeneity ranged from  $5.0 \times 10^{-4}$  to  $10^{-4} \text{ m/s}$ .



(a)

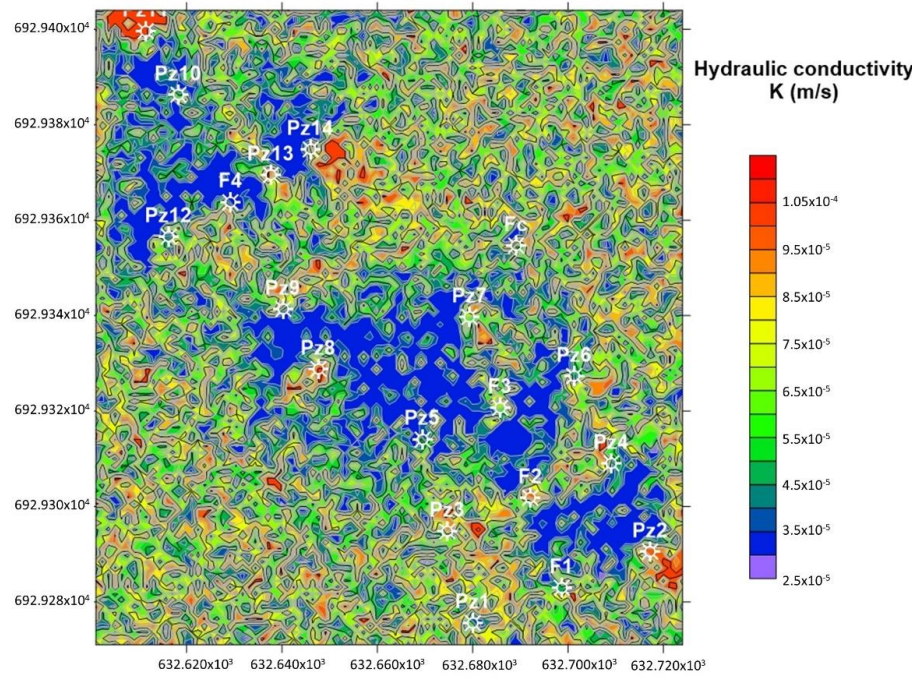
Figure 5. Cont.





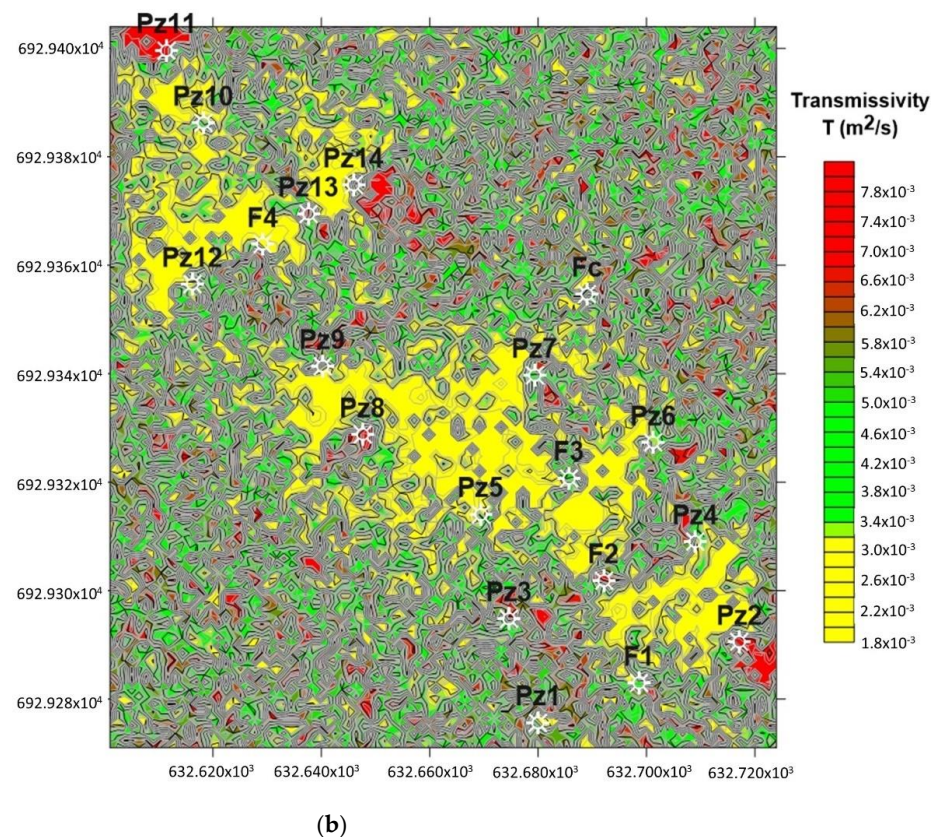
(b)

Figure 5. Relative (a) and absolute (b) errors deduced from the calibration.



(a)

Figure 6. Cont.



**Figure 6.** (a) Spatial distribution of the hydraulic conductivity and (b) repartition of the transmissivity in the chalk aquifer.

## 5. Discussion

The use of the Lattice Boltzmann method (LBM) in the numerical modeling of the groundwater allows to identify hydraulic characteristics of the chalk aquifer of the Hydrogeological Experimental Site of Beauvais, which revealed numerous transmissivity values ranging from 0.0025 to 0.007 m<sup>2</sup>/s. The spatial repartition of hydrogeological characteristics is accompanied by the anisotropic character of the chalk, which is characterized by the horizontal conductivity ( $10^{-14}$  to  $10^{-12}$  m/s) and the vertical conductivity ( $10^{-15}$  to  $10^{-13}$  m/s) [50]. The transmissivity and the hydraulic conductivity constitute principal characteristics allowing us to understand the groundwater flow. The variation of the transmissivity (0.0045 and 0.007 m<sup>2</sup>/s) and the hydraulic conductivity ( $5.0 \times 10^{-5}$  to  $10^{-4}$  m/s) is accompanied by the spatial distribution of the water content in the chalk aquifer, with maximum values of about 30–35% [46], and by the heterogeneity of lithological formations which are deduced from the analysis and the interpretation of drilling cuttings (dating from 2014 and with GPS coordinates: 49°27'36.88" N 2°04'17.38" E) belonging to the experimental hydrogeological site of Beauvais. It is mainly lithological deposits which are composed of clay with flint (Figure 7a), chalk with or without clay, and flint. This variation could be explained by primary and secondary hydraulic conductivities as defined by [50] and which are from  $10^{-8}$  to  $10^{-5}$  m/s and  $10^{-3}$  to  $10^{-1}$  m/s, respectively. The analysis and the interpretation of transmissivity values in different media were deduced from: the resonance magnetic sounding (Figure 7b,c); and the pumping test, especially in the 47 wells (chalk aquifer) and concerned plateau, dry valley, and humid valley with orders of magnitude about  $10^{-3}$  m<sup>2</sup>/s in plateau and an average transmissivity of about  $6.7 \times 10^{-3}$  m<sup>2</sup>/s [51]. From the geological structures point of view, value ranges which are deduced from the HYSFLO-LBM/CMA-ES code could be explained by the presence of fracture networks. Indeed in 2011, we realized near the hydrogeological experimental site 424 of fractures measurements. Three orientations have been identified: N-S, WNW-ESE, and NNE-SSW. In fact, numerous works highlighted a positive correlation



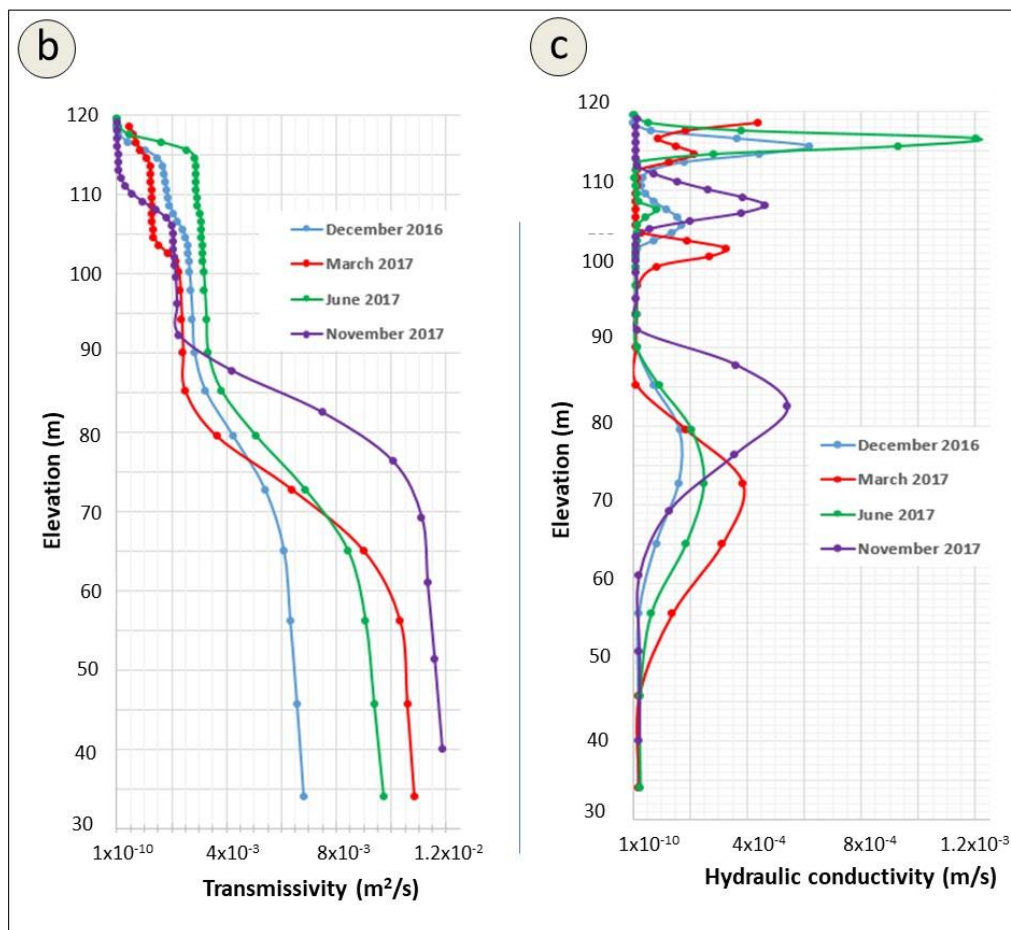
between fracture size and transmissivity [52] on the one hand and, on the other hand, the influence of the fracture network and the faulting mechanism on the spatial distribution of hydrogeological characteristics (transmissivity and hydraulic conductivity) and on the groundwater flow [45].

It is difficult to highlight in the field the relationship between fractures and the repartition of transmissivity values. Note the concordance transmissivity values resulting from hydrogeological experimentation (pumping tests and MRS) and the numerical modeling simulation by using the LBM. Hence, the accuracy of the results indicates that the use of the HYSFLO-LBM/CMA-ES code is recommended to simulate the flows governed by partial differential equations (PDEs) given by the model presented in this paper.

The integrated optimization algorithm HYSFLO-LBM/CMA-ES could be used in the industrial field (agricultural and food activities) as well as in the sector related to the supply of drinking water. The HYSFLO-LBM/CMA-ES code, as a numerical tool, constitutes a scientific opportunity to guide decision-makers towards favorable sectors for hydrogeological investigations and more transmissive areas in terms of water resources.

Ages	Depth	Lithology	Thickness	Description	
QUATERNARY			2 m	Embankments	
			6 m	Clay with flint	
	10 m		2 m	Clay with silex and chalk	
			2 m	Clay and more chalk deposits	
SENONIAN	20 m		2 m	Chalk less clay (20%)	
			2 m	50%	
	30 m		1 m	30%	
			3 m	40%	
			1 m	50%	
			2 m	30%	
			2 m	10%	
			3 m	40%	
			1 m	Clayey chalk with flint (30%)	
			1 m	Chalk with flint	
	40 m		5 m	Clayey chalk with flint (30 à 40%)	
			1 m	Chalk with flint (40%)	
		38 m		3 m	Clayey chalk with flint (20 à 30%)
				2 m	Chalk with flint (30%)
		50 m		3 m	Chalk with flint (20%)
				1 m	Chalk with flint (30%)
	1 m		Chalk with flint (10%)		
	7 m		Chalk with flint (20%)		

Figure 7. Cont.



**Figure 7.** (a) Lithological levels in the unsaturated/saturated zones of the chalk aquifer (Hydrogeological Experimental Site of Beauvais). (b) Synthesis of the repartition and evolution of the transmissivity and (c) the hydraulic conductivity obtained by resonance sounding techniques [47].

**Author Contributions:** Conceptualization, L.Z. and H.S.; methodology, L.Z. and H.S.; software, S.K. and H.S.; validation, S.K. and H.S.; formal analysis, L.Z.; investigation, L.Z., S.K., and H.S.; resources, S.K. and H.S.; data curation, L.Z.; writing—original draft preparation, L.Z., S.K., and H.S.; writing—review and editing, L.Z. and H.S.; visualization, L.Z. and H.S.; supervision, L.Z. and S.K.; project administration, L.Z.; funding acquisition, L.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data used to support the conclusions of this study are available on request and by agreement from the corresponding author.

**Acknowledgments:** The identification of transmissivity values will help us to know the hydrodynamic function of the chalk aquifer in the Oise region, especially in the Hydrogeological Experimental Site of Beauvais (HESB, northern part of the Paris basin). The construction of this site has benefited from the support of the European Regional Development Fund (FEDER), the Picardy region (Haut-de-France), and the Ministry of Higher Education and Research. This site and its hydrogeological platform are attached to the AGYLE Research team “Agroecology, Hydrogeochemistry, Environments, and Resources” of the Institut Polytechnique UniLassalle.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Srinivasa Raju, K.; Nagesh Kumar, D. Irrigation Planning Using Genetic Algorithms. *Water Resour. Manag.* **2004**, *18*, 163–176. [[CrossRef](#)]
2. Yeh, T.-C.J.; Liu, S.; Glass, R.J.; Baker, K.; Brainard, J.R.; Alumbaugh, D.; LaBrecque, D. A Geostatistically Based Inverse Model for Electrical Resistivity Surveys and Its Applications to Vadose Zone Hydrology. *Water Resour. Res.* **2002**, *38*, 1–13. [[CrossRef](#)]
3. Yang, Y.S.; Cronin, A.A.; Elliot, T.; Kalin, R.M. Characterizing a Heterogeneous Hydrogeological System Using Groundwater Flow and Geochemical Modelling. *J. Hydraul. Res.* **2004**, *42*, 147–155. [[CrossRef](#)]
4. Carrera, J.; Alcolea, A.; Medina, A.; Hidalgo, J.; Slooten, L.J. Inverse Problem in Hydrogeology. *Hydrogeol. J.* **2005**, *13*, 206–222. [[CrossRef](#)]
5. Nilsson, B.; Højberg, A.L.; Refsgaard, J.C.; Troldborg, L. Uncertainty in Geological and Hydrogeological Data. *Hydrol. Earth Syst. Sci. Discuss.* **2006**, *3*, 2675–2706.
6. Li, R.; Fang, L.; Liu, S. Hydrogeologic Parameters Inverse Analysis Based on Pumping Test by Comsol Multiphysics and Matlab. In Proceedings of the 2010 International Conference on Computer Design and Applications, Qinquangdao, China, 25–27 June 2010; Volume 2, pp. V2-160–V2-163.
7. Huang, S.-Y.; Wen, J.-C.; Yeh, T.-C.J.; Lu, W.; Juan, H.-L.; Tseng, C.-M.; Lee, J.-H.; Chang, K.-C. Robustness of Joint Interpretation of Sequential Pumping Tests: Numerical and Field Experiments. *Water Resour. Res.* **2011**, *47*. [[CrossRef](#)]
8. He, X.; Li, S.J.; Liu, Y.X.; Zhou, Y.P. Identification of Permeability Coefficient of Rock Massin Dam Foundation Based on Genetic Neural Network. *Chin. J. Rock Mech. Eng.* **2004**, *23*, 751–757.
9. Babazadeh, R.; Jolai, F.; Razmi, J.; Pishvae, M.S. Developing a Robust Programming Approach for the Responsive Logistics Network Design under Uncertainty. *Int. J. Ind. Eng. Theory Appl. Pract.* **2014**, *21*, 1–17.
10. Friedman, K.; Heaney, J.P.; Morales, M. Using Process Models to Estimate Residential Water Use and Population Served. *J. AWWA* **2014**, *106*, E264–E277. [[CrossRef](#)]
11. Ayvaz, M.T.; Elçi, A. Identification of the Optimum Groundwater Quality Monitoring Network Using a Genetic Algorithm Based Optimization Approach. *J. Hydrol.* **2018**, *563*, 1078–1091. [[CrossRef](#)]
12. Romero, C.E.; Carter, J.N. Using Genetic Algorithms for Reservoir Characterisation. *J. Pet. Sci. Eng.* **2001**, *31*, 113–123. [[CrossRef](#)]
13. Karpouzou, D.K.; Delay, F.; Katsifarakis, K.L.; de Marsily, G. A Multipopulation Genetic Algorithm to Solve the Inverse Problem in Hydrogeology. *Water Resour. Res.* **2001**, *37*, 2291–2302. [[CrossRef](#)]
14. Erickson, M.; Mayer, A.; Horn, J. Multi-Objective Optimal Design of Groundwater Remediation Systems: Application of the Niche Pareto Genetic Algorithm (NPGA). *Adv. Water Resour.* **2002**, *25*, 51–65. [[CrossRef](#)]
15. Zhang, Y.; Pinder, G.F.; Herrera, G.S. Least Cost Design of Groundwater Quality Monitoring Networks. *Water Resour. Res.* **2005**, *41*. [[CrossRef](#)]
16. Bayer, P.; Finkel, M. Evolutionary Algorithms for the Optimization of Advective Control of Contaminated Aquifer Zones. *Water Resour. Res.* **2004**, *40*. [[CrossRef](#)]
17. Elshall, A.S.; Pham, H.V.; Tsai, F.T.-C.; Yan, L.; Ye, M. Parallel Inverse Modeling and Uncertainty Quantification for Computationally Demanding Groundwater-Flow Models Using Covariance Matrix Adaptation. *J. Hydrol. Eng.* **2015**, *20*, 04014087. [[CrossRef](#)]
18. Rengers, F.; Lunacek, M.; Tucker, G. Application of an Evolutionary Algorithm for Parameter Optimization in a Gully Erosion Model. *Environ. Model. Softw.* **2016**, *80*, 297–305. [[CrossRef](#)]
19. Zhou, J.G. *Lattice Boltzmann Methods for Shallow Water Flows*; Springer: Berlin/Heidelberg, Germany, 2004; ISBN 978-3-540-40746-1.
20. Sukop, M.C.; Thorne, D.T. *Lattice Boltzmann Modeling: An Introduction for Geoscientists and Engineers*, 1st ed.; Springer: Berlin, Germany; New York, NY, USA, 2005; ISBN 978-3-540-27981-5.
21. Guo, Z.; Shu, C. *Lattice Boltzmann Method and Its Applications in Engineering*; Advances in computational fluid dynamics; World Scientific: Singapore, 2013; ISBN 978-981-4508-29-2.
22. Chen, Z.; Shu, C. *Simplified and Highly Stable Lattice Boltzmann Method: Theory and Applications: Theories and Applications*; World Scientific: Singapore, 2020.
23. Hansen, N.; Ostermeier, A. Completely Derandomized Self-Adaptation in Evolution Strategies. *Evol. Comput.* **2001**, *9*, 159–195. [[CrossRef](#)]
24. Darcy, H. Les Fontaines Publiques de la Ville de Dijon. Exposition et Application des Principes à Suivre et des Formules à Employer Dans Les Questions de Distribution D'eau: Ouvrage Terminé Par Un Appendice Relatif Aux Fournitures D'eau de Plusieurs Villes au Filtrage des Eaux et à la Fabrication des Tuyaux de Fonte, de Plomb, de Tôle et de Bitume; Dalmont. 1856. Available online: <https://lib.ugent.be/catalog/bkt01:000059712> (accessed on 19 March 2021).
25. Ciarlet, P.G. *The Finite Element Method for Elliptic Problems*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 2002; Volume 4. Available online: <https://www.elsevier.com/books/the-finite-element-method-for-elliptic-problems/ciarlet/978-0-444-85028-7> (accessed on 26 April 2021).
26. Dautray, R.; Lions, J.L. *Mathematical Analysis and Numerical Methods for Science and Technology*; Springer: Berlin/Heidelberg, Germany, 1990.
27. Rothman, D.H.; Zaleski, S. *Lattice-Gas Cellular Automata*; Cambridge University Press: Cambridge, UK, 1997; ISBN 0521 55201 X.
28. Agrawal, K.; Loezos, P.N.; Syamlal, M.; Sundaresan, S. The role of meso-scale structures in rapid gas–solid flows. *J. Fluid Mech.* **2001**, *445*, 151–185. [[CrossRef](#)]

29. Harris, S. *An Introduction to the Theory of the Boltzmann Equation*; Dover Publications Inc.: Mineola, NY, USA, 2004; ISBN 978-0-486-43831-3.
30. Zou, Q.; He, X. On Pressure and Velocity Boundary Conditions for the Lattice Boltzmann BGK Model. *Phys. Fluids* **1997**, *9*, 1591–1598. [[CrossRef](#)]
31. Bouzidi, M.; Firdaouss, M.; Lallemand, P. Momentum Transfer of a Boltzmann-Lattice Fluid with Boundaries. *Phys. Fluids* **2001**, *13*, 3452–3459. [[CrossRef](#)]
32. Miyagi, A.; Akimoto, Y.; Yamamoto, H. Well Placement Optimization for Carbon Dioxide Capture and Storage via CMA-ES with Mixed Integer Support. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*; Association for Computing Machinery: New York, NY, USA, 6 July 2018; pp. 1696–1703.
33. Li, C.; Heinemann, P.H. A Comparative Study of Three Evolutionary Algorithms for Surface Acoustic Wave Sensor Wavelength Selection. *Sens. Actuators B Chem.* **2007**, *125*, 311–320. [[CrossRef](#)]
34. Li, C.; Heinemann, P.; Reed, P. Genetic Algorithms (GAs) and CMA Evolutionary Strategy to Optimize Electronic Nose Sensor Selection. *Trans. ASABE* **2007**, *51*, 321–330. [[CrossRef](#)]
35. Bayer, P.; Finkel, M. Optimization of Concentration Control by Evolution Strategies: Formulation, Application, and Assessment of Remedial Solutions. *Water Resour. Res.* **2007**, *43*. [[CrossRef](#)]
36. Mersch, B.; Glasmachers, T.; Meinicke, P.; Igel, C. Evolutionary Optimization of Sequence Kernels for Detection of Bacterial Gene Starts. *Int. J. Neural Syst.* **2007**, *17*, 369–381. [[CrossRef](#)]
37. Sbalzarini, I.F.; Müller, S.D.; Koumoutsakos, P.D.; Cottet, G.-H. Evolution Strategies for Computational and Experimental Fluid Dynamic Applications. In *Proceedings of the Genetic And Evolutionary Computation Conference*, San Francisco, CA, USA, 6–9 July 2001; pp. 7–11.
38. Hamdani, H.; Radi, B.; Hami, A.E. Optimization of Solder Joints in Embedded Mechatronic Systems via Kriging-Assisted CMA-ES Algorithm. *Int. J. Simul. Multidisci. Des. Optim.* **2019**, *10*, A3. [[CrossRef](#)]
39. Hansen, N. The CMA Evolution Strategy: A Tutorial. Available online: <https://arxiv.org/abs/1604.00772> (accessed on 19 March 2021).
40. Vogel, C.R. *Computational Methods for Inverse Problems. Frontiers in Applied Mathematics*; SIAM: Philadelphia, PA, USA, 2002; p. 200.
41. Samarskii, A.A. *Vabishchevich, P.N. Numerical Methods for Solving Inverse Problems of Mathematical Physics*; Walter de Gruyt: Berlin, Germany, 2007; p. 453.
42. Tirat, M.; Belkessa, R.; Clément, J.P. *Données Géologiques et Hydrogéologiques Acquisées à La Date Du 31-12-67 Sur Le Territoire de La Feuille Topographique de Beauvais (N° 102)*; OISE: Toronto, ON, Canada, 1969; p. 92.
43. SLB. *Diver. Manual*; Schlumberger Water Service: Giesbeek, The Netherland, 2014; p. 33.
44. Dagan, G.; Fiori, A.; Janković, I. Transmissivity and Head Covariances for Flow in Highly Heterogeneous Aquifers. *J. Hydrol.* **2004**, *294*, 39–56. [[CrossRef](#)]
45. Zouhri, L.; Gorini, C.; Mania, J.; Deffontaines, B.; Zerouali, A.E.H. Spatial Distribution of Resistivity in the Hydrogeological Systems, and Identification of the Catchment Area in the Rharb Basin, Morocco/Répartition Spatiale de La Résistivité Dans Les Systèmes Hydrogéologiques et Détection Des Zones de Captages Dans Le Bassin Du Rharb, Maroc. *Hydrol. Sci. J.* **2004**, *49*, 398. [[CrossRef](#)]
46. Zouhri, L.; Lutz, P. Magnetic Resonance Sounding for the Estimate of Hydrogeologic Characteristics of Chalk Aquifers (North of France). *Arab. J. Geosci.* **2020**, *13*, 1064. [[CrossRef](#)]
47. Smaoui, H.; Zouhri, L.; Ouahsine, A.; Carlier, E. Modelling of Groundwater Flow in Heterogeneous Porous Media by Finite Element Method. *Hydrol. Process.* **2012**, *26*, 558–569. [[CrossRef](#)]
48. Zouhri, L.; Armand, R. Groundwater Vulnerability Assessment of the Chalk Aquifer in the Northern Part of France. *Geocarto Int.* **2019**, 1–24. [[CrossRef](#)]
49. Bault, V.; Borde, J.; Follet, R.; Laurent, A.; Tourliere, B. Atlas Hydrogéologique Numérique de l’Oise. Phase 3: Notice. 81 Ill., 55 Tab., 2 Ann., 1 Cd-Rom, 1 Carte A0. Avec La Collaboration de DE LEVEAU E. Et WILLEFERT V. 2012, p. 320. Available online: <https://infoterre.brgm.fr/rapports/RP-61081-FR.pdf> (accessed on 19 March 2021).
50. Domenico, P.A.; Schwartz, F.W. *Physical and Chemical Hydrogeology*; Wiley: New York, NY, USA; Chichester, UK, 1998; ISBN 978-0-471-59762-9.
51. Bault, V.; Follet, R. Atlas Hydrogéologique Numérique de l’Oise. Phase 2. Notice. Rapport Avancement. 2011, p. 172. Available online: <https://infoterre.brgm.fr/rapports/RP-59757-FR.pdf> (accessed on 19 March 2021).
52. Hyman, J.D.; Aldrich, G.; Viswanathan, H.; Makedonska, N.; Karra, S. Fracture Size and Transmissivity Correlations: Implications for Transport Simulations in Sparse Three-Dimensional Discrete Fracture Networks Following a Truncated Power Law Distribution of Fracture Size. *Water Resour. Res.* **2016**, *52*, 6472–6489. [[CrossRef](#)]