



**HAL**  
open science

# Identification of Hydrodynamic Dispersion Tensor by Optimization Algorithm Based LBM/CMA-ES Combination

Hassan Smaoui, Lahcen Zouhri, Sami Kaidi

► **To cite this version:**

Hassan Smaoui, Lahcen Zouhri, Sami Kaidi. Identification of Hydrodynamic Dispersion Tensor by Optimization Algorithm Based LBM/CMA-ES Combination. *Water*, 2022, 14, 10.3390/w14010125 . hal-03542465

**HAL Id: hal-03542465**

**<https://hal.science/hal-03542465>**

Submitted on 25 Jan 2022

**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

# Identification of Hydrodynamic Dispersion Tensor by Optimization Algorithm Based LBM/CMA-ES Combination

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

<sup>1</sup> CEREMA Risques Eaux et Mer (REM) EPR HA, 134 Rue de Beauvais, 60280 Margny-Les-Compiègne, France; sami.kaidi@cerema.fr

<sup>2</sup> AGHYLE, Institut Polytechnique UniLaSalle Beauvais, SFR Condorcet FR CNRS 3417 19 Rue Pierre Waguet, 60026 Beauvais, France; lahcen.zouhri@unilasalle.fr

\* Correspondence: hassan.smaoui@cerema.fr

**Abstract:** The hydrodynamic dispersion tensor (HDT) of a porous medium is a key parameter in engineering and environmental sciences. Its knowledge allows for example, to accurately predict the propagation of a pollution front induced by a surface (or subsurface) flow. This paper proposes a new mathematical model based on inverse problem-solving techniques to identify the HDT (noted  $\bar{D}$ ) of the studied porous medium. We then showed that in practice, this new model can be written in the form of an integrated optimization algorithm (IOA). The IOA is based on the numerical solution of the direct problem (which solves the convection–diffusion type transport equation) and the optimization of the error function between the simulated concentration field and that observed at the application site. The partial differential equations of the direct model were solved by high resolution of ( $\Delta x = \Delta y = 1$  m) Lattice Boltzmann Method (LBM) whose computational code is named HYDRODISP-LBM (HYDRO-DISPersion by LBM). As for the optimization step, we opted for the CMA-ES (Covariance Matrix Adaptation-Evolution Strategy) algorithm. Our choice for these two methods was motivated by their excellent performance proven in the abundant literature. The paper describes in detail the operation of the coupling of the two computer codes forming the IOA that we have named HYDRODISP-LBM/CMA-ES. Finally, the IOA was applied at the Beauvais experimental site to identify the HDT  $\bar{D}$ . The geological analyzes of this site showed that the tensor identified by the IOA is in perfect agreement with the characteristics of the geological formation of the site which are connected with the mixing processes of the latter.

**Keywords:** groundwater; numerical modeling; lattice Boltzmann method; relaxation time; hydrodynamics; isotropy; inverse problem; optimization; CMA-ES algorithm



**Citation:** Smaoui, H.; Zouhri, L.; Kaidi, S. Identification of Hydrodynamic Dispersion Tensor by Optimization Algorithm Based LBM/CMA-ES Combination. *Water* **2022**, *14*, 125. <https://doi.org/10.3390/w14010125>

Academic Editor: Alessandra Feo

Received: 17 November 2021

Accepted: 30 December 2021

Published: 5 January 2022

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



**Copyright:** © 2022 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

Groundwater that supplies phreatic aquifers is one of the most important freshwater resources. However, one of the supposed dramatic consequences of global climate change points to a scarcity in the “near” future of this resource across the world. It, therefore, becomes urgent to take up this challenge and propose plans for the development, use, and protection of groundwater in a healthy, sustainable, and responsible manner. Given the importance and the interest of these developments, the development of these plans must be based on a quality scientific approach. In other words, the management of water resources must involve a good understanding of the complex physical phenomena involved in the transformation cycle of groundwater. Among these phenomena, we evoke the mass transport by advection–diffusion in porous media which is present in many sectors of socio-economic activity in relation to flows. Examples include medicine (flow of fluid through organs), geology (thermal rock, thermal energy management), environment (contamination of the water table, radioactive waste), chemistry (catalytic reactors, filtration . . . ), petroleum (production of natural gas, flow of petroleum), mechanics (nuclear reactors, insulation,

combustion . . . ), and many others. The phenomena of dispersion can be of various origins. We talk about mass diffusion which is due to the mixing processes between two miscible fluids. This dispersion is estimated from the determination of the concentration fields of a passive tracer traveling from its point of injection. If, in addition, these fields present a gradient, we then see a new dispersion appear (even in the absence of flow). This is called molecular diffusion. It should be noted that the timescales associated with molecular diffusion are sufficiently large so that the latter can induce dispersion at the macroscopic scale. Finally, comes turbulent diffusion as its name suggests characterizing the turbulent flow in the porous media studied. It acts in the same way as molecular diffusion. Other phenomena can also be at the origin of diffusion by mixing processes. Mention is made, for example, of the mixture due to the various obstructions that a porous media may contain. This type of mixing is then caused by the tortuosity of the porous media studied. Likewise, the geometry of the porous media produces recirculation of the flow which in turn can induce the dispersion process. Furthermore, the connectivity of the porous media is likely to force the flow to borrow certain parts and not others of the media, thus creating a mixing zone from which the diffusion phenomenon can result. Finally, comes the hydrodynamic dispersion which is due to the existence of a speed gradient of the fluid phase. In a porous media, this speed gradient is consequent on the non-slip condition at the level of the solid wall of the pores. In summary, all these diffusion process types described above show that the mixing process in porous media is very difficult to apprehend since it is the consequence of the combination of a large number of phenomena of different time and space scales. This difficulty thus makes the determination of the dispersion coefficient extremely difficult.

From the previous paragraphs, we also understand that the dispersion coefficient is considered an important parameter in the knowledge of solute transport in the hydrogeological system. Therefore, its estimation with good precision is inevitable to guarantee a fine and successful description of the studied porous media. As illustrated above, the identification of the dispersion coefficient concerns most of the socio-economic sectors.

Indeed, the identification of the diffusion has been studied for industrial activities, such as the technical aspect of battery systems [1] and especially on sulfuric acid. Likewise, analytical experiments have made it possible to determine the dispersion coefficients by implementing scanning fluorescence correlation spectroscopy, in particular for molecules in solution [2]. In terms of transport in environmental media, Monteith and Unsworth [3] identified the dispersion coefficient for mass and heat transfer (water vapor and CO<sub>2</sub> transfer). In atmospheric engineering, Schnelle [4] carried out experiments to determine the dispersion coefficient of the study media with the aim of performing simulations for the prediction of accidental atmospheric pollution. In hydrogeology, the use of passive tracers, particularly in understanding the transfer of solutes, provides relevant information on hydrodynamic groundwater resources and to complete hydro-dispersive systems. In fact, passive tracers provide more detailed information on the dispersion coefficients, but the measurement campaigns require serious preparation on the fields and analysis protocols. The success will depend on the duration of the campaign, the choice of tracers, and the mass injected into the hydrogeological systems [5]. But with the spectacular progress of computing powers, the identification of the dispersion coefficient can be processed by numerical modeling which constitutes a reliable and inexpensive alternative to identification by experimentation. So this problem has been formulated as an optimization problem subject to constraints. Indeed, as specified above, the use of numerical simulations becomes essential for estimating mass transfer in groundwater. These models (called direct model) require the data of a set of parameters (called input), such as diffusivity, source/sink terms, and boundary and initial conditions. However, in practice, we only have a limited amount of these data to carry out reliable simulations to characterize the studied hydrogeological system. To remedy this difficulty, we resort to inverse modeling. Parameter identification occurs in many engineering disciplines: oceanography, meteorology, medicine, mechanical, and thermal industry, environment, hydrogeology, and many more.

Inverse models are used to identify the input parameters of the direct model by incorporating in-situ or laboratory observations. Of course, solving the inverse problems (IP) may also encounter difficulties related to the quality of the in-situ measurements that will be incorporated into the model. Indeed, the uncertainties (noise) attached to the measurements are at the origin of the unstable solutions thus rendering the IP ill-posed. In addition to unstable solutions, it is well known that inverse problems also suffer from the non-uniqueness of the solution. It is then necessary to find a precise mathematical framework to apply the methods of solving inverse problems. It means choosing the right approximation spaces and adopting the right regularization methods. It is for these reasons that the IP have been extensively studied both on theoretical and applied aspects [6–11]. In addition, several research works have been published in the field of hydrogeology and, in particular, for the determination of the transmissivity and/or dispersion coefficients necessary for the study of the propagation of accidental pollution. Obviously, we cannot list all of the work on IP, but we can cite some relevant references on the subject. These are the excellent book of Sun [12], the review of Zhou [13,14].

The aim of this paper is the application of the inverse problem techniques solution in hydrogeology and more particularly for the identification of the isotropic HDT of the experimental site “Beauvais LaSalle (North of France)”. Indeed, this identification will be based on the numerical solution of the inverse problem. This step will be accomplished by the direct problem (DP) solution by LBM, followed by a step of numerical optimization based on a metaheuristic algorithm called CMA-ES. For the direct model, we solve the anisotropic advection diffusion type transport equation. Several methods of solving this type of equation based on classical discretization methods (finite differences, finite elements, finite volumes) have been proposed in the literature. In recent years, we have seen the proposal of several algorithms for solving ADE by LBM and its variants [15–19]. These investigations have shown that LBM provides excellent results in terms of precision, efficiency, and applicability (highly parallelizable in a natural way). As far as we are concerned to solve the ADE constituting the direct problem, our choice fell on the use of the LBM because of its high simplicity of implementation and its easy parallelization. Thus, the final objective of this paper is the presentation of an integrated optimization algorithm for the identification of an anisotropic HDT (noted HYDRODISP-LBM/CMA-ES). This algorithm combines a hydro-dispersive calculation code by LBM (HYDRODISP-LBM) and an optimization code by the metaheuristic method CMA-ES.

This paper is organized as follows. Section 1 is an introduction presenting the state of the art on the numerical solution of ADE and its interest in the various engineering sciences. The second section presents the different mathematical tools that lead to the integrated optimization code HYDRODISP-LBM/ CMA-ES. Section 3 is devoted to the application of the HYDRODISP-LBM/CMA-ES algorithm to the realistic case of the Beauvais LaSalle site to identify the anisotropic HDT characterizing this site in the North of France. Finally, Section 4 discusses the results obtained, presents a general conclusion of the paper, and proposes some perspectives for possible improvements of the proposed integrated optimization algorithm HYDRODISP-LBM/CMA-ES.

## 2. Mathematical Formulation of the Model

### 2.1. Governing Equations

The mixing processes by dispersion in a media (also porous) denoted by  $\Omega$  characterized by a HDT  $\bar{\bar{D}}$  and crossed by a flow of velocity  $\vec{v}$  are governed by a scalar equation of the convection-diffusion type formulated by:

$$\frac{\partial C}{\partial t} + \nabla \cdot (\vec{v}C) = \nabla \cdot (\bar{\bar{D}}\nabla C) + S_C \quad (1)$$

where  $C$  is the diffused concentration,  $t$  is the time,  $\nabla \cdot$  and  $\nabla$  are respectively the divergence and the gradient operators and  $S_C$  is the term source or sink term. It should be noted that the velocity field  $\vec{v}$  denotes the velocity of the fluid flow carrying the concentration  $C$

through the studied porous medium. This velocity was calculated beforehand by the numerical solution of Darcy’s law by the LBM method implemented in the computer code of HYSFLO-LBM [20]. As specified in the introductory section, here, we are interested in the isotropic HDT, therefore in the 2D modeling, the tensor  $\bar{D}$  will be of the form:

$$\bar{D} = (D_{ij})_{\substack{1 \leq i \leq 2 \\ 1 \leq j \leq 2}} = D\delta_{ij} \quad \text{with} \quad \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (2)$$

To maintain isotropy, the diffusion coefficient  $D$  must be taken in the directions of the flow direction, by considering the following expression of the  $D$  :

$$D(x, y) = D_m + K_T(x, y) \|\vec{v}\| \quad (3)$$

where  $D_m$  is the molecular diffusion,  $K_T$  is the transversal diffusivity and is the  $\|\vec{v}\|$  modulus of the flow velocity. It should also be noted that the coefficient  $D$  is dependent on space and time through the parameters  $K_T$  (transversal dispersivity) and  $\|\vec{v}\|$ . Therefore, the identification of  $D$  is equivalent to identifying  $K_T$  (variable in space). Thus, the optimization algorithm implemented for this study is based on the identification of the coefficient  $D$  deduced from the Expression (3). Equation (1) expresses the partial differential equation (PDE) to be solved to estimate the concentration  $C$ . This PDE is not complete without the specification of the boundary and initial conditions. By introducing these conditions on an arbitrary 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 complete mathematical model is given by:

$$(DP) = \begin{cases} \frac{\partial C}{\partial t} + \nabla \cdot (\vec{v}C) = \nabla \cdot (\bar{D}\nabla C) + S_C & \text{in } \Omega \\ C = C_D & \text{on } \Gamma_D \\ \vec{n} \cdot (\vec{v}C - \bar{D}\nabla C) = C_N & \text{on } \Gamma_N \end{cases} \quad (4)$$

where  $C_D$  is a known function for imposing the Dirichlet condition  $\Gamma_D$ ,  $C_N$  is also a known function for imposing the Neuman condition on the domain boundary  $\Gamma_N$  and the vector  $\vec{n}$  denotes the normalized normal vector to  $\Gamma_N$  oriented towards the outside of the domain.

For a given HDT  $\bar{D}$ , the mathematical Model (4) describes the spatio-temporal evolution of a concentration field  $C$  transported by the advection and diffusion processes. In this case, we agree to say that we solve the direct problem (DP). Conversely, if the concentration  $C$  is known at a few points in the  $\Omega$  (from the measured values, for example), the Model (4) allows, in this case, determining the HDT  $\bar{D}$ . In this case, we talk about the solution of the inverse problem noted (IP).

It is well known that the numerical resolution of the inverse problem supposes certain conditions of regularity. In this context, Ciarlet [21] studied this problem from a mathematical point of view and proposed certain conditions which ensure the existence and uniqueness of the problem formulated by the system of Equation (4). The reader interested in the mathematical analysis aspect of the inverse problem will find rich information in Ramm’s [22] and Kirsch’s [23] books. Other authors are also interested in solving the inverse problem and have shown that in hydrogeology, this problem is poorly posed. To obtain physical solutions a step of regularization is necessary before starting the numerical resolution. In practice, this step consists of, among other things, modifying the objective function so that the optimization phase tends towards a physical solution (existence, uniqueness, and stability). We come back to this part in the section which describes the DP/optimization coupling.

## 2.2. Lattice Boltzmann Method

The numerical resolution of the transport by the advection–diffusion equation (ADE) constitutes a major stake in environmental engineering. Classical discretization methods, such as the finite differences method (FDM), the finite elements method (FEM), and the finite volumes method (FVM), have been and are still widely used to solve this type of equation. However, for two decades now, LBM has started to establish itself as a serious alternative to the classic methods (FDM, FEM, FVM). LBM’s mathematical formalism is based on the kinetic theory of gases. Unlike the concept of continuous media, LBM considers the medium as a discrete space of particles in perpetual reaction with each other. LBM is also based on the principles of statistical physics since it is interested in a distribution function (DF) of the velocities  $f$  rather than the average velocity of the representative elementary volume (REV) on which are based the equations of the mechanics of continuous media. Therefore, LBM solves the Boltzmann equation (BE) that governs the advection of  $f$  in the presence of a source term. This equation which describes the macroscopic behavior of the fluid via the DF  $f(\vec{x}, \vec{c}, t)$  is expressed by the following equation [24,25]:

$$\frac{\partial f}{\partial t} + \vec{c} \cdot \nabla f = \Omega(f) \quad (5)$$

where the function  $f$  is the probability that at the time  $t$  a particle is at position  $\vec{x}$  with velocity  $\vec{c}$ .  $\Omega$  is an operator that models collisions between particles. We can understand that LBM solves the scalar Equation (5) for the DF  $f$  instead of the ADE Equation (4) for the concentration  $C$ . To deduce the macroscopic variable  $C$ , we go through the calculation of different orders of the DF moments. For example, if we use LBM for Navier–Stokes, the density variable is deduced from the DF moment of order zero, the fluid velocity is deduced from the DF moment of order one, and finally the kinetic energy is deduced from the DF moment of order two, and so on.

The simplest collision operator is undoubtedly proposed by Bhatnagar–Gross–Krook (BGK) [26]. However, several authors have shown that this operator may prove insufficient to correctly cover the macroscopic variables from the resolution by BE. In this paper, we have used the BGK collision operator which is largely sufficient to cover the macroscopic variables when solving the ADE by LBM. This operator is written by:

$$\Omega(f) = -\frac{1}{\tau} (f - f^{(eq)}) \quad (6)$$

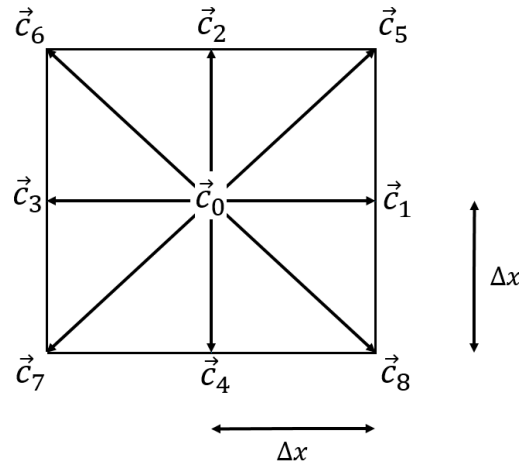
where  $\tau$  is the relaxation time,  $f^{(eq)}$  is the thermodynamic equilibrium distribution function (EDF). All the basic books that introduce the LBM describe the entire approximation steps to obtain an analytical form of this function. Zouhri et al. [20] proposed a synthetic presentation of the expression of  $f^{(eq)}$ .

## 2.3. Solution of the ADE by LBM

Like any PDE, the resolution of the BE (5) begins with a discretization of the variables of the equation (here it is the space  $\vec{x}$  and speed  $\vec{c}$ ). In 2D spatial discretization, we use a Cartesian computing grid of size  $\Delta x$ . For the speeds, the discretization indicates that at a point  $(i, j)$  of the lattice, a speed is allowed to take a predefined number of directions. Thus in 2D, if we choose a discretization of 9 directions, one speaks about the “D2Q9 lattice”. Figure 1 shows the layout of this type of lattice.

This configuration indicates that after a collision the particle at the point  $(i, j)$  can only move in one of the nine directions (staying at the point  $(i, j)$  is considered as a direction of velocity  $\vec{c}_0$ ). For the space discretization, several schemes have been proposed for this PDE type, but LBM uses a particular characteristic scheme. In fact, characteristic schemes are always accompanied by an interpolation step to estimate the speed when the endpoint of the characteristic does not exactly coincide with a node of the lattice. LBM then chooses

a time step so as to avoid this interpolation. With this particularity, the LBM algorithm consists of two stages: the collision and the propagation (called also streaming). These two stages make LBM highly parallelizable.



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

The discretization of BE must reproduce the collision and propagation stages as well as possible. After the collision, begins the propagation at which a fictitious particle is allowed to orient itself in the lattice only towards one of the 8 direct neighbors with a predefined speed  $c = \Delta x / \Delta t$ . Thus, the discretization of the propagation step (streaming) is expressed by:

$$f_i(\vec{x} + \vec{c}_i \Delta t, t + \Delta t) = f_i^*(\vec{x}, t) \quad \text{for } i = 0, \dots, 8 \quad (7)$$

where  $f_i^*$  is the value of  $f_i$  before the streaming and the 9 speeds are given explicitly by the expressions below

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

At the end of the propagation, the particles settle in the lattice to interact with the other particles (this is the collision) and start again in other directions and at other speeds according to diffusion rules. Thus, the discretization of the collision step is expressed by:

$$f_i^*(\vec{x}, t) = f_i(\vec{x}, t) - \frac{\Delta t}{\tau} (f_i(\vec{x}, t) - f_i^{(eq)}(\vec{x}, t)) \quad \text{for } i = 0, \dots, 8 \quad (9)$$

Finally, by combining expressions (8) and (9), we find the most popular form of discretization of BE used for fluid flow simulations based on the LBM.

$$f_i(\vec{x} + \vec{c}_i \Delta t, t + \Delta t) - f_i(\vec{x}, t) = -\frac{\Delta t}{\tau} (f_i(\vec{x}, t) - f_i^{(eq)}(\vec{x}, t)) \quad \text{for } i = 0, \dots, 8 \quad (10)$$

To complete the solution of BE, the algorithm of LBM needs to specify the ED  $f^{(eq)}$  and the relaxation time  $\tau$ . It should be noted that the  $f^{(eq)}$  is defined as the integral of the Maxwell–Boltzmann distribution over the entire velocity space. As the exact calculation of such an integral is difficult, it is customary to approach it numerically using high-order quadrature formulas. Thus, the approximation of  $f^{(eq)}$  used for LBM is estimated by the Gauss–Hermite quadrature formula which is found to have excellent accuracy.

It is quite clear that the expression of  $f^{(eq)}$  depends on the type of the PDE to be solved. Indeed, any PDE is formulated for a specific physical phenomenon. This phenomenon is a combination of several other phenomena of different physical magnitudes. On the other hand, we remind that the  $f^{(eq)}$  expression was derived from a Taylor expansion of the Maxwell–Boltzmann distribution in the neighborhood of  $\frac{\|\vec{v}\|}{c_s} < 1$ , where  $c_s$  is the isothermal sound velocity (for details of derivation, see [27]). One deduces that it is the order of this development which allows expressing the analytical expression of  $f^{(eq)}$ . This expression must cover and represent all orders of the physical quantities included in the phenomenon that we wish to model. For example, to solve the Navier–Stokes PDEs, we have to perform the two-order Taylor expansion to obtain a  $f^{(eq)}$  which will correctly reproduce the velocities and the pressure of the studied flow. On the other hand, to solve ADE by LBM, a Taylor development of order 1 is sufficient to find the expression of  $f^{(eq)}$  which will reproduce the phenomena of advection and diffusion. As mentioned above, in this paper, we used the LBM D2Q9 model to solve Equation (4). In this case, the ED  $f^{(eq)}$  is given by:

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

where  $c_s = 1/\sqrt{3}$  and  $(\omega_i)_{0 \leq i < 8}$  are the weight coefficients resulting from the Gauss–Hermite type quadrature formula which was adopted to approximate the calculation of the integral. These coefficients are given by:

$$\omega_0 = \frac{4}{9} \quad \omega_{1,2,3,4} = \frac{1}{9} \quad \omega_{5,6,7,8} = \frac{1}{36} \tag{12}$$

Details on how to compute the weighting coefficients  $\omega_i$  can be found in Succi [25].

As mentioned previously, the relaxation time  $\tau$  must be also specified. In fact, the multi-scale Chapman–Enskog Expansion allows proposing a constraint linking the dispersion coefficient  $D$  and the relaxation time  $\tau$ . Therefore, during the execution of the LBM algorithm, this parameter will be calculated explicitly from the  $D$  coefficient by:

$$D = \frac{\Delta x^2}{3\Delta t} \left( \frac{1}{\lambda} - \frac{1}{2} \right) \tag{13}$$

$\lambda$  is the dimensionless relaxation time ( $\lambda = \Delta t / \tau$ ).

At this stage, we have defined all parameters necessary for the execution of the D2Q9 LBM algorithm which can be summarized in the following two steps:

$$\text{collision : } f_i(x, y, t + \Delta t) = (1 - \lambda) f_i(x, y, t) + \lambda f_i^{(eq)}(x, y, t) \quad i = 0, \dots, 8 \tag{14a}$$

$$\text{streaming : } f_i(x + \Delta x, y + \Delta y, t + \Delta t) = f_i(x, y, t + \Delta t) \quad i = 0, \dots, 8 \tag{14b}$$

It should be noted that the different types of boundary conditions are imposed on the unknown  $C$  of the PDE (4). However, the LBM method solves the PDE (5) of unknown the DF  $f$ . It is then necessary to transform the boundary conditions relating to  $C$  variables to new ones which will be imposed on the DF  $f$ . In [28,29] one can find various transformations which make it possible to impose the various types of boundary conditions (Dirichlet, Neuman, Robin, etc.). To finish this subsection, we give the LBM algorithm which was implemented to solve the PDE (4) to build the computation code HYDRODISP-LBM (Algorithm 1).



**Algorithm 1: The pseudo-code of LBM**


---

```

set numerical parameters ( $\omega_i, c_s, D(\text{or } K_T), \vec{v}, \dots$ )
initialization
initialize  $C_0$  by the measured values
Compute  $f_k^{(eq)}$  associated to  $C_0$  from (Equation (11))
loop:  $t + \Delta t$ 
    compute  $\lambda_{xy}$  from (Equation (13))
    compute  $f_k^{(eq)}$  from (Equation (11))
    collision from (Equation (14a))
    streaming from (Equation (14b)) (after 14a and 14b we have  $(f_k)_{0 \leq k \leq 8}$ )
    set boundary condition
    compute  $C = \sum_{k=0}^8 f_k$ 
    test  $\|C - C_0\| / \|C_0\| < tol$  exit
    set  $C_0 \leftarrow C$ 
go to the next time step

```

---

One can remark that the velocity field  $\vec{v}$  is a datum for the DP algorithm. This velocity field can come from any numerical model solving the Darcy's equation in the concerned study area. For the application presented in this paper, we used the HYSFLO-LBM model already presented in a published paper in this journal [20].

#### 2.4. CMA-ES Algorithm

For the optimization step that constitutes the numerical resolution of an inverse problem, we have chosen to adopt a method in the category of metaheuristic algorithms. It is more precisely the evolution strategy algorithm commonly called CMA-ES (Covariance Matrix Adaptation–Evolution Strategy) [30] (Hanssen, 2016). This choice was motivated by several reasons, namely that this algorithm is very well adapted to the cases of objective functions which are not known in an explicit way (this is our case in this study, since the values of our objective functions are calculated by the resolution of the direct problem). Based on stochastic processes, CMA-ES is particularly suited to the optimization of non-convex functions whose values are noisy. This algorithm is even recommended for poorly conditioned problems. Contrary to the algorithms with a gradient that converges towards the local optimums, CMA-ES ensures a convergence towards the global optimum. Finally, this algorithm is qualified as robust since it has been used successfully in various disciplines of engineering and environmental sciences.

As mentioned previously, CMA-ES is an evolutionary strategy type algorithm with a functioning similar to that of genetic algorithms. That is to say, this algorithm acts on a population of individuals in four stages: initialization, selection, recombination, and mutation. During its execution, the CMA-ES algorithm connects the 4 steps to produce “the best individuals” for the next iteration (also called generation). If we denote by  $\mu$  the number of individuals (also called parents) in the population, at each generation, CMA-ES operates on these  $\alpha$  individuals (also called children) to produce  $\beta$  “best individual” ( $\beta \leq \alpha$ ) among the  $\alpha$  individuals. This particular CMA-ES algorithm is then called CMA-ES ( $\alpha, \beta$ ).

Mutation is arguably the main step in the CMA-ES algorithm. The parameters of this step (called strategy parameters) direct this algorithm for the covariance matrix transformations and their adaptation according to the generated population. Strategy parameters are automatically adapted (without user intervention) based on information from the previous generation [31,32]. This advantage makes the CMA-ES algorithm more and more attractive for the most complex optimization problems. In order not to overload the presentation, we will not give here more details on CMA-ES, but the interested reader will consult our previous work [20,33,34]. To finish this subsection, we give the flowchart of the operation of the CMA-ES algorithm (Algorithm 2).

---

**Algorithm 2: The pseudo-code of CMA-ES ( $\alpha, \beta$ )**

---

set numerical parameters ( $d, \alpha$ )  
 initialization  $\sigma, m, \bar{C} \dots$ , with  $(m)_{1 \leq i \leq \alpha} \in \mathbb{R}^d$   
 generation:  $g + 1$   
     selection  $\rightarrow X, \beta, m \dots$ , with  $(X)_{1 \leq i \leq \alpha} \in \mathbb{R}^d$   
     mutation  $\rightarrow X = m + \sigma \mathcal{N}(\vec{0}, \bar{C})$   
     evaluation  $\rightarrow F(X)$   
     test exit  $\parallel \rightarrow F(X) \parallel tol$  exit  
     adaptation  $\rightarrow \sigma, \bar{C}$   
 go to the next generation

---

where  $d$  is the size of the optimization variable,  $\alpha$  is the size of the population. Hansen [35] suggests  $\alpha = 4 + 3 \ln(d)$ ,  $\sigma$  is a real positive parameter,  $\bar{C}$  is the covariance matrix (symmetric and positive definite).  $(X)_{1 \leq i \leq \alpha} \in \mathbb{R}^d$  is the individual's population,  $\beta$  is the number of the "best individual" according to the fitness value (i.e.,  $F((X)_{1 \leq i \leq \alpha})$ ),  $F$  is the objective function,  $m$  is the weighted average of  $(X)_{1 \leq i \leq \beta}$  vectors on these  $\beta$  individuals,  $\mathcal{N}(\vec{0}, \bar{C})$  denotes a vector of independent normal random numbers with zero mean and covariance matrix  $\bar{C}$ . Finally, we remind that  $\sigma$  and  $\bar{C}$  are the strategy parameters that the algorithm CMA-ES self-adapts during each generation.

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

For transport phenomena in porous media by advection–diffusion mixing processes, the HDT  $\bar{D}$  is a main parameter for the hydro–dispersive groundwater modeling. It is, therefore, necessary to know this tensor with high accuracy in order to obtain correct results: for example, for a groundwater front pollution propagation. However, this tensor is an intrinsic parameter of the studied porous medium and it is generally unknown. Several techniques have been proposed for its determination and they are essentially experimental. In fact, to determine the tensor  $\bar{D}$  experimentally, a passive tracer is injected at a point in the domain (called a well), and we follow during the time its movement by taking concentrations of the tracer at another point distant from the injection point. By using the basic statistical tools applied to the restored curve, we can then deduce the dispersion coefficient of the studied porous media. However, this technique requires prior knowledge of the direction of flow which will guide the choice of sampling point, but generally, this information is difficult to obtain without going through the hydrodynamic study of the study area. In addition, this method has also been proven ineffective when the medium is characterized by anisotropy.

In this paper, we propose to identify the HDT  $\bar{D}$  of a porous medium in a numerical way. Indeed, to identify  $\bar{D}$ , we formulated the problem from a mathematical point of view as a solution to an inverse problem with constraints. In another way, if one has the measurements of concentrations in the field denoted by  $C_{obs}$ , this problem is then equivalent to finding the tensor  $\bar{D}$  which minimizes the difference between the observed concentrations and those computed by solving the DP that we denote  $C_{comp}$ . In this case the IP that must be resolved can be formulated by:

$$(IP) = \begin{cases} \text{Minimize } F(\bar{D}), & \bar{D} \in \mathcal{A}_D \\ \text{subject to} & \\ (DP) = \begin{cases} \frac{\partial C}{\partial t} + \nabla \cdot (\vec{v} C) = \nabla \cdot (\bar{D} \nabla C) + S_C & \text{in } \Omega \\ C = C_D & \text{on } \Gamma_D \\ \vec{n} \cdot (\vec{v} C - \bar{D} \nabla C) = C_N & \text{on } \Gamma_N \end{cases} \end{cases} \quad (15)$$

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

The inverse problem (15) is the formulation of the integrated optimization algorithm (IOA) which will be implemented to build the HYDRODISP-LBM/CMA-ES computation code. It is an iterative algorithm that couples the step of solving the DP and the step of optimization by CMA-ES. At convergence, this algorithm provides the HDT that we wish to identify  $\bar{\bar{D}}_{opt}$  such that:

$$\bar{\bar{D}}_{opt} = \underset{\bar{\bar{D}} \in \mathcal{A}_D}{\min} \|C_{obs} - C_{comp}(\bar{\bar{D}})\| \quad (16)$$

Obviously and like any numerical method, the numerical resolution of the IP is also confronted with certain difficulties, such as the stability, the uniqueness of the solution, and the convergence of the optimization algorithms, and, in particular, if it is the methods with gradients which have been adopted. All these difficulties have been massively studied by various researchers around the world from both a mathematical and numerical point of view. Several techniques have been proposed to circumvent these difficulties. In order to not repeat all of these techniques here, the interested reader will find a summary in the Reference [20], but for more information, the reader may also consult the reference books on inverse problem topics. Par example [7,12,36–39].

It should be noted that in hydrogeology, if one has the measurements of the hydraulic potential  $\varphi$ , the problem of identifying the transmissivity tensor  $\bar{\bar{T}}$  (Darcy's equation) is formulated in a similar way to problem (15) [20]. It is considered as an ill-posed problem in the sense of Hadamard [40]. Consequently, the various techniques of numerical solution of this IP do not ensure the uniqueness of the solution (denoted  $\bar{\bar{T}}_{opt}$ ). To overcome this difficulty, it is then necessary to make additional hypotheses called a regularization strategy [39,41,42]. In practice, the most used regularization method for this type of problem consists in introducing a priori information on the tensor  $\bar{\bar{T}}$  (for example, all kinds of measurements of  $\bar{\bar{T}}$  available on the study domain). This strategy is called the Tikhonov regularization method [43] which consists of modifying the objective function  $G$  of the IP to become:

$$G(\bar{\bar{T}}) = \|\varphi_{obs} - \varphi_{comp}(\bar{\bar{T}})\| + \|\lambda \bar{\bar{T}} - \bar{\bar{T}}_*\| \quad (17)$$

where  $\bar{\bar{T}}_*$  is the tensor containing a priori information (values of measurements), and  $\lambda$  is the non-negative regularization parameter. In the book of Vogel [7], one can find the different choices of the regularization parameter  $\lambda$  that can significantly improve the efficiency (in the sense of convergence) of the IOA formulated by (15). All the choices for this parameter have been made in order to preserve the consistency with the inaccuracy of the input data of the problem. Finally, let us note that in the absence of a priori information, the tensor  $\bar{\bar{T}}_*$  must be taken equal to zero.

As far as we are concerned, for the identification of the HDT  $\bar{\bar{D}}$ , we do not have the a priori information  $\bar{\bar{D}}_*$ . On the other hand, the fact of writing  $\bar{\bar{D}}$  in a precise form according to the flow velocity via expression (3) is then equivalent to a priori information. Therefore, in our case, we do not need to specify the value of the tensor  $\bar{\bar{D}}_*$ . In addition, and for more regularity, during the optimization step of the IOA, we will use the objective function  $F$  modified by the Tikhonov regularization method, such as

$$F(\bar{\bar{D}}) = \|C_{obs} - C_{comp}(\bar{\bar{D}})\| + \mu \|\bar{\bar{D}}\| \quad (18)$$

where  $\mu$  is non-negative regularization parameter.

To finish this section, we describe the successive operations of the computer code HYDRODISP-LBM/CMA-ES implemented for the IOA. The code starts with an initial HDT  $\bar{\bar{D}}^{(0)}$  that the CMA-ES algorithm also uses as the tensor value at generation  $g = 0$ . CMA-ES performs the selection and mutation steps to provide a new HDT  $\bar{\bar{D}}$ . With this tensor, it is the computer code HYDRODISP-LBM of the direct problem in (15) which runs in turn to provide the value of the concentration  $C_{comp}(\bar{\bar{D}})$ . At this stage of the IOA, the objective function  $F$  is constructed according to expression (18), then CMA-ES proceeds to the evaluation step (fitness) in order to be able to test the convergence. If the convergence is obtained, CMA-ES returns the optimal HDT  $\bar{\bar{D}}_{opt}$ . Otherwise, CMA-ES goes to the step of adapting the strategy parameters  $(\sigma, \bar{\bar{C}})$  to propose the new generation  $(g + 1)$ . Thus, if necessary, these steps will be repeated until the convergence of the IOA towards the optimal value  $\bar{\bar{D}}_{opt}$ . Algorithm 3 summarizes the different steps of the IOA implemented by the HYDRODISP-LBM/CMA-ES computer code in this study to solve numerically the IP formulated by (15).

---

**Algorithm 3: The pseudo-code of HYDRODISP-LBM/CMA-ES**

---

```

set the observations values vector  $C_{obs}$ 
generation:  $g$ 
initialization  $\bar{\bar{D}}^{(g)}$ 
generation:  $g + 1$ 
    selection from  $\bar{\bar{D}}^{(g)}$  + mutation  $\rightarrow \bar{\bar{D}}^{(g+1)}$ 
    solve the (DP) by HYDRODISP-LBM  $\rightarrow C_{comp}(\bar{\bar{D}}^{(g+1)})$ 
    construct the objective function according (18)
    evaluation  $\rightarrow F(\bar{\bar{D}}^{(g+1)})$ 
    test if  $\|F(\bar{\bar{D}}^{(g+1)})\| < tol$  then convergence to the  $\bar{\bar{D}}_{opt}$  and exit
    adaptation  $\sigma^{(g+1)}, \bar{\bar{C}}^{(g+1)}$ 
go to the next generation

```

---

### 3. Realistic Case: The Experimental Hydrogeological Site of Beauvais (Unconfined Aquifer)

Hydrogeological models on chalk aquifers (North of France) are rare and focus especially on hydrodynamics approaches. The knowledge of the groundwater flow and the geometrical of geological formations in the Oise department (northern part of France) led us to build the first model by using the Algorithm which is based on LBM/CMA-ES combination and thus making it possible to propose a water resources management and forecasting plan [20]. This model constitutes a fundamental basis and an opportunity to develop a new numerical methodology in order to determine the distribution of the isotropic HDT in the cretaceous formation which constitutes the principal aquifer in the north of France. The chalk lithological levels which are presented in the hydrodynamics simulations [20] represent the groundwater chalk aquifer in the Hydrogeological Experimental Site of Beauvais (HESB) and where the distribution of the hydraulic conductivity and transmissivity are deduced from the geophysics, pumping tests, and numerical processes. In order to complete the anterior study, it was necessary to develop the hydro-dispersive numerical model for identifying the isotropic dispersion coefficients which are unknown in chalk aquifers and based on hydro-chemical measurement campaigns carried out on this site.

The HESB located at (49° 27' 35.72" N, 2° 04' 5.51" E) was set up in 2015 and it has benefited from the support of the Ministry of Higher Education and Research, Haut-de-France region, the European Regional Development Fund (FEDER), and the Institut Polytechnique UniLaSalle Beauvais. The HESB is equipped with twenty wells with 110 m of depth. Measurements of water levels, temperature, and electrical conductivity are provided by the

CTD DIVERS type sensor with automatic acquisition which was developed by Schlumberger water Service [44]. Hydro-chemical analysis of water samples is carried out in the Institut Polytechnique UniLaSalle by using the techniques of Dionex High Performance Ion Chromatography especially anions and cations analyses ( $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{HCO}_3^-$ ,  $\text{Cl}^-$ ,  $\text{NO}_3^-$ ,  $\text{PO}_4^{2-}$  and  $\text{SO}_4^{2-}$ ). Finally, in order not to overload the content of this paper, we have given here only a brief description of the studied domain, omitting to describe the geological formation of the site. However, the reader interested in this aspect of the paper can find rich information on HESB in our previous paper [20].

#### 4. Discussion of the Modeling Results

Before starting the interpretation and discussion of the simulation results of the IOA, we give a description of the conditions of these simulations. The IOA was applied to the Beauvais site (North of France) which has been described in detail in our article published in this journal [20]. It is an unconfined aquifer-type site with continuous instrumentation since 2015. The site is rectangular in shape with a length of 133 m and a width of 123 m. For a high spatial resolution, we adopted a computational grid of  $\Delta x = \Delta y = 1$  m, that is to say a computational lattice for the DP of  $123 \times 133$ . On this site, several sensors have been deployed to measure the various parameters of the aquifer. Thus, we had 20 points of measuring the concentration of different chemical substances. Among the ions analyzed, we chose for our simulations in this paper to use the chloride ion ( $\text{Cl}^-$ ) concentration as the tracer. From these 20 concentration values, we deduced the concentration fields over the entire computational domain by Kriging type interpolation. Figure 2 shows the concentration obtained; the interpolated concentration is denoted  $C_{obs}$  which was used as the reference field necessary to the IOA for the construction of the objective function  $F$  expression (18). Finally, we note that the numerical solution of the DP required the knowledge of the values of the concentration on the four edges of the computational domain. For all the simulations presented in this study, we imposed the DP Dirichlet type boundary conditions. The values of the  $C_D$  function were extracted from those of  $C_{obs}$  on the four edges. Thus, if we denoted by  $\Gamma_D$  the set of the four edges, then we have  $C_D = C_{obs}(\Gamma_D)$ .

As specified previously, the direct problem-solving ADE require knowledge of the velocity field  $\vec{v}$  necessary of for the advective transport phase of the chlorine concentration. In fact, before proposing this paper which studies the hydro-dispersive aspect of HESB, we presented a purely hydrodynamic study of this site [20]. Indeed, since we have 20 hydraulic head measurement points (Figure 2), we then proposed a similar approach (IOA: HYSFLO-LBM/CMA-ES) in order to identify the transmissivity tensor  $\bar{T}$  of the water table [20]. Once the tensor  $\bar{T}$  was identified, by using Darcy's law we were able to estimate the velocity field of the studied site (Figure 3). This figure shows a complex structure of the underground flow of the HESB but highlights several wells and sources in the field consistent with the measurements of the hydraulic head taken on the HESB.

The IOA HYDRODISP-LBM/CMA-ES developed in this study was applied to the HESB site to identify the HDT  $\bar{D}$ . As explained by algorithm 3, a simulation by IOA begins with the first step by solving the DP by HYDRODISP-LBM, then performs the optimization by the CMA-ES algorithm in the second step. The computational time of the DP is insignificant (12 s on an Intel Xeon E5520 @ 2.27 GHz CPU with 32 GB of RAM). On the other hand, it is the optimization stage that consumes the most computing time. In fact, in the CMA-ES algorithm, the individual mutation stage explores all the elements of the research space to hope to propose the "best" population of  $\beta$  individuals among the  $\alpha$  parents. The other steps of CMA-ES, such as selection, recombination, and adaptation, do not consume much time compared to the mutation. For information, a complete simulation until convergence by HYDRODISP-LBM/CMA-ES consumes a little more than 2 days in calculation time for a tolerance  $\varepsilon = 10^{-4}$ . Moreover, it is important to note that the CMA-ES algorithm is based on stochastic processes for the research stage and consequently, the

individuals of the populations proposed for each generation may be affected by certain errors. To reduce the impact of the propagation of these errors in the final results of the identified tensor, we carried out about fifteen successive simulations, and we considered the optimal HDT  $\bar{\bar{D}}_{opt}$  as the arithmetic mean of these fifteen simulations. Before starting the interpretation of the modeling results, it is useful to define the numerical quantity which helps us to quantify the errors. These are the absolute error ( $E_a$ ) and the relative error ( $E_r$ ) defined by:

$$E_a = \|C_{obs} - C_{comp}(\bar{\bar{D}}_{opt})\| \quad \text{and} \quad E_r = \frac{\|C_{obs} - C_{comp}(\bar{\bar{D}}_{opt})\|}{\|C_{obs}\|} \quad (19)$$

It is interesting to note that the application of the IOA to a site can meet the double objectives: it allows both the calibration of the model of the DP (we speak about the automatic calibration) and the identification of the HDT of the studied site. Figure 4a,b respectively show the field of the measured  $Cl^-$  concentration field and the simulated one by the model HYDRODISP- LBM by considering the identified HDT  $\bar{\bar{D}}_{opt}$  as data of the DP. In a visual comparison of these two figures, one can conclude an excellent agreement between these two fields of concentration. Moreover, to quantify the small remaining errors, we must examine Figure 5a, b. From these two figures, we read that the relative error did not exceed 4% while the absolute error hardly reached 1 mg/L.

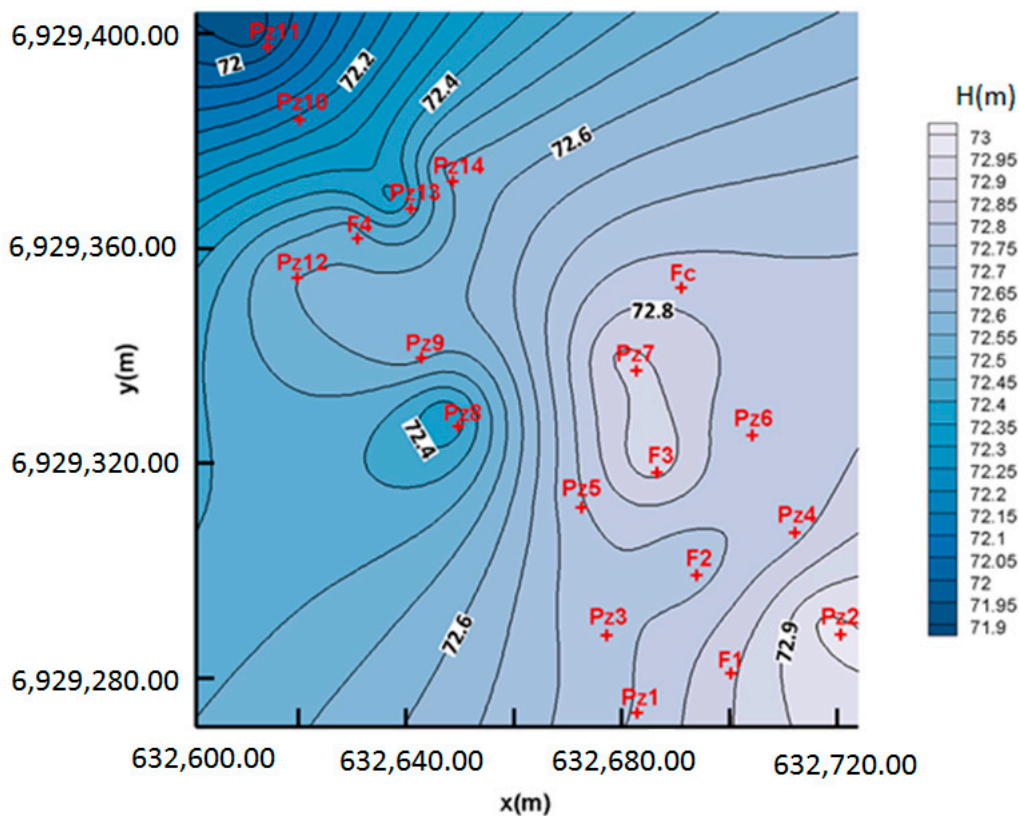


Figure 2. Interpolated Chloride concentration ( $C_{obs}$ ) and location of measurements points.

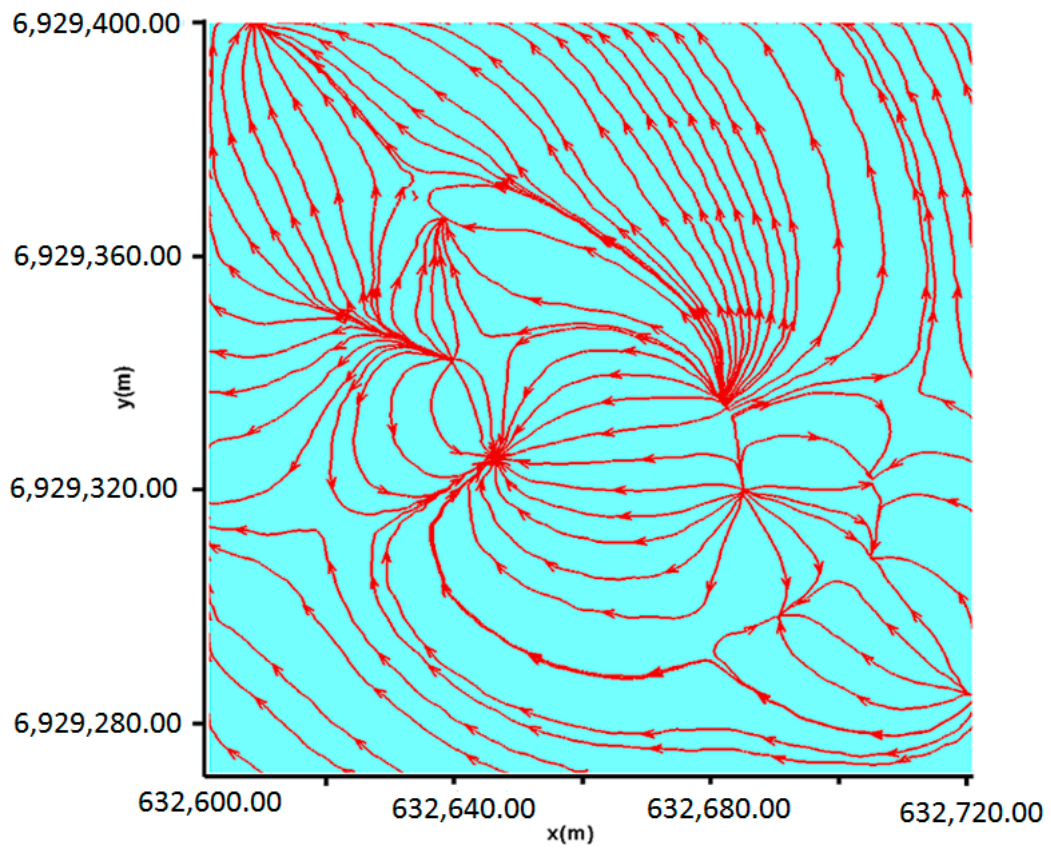


Figure 3. The computed streamlines by HYSFLO-LBM computer code from Darcy’s law [20].

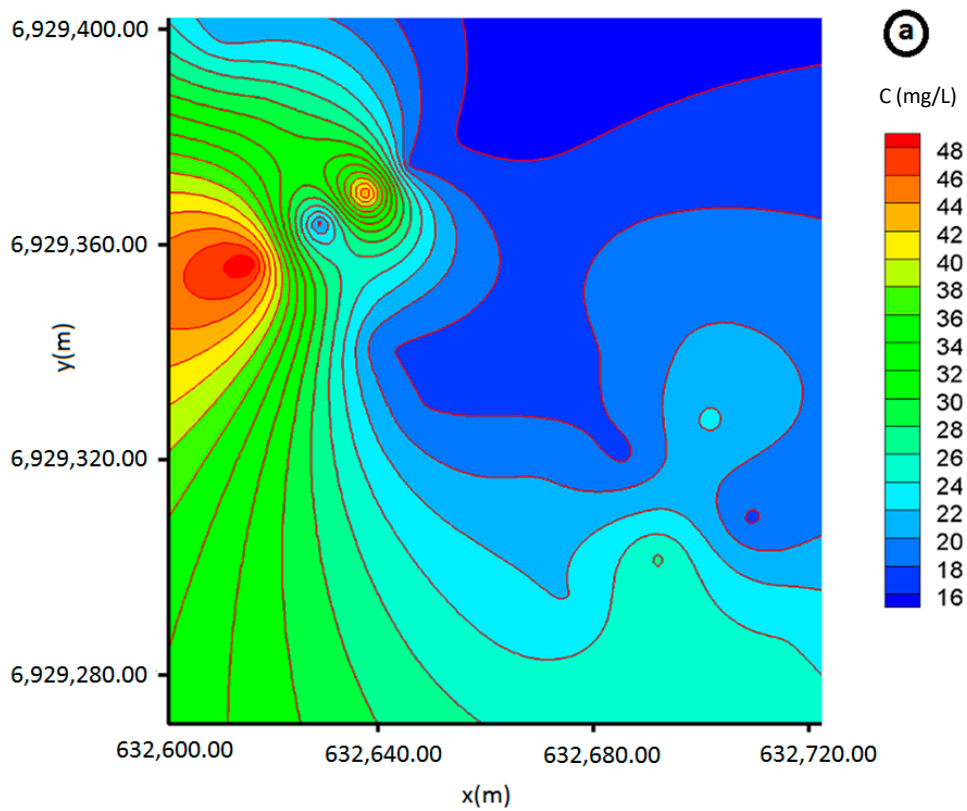
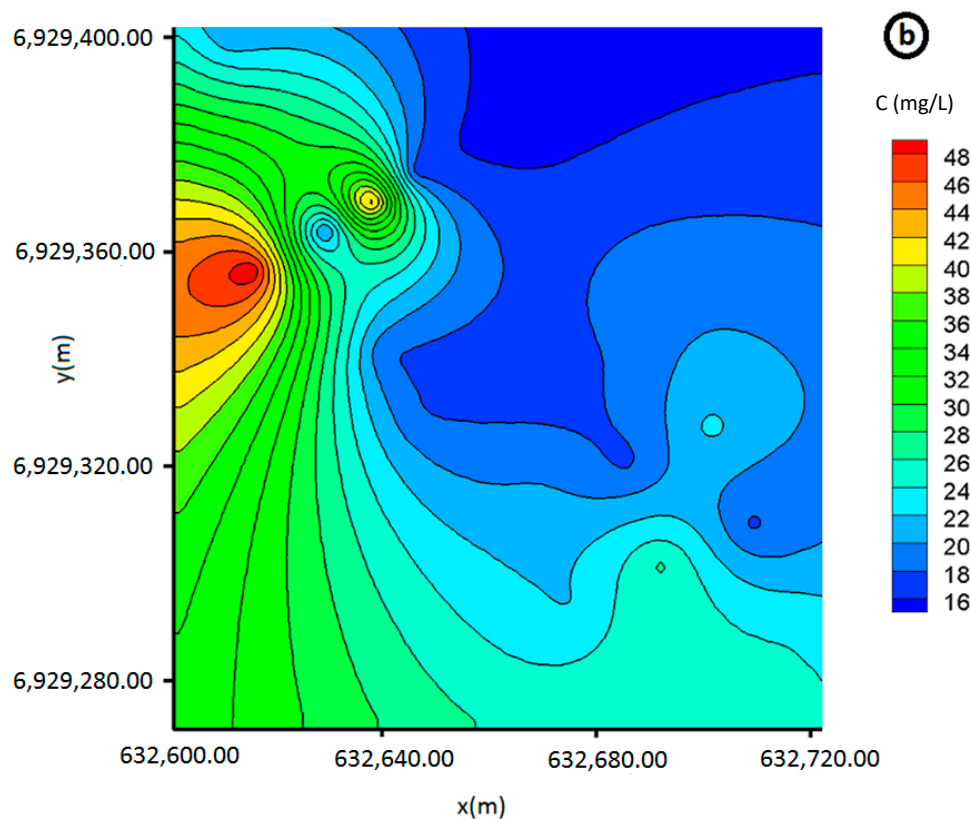


Figure 4. Cont.

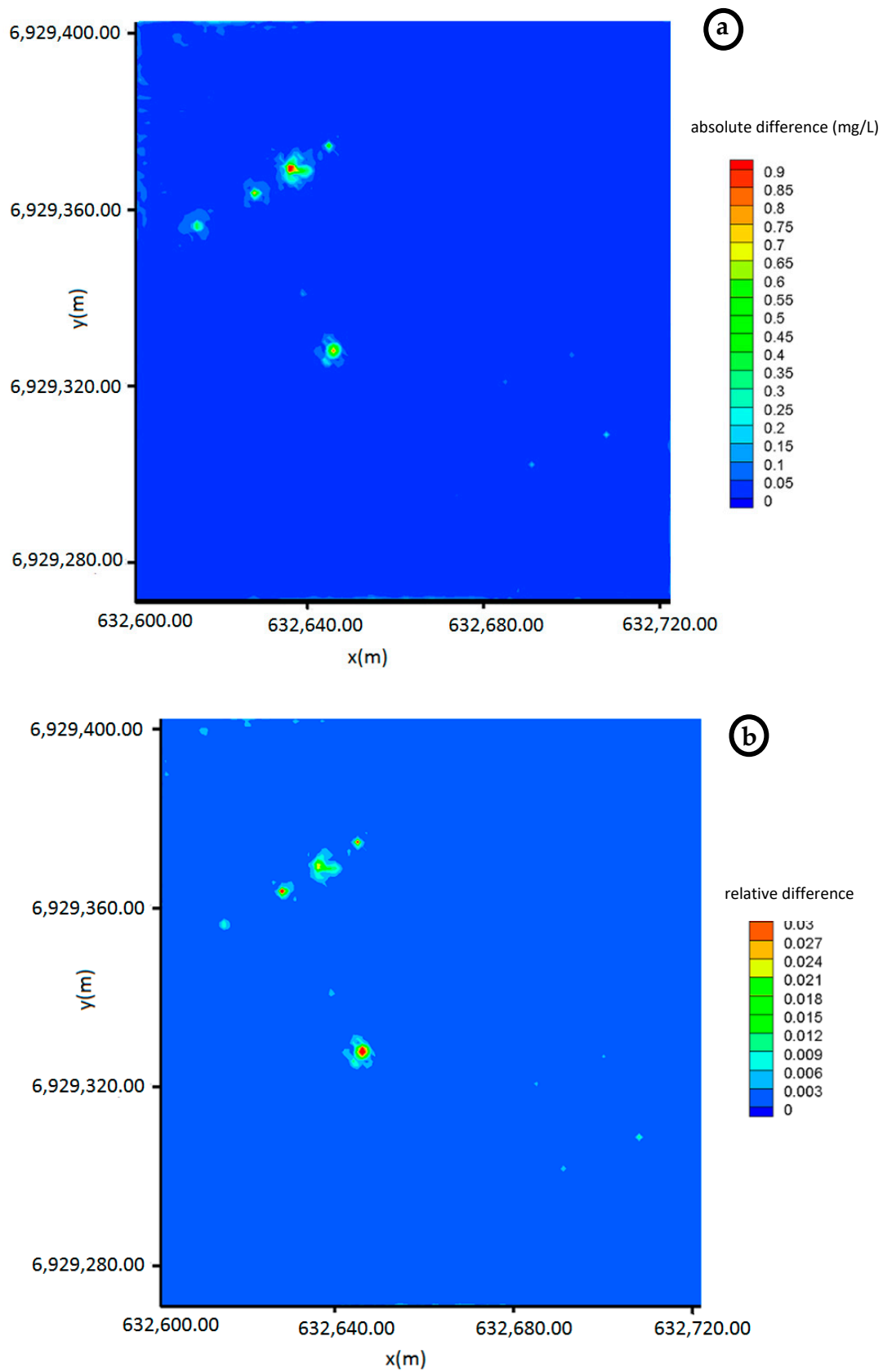


**Figure 4.** (a) The observed Chloride ion concentration  $C_{obs}$  (mg/L); (b) The computed Chloride ion concentration  $C_{comp}$  (mg/L).

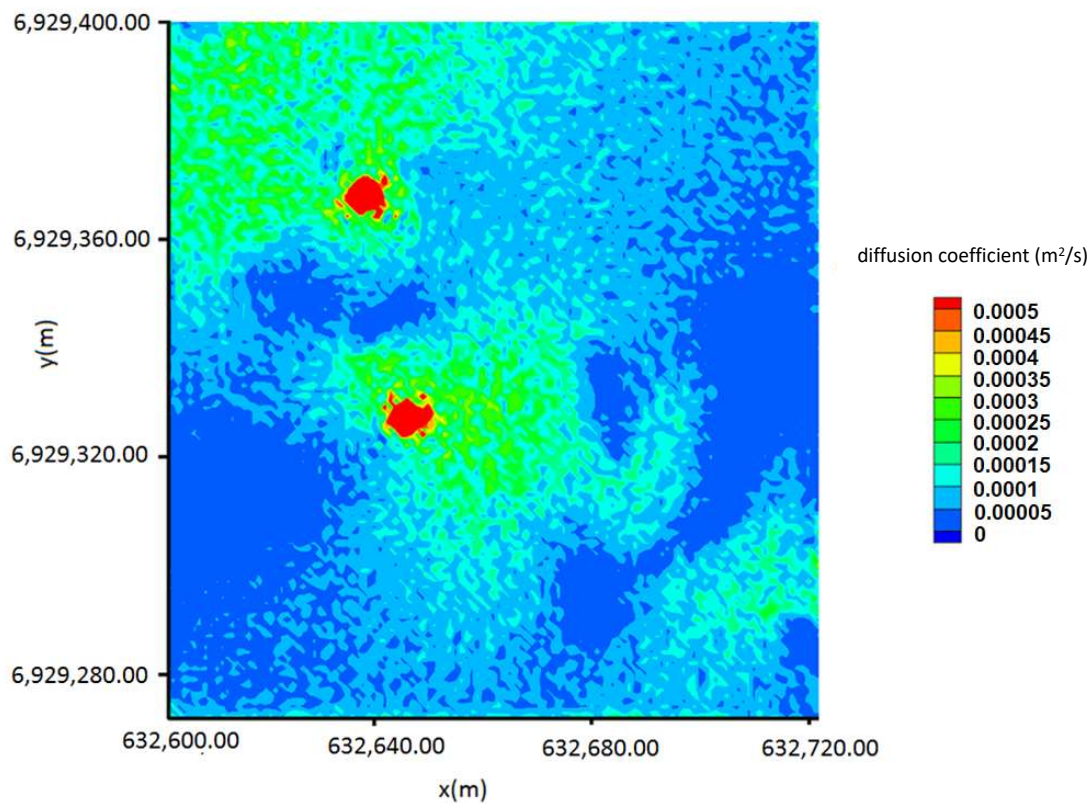
If we refer to Figure 2 which references the 20 measurement points, we observe that the most important errors were located at the piezometric points Pz8 Pz12, Pz13 and at the well point F4. In fact, this observation is not surprising, since, in our previous study of the hydrodynamics of the same area [20], we showed that these parts of the site are characterized by a complex geological and flow structure. These remarks are confirmed if we examine in addition Figure 3, which presents the flow streamlines. This figure shows that in the vicinity of these points, we observe the presence of a succession of sources (divergent streamlines) and of wells (converging streamlines) which induces a hydrodynamic difficulty to apprehend. Finally, if we admit the complexity of the geology of the site and the errors made on the measurements (either by the accuracy of the deployed sensors or by the difficulty of access to the measurement point), we can then conclude that the proposed coupling in this work (HYDRODISP-LBM/CMA-ES provides an excellent result for the calibration of models solved by the DP.

As mentioned above, the application of the HYDRODISP-LBM/CMA-ES coupling to the site also makes it possible to identify the HDT of the  $\text{Cl}^-$  tracer. Figure 6 shows the spatial distribution of the HDT. The analysis and the interpretation of the dispersion map (Figure 6) revealed higher dispersion in especially in the central and the north sectors of the HESB (about  $5 \times 10^{-4} \text{ m}^2/\text{s}$ ). In parallel, these sectors were characterized by the spatial distribution of the dispersion coefficients which varied between  $4 \times 10^{-5} \text{ m}^2/\text{s}$  and  $1.5 \times 10^{-4} \text{ m}^2/\text{s}$ . This variation shows the complexity of the chalk formation and its heterogeneity according to the groundwater flow, hydrogeological characteristics (permeability and transmissivity), porosity, and the presence of fractures. Figure 6 shows also that the maximum dispersion was obtained at points Pz8 and Pz14. This remark is not surprising since from a hydrodynamic point of view these two points behave like wells with maximum velocity (Figure 2). In view of Expression (3) which linearly links the dispersion to the velocity magnitude, we can then predict that the maximum dispersion will also be obtained at these two points.





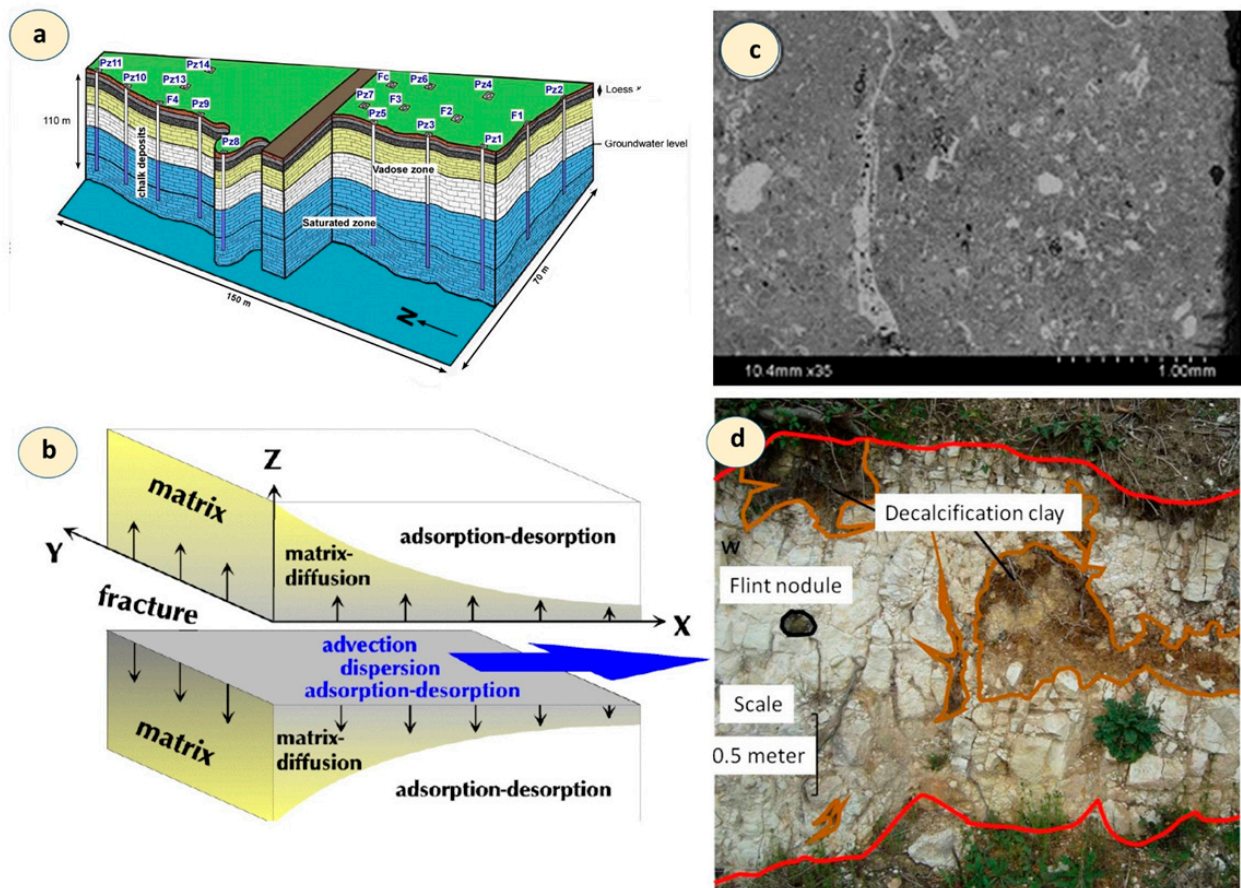
**Figure 5.** (a) The absolute error between  $C_{obs}$  and  $C_{comp}$  (mg/L); (b) The relative error between  $C_{obs}$  and  $C_{comp}$ .



**Figure 6.** The identified HDT of the Beauvais hydrogeological experimental site.

We specify that there are no references on the HDT concerning our study site HESB. The investigations presented in this paper are the first and new hydro-dispersive investigation results for this site. Consequently, we have no means of verifying the orders of magnitude of the values that we announce. Nevertheless, we believe that the approach of the identification of the tensor  $\overline{D}$  is commendable. Of course, there are a number of investigations [45–48] (most of them are experimental in the laboratory) that determine the diffusion coefficient of chloride. However, not the transverse dispersivity which we are interested in in this paper, and consequently, these works cannot inform us about the range of expected values of the dispersivity  $K_T$ . On the other hand, all these authors agree on the fact that the determination of this coefficient is a difficult exercise to carry out since it strongly depends on several parameters controlling the mixing processes in a given medium. Indeed, these mixing processes result from the two different transport mechanisms. The first is due to matrix diffusion and the second is governed by the transfer of solutes by the permeable environment flux which is characterized by fractures. According to [49], these solutes are transferred by the concentration gradient from the fluid parts contained in the permeable support (fractures) to the non-fluid parts (matrix) (Figure 7b). In the chalk environment (Figure 7a), the transport media required the first time for the characterization of the petrophysical characteristics of the chalk formation. Indeed, the use of the Scanning Electron Microscopy (SEM) (Figure 7c) analysis of the chalk samples coming from outcrops which are located near to our site HESB allows defining the porosity of the chalk. The carbonate matrix is characterized by a sedimentological wackstone texture. The SEM (Figure 7d) method provided an intra-granular porosity that varied between 25 and 30% [50]. This porosity could be classified according to the degrees of fissuration and to the vertical distribution allowing to deduce the primary and secondary porosity with 0.15–45% and 0.005–0.02, respectively [51]. The variation of diffusion coefficient could be related to the relationship between geometrical properties of fractures and matrix [52]. This complexity of the chalk aquifer and especially the variation of diffusion coefficients have

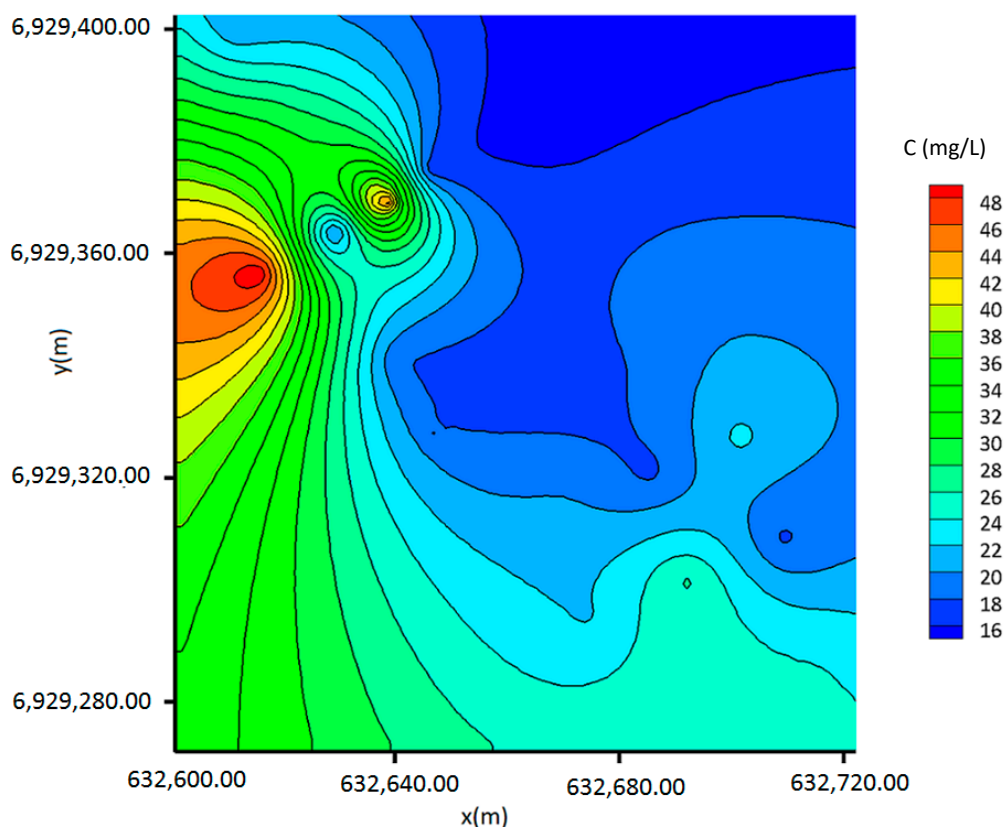
been evaluated in the Negev Desert [47] as a function of the porosity, permeability, the depth (Eocene section of the 285 m), and mineralogy characteristics of this formation. The estimated diffusion coefficient of the Negev Desert is very low (orders of magnitude about  $10^{-10}$  m<sup>2</sup>/s), but the most important results reveal that the diffusion coefficients of Eocene chalk depend especially on the porosity.



**Figure 7.** (a) Location of hydrogeological wells of HESB; (b) relationship matrix and fractures; (c) porosity analysis by SEM techniques; (d) Chalk outcrop used for SEM samples.

To assess the degree of importance of the heterogeneity of the dispersion tensor  $\bar{\bar{D}}$ , we performed a new simulation, but with a constant dispersion coefficient  $D_{mean}$ . This coefficient was obtained by the spatial mean of the identified tensor  $\bar{\bar{D}}_{opt}$  ( $D_{mean} = \frac{1}{mn} \sum_{\substack{0 \leq i \leq n \\ 0 \leq j \leq m}} \bar{\bar{D}}_{opt}(i,j)$ ).

Figure 8 shows the concentration field obtained by adopting the constant dispersion coefficient (in this study we find  $D_{mean} = 1.18 \times 10^{-4}$  m<sup>2</sup>/s). By comparing it with the Figure 4b, which was obtained by the heterogeneous tensor, we observe a quasi-perfect similarity except in the vicinity of the point Pz8 and the point Pz13 where the isocontours were, this time, off-center compared to the position of the source. This finding is not surprising since we mentioned previously that these zones of the studied area are characterized by a complex geological formation inducing a flow that is difficult to apprehend. In view of these results, we can conclude that the heterogeneity of the dispersion tensor does not play an important role in the dispersion of the tracers in this study area.



**Figure 8.** The computed Chloride ion concentration  $C_{comp}$  (mg/L) with constant dispersion coefficient ( $D_{mean} = 1.18 \times 10^{-4} \text{ m}^2/\text{s}$ ).

## 5. Conclusion and Perspectives

In this paper, we presented a new IOA called HYDRODISP-LBM/CMA-ES in order to perform the identification of the dispersion tensor problem. We reformulated this problem into an inverse problem. The innovative aspect of the new code lies in the introduction of LBM as a DP solver, but the association with a metaheuristic algorithm for the optimization phase is also original. LBM implementation for the DP solution is extremely fast in computing time. This gain of time allowed us to process simulations in high spatial resolution ( $\Delta x = \Delta y = 1\text{m}$ ) and also to compensate for the exorbitant calculation time required for the optimization step with a metaheuristic algorithm like CMA-ES. As an example, we proposed an IOA which uses the Finite Element Method [33], Control Volume Finite Element Method [34] for the DP solver and CMA-ES for the optimization (a less fine spatial resolution that we deal with in this paper). To identify the transmissivity tensor, we needed no less than 5 days of calculation. While the case treated here with a high-resolution LBM required only 2 days of calculation.

HYDRODISP-LBM/CMA-ES was successfully applied to the real case of the HESB. It gave excellent results in calibration mode. Likewise, the high spatial resolution that we were able to adopt thanks to the high performance of LBM, allowed us to highlight the heterogeneity of HDT. On the other hand, HYDRODISP-LBM/CMA-ES in identification mode seems to give satisfactory results, but the lack of data on the hydro-dispersive aspect of the site, did not allow us to criticize the approach. As perspectives, we are considering new investigations to improve and gain in the robustness of the proposed IOA by carrying out the following actions:

- 1 Conduct new experiments on the study site on tracking a tracer to build a database which we will use to validate the proposed IOA for identification mode.
- 2 Introduce anisotropy in LBM taking into account both longitudinal and transverse dispersivity ( $K_L$ ,  $K_T$ ), which amounts to writing  $\bar{D}$  in the form  $D_{ij}(x, y) = D_m + K_T \frac{\vec{v}_i \vec{v}_j}{\|\vec{v}\|} + (K_L - K_T) v_i v_j / \|\vec{v}\|$  to identify two parameters  $K_L$  and  $K_T$ .
- 3 Develop a strong coupling between hydrodynamics and dispersion. In fact, the IOA proposed here assumes knowledge of the velocity field  $v$ . However, it is not generally known, but calculated by solving Darcy's equation. The strong coupling, we propose consists of simultaneously solving Darcy's equation and the ADE transport equation.
- 4 Implement a parallel version of LBM to reduce computing time.

## 6. Patents

There are no patents resulting from the work.

**Author Contributions:** Conceptualization, H.S. and L.Z.; methodology, H.S. and L.Z.; software, H.S. and S.K.; validation, H.S. and S.K.; formal analysis, L.Z.; investigation, H.S., L.Z. and S.K.; resources, S.K. and H.S.; data curation, L.Z.; writing—original draft preparation, H.S., L.Z., S.K.; writing—review and editing, H.S. and L.Z.; visualization, H.S., L.Z., S.K.; supervision, H.S. and L.Z.; 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.

**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:** This study will help us to understand the mechanism of mixing processes of the chalk aquifer in the Oise region especially in the Hydrogeological Experimental Site of Beauvais (HESB, northern part of the Paris basin). The identification of HDT will contribute to the knowledge of the hydro-dispersive characteristics of the site in order to enrich the database on the studied site. Likewise, this study will strengthen our expertise on the site and in particular in the event of accidental and catastrophic pollution. 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 AGHYLE Research team "Agroecology, Hydrogeochemistry, Environments and Resources" of the Institut Polytechnique UniLasalle.

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

## References

1. Torabi, F.; Ahmad, P. *Simulation of Battery Systems: Fundamentals and Applications*; Academic Press/Elsevier: London, UK, 2019; 420p.
2. Petrášek, Z.; Schwille, P. Precise Measurement of Diffusion Coefficients using Scanning Fluorescence Correlation Spectroscopy. *Biophys. J.* **2008**, *94*, 1437–1448. [[CrossRef](#)] [[PubMed](#)]
3. Monteith, J.; Unsworth, M. *Principles of Environmental Physics: Plants, Animals, and the Atmosphere*, 4th ed.; Academic Press/Elsevier: Warsaw, Poland, 2013; pp. 679–705.
4. Schnelle, K.B., Jr. Atmospheric Diffusion Modelling. In *Encyclopedia of Physical Science and Technology: "Atmospheric Science"*, 3rd ed.; Meyers, R.A., Ed.; Elsevier: Amsterdam, The Netherlands, 2003; 396p.
5. Leibundgut, C.; Malozewski, P.; Külls, C. Tracers in Hydrology. In *Treatise on Water Science*; Wilderer, P., Ed.; Elsevier/IWA: London, UK, 2011; Volume 2, pp. 215–236.
6. Bellout, H.; Freidman, A.; Isakov, V. Inverse Problem in Potential Theory. *Trans. AMS* **1992**, *332*, 271–296. [[CrossRef](#)]
7. Vogel, C.R. Computational Methods for Inverse Problems. In *Frontiers in Applied Mathematics*; SIAM: Philadelphia, PA, USA, 2002; 200p.
8. Sun, K.K.; Isaac, M.D. Solution to Inverse Heat Conduction Problem in Nanoscale Using Sequential Method. *Numer. Heat Transf. Fundam.* **2003**, *44*, 439–456.
9. Belhachmi, Z.; Karageoghis, A.; Taous, K. Identification and Reconstruction of a Small Leak Zone in a Pipe by a Spectral Element Method. *J. Sci. Comput.* **2006**, *27*, 111–122. [[CrossRef](#)]
10. Choulli, M.; Yamamoto, M. Generic Well-Posedness of an Inverse Parabolic Problem—The Hölder-space approach. *Inverse Probl.* **1996**, *12*, 195–205. [[CrossRef](#)]

11. Choulli, M. An Inverse Problem in Corrosion Detection: Stability Estimates. *J. Inverse Ill-Posed Probl.* **2004**, *12*, 349–367. [[CrossRef](#)]
12. Sun, N.-Z. Inverse Problems in Groundwater Modeling. In *Theory and Applications of Transport in Porous Media*; Springer: Dordrecht, The Netherlands, 2013; Volume 6, 338p.
13. Zhou, H.; Gómez-Hernández, J.J.; Li, L. Inverse Methods in Hydrogeology: Evolution and Recent Trends. *Adv. Water Resour.* **2014**, *63*, 22–37. [[CrossRef](#)]
14. Carrera, J.; Alcolea, A.; Medina, A.; Hidalgo, J.; Slooten, L.J. Inverse Problem in Hydrogeology. *Hydrogeol. J.* **2005**, *13*, 206–222. [[CrossRef](#)]
15. Jawerth, B.; Lin, P. Lattice Boltzmann Models for Anisotropic Diffusion of Images. *J. Math. Imaging Vis.* **1999**, *11*, 231–237. [[CrossRef](#)]
16. Merks, R.H.; Hoekstra, A.G.; Sloot, P.M.A. The Moment Propagation Method for Advection-Diffusion in Lattice Boltzmann Method: Validation and Péclet Number Limit. *J. Comput. Phys.* **2002**, *183*, 563–576. [[CrossRef](#)]
17. Zhang, X.; Bengough, A.G.; Crawford, J.W.; Young, I.M. A Lattice BGK Model for Advection and Anisotropic Dispersion Equation. *Adv. Water Resour.* **2002**, *25*, 1–8. [[CrossRef](#)]
18. Ginzburg, I. Equilibrium-type and Link-type Lattice Boltzmann Models for Generic Advection and Anisotropic Dispersion Equation. *Adv. Water Resour.* **2005**, *28*, 1171–1195. [[CrossRef](#)]
19. Zhou, J.G. Lattice Boltzmann Method for advection and Anisotropic dispersion equation. *J. Appl. Mech.* **2011**, *78*, 021007. [[CrossRef](#)]
20. Zourhi, 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. [[CrossRef](#)]
21. Ciarlet, P.G. The Finite Element Method for Elliptic Problems. In *Studies in Mathematics and its Applications*; Lions, J.L., Papanicolaou, G., Rockafellar, R.T., Eds.; Elsevier: Amsterdam, The Netherlands, 1978; Volume 4, 548p.
22. Ramm, A.G. Inverse Problems. In *Mathematical and Analytical Techniques with Applications to Engineering*; Jeffrey, A., Ed.; Springer: Boston, FL, USA, 2005; 462p.
23. Kirsch, A. Inverse Problems with Applications in Science and Engineering. In *Applied Mathematical Sciences*, 2nd ed.; Springer: New York, NY, USA, 2011; Volume 120, 323p.
24. Wolf-Gladrow, D.A. Lattice Gas Cellular Automata and Lattice Boltzmann Models. In *Lecture Notes in Mathematics*; Springer: Berlin/Heidelberg, Germany, 2000; 1725; 315p.
25. Succi, S. The Lattice Boltzmann Equation for Fluid Dynamics and Beyond. In *Numerical Mathematics and Scientific Computation*; Oxford University Press: Oxford, MS, USA, 2001; Volume 22, 304p.
26. Batnagar, P.L.; Gross, E.P.; Krook, M. A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems. *Phys. Rev.* **1954**, *94*, 511–525. [[CrossRef](#)]
27. Mohamad, A.A. *Lattice Boltzmann Method “Fundamentals and Engineering Applications with Computer Codes”*, 2nd ed.; Springer: London, UK, 2019; 228p.
28. Zou, Q.; He, X. On Pressure and Velocity Boundary Conditions for the Lattice Boltzmann BGK Model. *Phys. Fluids* **1997**, *9*, 1591–1598. [[CrossRef](#)]
29. Bouzidi, M.; Firdaouss, M.; Lallemand, P. Momentum Transfer of a Boltzmann-Lattice Fluid with Boundaries. *Phys. Fluids* **2001**, *13*, 3452–3459. [[CrossRef](#)]
30. Hansen, N. The CMA Evolution Strategy: A Tutorial. *arXiv* **2016**, arXiv:1604.00772.
31. Bayer, P.; Finkel, M. Evolutionary Algorithms for the Optimization of Advective control of Contaminated Aquifer Zones. *Water Resour. Res.* **2004**, *40*, 1–19. [[CrossRef](#)]
32. Bayer, P.; Finkel, M. Optimization of Concentration Control by Evolution Strategies: Formulation, Application and Assessment of Remedial Solutions. *Water Resour. Res.* **2007**, *43*, 1–19. [[CrossRef](#)]
33. Smaoui, H.; Zouhri, L.; Kaidi, S.; Carlier, E. Combination of FEM and CMA-ES algorithm for transmissivity identification in aquifer systems. *Hydrol. Processes* **2018**, *32*, 264–277. [[CrossRef](#)]
34. Smaoui, H.; Maqsoud, A.; Kaidi, S. Transmissivity Identification by Combination of CVFEM and Genetic Algorithm: Application to the Coastal Aquifer. *Math. Probl. Eng.* **2019**, *2019*, 3463607. [[CrossRef](#)]
35. Hansen, N.; Ostermeier, A. Completely Derandomized Self-Adaptation in Evolution Strategies. *Evol. Comput.* **2001**, *9*, 159–195. [[CrossRef](#)]
36. Choulli, M. Une introduction aux Problèmes Inverses Elliptiques et Paraboliques. In *Mathématiques & Applications*; Allaire, G., Benaïm, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 65, 266p.
37. Kern, M. Numerical Methods for Inverse Problems. In *Mathematics and Statistics*; Limnios, N., Ed.; ISTE/John Wiley & Sons, Inc.: London, UK; Hoboken, NJ, USA, 2016; 234p.
38. Samarskii, A.A.; Vabishchevich, P.N. Numerical Methods for Solving Inverse Problems of Mathematical Physics. In *Inverse and Ill-Posed Problems Series*; Kabanikhin, S.I., Ed.; Walter De Gruyter: Göttingen, Germany, 2008; 453p.
39. Shuai, L.; Pereverzev, S.V. Regularization Theory for Ill-Posed Problems. In *Inverse and Ill-Posed Problems Series*; Kabanikhin, S.I., Ed.; Walter De Gruyter: Göttingen, Germany, 2013; Volume 58, 304p.
40. Hadamard, J. *Lectures on Cauchy’s Problem in Linear Partial Differential Equations*; Yale University Press: London, UK, 1923; 334p.
41. Engl, H.W. Regularization methods for the stable solution of inverse problems. *Surv. Math. Ind.* **1993**, *3*, 71–173.
42. Engl, H.W.; Hanke, M.; Neubauer, A. Regularization of Inverse Problems. In *Mathematics and Its Applications*; Hazewinkel, M., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1996; 322p.
43. Tikhonov, A.N. On the solution of ill-posed problems and the method of regularization. *Dokl. Akad. Nauk. SSSR* **1963**, *151*, 501–504.
44. SCHLUMBERGER: *Diver Sensor Manual*; Schlumberger Water Services: Delft, The Netherlands, 2014; 33p.

45. Fretwell, B.A. Distribution of Contaminants in the Seasonally Unsaturated Zone of the Chalk Aquifer. Ph.D. Thesis, University College London, London, UK, 1999; 560p.
46. Barone, F.S.; Rowe, R.K.; Quigley, R.M. Estimation of Chloride Diffusion Coefficient and Tortuosity Factor for Mudstone. *J. Geotech. Eng.* **1992**, *118*, 1031–1046. [[CrossRef](#)]
47. Polak, A.; Nativ, R.; Wallach, R. Matrix diffusion in northern Negev fractured chalk and its correlation to porosity. *J. Hydrol.* **2002**, *268*, 203–213. [[CrossRef](#)]
48. Erdogdua, S.; Kondratovab, I.L.; Bremner, T.W. Determination of chloride diffusion coefficient of concrete using open-circuit potential measurements. *Cem. Concr. Res.* **2004**, *34*, 603–609. [[CrossRef](#)]
49. Ahmed, F.A.A. Contaminant Transport in Fractured Chalk Aquifer at Sigerslev, Denmark as Characterized by Tracer Techniques. Ph.D. Thesis, Rheinischen Friedrich-Wilhelms-Universität, Bonn, Germany, 2004; 105p.
50. Zouhri, L.; Lutz, P. Hydrogeophysical Characterization of the Porous and Fractured Media (Chalk Aquifer in the Beauvais, France). *Environ. Earth Sci.* **2016**, *75*, 343. [[CrossRef](#)]
51. Domenico, P.A.; Schwartz, F.W. *Physical and Chemical Hydrogeology*; Wiley & Sons: New York, NY, USA, 1998; 505p.
52. Polak, A. Chemical diffusion between a fracture and surrounding matrix: Measurement by Computed Tomography and Modelling. *Water Resour. Res.* **2003**, *39*, 1106. [[CrossRef](#)]