Are oscillatory pumping tests a suitable tool for characterizing aquifer's heterogeneity?

Mara Meggiorin

Thesis to obtain the Master of Science Degree in

Environmental Engineering

Supervisors: Prof. Alberto Bellin
Prof. Luis Filipe Tavares Ribeiro

Examination Committee

Chairperson: Prof. António Jorge Silva Guerreiro Monteiro

Supervisor: Prof. Luis Filipe Tavares Ribeiro

Member of the Committee: Prof. Rodrigo de Almada Cardoso Proença de Oliveira

July 2016

Everything should be made
as simple as possible,
but no simpler.

Albert Einstein

Contents

List of Figures

List of Tables

Abstract

The subsurface water movement is an important step of the water cycle which fully understanding is not yet achieved. Indeed, groundwater flow strongly depends on encountered geological underground formations and it is difficult to reproduce not being directly observable. Many tests and models are applied to determine formation properties averages. Usually they are time-consuming and assume a homogeneous field, or only vertically heterogeneous. This way they do not consider the possible presence of fast pathways which considerably influence contaminant movement. Therefore, this paper proposes to use the steady-periodic model for oscillatory pumping test in order to characterize hydrogeological formations evaluating local properties: conductivity and storativity. The test stimulates aquifers with several periodic signals of different frequencies, the model elaborates data by a modified Bayesian inversion and the pilot point approach in which governing equations of subsurface movement are used in the Fourier space. Different typology of results and simulations are carried out in order to assess capability to infer hydraulic properties variations with limited prior information. In conclusion, the average of Monte Carlo simulations is able to infer the heterogeneity pattern of both properties but it does not correctly evaluates extreme values. While, the best simulation captures local properties values and their statistics with acceptable but variable accuracy. Regarding contaminant transport, connectivity results approximatively simulated. The most important observation is the higher sensibility of the model to storativity than to conductivity. Models currently used usually focus on conductivity and storativity is not accurately reproduced. While this work provides a sensible measurement for this lacking property of the field.

Sommario

Il movimento dell'acqua nel sottosuolo é un passaggio importante del ciclo della acqua, é anche uno dei passaggi piú difficili da modellare dato che il flusso acquifero varia fortemente a seconda delle formazioni geologiche incontrate. Visto la sua variabilitá e l'impossibilitá di osservare direttamente lo scorrimento, l'acqua sotterranea risulta difficile da simulare correttamente. Attualmente vi sono molti test e modelli finalizzati alla determinazione delle proprietá medie dellaquifero, ma solitamente sono costruiti sull'assunzione di assoluta omogeneitá del campo, escludendo quindi la possibile presenza di percorsi veloci che sono estremamente importanti nel trasporto di contaminanti. Questa tesi quindi vuole caratterizzare l'eterogeneitá delle formazioni usando il modello steady-periodic per un test di pompaggio alternato. Il pompaggio stimola laquifero con segnali periodici di diverse frequenze ed il modello elabora le osservazioni. Tale elaborazione consiste nell'usare le equazioni del moto nello spazio di Fourier, un'inversione simile alla Bayesiana, i metodi dei punti pilota e Monte Carlo. Il modello é stato provato con diversi test e molte tipologie di risultati sono riportate per una comparazione che facilita la completa comprensione dell'efficacia del modello nel riprodurre sia conduttivitá e storativitá locali che il trasporto di contaminanti. I test riportati in questo studio dimostrano che la media delle simulazioni Monte Carlo riesce ad individuare il pattern di eterogeneitá ma non riconosce i valori massimi, mentre la migliore simulazione vi riesce individuando quindi migliori statistiche. Il trasporto di inquinante non riesce invece ad essere riprodotto, solo la funzione di connettivitá viene approssimativamente individuata. Infine, l'aspetto piú importante del modello proposto é la sua sensibilitá alla storativitá. Considerato che i modelli solitamente si concentrano nella conduttivitá del campo, questo modello puó essere efficientemente utilizzato per la riproduzione della proprietá usualmente meno accurata.

Resumo

O deslocamento da água subterrâenea é uma fase importante do ciclo da água e sua compreensão plena ainda não foi alcançada. De facto, o fluxo das águas subterrâneas depende fortemente da formação geológica do aquifero e é dificil a sua reprodução não sendo directamente observável. Muitos testes e modelos são aplicados para determinar as propriedades médias de formação. Normalmente, assumem-se um campo homogéneo ou heterogéneo apenas na vertical. Dessa forma, eles não consideram a possivel presença de caminhos mais rápidos que influenciam consideravelmente o movimento de contaminantes. Portanto, este trabalho propõe a utilização do modelo estacionário periódico para teste de bombeamento oscilatório de modo a caracterizar formações hidrogeológicas avaliando propriedades locais: condutividade e capacidade de armazenamento. Os testes estimulam aquiferos com vários sinais periódicos de diferentes frequências, o modelo de dados elaborado para uma inversão semelhante ao Bayesiana e a abordagem do ponto piloto em que as equações que governam o movimento do subsolo são usadas no espaço de Fourier. Diferentes tipologias de resultados e simulações são realizadas a fim de avaliar a capacidade de inferir nas variações das propriedades hidráulicas com informação prévia limitada. Em conclusão, tem-se que a média das simulações de Monte Carlo é capaz de deduzir o padrão de heterogeneidade de ambas as propriedades, mas ela não avalia correctamente os valores extremos. Embora a melhor simulação capte valores das propriedades locais de estatisticas com precisão aceitável, ela é variável. Em relação ao transporte de contaminantes, a conectividade é aproximadamente simulada. A observação mais importante é a maior sensibilidade do modelo para capacidade de armazenamento do que a condutividade. Modelos atualmente usados geralmente concentram-se em condutividade mas capacidade de armazenamento não é reproduzido com precisão. Enquanto que este trabalho fornece uma medida sensivel para esta proprietade de campo.

Introduction

It is generally known how water continuously moves following the water cycle: evaporating, raining, moving on the soil and through it. It is also commonly understood how much water and its cycle balance is important in our life for both humans, flora, fauna and the whole Earth. Water importance and vulnerability is amply discussed across fields. Society is moving with actions and precaution to a high respect of this life resource. However, in order to make these efforts more effective it is essential that science fully comprehends the hydrological cycle in its entirety. Up to this point, hydrologists can describe quite in detail all steps of the water cycle that are visible and directly experienced. Mainly two points remain uncertain: the transpiration step and the subsurface water movement. The former is the process of water movement through a plant and its evaporation from aerial parts, the latter is the flow of water beneath earth's surface. These are indeed steps difficult to observe and model: the first because it is a micro-scale cellular process, the latter because takes place completely underground. Groundwater is defined as water contained in soil pores and in the fractures of rocks. It is not stopped, it slowly moves downwards recharging rivers, sustaining ecosystems and yielding water to wells. Indeed, subsurface water is often preferred as water source for human uses because it is usually cheaper to make drinkable and less vulnerable to pollution than surface water. On the other hand, when it is polluted, it is difficult to timely notice and to clean up.

These are some reasons for which groundwater is the topic of this work. And, specifically, this paper proposes to use the steady-periodic model for oscillatory pumping test in order to characterize hydrogeological formations, mainly defining their heterogeneity. A deep and precise comprehension of underground heterogeneity can be very useful in several situations, such as to better understand the water resource extraction's effect or the pollution plume dynamic and extension. Indeed, it is well-known that infiltration process, groundwater storage and contaminant spreading are controlled by natural heterogeneity and the by large spatial variability of hydraulic parameters. However, heterogeneity is great and can often have multiple spatial scales of variability, leading to different levels of characterization depending on the hydrogeological objectives. The model proposed in this paper provides a deep understanding of the spatial distribution of hydraulic properties that is generally needed, for example, in contaminant transport studies at the local scale.

Many models exist and are applied in order to determine the properties' averages and, therefore, the average response of the field to water movement. The traditional way to proceed is by assuming that hydraulic properties are spatially constant; when this assumption does not lead to fit hydrological data, only the vertical heterogeneity is examined. Unfortunately, when this approach is taken, heterogeneity gets characterized very approximatively. In the field, pumping tests analyze large areas with low precision, while other tests are able to produce detailed results for the closer area but inaccurate estimations are given for areas only few tens meters away. Oscillatory pumping tests seem to have the potential to overcome these limitations. The presented model elaborates observational data of this test and achieves an improved characterization in the area. Comparing to other existing post-processing codes, the major contribution is the use of governing equations of groundwater movement in the Fourier space. Thus, computational time is considerably reduced and heterogeneity information is fully used.

Aim of the study is to assess capability of oscillatory pumping tests to infer hydraulic properties variations. This is done in three stages. First, the study presents and describes the pumping test and the model. Then a preliminary sensitivity analysis is carried out to determine best parameters to use in the following simulations. Finally the model is tested and validated by the mean of simulations, starting from a simple one and concluding with a random field characterized by high variances.

1

Hydrogeological formations characterization

Underground formations are difficult to study and fully comprehend because we cannot directly see them. Anyway, there are some tests useful to determine some important characteristics of hydrological formations. These properties, expressed by certain parameters, help us to better understand the groundwater behavior.
This chapter aims to give the basic knowledge to understand all the following passages. In the first section, the general underground structure is described, the types of aquifers and important concepts are explained. In the second one, parameters useful to describe groundwater movement are listed and explained in detail. Following, the most applied tests to determine these parameters are reported and described, the general scope is to understand their applicability, advantages and disadvantages. In the last section, the test coded in PhasONE, the oscillatory pumping test, is specifically described.

1.1 General structure

Underground there is water coming from infiltration processes of surface streams and runoff. This water infiltrates and, depending on the geological structures it runs into, it moves downwards following different paths. Indeed, geological formations have a broad range of permeability to water: they can be completely saturated with water and groundwater can flow easily or they can be completely impermeable and water cannot flow through them.

1.1.1 Hydrogeological formations

As just mentioned, subsurface water flows where there are geological structures able to transmit it, therefore water tends to channel into permeable formations. However, between these formations there are several grades of permeability to water and it is possible to classify them in several hydrogeological formations:

- Aquifer: single or group of geologic saturated formations able to transmit and yield significant quantities of water. They consist of materials at high permeability, such as sands and porous sandstones. This high value is attributed to connected pores between sand grains or fractures in rocks. They are economically suitable to extract water from wells or springs.

- Aquitard: confining bed that retards the flow of water but it does not stop it. They transmit small amounts of water that are significant for the regional groundwater flow but they are not convenient to supply water. Materials with low permeability are present, such as siltstone or mudstone.

- Aquiclude: hydrogeologic unit not able to transmit water at a valuable rate, although porous and capable of storing water. It usually acts as a barrier to the regional groundwater flow. They are characterized by rocks such as clays, shales and metamorphic rocks.

- Aquifuge: hydrogeologic unit completely isolated from any aquifer or infiltration, hence it cannot store or transmit water.

1.1.2 Aquifers

In order to extract water for any supply, technicians look for aquifers because they are the right kind of formations to store and release water to wells. Moreover, if there is underground pollution, usually it dissolves in the nearest aquifer and moves with it reaching also far distances with the necessary time. Focusing therefore on aquifers, there are many types of them depending mainly on boundaries: unconfined, perched, confined, leaky and multiple aquifers.
Unconfined aquifers are also called water-table or phreatic aquifers. They are characterized by an impermeable layer or an aquitard on the bottom and by a free surface on the upper boundary. This means that the aquifer thickness is simply

the whole saturated area and the free surface of the aquifer is under atmospheric pressure. Usually a capillary fringe is present above the water table. In this upper zone, groundwater fills pores above the water table seeping up by capillary action. The lower part of this fringe is completely saturated as below the water table, but there the relative pressure is negative, that means it is lower than the atmospheric one. It is important to underline that normally the water under capillarity actions is not extracted through the well and that usually the hydraulic effects of this zone are neglected in studies. Generally, a way to recognize an unconfined aquifer is looking at the water level in a well: the level is equal to the water-table elevation at the same location. Moreover, the well is defined a water-table well.

Perched aquifers are a particular type of the unconfined typology. They are still under the water table conditions but they locate above the main water-table. It happens if a relatively small aquitard is present between the main water-table and the ground surface.

Confined aquifers are also called pressure or artesian aquifers. They are bounded both on the bottom and on the top by impermeable or slightly permeable layers. Also this type can be recognized looking at the water level in a well: it is above the upper boundary. This water level is called potentiometric or piezometric surface, it represents an imaginary surface defining the level to which water would rise were it completely pierced with wells. If this piezometric surface in the well is higher than the ground, the well is defined as a flowing well. Otherwise, it is called an artesian well.

Leaky aquifers are those that lose or gain water through confining aquitards. The term nonleaky is also used in opposite conditions.

Generally, aquifers are part of a more complex hydrogeologic system where aquifers and separating confining layers are present. The whole system is called a system of multiple aquifers.

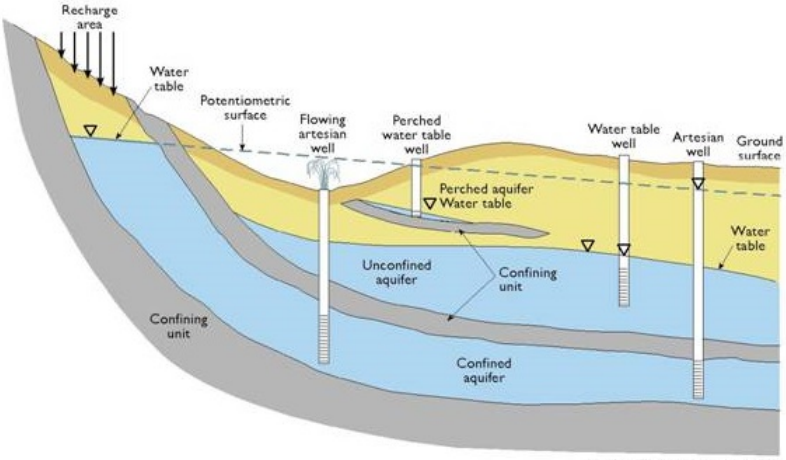

Figure 1.1.1: Types of aquifers. Source: Hartman et al. (1969)

1.2 Main parameters

The groundwater movement is mainly driven by variable hydraulic head and influenced by the permeability of the aquifer and its ability to yield water. For a steady-state flow condition, the hydraulic conductivity is the parameter that describes the fluid mobility and quantitatively relates pressure gradients to its velocity. While, for transient flow conditions, the formation capacity to store or to release fluid becomes also relevant. There are some hydraulic properties both of soil and fluid that determine the behavior of the fluid passing through the soil system under certain conditions.

1.2.1 Hydraulic head and drawdown

In the previous section, the piezometric surface was already anticipated classifying aquifers. Here, it is explained more in detail its basic concept: the hydraulic head, piezometric head or simply head. It is a combined measure of elevation and water pressure at a certain point in an aquifer, as shown in figures 1.2.1 and 1.2.2. It is a measure of the total energy of the water. The head in a certain point is simply the water level in a piezometer, which is a small tube with the bottom located in the aquifer. In an unconfined aquifer, the level is the same as the water table elevation;

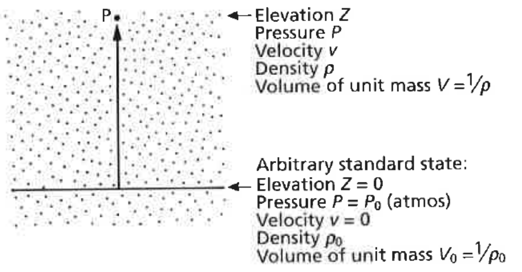

Figure 1.2.1: Work done in moving a unit mass of fluid from the standard state to a point P. Source: Hiscock and Bense (2014)

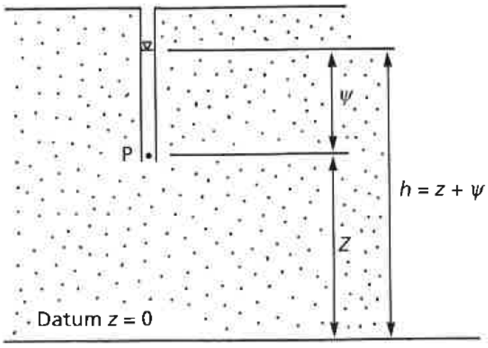

Figure 1.2.2: Relation between hydraulic head h, pressure head Ψ and elevation z at point P in a porous material. Source: Hiscock and Bense (2014)

while in a confined aquifer, the water level is higher than the upper boundary and it is equal to the piezometric surface elevation. The piezometric head can be easily calculated as:

$$H = z + \frac{P(x,y,z,t)}{\gamma} = z + \frac{P(x,y,z,t)}{\rho g} \tag{1.2.1}$$

where H is the hydraulic head $[L]$, z is the elevation at the piezometer bottom $[L]$, $P(x,y,z,t)$ is the gauge pressure in that specific point and instant $[M/(LT^2)]$ and γ is the liquid unit weight $[M/(L^2T^2)]$. The term $P(x,y,z,t)/\gamma$ is called the pressure head and it is the elevation of the water column in the piezometer. This relationship is valid in both saturated and unsaturated zones of porous material, though it is necessary to underline that the pressure head is positive in saturated zones while it is negative in unsaturated ones due to the assumption of zero value atmospheric pressure.

The trend of hydraulic head through an aquifer is essential to understand how groundwater flows. Indeed, the head is the main flow driver. If it is constant, there is no flow in the formation; while if it varies, the water moves in direction of the lower head (toward lower energy). Therefore, actually the main driver is the head difference, the hydraulic gradient. It is also possible to conclude that groundwater can move uphill if the hydraulic gradient drives this way.

Another important concept, linked to the hydraulic head, that is important to know is the drawdown, s $[L]$. Indeed, the piezometric head varies under water extraction or injection conditions and the vertical distance between the initial head and the final head in a certain point is called drawdown. In unconfined aquifers, it is the drop of the water table; in confined aquifers, it is the drop of the piezomtric surface. Drawdown is indissolubly linked with the presence of an operating well extracting or injecting water or of natural phenomena, such as tides and other seasonal variations. Moreover, in extraction or injection conditions it is possible to define the area of influence as the nearby zone affected by a drawdown. For example, a pumping well in a homogeneous formation will have a perfectly circular area of influence, which radius depends on the discharge extracted or injected. The following figure shows the drawdown of a well, namely its cone of depression.

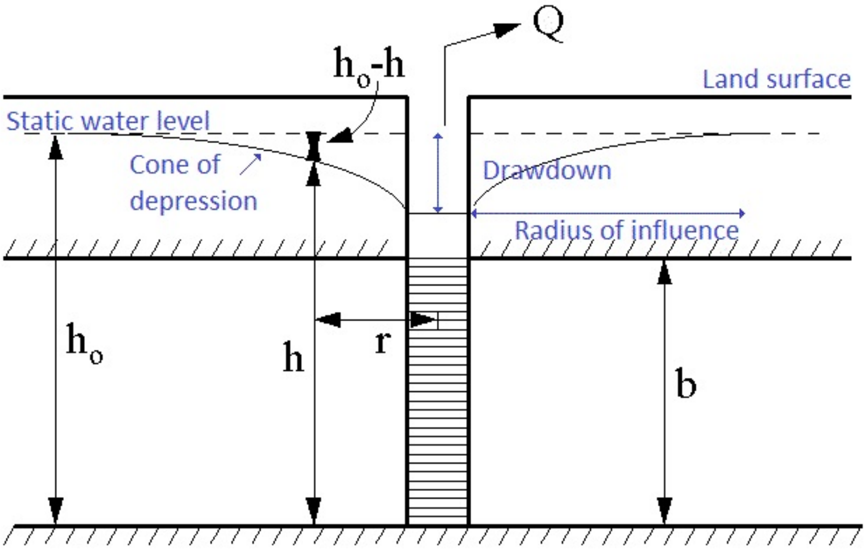

Figure 1.2.3: Section of a pumping well showing drawdown, cone of depression and radius of influence

1.2.2 Hydraulic conductivity and transmissivity

The main and most indicative parameter to characterize an aquifer is the hydraulic conductivity K, one of the hydraulic properties of both soil and fluid: it measures the soil's capacity to transmit water in presence of an hydraulic gradient, that is the ease of movement of water through a porous material. This parameter is defined as a constant of proportionality relating the specific discharge under a unit gradient in Darcy's law:

$$q_i = -K \frac{dH}{dx_i} \tag{1.2.2}$$

where q_i is the specific discharge in the porous medium in the i-th direction, also called Darcy's velocity, and dx_i is the distance in the analyzed direction. Basing on the equation 1.2.1, the hydraulic conductivity is the ratio of the Darcy's velocity to the hydraulic gradient applied. Therefore, K has the same dimension of the velocity $[L/T]$. This parameter depends obviously on the soil characteristics: such as the grain size, the structure of the soil matrix, the type of fluid and the saturation of the matrix. The most relevant properties of the soil matrix are porosity, pore size distribution, pore shape, specific surface and tortuosity; while the most relevant ones of the fluid are its density and viscosity. Indeed, the hydraulic conductivity for a subsurface saturated system can be expressed also as (Hiscock and Bense, 2014):

$$K = \frac{k\rho g}{\mu} \tag{1.2.3}$$

where k is the intrinsic permeability of the soil $[L^2]$, g is the gravitational acceleration $[L/T^2]$, ρ and μ are respectively the density $[M/L^3]$ and the viscosity $[M/(LT)]$ of the fluid. The saturated hydraulic conductivity can largely varies in soils, its potential range expands for several orders of magnitude, depending on the soil material ($10^{-12} m/s$ for tight clays to $10^{-1} m/s$ for coarse gravels) and its variability has many different spatial scales, from less than m^3 to greater than km^3. Following, in table 1.2.2 are reported the representative values for different soil textures.

Another property of the aquifer, related to the hydraulic conductivity, is the transmissivity T. It is the rate of water under a unit hydraulic gradient through a unit width of aquifer of given saturated thickness. This aquifer's property is linked to the aquifer's hydraulic conductivity as shown following:

$$T = Kb \tag{1.2.4}$$

where b is the aquifer thickness $[L]$, therefore the transmissivity dimension is $[L^2/T]$. It is here necessary to specify that usually this parameter refer to a confined aquifer

Table 1.1: Representative values of saturated hydraulic conductivity of different soil textures. Source: Clapp and Hornberger (1978).

Texture	Saturated Hydraulic Conductivity	
	[m/yr]	[m/s]
sand	$5.50\ 10^{3}$	$1.74\ 10^{-4}$
loamy sand	$4.93\ 10^{3}$	$1.56\ 10^{-4}$
sandy loam	$1.09\ 10^{3}$	$3.46\ 10^{-5}$
silty loam	$2.27\ 10^{2}$	$7.20\ 10^{-6}$
loamy sand	$2.19\ 10^{2}$	$6.94\ 10^{-6}$
sandy clay loam	$1.99\ 10^{2}$	$6.31\ 10^{-6}$
silty clay loam	$5.36\ 10^{1}$	$1.70\ 10^{-6}$
clay loam	$7.73\ 10^{1}$	$2.45\ 10^{-6}$
sandy clay	$6.84\ 10^{1}$	$2.17\ 10^{-6}$
silty clay	$3.21\ 10^{1}$	$1.02\ 10^{-6}$
clay	$4.05\ 10^{1}$	$1.28\ 10^{-6}$

of thickness b, however it can be also used for unconfined aquifers where b is the saturated thickness or, equivalently, it is the height of the water table above the low permeable or impermeable bottom.

1.2.3 Specific storage, specific yield and storativity

Regarding the aquifer capacity to store or release water, it is strongly linked to the formation compressibility and elasticity. Indeed, aquifers are compressible and elastic as all other solid mediums, however they are composed by solid grains and water filled pores. Therefore, when compressed, an aquifer yields water. There are many parameters to describe this release: the specific storage S_s, the specific yield S_y and the storativity S. Whereas the last refers to water stored in the entire thickness of an aquifer, others refer to the storage in a specified small volume.
First of all, it is here important to define exactly the porosity n. It is the fraction of void space in the formation:

$$n = \frac{V_v}{V_t} \tag{1.2.5}$$

where V_v is void volume $[L^3]$ and V_t is total volume $[L^3]$. However, the water cannot move through all these pores, it can move only through the connected pores. In order to take this into consideration, the effective porosity n_e is defined as the ratio of the connected void volume to the whole volume of the formation.

The specific storage $[L^{-1}]$ takes into consideration the compressibility of soil and water as follows:

$$S_s = \rho g(\alpha + n_e \beta) \tag{1.2.6}$$

where α is formation compressibility $[T^2 L/M]$, n_e is effective porosity $[-]$ and β is water compressibility $(4.410^{-10} ms^2/kg)$ $[T^2 L/M]$. Usual ranges are reported in the following table for some soil textures.

The specific yield, sometimes called effective porosity, is defined only for unconfined aquifers. It is the volume of water released from storage per unit surface area and per unit decline of the water table. It is dimensionless and ranges in value from 0.01 to 0.30. It can be expressed as:

Table 1.2: Representative values of Specific Storage of different soil textures. Source Batu (1998)

Material	$S_S[m^{-1}]$		
Plartic clay	$2.56\ 10^{-3}$	to	$2.03\ 10^{-2}$
Stiff clay	$1.28\ 10^{-3}$	to	$2.56\ 10^{-3}$
Medium hard clay	$9.19\ 10^{-4}$	to	$1.28\ 10^{-3}$
Loose sand	$4.92\ 10^{-4}$	to	$1.02\ 10^{-3}$
Dense sand	$1.28\ 10^{-4}$	to	$2.03\ 10^{-4}$
Sdense sandy gravel	$4.92\ 10^{-5}$	to	$1.02\ 10^{-4}$
Fissured rock	$3.28\ 10^{-6}$	to	$6.89\ 10^{-5}$
Sound rock		$<$	$3.28\ 10^{-6}$

$$n = S_y + S_r \tag{1.2.7}$$

where n is the total porosity $[-]$ and S_r is the specific retention $[-]$. The latter is the amount of water retained against the gravity force by capillarity forces.

The storativity S is differently defined for confined and unconfined aquifers. In confined aquifers, the storativity is equal to:

$$S = S_s b \tag{1.2.8}$$

where b is the aquifer thickness $[L]$. The usual range of this parameter is from 0.005 to 0.00005. While in unconfined aquifers, the storativity is equal to:

$$S = S_y + S_s b \tag{1.2.9}$$

However, usually $S_s b$ is much smaller than the specific yield, therefore, in this case, the storativity is simply equated with specific yield.

1.2.4 Diffusivity

Diffusivity, D $[L^2/T]$, is a measure of the rapidity with which a pressure disturbance disseminates through the formation. Although not commonly used, it is significant because aquifers characterized by a large hydraulic diffusivity respond quickly to changed hydraulic conditions at one location, transmitting to nearby regions. Diffusivity is a combination of previous ones: the formation transmissivity and storage as follows:

$$D = \frac{T}{S} = \frac{K}{S_s} \tag{1.2.10}$$

1.2.5 Homogeneity and isotropy

All aquifer properties just described usually do not conform to the idealized uniform porous material. Homogeneity, heterogeneity, isotropy and anisotropy are concepts used to describe spatial variations and directional trends in aquifer properties. A hydrogeological formation is defined homogeneous for a property if this is independent on the position within the formation, while it is heterogeneous if the property value changes from place to place. Heterogeneity can characterize the horizontal plan as the vertical section depending on the environment of formation of the deposit or the rock type. For example, layered heterogeneity is common in sedimentary formations where each stratum has its own properties, while trending heterogeneity is usually present where there was a sorting and grading of the deposit, such as in alluvial fans and glacial outwash plains.

A hydrogeological formation is defined isotropic for a specific property if the characteristic is independent on the direction of measurement at a point. If the property value in a point varies with the direction, the formation is defined as anisotropic in that point. In the latter case, two principal directions of anisotropy can be identified in correspondence to the maximum and minimum values and are usually at right angles to each other. Anisotropy could characterize the formation both at small scale with the particles orientation and at large scale, for example, with fractures. Considering an anisotropic formation and talking about the saturated hydraulic conductivity, usually the vertical component is smaller of one to two orders of magnitude than the horizontal component.

To make these concepts clearer, figure 1.2.4 shows the four possible combinations of homogeneity and isotropy in a formation.

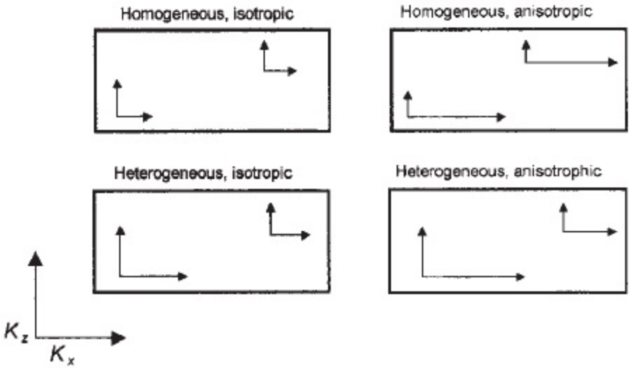

Figure 1.2.4: Four possible combinations of homogeneity and isotropy for hydraulic conductivity of a porous material. Source: Freeze and Cherry (1979)

Usually aquifers are heterogeneous formations with variable parameters in space, specifically strata are near-isotropic in the horizontal plane and very anisotropic in the vertical direction. Talking mainly about hydraulic conductivity, it is not possible to map exactly its spatial distribution because of the huge effort in data acquisition and a probable alteration for too much testing. Moreover, this detail is not required for all applications. Sometimes the average response is enough for management and decision making.

In order to determine this average response described by the effective hydraulic conductivity K_{ef}, the Darcy's law 1.2.2 is averaged by the expected value operator.

$$< q_i >= -K_{ef} < \frac{dH}{dx_i} > \tag{1.2.11}$$

Effective hydraulic conductivity provides primary statistic information. It can be estimated by forms of averaging of measured hydraulic conductivities, depending on the flow dimensionality. If the flow is one-dimensional, it corresponds to the harmonic mean: $K_H = (< K^{-1} >)^{-1}$; while if it is two-dimensional (or lognormal distribution), it corresponds to the geometric mean: K_G, so that $ln(K_G) =< lnK >$.

1.2.6 Connectivity

The last important property of an aquifer is its connectivity. This parameter is essential for solute transport in heterogeneous formations because it describes the ease with which water can move between geological formations. Indeed, if the average hydraulic conductivity is relatively low, it is expected a long travel time of a pollutant to move from one point to another. Anyway, the heterogeneity can involve the presence of connected areas at high conductivity that channel water and pollutant speeding its movement. Moreover, this acceleration implicates not only a early solute arrival, therefore a higher risk in the destination point, but also a lower grade of degradation of the pollutant, if organic, due to the time and the usual higher oxygenation in these pathways.

Following Bianchi et al. (2011), there are several useful indicators to quantify the connectivity in field and they are divided mainly in two groups depending on the set of parameter used. The first class takes into considerations parameters of the spatial characteristics of formations, the second one considers flow and transport parameters. Below, the indicator CI of this second group is reported and used in simulation later on.

$$CI = \frac{t_{av}}{t_{5\%}} \tag{1.2.12}$$

that is the ratio between average arrival time and arrival time of the faster 5% of particles. A high value indicates that water is channeling and therefore the presence of preferential paths.

An alternative and more precise approach to evaluate this parameter in a field is to determine the connectivity function $\tau(d)$. Following the approach used in Western et al. (2001), the connectivity of the spatial pattern G is evaluated thresholding the original pattern (pixels above the threshold belong to A) and considering two pixel connected if there is a continuous path between them of neighboring pixels above the threshold. The function $\tau(d)$ is the lag-dependent probability that a pixel above the threshold in x is connected to another pixel in $x+d$. It is expressed as

$$\tau(d) = P(x \leftrightarrow x+d | x \in A, x+d \in G) \tag{1.2.13}$$

The connectivity function is calculated in the following simulations chapter 4 following the same general steps applied in the Western et al. (2001) work, where the algorithm used is described more in detail. Firstly, the data is thresholded into HIGH or LOW basing on the desired percentile value (50%, 75% and 90%). Then, each continuous cluster is labeled by looping through all pixels and whenever an unlabeled HIGH pixel is encountered, it and its HIGH neighbors are uniquely labeled. Finally, the omnidirectional connectivity function is determined for each separation bin as the ratio between the number of connected pairs and the number of pairs.

1.3 Tests used

As just said, many properties of hydrogeological formations influence the subsurface water movement. Given their heterogeneity in space and their multiple spatial scales of variability, depending on the objective to estimate different levels of characterization are needed and therefore different investigation methods.

1.3.1 Measurement acquisition approaches

Methods can be roughly classified in various measurement acquisition approaches depending on the investigation scale at which they are performed: laboratory scale ($\sim 10^{-4}$ to 1 m), local scale ($\sim 10^{-1}$ to 10^{2} m) and regional scale ($\sim 10^{1}$ to 10^{5} m). As it can be seen in figure 1.3.1, typically there is a trade-off between the method scale and its resolution. For instance, benchtop measurements, that are for example core, column and tank measurements, provide information at a very high resolution but for a really small volume of soil, on the contrary well tests provide informations at a lower resolution but over larger areas.

More than the objective to characterize, the spatial scale and the resolution, many factors have to be taken into consideration, such as: conditions at the site, availability of other data, technicians experience, time funds and computational resources. Usually, firstly a reconnaissance investigation is carried out with low-resolution approaches, then a more detailed one with a higher-resolution approach.

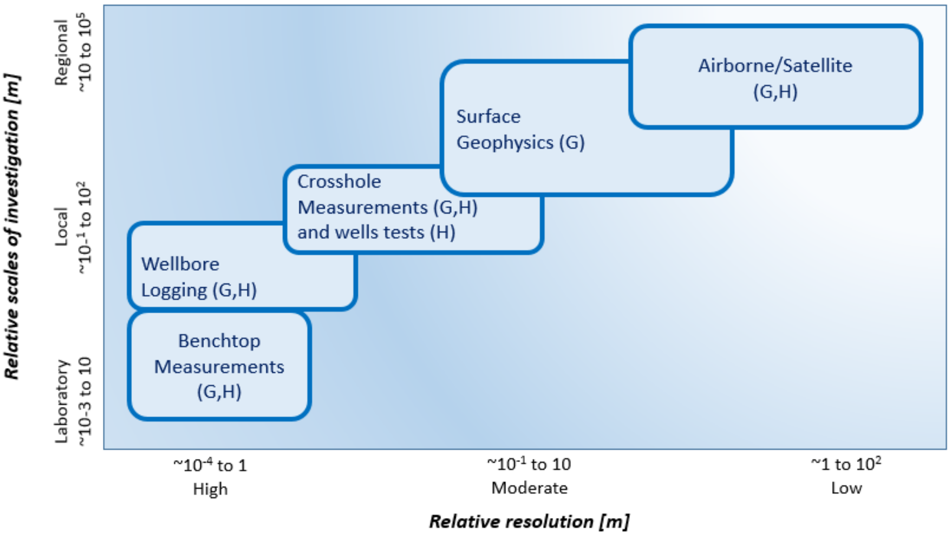

Figure 1.3.1: Chart showing trade-off between the relative resolution of the information obtained using different geophysical(G) and hydrological (H) measurement acquisition approaches and the relative scale of the investigations for which those acquisition geometries are typically used. Source: Rubin and Hubbard (2005)

In this work, the objectives are hydrological parameters, first of all the hydraulic conductivity, and they are mainly investigated thanks to hydrological methods, such as: hydraulic tests, hydraulic tomography and sieves. These methods are summarized in table 1.3 and following described in more detail.

Generally, it is possible to divide these methods into laboratory and fields tests. First ones analyze small samples of soil collected during core-drilling programs. Given the small sample size, results are usually considered only as a point representation of soil properties for a specific direction. On the contrary, field methods involve a relatively large region of the soil. Therefore, results should represent the averaged response of the whole area considering both vertical and horizontal directions.

Table 1.3: Common hydrological characterization methods used for hydrogeological investigations. Source: Rubin and Hubbard (2005)

Acquisition approaches	**Characterization methods**	**Attributes typically obtained**	**Examples of hydrogeological objectives**
Well tests	pumping test slug tests single wellbore hydraulic test	drawdown	hydraulic conductivity specific storage
Crosshole	hydraulic tracer test	tracer concentration	hydraulic conductivity dispersivity fast flow paths
	hydraulic tomography	drawdown	hydraulic conductivity
Wellbore	flowmeter test	water flow	hydraulic conductivity
	neutron probes	back-scattered neutron counts	water content
Benchtop	permeameters	hydraulic head	hydraulic conductivity
	sieves	grain size distribution	hydraulic conductivity
	time domain reflectometer	dielectric constant	water content

1.3.2 Laboratory tests

Describing first laboratory tests, saturated hydraulic conductivity can be estimated using permeameter methods or particle-size analyses. In the first category, water flows through the core sample under either constant or variable hydraulic gradient. First type is preferred for moderate or high conductivities, otherwise the second type is more effective. These tests are usually carried out in vertical position to evaluate the vertical component of the conductivity and they have to be carried out carefully because they are largely dependent on some experimental procedures errors, such as entrapped air and non-Darcian flow. Moreover, it is stated in cite-Hydrogeophisics that also estimates of original samples and repacked cores differ: usually latter estimates are higher and that this one is lower than other field tests. This is probably due to heterogeneity, imperfect laboratory and recovery procedures. The second category bases on the relationship between hydraulic conductivity and sample's physical properties, such as particle-size statistics linked to pore-size distribution. Many empirical and theoretical relationships exist but it is difficult to determine their appropriateness for a specific site.

1.3.3 Field tests

Starting with field tests, pumping tests are the most commonly used methods to have information about transmission capability of an aquifer. They are mainly used for water-supply purposes. It consists of a central well pumping at a constant rate and the drawdown is measured at that well and other surrounding wells. These collected data are then analyzed with different models to obtain an estimated hydraulic conductivity and specific storage over a relatively large volume. As pointed out in citeHydrogeophysics, main limitations are lacking information about conductivity variations in space and the movement of significant quantities of water. This is undesiderable in contaminated sites because of costs of treatment of the pumped water and the risk of altering contaminant distripution.

Slug tests consist of an instantaneous change in head in a well and then of measuring the recovering head in that well. Later on, collected data is analyzed with various models and conductivity is estimated with a large influence of the closer material. They are quite easy to apply in practice and they are used both in water-supply and water-quality investigations. If wells network is extensive in the site, this method can be very useful on describing the spatial variation of hydraulic conductivity, potentially identifying preferential pathways or barriers. However, quality of results are strongly dependent on the well-development process, it has to be done with particular care.

A better estimation of vertical variations are obtained by multilevel slug tests which goal is to characterize vertical variations of hydraulic conductivity. The only difference is that slug tests are performed at different depths. Results can be altered by a packer circumnavigation happening in certain conditions, but these are known and some mitigation measures are present.

Geophysical logging are useful to determine hydraulic conductivity variations in unconsolidated formations by looking at variation in clay content. However, given that averaging volume is relatively large, it is difficult to characterize small variation within a single unit. The two most common types use natural gamma and electrical conductivity. Natural-gamma logs record radiation versus depth: a high reading is usually indicative of a clay-rich layer, while a low reading indicates sand and gravel. Electrical-conductivity logs detect variations in clay content, filled porosity and water chemistry. Unlike natural-gamma logs, logging speed does not affect results.

The dipole-flow test is a hydraulic test with an elaborated mechanism in the single-well and it allows to estimate the radial component of hydraulic conductivity. It is evaluated from the head change in chambers in steady state thanks to an empirical equation valid in most field applications. Anyway, this equation includes the anisotropy ratio that is rarely known a priori and it is generally considered equal to 1. Rubin and Hubbard (2005) underlines as some errors in the estimation can be due to the insufficient testing and to the assumption of homogeneous formation and well-development. On the contrary, some important advantages are no water is injected or extracted, possibility to define the spatial scale of the test and reducing well losses in high-conductivity media by using low flow rates.

The borehole-flowmeter test is considered the most efficient in estimating spatial variations in conductivity. It consists of a well pumping at a constant rate while a downhole flawmeter measures the vertical flow from bottom to top. The cumulative vertical flow is then processed to obtain the vertical profile of hydraulic conductivity. The main limitatios are flow bypassing the flowmeter mainly in high-conductivity formations and the necessity of a long-screened well.

Most of previous tests are performed in permanent wells and this limits their capability to detect spatial variability, in unconsolidated formations an alternative is direct-push methods. Many methods have been developed but the approximation is relatively poor, most promising methods are hydraulic tests in direct-push equipment, such as direct-push slug test and direct-push permeameter. The first consists of a slug test performed in a small-diameter pipe and at multiple levels in a single probehole. Many slightly different procedures exist but all are limited by the time to steady state and other logistical issues, moreover estimates are critically dependent on well-development procedures. On the contrary, the direct-push permeameter is not dependent on development phases. Conductivity estimates are acquired at the desired depth by injecting water through a screen, monitoring pressure changes, then the spherical form of Darcy's law in steady state is applied. This method is fast and efficient particularly for low conductivities, but with adequate equipment it can be more efficient in a broad range.

Except for the pumping test, all described tests give information about properties very close to well or probehole. Though, this is restrictive for solute transport that strongly depends on the connectivity. Multiwell tracer tests are able to provide informations about the changes of hydraulic conductivity between wells by injecting in a well a tracer of a specific type at certain concentration and by checking its presence in surrounding wells. An ideal tracer should be non-toxic and easily measurable at low concentrations, it should not be naturally present in the studied groundwater system, or at really low concentrations, it should not react or be absorbed and it should follow same pathways as the investigated substance. There is not a tracer with such characteristics. Moreover, they are time consuming in intergranular materials, expensive and demand a big effort for routine procedures.

Another option is the hydraulic tomography. This method consists of a series of short-term pumping tests in which the stressed interval varies between tests. This stress is read by nearby observation wells. Then, the collected drawdown data is analyzed by inverting these data with several procedures. This method potentially can provide detailed information about properties between wells, not only in the immediate vicinity of the pumping well. However, Cardiff et al. (2013a) states that tomographic analysis for traditional pumping tests requires either a long test time because the steady state has to be reached or high computational costs for the analysis of the collected transient data. Moreover, for Bohling and Butler (2010), a unique solution to the inverse problem cannot be obtained without introducing some form of regularization which in effect biases or constrains the solution.

Thanks to new techniques and modern instrumentation, borehole and surface geophysical techniques are used in the last time. Usually surface techniques are used for a preliminary investigation while borehole ones are used for more detailed information. Some examples of these new methods are acoustic televiewer, electrical resistance, thermal techniques, electromagnetic induction and seismic profiling.

Another modern technique is the remote sensing. These methods are very useful to collect data on a regional scale and to assess groundwater resources' state. Of course they are not detailed, but they have a global coverage and a huge amount of data is available, moreover they can provide error statistics and meaningful spatial avarages. For example, aerial thermal infrared imaging is increasingly used to map groundwater discharges, while altimetry measurements in long time can detect subsidence occurrence and high-precision satellite gravimetry detects changes in gravity that are mainly caused by changes in total water storage in atmosphere, at and below the ground surface.

1.3.4 Oscillatory pumping test

Oscillatory or periodic pumping test is a pumping test in which fluid is extracted during half a period and then reinjected. This is generally used to estimate effective aquifer properties, reducing the amount of water moved by a classical pumping test, no net water extraction, and robust signal measurements. Moreover, intuitively they cause less contaminant plume movement and, given that the frequencies are known, it is possible to separate responses from other hydrologic processes or noises. Thus, pros of this test are the possibility to use it for long-term monitoring without any alteration for water extraction and injection and the short test time, given that multiple stimulations at several locations and frequencies can be done at once and following separated through frequency-domain decomposition.

Depending on the period used, the response of the aquifer is different. Knowing that, Cardiff et al. (2013a) uses several signals of different periods to stress the aquifer, response are analyzed through inverse modeling. This procedure is called multifrequency oscillatory hydraulic imaging (M-OHI). This way computational cost is drastically reduced and heterogeneity information obtained from responses to the different frequencies is fully used.

2

Code approach

All tests listed in the previous chapter have pros and cons: some of them are useful for a general characterization objective, some for a more local one and others are useful for a broad characterization of the heterogeneity. For most of previously described field tests is necessary to process the collected data in order to obtain the values of conductivity and storativity. The code here described and analyzed is specific to obtain local values from an oscillatory pumping test. This chapter aims mainly to describe applied theory and followed logic, moreover the last section lists all important conditions assumed in the code.

2.1 Innovative approach

Usually for periodic pumping test analysis, responses to the several stimulations of different frequencies are analyzed through a fully transient numerical model but it is a very long computational time user. Therefore, the idea is to use a steady-periodic model formulation, used previously by Cardiff et al. (2013a) in a simple field case. This strategy is named Multifrequency Oscillatory Hydraulic Imaging (M-OHI) and its main advantages are reducing computational cost and fully utilizing heterogeneity information provided by responses at different frequencies. The equivalence on results for fully transient and steady-periodic numerical models as well as the different heterogeneity sampled by oscillatory signals at several frequencies are already showed in Cardiff et al. (2013a).
In the present thesis we followed the approach suggested by Castagna and Bellin (2009) and Castagna et al. (2011), by which a large number of spatial random fields,

representing the hydraulic conductivity and the storage coefficient, are created, all respecting the overall statistics dictating the type of spatial variability of the site and conditional to the information provided by the pumping test. With the term conditional we mean here that only the random fields providing a good reproduction of the pumping tests are accepted, the remaining are removed. For example in a contamination study, this approach could be very useful because results include several fields leading to different plume's shapes and extensions. This way, with a post-processing, it is possible to determine minimum and maximum extension of the contaminant plume, the average plume and the most probable plume. It can reduce the uncertainty in this kind of studies and it can help to consider all possible realizations.

An important note has to be made about the contribution of the Hydrology group of Trento University who implemented the code. The present study builds on the existing code in two ways: understand how much the code's inference is able to correctly reproduce the real field and make few small adjustments to the code to enable a more efficient estimation and post-processing.

2.2 Basic Theory

The topic of this section is to explain the basic theory and logic of the model implemented. The steady-periodic model uses governing equations of groundwater, written in the Fourier space. In order to understand the model, firstly the fully transient model is described too. Then, the methodology itself is described. It starts by generating random fields that are then conditioned by the observation data, conditioned fields are following inverted and optimized in order to converge versus a field that optimally respect statistics. Furthermore, in order to have a range of possible fields, many simulations are done to follow the Monte Carlo framework.

2.2.1 Mathematical model

Transient model

In order to understand the steady-periodic model, firstly is described the most common transient model in which the mass conservation law requires equivalence between the net rate of fluid mass flow into the control volume and the time rate of change of fluid mass storage within the volume. Taking into consideration a confined aquifer or unconfined, assuming standard linearized water table approximation, it is possible to write the following governing equations:

$$\begin{aligned} S_s\frac{\partial h}{\partial t} &= \nabla\cdot(K\nabla h)+q \qquad && \forall \mathbf{x}\in\Omega, t\geq 0 \\ h &= 0 && \forall t\geq 0, \mathbf{x}\in\Gamma_d \\ \nabla h\cdot\mathbf{n} &= 0 && \forall t\geq 0, \mathbf{x}\in\Gamma_n \\ K\nabla h\cdot\mathbf{n} &= S_y\frac{\partial h}{\partial t} && \forall t\geq 0, \mathbf{x}\in\Gamma_w \end{aligned} \tag{2.2.1}$$

where $\mathbf{x}$ is the vector with the spatial coordinates $[L]$, t is time $[T]$, $\mathbf{n}$ is the outward normal of a certain boundary, $S_s(\mathbf{x})$ is the specific storage $[1/L]$, $S_y(\mathbf{x})$ is the specific yield $[-]$, $K(\mathbf{x})$ is the isotropic hydraulic conductivity $[L/T]$, $h(\mathbf{x},t)$ is the change in head from an initial steady condition $[L]$, Ω is the domain of interest, Γ_d, Γ_n and Γ_w are respectively Dirichlet, Neumann and linearized water table boundaries and $q(\mathbf{x})$ is volumetric water source $[(L^3/T)/L^3]$.

Same equations can be written for aquifers with horizontal confining layers, considering it as a 2D aquifer averaged in depth. In this case, transmissivity $T[L^2/T]$ and storativity $S[-]$ are used as shown in equation 2.2.2:

$$\begin{aligned} S\frac{\partial h}{\partial t} &= \nabla\cdot(T\nabla h)+q \qquad && \forall \mathbf{x}\in\Omega, t\geq 0 \\ h &= 0 && \forall t\geq 0, \mathbf{x}\in\Gamma_d \\ \nabla h\cdot\mathbf{n} &= 0 && \forall t\geq 0, \mathbf{x}\in\Gamma_n \end{aligned} \tag{2.2.2}$$

Moreover, in order to consider a periodic pumping, q is assumed to cary in time as follows:

$$q = Q(\mathbf{x})cos(\omega t) \tag{2.2.3}$$

where $Q(\mathbf{x})$ is the peak flow rate in $[(L^3/T)/L^3]$ and ω is the frequency in $[rad/T]$.

Steady-periodic model

After some time in which aquifer response still depends on initial conditions, it is expected that domain responds at the same input frequency, even if phase offset and magnitude of oscillations vary. Cardiff et al. (2013a) passes to a steady-periodic model considering the oscillator and head changes as real parts of complex values.

$$\begin{aligned} q &= Re[Q(\mathbf{x})exp(i\omega t)] \\ h(\mathbf{x},t) &= Re[\Phi_\omega(\mathbf{x})exp(i\omega t)] \end{aligned} \tag{2.2.4}$$

where $\Phi_\omega(\mathbf{x})$ is the phasor, a complex-valued field variable that include both amplitude and phase offset of the signal wave at each location. The subscript ω remembers that different phasor solutions are obtained depending on the input frequency. Following, governing equations 2.2.5 are obtained by applying definitions 2.2.4 into equations of the transient model 2.2.1 and dividing by $exp(i\omega t)$ given that equations must apply for all times.

$$\begin{aligned} i\omega S_s\Phi_\omega &= \nabla\cdot(K\nabla\Phi_\omega)+Q && \forall\mathbf{x}\in\Omega \\ \Phi_\omega &= 0 && \forall\mathbf{x}\in\Gamma_d \\ \nabla\Phi_\omega\cdot\mathbf{n} &= 0 && \forall\mathbf{x}\in\Gamma_n \\ K\nabla\Phi_\omega\cdot\mathbf{n} &= i\omega S_y\Phi_\omega && \forall\mathbf{x}\in\Gamma_w \end{aligned} \tag{2.2.5}$$

Solving previous equations, Φ_ω is calculated and then the real part of $\Phi_\omega exp(i\omega t)$ is the solution of h for the steady-periodic response.

Dimensionless steady-periodic model

Considering the first equation of 2.2.5, it is possible to re-write it dimensionless. The result will be later used for the frequency sensitive analysis. Let's start from rewriting the equation explicitly 3D where symbol $*$ marks dimensional parameters.

$$i\omega^* S_s^*\Phi_\omega^* = \frac{\partial}{\partial x_1^*}\left[K^*\frac{\partial\Phi_\omega^*}{\partial x_1^*}\right]+\frac{\partial}{\partial x_2^*}\left[K^*\frac{\partial\Phi_\omega^*}{\partial x_2^*}\right]+\frac{\partial}{\partial x_3^*}\left[K^*\frac{\partial\Phi_\omega^*}{\partial x_3^*}\right]+Q^* \tag{2.2.6}$$

In order to obtain a dimensionless equation, some parameters are divided by spatial scales of two directions I_{Yh} and I_{Yv} and the conductivity is rewritten.

$$\begin{aligned}
&x_{1,2} = \frac{x^*_{1,2}}{I_{Yh}} \\
&x_3 = \frac{x^*_3}{I_{Yv}} \\
&\Phi_\omega = \frac{\Phi^*_\omega}{I_{Yh}} \\
&K^* = e^{<Y>+Y'} = e^{<Y>}e^{Y'} = K_G e^{Y'}
\end{aligned} \tag{2.2.7}$$

Therefore, it is possible now to write:

$$i\omega^* S^*_s \Phi_\omega I_{Yh} = \frac{1}{I_{Yh}}\frac{\partial}{\partial x_1}\left[K_G e^{Y'}\frac{\partial \Phi_\omega}{\partial x_1}\right] + \frac{1}{I_{Yh}}\frac{\partial}{\partial x_2}\left[K_G e^{Y'}\frac{\partial \Phi_\omega}{\partial x_2}\right] + \frac{1}{I_{Yv}}\frac{\partial}{\partial x_3}\left[K_G e^{Y'}\frac{I_{Yh}}{I_{Yv}}\frac{\partial \Phi_\omega}{\partial x_3}\right] + Q^* \tag{2.2.8}$$

$$i\omega^* S^*_s \Phi_\omega I_{Yh} = \frac{K_G}{I_{Yh}}\left\{\frac{\partial}{\partial x_1}\left[e^{Y'}\frac{\partial \Phi_\omega}{\partial x_1}\right] + \frac{\partial}{\partial x_2}\left[e^{Y'}\frac{\partial \Phi_\omega}{\partial x_2}\right] + \left(\frac{I_{Yh}}{I_{Yv}}\right)^2\frac{\partial}{\partial x_3}\left[e^{Y'}\frac{\partial \Phi_\omega}{\partial x_3}\right]\right\} + Q^*\frac{K_G}{I_{Yh}}\frac{I_{Yh}}{K_G} \tag{2.2.9}$$

$$i\frac{\omega^* S^*_s I^2_{Yh}}{K_G}\Phi_\omega = \frac{\partial}{\partial x_1}\left[e^{Y'}\frac{\partial \Phi_\omega}{\partial x_1}\right] + \frac{\partial}{\partial x_2}\left[e^{Y'}\frac{\partial \Phi_\omega}{\partial x_2}\right] + \frac{1}{e^2}\frac{\partial}{\partial x_3}\left[e^{Y'}\frac{\partial \Phi_\omega}{\partial x_3}\right] + Q^*\frac{I_{Yh}}{K_G} \tag{2.2.10}$$

where e is the anisotropy ratio I_{Yh}/I_{Yv}. It is easily possible to see how Q^* and ω^* can be written dimensionless as:

$$\begin{aligned}
&\omega = \frac{\omega^* S^*_s I^2_{Yh}}{K_G} = \frac{\omega^* I^2_{Yh}}{D} \\
&Q = Q^*\frac{I_{Yh}}{K_G}
\end{aligned} \tag{2.2.11}$$

Finally the completely dimensionless equation is:

$$i\omega\Phi_\omega = \frac{\partial}{\partial x_1}\left[e^{Y'}\frac{\partial \Phi_\omega}{\partial x_1}\right] + \frac{\partial}{\partial x_2}\left[e^{Y'}\frac{\partial \Phi_\omega}{\partial x_2}\right] + \frac{1}{e^2}\frac{\partial}{\partial x_3}\left[e^{Y'}\frac{\partial \Phi_\omega}{\partial x_3}\right] + Q \tag{2.2.12}$$

2.2.2 COCR solver

In order to solve the governing equation 2.2.5 of the steady-periodic model, a solver for complex symmetric linear systems is necessary. The method used is the Conjugate A-Orthogonal Conjugate Residual (COCR), a variation of the first-order

optimization algorithm 'gradient descent'. It is chosen because in Sogabe and Zhang (2007) it is showed as usually this method gives smoother convergence behavior than other methods, such as Conjugate Orthogonal Conjugate Gradient (COCG). Moreover, it sometimes reach convergence sooner than the Quasi-Minimal Residual method (QMR) in terms of number of iterations.
Following the algorithm is reported to solve a linear system of the form $A\mathbf{x} = \mathbf{b}$, more details can be found in Sogabe and Zhang (2007).

$$\begin{aligned}
&\mathbf{x}_0 \text{ is the initial guess}\\
&\mathbf{r}_0 = b - A\mathbf{x}_0\\
&\text{set } \mathbf{p}_{-1} = 0, \beta_{-1} = 0\\
&\text{for n=0,1,.... until } ||\mathbf{r}_n|| \leq \epsilon ||\mathbf{b}|| \text{ do:}\\
&\mathbf{p}_n = \mathbf{r}_n + \beta_{n-1}\mathbf{p}_{n-1}\\
&\alpha_n = \frac{(\overline{\mathbf{r}}_n, A\mathbf{r}_n)}{(\overline{A}\overline{\mathbf{p}}_n, A\mathbf{p}_n)}\\
&\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n \mathbf{p}_n\\
&\mathbf{r}_{n+1} = \mathbf{r}_n - \alpha_n A\mathbf{p}_n\\
&\beta_n = \frac{(\overline{\mathbf{r}}_{n+1}, A\mathbf{r}_{n+1})}{(\overline{\mathbf{r}}_n, A\mathbf{r}_n)}
\end{aligned} \tag{2.2.13}$$

where A is a non-Hermitian symmetric matrix ($A \neq \overline{A}^T$ and $A = A^T$), $\mathbf{x}_n$ is the n-th approximate solution, $\mathbf{r}_n$ is the residual vector ($:= \mathbf{b} - A\mathbf{x}_n$), $\mathbf{p}_n$ is the search direction, α and β are two parameters determined by two orthogonality conditions ($\mathbf{r}_n \perp W$ and $A\mathbf{p}_n \perp W$).

This method is applied considering all mesh nodes and it calculates approximate solutions that assign a value to each node. This assignation is done by using a linear basic function or a parabolic one to interpolate, depending on the wanted resolution order. Indeed, a first order solver linearly interpolates values on nodes. While the second order solver locates another imaginary node between considered nodes and it creates parabolic elements for the interpolation. The second order is usually more precise but it considerably increases the computational time because it doubles the nodes to evaluate.

2.2.3 Inversion

In order to evaluate spatial variability of conductivity and storativity in a multiple well periodic pumping test, collected transient head signals have to be inverted. This way a realization of hydraulic parameters, compatible with all observations, is inferred. Inversion is particularly suitable for multiple well pumping tests because redundant measurements facilitate it. Redundancy helps to separate noise from signal.

Usually, inversion process needs to assume a priori the model of spatial variability of conductivity in order to avoid the extraction of the scale of spatial variability from data. However, this leads to neglect connectivity aspects.

In this model the Bayesian approach is used for inversion, it provides a robust framework to sort out admissible solutions and retain only those compatible with observations. Castagna and Bellin (2009) already demonstrates its effectiveness in synthetic experiments for constant pumping rates. Furthermore, transmissivity and storativity are here both spatially variable and jointly inverted.

Basic assumption is Y and Z as independent normally distributed Random Space Functions (RSFs) and they are defined respectively as natural logarithms of transmissivity and storativity.

$$\begin{aligned} Y &= ln(T) \\ Z &= ln(S) \end{aligned} \tag{2.2.14}$$

They have constant means (m_Y, m_Z) and variances (σ_Y^2, σ_Z^2) and the following axisymmetric exponential covariance function:

$$C_Y(r_1, r_2, r_3) = \sigma_Y^2 exp\left[-\sqrt{\frac{r_1^2 + r_2^2}{I_{Yh}^2} + \frac{r_3^2}{I_{Yv}^2}}\right] \tag{2.2.15}$$

where $\mathbf{r}$ is the distance vector between two points. This approach is particularly interesting because it does not assumes statistics of the field and it considers structural parameters, such as mean, variance and integral scales, as unknown in addition to local values of the properties. In the Bayesian approach, their spatial distribution is determined minimizing a fitness function 2.2.16, more details are given in Castagna and Bellin (2009).

$$L(\mathbf{a},\theta) = \left[\mathbf{z} - \Im(\mathbf{a},\theta)\right]^T C_v^{-1} \left[\mathbf{z} - \Im(\mathbf{a},\theta)\right] + \left[\bar{\mathbf{a}} - \mathbf{a}\right]^T C_a^{-1} \left[\bar{\mathbf{a}} - \mathbf{a}\right] \tag{2.2.16}$$

where $\mathbf{a}$ is the vector containing the unknown K and S_Y values, θ is the vector of unknown structural parameters of the geostatistical model of spatial variability, $\bar{\mathbf{a}}$ are prior estimates, $\mathbf{z}$ contains head measurements, $\Im$ is the forward nonlinear operator. C_v and C_a are covariances functions where $\mathbf{v}$ is the vector of the measurement errors. In this work, the fitness function is slightly different because it is not included the penalization factor. This factor that is considered by the second block in 2.2.16 counts for values that are particularly far from a prior estimation. However, in this case we are not considering any prior notion, therefore this factor is not considered. Consequently, the fitness function to minimize counts for the first block in equation 2.2.16, as shows 2.2.17.

$$L(\mathbf{a},\theta) = \left[\mathbf{z} - \Im(\mathbf{a},\theta)\right]^T C_v^{-1} \left[\mathbf{z} - \Im(\mathbf{a},\theta)\right] \tag{2.2.17}$$

Equivalently it is here used and maximized the modified Nash-Sutcliffe coefficient, a variation of the Nash-Sutcliffe efficiency that is very appropriate for periodic signals. It is defined as:

$$NSEE = \frac{1 - \frac{\sum_{t=1}^{ntime} |H_{sim} - H_{obs}|}{\sum_{t=1}^{ntime} |H_{obs} - \overline{H_{obs}}|}}{n_{sample}} \tag{2.2.18}$$

where $ntime$ is the number of time steps, H_{sim} are head variations in observing wells locations for tested frequencies, H_{obs} are corresponding observed head variations, $\overline{H_{obs}}$ is the mean of observed values, n_{sample} is the number of observations. The efficiency ranges from minus infinity to plus one. It reaches the maximum of one if time series are identical, while it assumes negative values if the model is worse than the constant model reproducing the average of observations. This error function provides a balanced evaluation of different characteristics and it well fits periodic signals.

To minimize the fitness function 2.2.17, the equation is usually linearized with different approaches, but they all require the computation of the sensitivity matrix that is the most computational demanding step. In order to avoid this linearization, in this work a different approach is applied that is an efficient genetic algorithm based on particle swarming.

2.2.4 Pilot Point Method

There is a delicate balance between model complexity and robustness of the inversion technique. Indeed, unknown parameters in $\mathbf{a}$ have not to be too much otherwise the inversion problem becomes ill posed.
If it would be necessary to include all values of hydraulic properties for each grid cell in the unknown vector $\mathbf{a}$, the inverse problem would be surely ill-posed due to the large number of unknowns. A solution could be to coarse grid reducing considerably the unknowns number, anyway this leads to a less precise and detailed spatial heterogeneity. Another way, here applied, is to consider as unknown only hydraulic properties in specific strategic locations, the so called pilot points (Certes and de Marsily (1991)), and generate several alternative realizations, all conditional to pilot points values.
The objective of the inversion is therefore to optimally evaluate properties values at pilot points locations (N_{pp} unknowns) and, obviously, estimate the structural statistic parameters (means, variances and integral scales) that identify geostatistical models of spatial variability.
In order to do that, firstly two unconditional Y and Z fields are randomly generated, they are then perturbed in pilot point locations in such a way to minimize the fitness function 2.2.17 (see the work by Castagna and Bellin (2009)). Given that fields are spatially correlated, this local perturbation causes changes in surrounding locations with respect to the unconditional field. Modifications are more significant close to the pilot point location and it decreases moving far away. Therefore, the perturbation propagation generates conditioned fields.
Castagna and Bellin (2009) already showed that there is an optimal number of pilot points and that it depends on the amount of information present as observational data. Specifically, the number of pilot points is proportional to observations: if there are too many pilot points, the number of unknowns increases and the inversion procedure becomes problematic; on the contrary, if they are too few, realizations are less conditioned and therefore less realistic.

2.2.5 Particle Swarm Optimizer

As just mentioned perturbations at pilot points propagate on fields, but it is necessary to optimize these perturbations in order to have conditioned fields compatible with observations. The problem of the analyzed parameter space is that it has many local minima and the initial conditions are completely random, therefore they should not be too conditioning and relevant in the optimization process.

For these reasons, Particle Swarm Optimizer is applied to obtain the minimize of the objective function 2.2.17. This optimization technique, inspired by animal social behaviour, is based on an iterative procedure. It is applied in the PhasONE code because it is self-adaptive, it well performs in a space with many local minima, it is quite insensitive to initial conditions and also to the objective function's shape.

This optimization procedure considers a set of candidate solutions, called particles, and moves them around in the parameter space according to a formula (equation 2.2.19), function of particle's position and velocity. The movement of the particle depends on three main components: a spatially constant drift v_i^k (where the index i refer to the particle and k to the iteration), a random movement whose amplitude is related to the distance between the particle and the best position it identified during its wandering (the partial best position $p_{best,i}^k$) and another random movement similarly dependent on the distance between the particle and the best position explored by all particles (the global best position g_{best}^k). Both best positions and the particle position (x_i^k) are updated at each iteration. This way, it is expected a general movement toward best solutions.

$$\begin{cases} v_i^k = \omega_{inertia} v_i^k + c_1 r_1^k (p_{best,i}^k - x_i^k) + c_2 r_2^k (g_{best}^k - x_i^k) \\ x_i^k = x_i^{k-1} + v_i^k \end{cases} \tag{2.2.19}$$

where $\omega_{inertia}$ is an inertia weight whose value decreases moving forward on iterations (from 0.9 to 0.4), c_1 is called the constant cognitive learning factor and c_2 is the social one (both equal to 2), r_1^k and r_2^k are arrays of uniformly distributed random numbers bounded between 0 and 1. Furthermore, if a particle hits the boundary of the search space, its velocity component which is normal to that boundary is set to zero, applying so the absorbing wall boundary condition.

Particle Swarm Optimizer is an effective algorithm because it does not assume anything about the problem and it can search in a very large space of admissible solutions.

In this specific work, the dimension of the search-space n, therefore the dimension of vectors $\mathbf{x}$, $\mathbf{v}$, $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{p}$ and $\mathbf{g}$, is equal to $2N_{pp} + N_\theta$, where N_{pp} is the number of pilot points and N_θ is the number of parameters of the geostatistical model of spatial variability. Therefore, the coordinates of a solution particle are obtained by merging vectors of unknowns, $\mathbf{a}$ and θ.

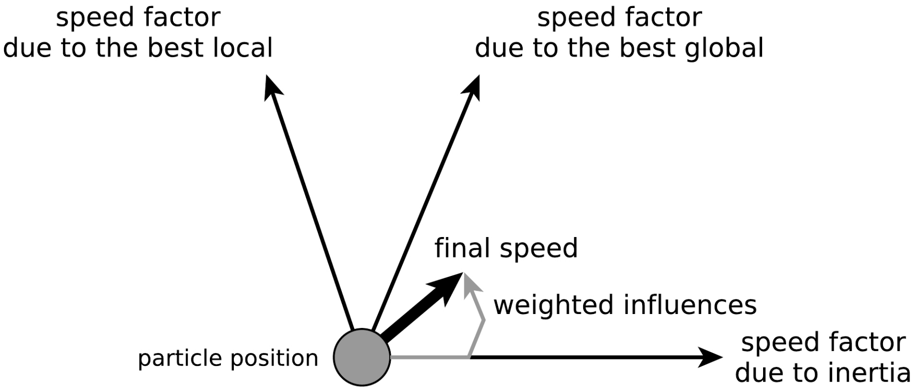

Figure 2.2.1: Velocity's components of particle at time t of the PSO process: partial best, global best and inertia influences.

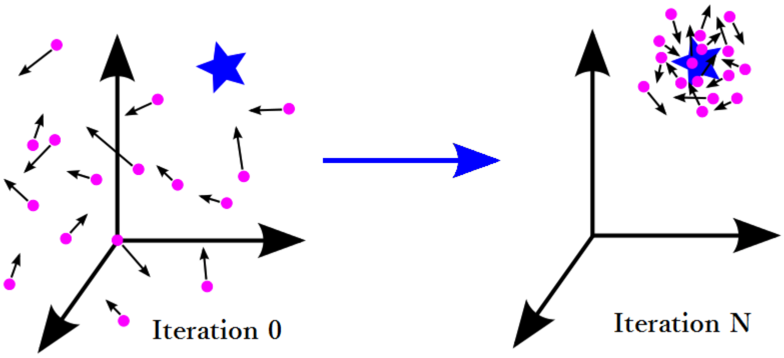

Figure 2.2.2: Representation of PSO precess: starting from random locations and moving towards the global best.

2.2.6 Monte Carlo framework

PSO is used in order to find a realization compatible with observations, but it is not ensured its veracity. Indeed, there could be many realizations compatible with observation data. In order to consider a broad range of possible realizations, the Monte Carlo framework is applied.

The solution with higher efficiency is the one that better fits observation data and it is called the best solution. However, thanks to the several realizations multiple optimal values for each parameter are identified, from which it is possible to define a posteriori their statistics: the pdf, the mean value and its variance. With such approach, it is not obtain a certain solution but an idea of realization with its uncertainty.

2.3 Code in detail

Now that the basic theory and the approach used were described, main implementation steps of the code developed by the Hydrology group of Trento University are summarized and main conditions and assumptions are listed.

2.3.1 Implementation steps

In this preliminary work on the code, observational data are obtained from a synthetic field by simulating a tomography test in known properties fields. Given means, variances and integral scales of the two properties distributions, random fields are generated by HYDROGEN, a random field generator developed by Bellin and Rubin (1996).

$$\begin{aligned} Y(\mathbf{x}) &= m_Y + Y'(\mathbf{x}) \\ Z(\mathbf{x}) &= m_Z + Z'(\mathbf{x}) \end{aligned} \tag{2.3.1}$$

where Y' and Z' are the fluctuations around means. Then, locally they are transformed to create the reference conductivity and storativity fields.

$$\begin{aligned} \overline{K}(\mathbf{x}) &= exp[Y(\mathbf{x})] \\ \overline{S}(\mathbf{x}) &= exp[Z(\mathbf{x})] \end{aligned} \tag{2.3.2}$$

Following, the hydraulic tomography test is simulated on reference fields. The governing equation 2.2.5 is solved for each well pumping and for each frequency of the oscillatory pumping test. Indeed, a single test is considered as the signal emission from a certain well that is pumping periodically with a certain frequency and the contemporaneous transient head recording in other non-pumping wells. Therefore, observational data will consist of $N_{ws} - 1$ values recorded at non-pumping wells for each test. The number of tests N_{tests} is equal to the number of wells that are pumping in sequence N_{ws} multiplied for the frequencies used N_{ω}.

Then, the inversion process is followed to determine realizations compatible with the observations of reference fields. This procedure, which s the same used by Castagna and Bellin (2009), consists of five steps, where Pr represents both properties of log-storativity and log-conductivity:

1. Unconditional fields, Pr^u, are randomly generated by HYDROGEN.

2. Unconditional fields are conditioned to log of the measurements of hydraulic property $Pr_k(\mathbf{x}_{k,j})$, where j indicates the measurement location number, it ranges from 1 to the number of measurements N_k. Conditioning is done by distributing residuals, $Pr_k(\mathbf{x}_{k,j}) - Pr^u(\mathbf{x}_{k,j})$, to grid points by kriging and then added to unconditional fields Pr^u. These fields are conditioned only to measurements and they will be following called prior log fields $Pr^{c,p}$. At this step, a preliminary set of structural parameters θ is generated respecting some predefined bounds.

3. Prior log fields are further conditioned basing on pilot points values, $Pr_{pp}(\mathbf{x}_{pp,j})$ where $j = 1, ...N_{pp}$, with N_{pp} being the number of pilot points. These values are randomly generated at first and then they are obtained by the PSO by minimizing the fitness function 2.2.17. The procedure of this conditioning step is similar to the previous one: residuals, $Pr_{pp}(\mathbf{x}_{pp,j}) - Pr^{c,p}(\mathbf{x}_{pp,j})$, are distributed by kriging in the whole domain. These fields are therefore conditioned both to measures and to pilot points and they will be called conditioned fields, Pr^c.

4. In this step, the governing equation 2.2.5 is solved for conditioned fields in order to compute transient heads at observation wells. These heads will be then compared to observations.

5. The fitness function 2.2.17 is evaluated with z being the amplitude and phase shift of head oscillations at all observation (wells) points. Therefore another

set of parameters is obtained by using the PSO searching algorithm and the steps from 2 to 5 repeated.

The iterations end when all particles are stacked in the same position, or the maximum number of iterations is reached, this number is typically assigned in order to contain the simulation time within acceptable bounds.
Moreover, the procedure just described is repeated in the Monte Carlo framework where unconditional fields are randomly generated with a different seed.

2.3.2 Geometry, grid and conditions

Focusing on the domain studied in this work, the work aims to characterize an area of 10 $m \times 20$ m in which seven wells are distributed with the geometry shown in figure 2.3.1 ($N_w = 7$). From the figure, it is possible to notice that twelve pilot points are located in strategic positions between wells ($N_{pp} = 12$). Moreover, no measurement points are located in this study ($N_k = 0$). Wells are placed in such a way to distance from a minimum of 5.5 m to a maximum of 16 m.

Furthermore, the two main anisotropy axes are assumed to coincide with the reference system utilized for solving the problem by using governing equations 2.2.5. The mesh used is square-shaped and the square sides dx are 1 m long. The shape was chosen just for its simplicity. Regarding the size, it was established in order to guarantee that dx stays at least three times in the integral scales, however it has not to generate a mesh too much refined that would lead to long computational times. Another important issue is how assignation of heterogeneous hydraulic parameters (Y and Z) happens. Firstly, the two independent random functions are generated on this regular grid by $HYDRO_GEN$ (Bellin and Rubin (1996)). The RSF generator assigns a value to the center of the grid. However, because the solver works on cells' corners, the center value is assigned to the bottom-left corner.

It is now necessary to define some boundary conditions for the domain. However, to avoid the influence of these boundaries reaching the area of interests, conditions of constant head are assigned to an external area sized 60 $m \times 60$ m. This way, the change in head h faraway from the interested area is blocked equal to 0.
Moreover, drawdown data are not affected by the head gradient of the domain be-

cause of the linearity of governing equations 2.2.5. Therefore, the head gradient is completely neglected and not assigned at all.

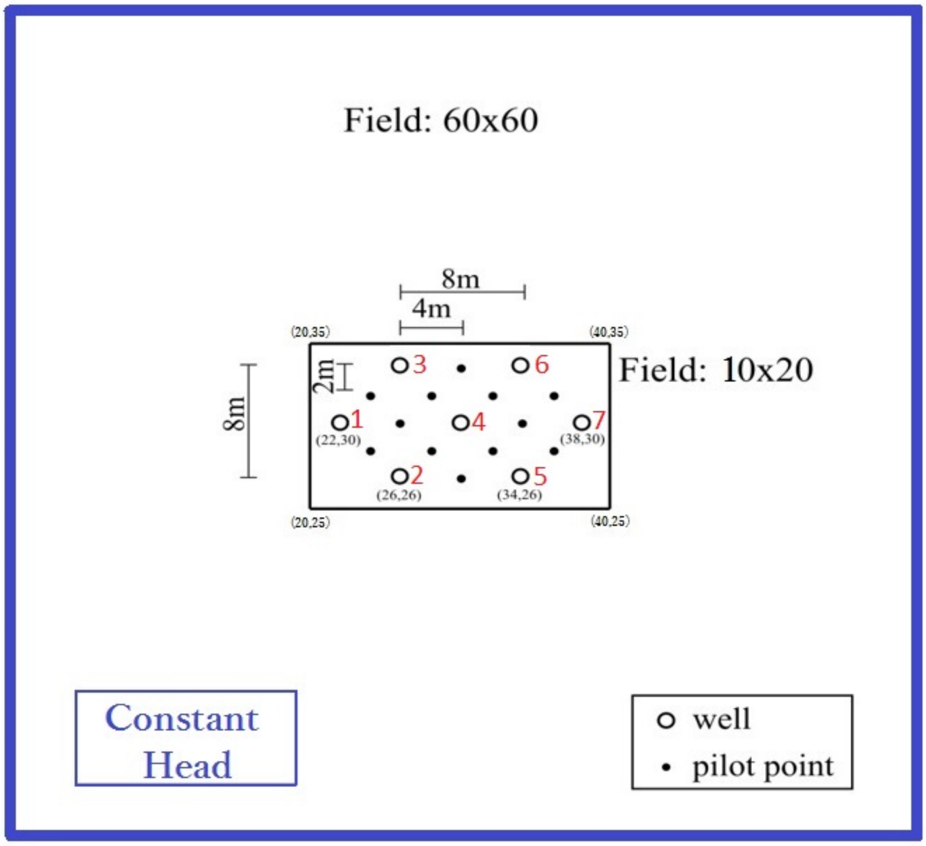

Figure 2.3.1: Geometry of the studied domain, where is evident the central area of interest.

3

Preliminary sensitivity analysis

In order to use correctly the code, it is important to fully understand how the code works varying some characteristics both of code and field.
These variable properties are order of the code solver, average diffusivity of the field, discharge and frequencies used for the simulated pumping test. The first one is important for the solution accuracy and the computational time. While others are relevant because in an oscillatory pumping test it is essential that observing wells are affected by the pumping one, therefore the influence extension is crucial. Consequently, effected area is used as comparison in field's and test's properties analyses. The influence area is here defined as the area in which the head variation has a value major or equal to 10% of the maximum change in the pumping well location. Areas of influence of a pumping well are particularly relevant because they should enclose surrounding observing wells to make the pumping test valid. Anyway, they cannot be too large otherwise they touch field boundaries that is not computationally correct.
Following, the sensitivity analysis is carried out starting from two isotropic main scenarios characterized by the basic parameters reported in table 3.1, they differ only on variances. The so called Scenario 1 has a variance equal to 2 that is a usual high value in field (see tables in the Handbook edited by Jaques Deller (1998)), while Scenario 2 has a higher variance equal to 4. This is quite a high value given that the MADE site is considered highly heterogeneous with a large-scale variance of hydraulic conductivity equal to 4.5, in which it was determined a higher small-scale variance of 5.45 (Zheng et al. (2011) and Harvey and Gorelick (2000)). The other parameters are selected within their boundaries. The frequencies used in the

following analysis are reported in table 3.2.

Table 3.1: Parameters of the two main scenarios

	m_Y $[m]$	σ_Y^2 $[m^2]$	m_Z $[m]$	σ_Z^2 $[m^2]$	I_x $[m]$	I_y $[m]$	Q $[l/s]$
Scenario 1	-8	2	-8	2	4	4	1
Scenario 2	-8	4	-8	4	4	4	1

Table 3.2: Frequencies used in main scenarios

Period $[s]$	Frequency $[s^{-1}]$
5	1.2566
10	0.6283
15	0.4189
20	0.3142
50	0.1256

3.1 Code analysis: Order 1 and 2

The first analysis is about the order of the code solver. Indeed, generally higher is the order, higher is the accuracy of the result, with some limits due to the stability of the solution. However, higher is the order and longer is the computational time. Of course, a highly time consuming code is not suitable to use for technicians if it is slightly better than a faster solver, and this is what this section aims to check: the solution's improvement due to an increased order.

In order to compare the two orders, two experiments are done: check head variation along the transect of the pumping well at certain times and check head variation in time for a point close to the pumping well. All simulations are carried out for each well and for each frequency but here are reported only results for the middle well and medium frequency.

For the first experiment, as it is possible to notice from figures 3.1.1, there is a discrepancy between the two orders mainly at the pumping well position. Already one meter away from it, solutions are very similar and two meters away are almost equal. Remembering that the code only use values at other wells' positions, more than 4 m far, what is important to observe in this step is the similarity of simulations at greater distances. Therefore, this check is completely fulfilled.

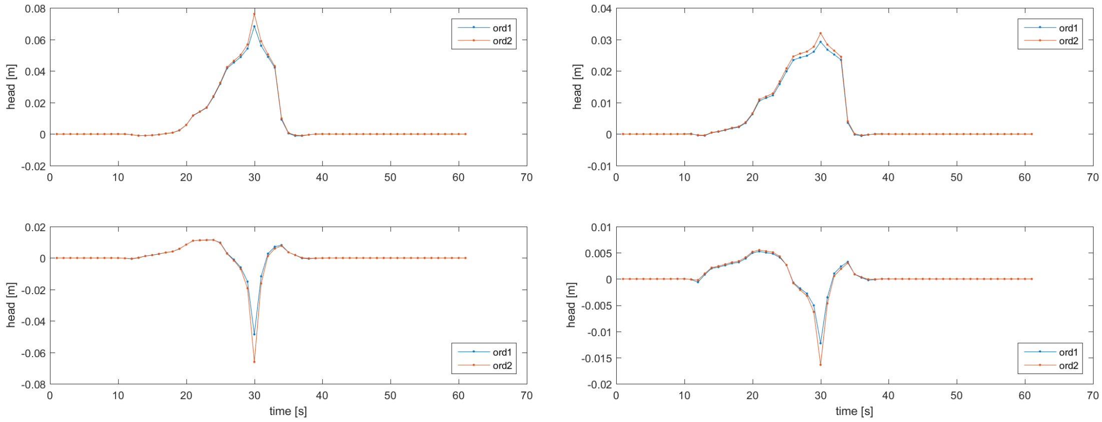

Figure 3.1.1: Transect in $T/4$ and $T/2$ - pumping well (30,30) and frequency 0.4189 s^{-1} - Scenarios 1 and 2

As just mentioned, the main interest is on the discrepancy at distances greater than 4 m. Observing that differences go smaller increasing the distance, the second experiment focuses on head at a point exactly 4 m far from the pumping well. Figures 3.1.2 show the head variation in time at location (34,30) while the middle well is pumping with a frequency of 0.4189 s^{-1}. It is noticeable that results are very similar and slightly different on peaks. However, it is necessary to report that not all results for this experiment are like that. Firstly, for higher frequencies simulations are quite different for the two orders, order 1 tends to underestimate the head. Anyway, the head oscillation is orders of magnitude smaller than the reported result, therefore even if they seem completely out, the absolute error is not so different from before. Secondly, there are some wells that have a worst comparison than others for each frequency, this can be due to the diffusivity in the specific position, indeed results depend on signal transmission and informations arriving to observation wells.

After having shown that orders slightly differ on solutions, let's compare approximatively the computational time. For the same operation of creation of the synthetic

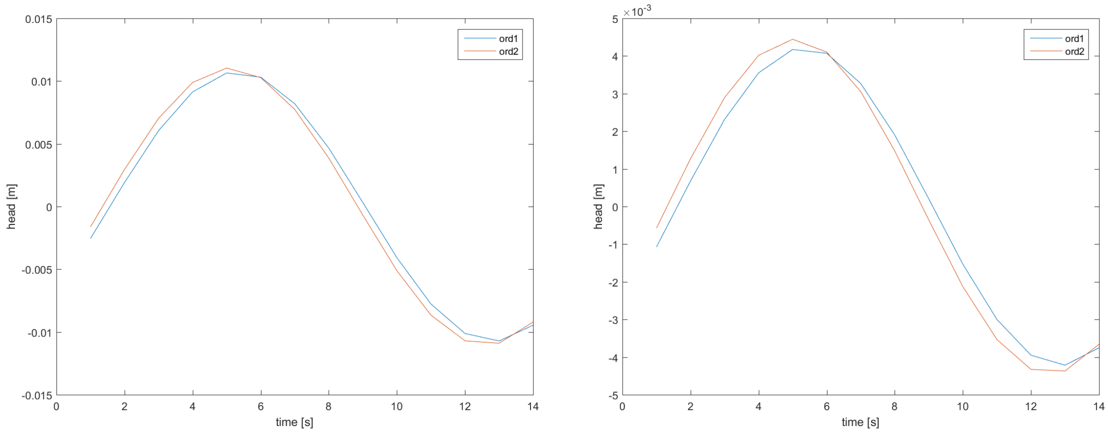

Figure 3.1.2: Head variations in time in point (34,30) - pumping well (30,30) and frequency 0.4189 s^{-1} - Scenarios 1 and 2

field, the 2^{nd} order code takes 16 min 42 s while the 1^{st} order takes only 1 min 21 s, twelfth of time. Even if the PSO process is much more variable in time depending on starting position of particles, the saved time is surely remarkable. Therefore, it is possible to conclude that the first order is not lacking in resolution and much more convenient in time.

3.2 Field analysis: Diffusivity

The second analysis aims to understand the importance of a simulated field property: the diffusivity. As already mentioned, diffusivity describes the ability of the field to transmit a signal, higher it is, faster is the aquifer response and larger is the influenced area.

This analysis comprehends three cases which differ in the mean of the log-storativity Z: D0 where the diffusivity is exactly equal to 1 m^2/s; D1 where it increases to 2.72 m^2/s; D2 where it reaches 7.39 m^2/s. The following table 3.3 reports the main chosen parameters, averages of Y and Z, the correlated average conductivities and storativities, K and S, and their radius of influence determined approximatively as:

$$R = \sqrt{TD} \tag{3.2.1}$$

where T is the period that varies from 5 s to 50 s as shown in 3.2. Corresponding to these period values, minimum and maximum radii of influence are calculated. These two last columns are an approximate indication of the extension of the area of influence.

Table 3.3: Diffusivities used in the sensitivity analysis

Case	m_Y	$mean(K)$	m_Z	$mean(S)$	Diffusivity	Radius of influence min	 max
	$[-]$	$[m/s]$	$[-]$	$[1/m]$	$[m^2/s]$	$[m]$	$[m]$
D0	-8	$3.35\ 10^{-04}$	-8	$3.35\ 10^{-04}$	1.00	2.24	7.07
D1	-8	$3.35\ 10^{-04}$	-9	$1.23\ 10^{-04}$	2.72	3.69	11.66
D2	-8	$3.35\ 10^{-04}$	-10	$4.54\ 10^{-05}$	7.39	6.08	19.22

In this analysis, the area of influence of a pumping well is the key tool of comparison. Indeed, what is important to check is that it reaches nearby observation wells because only in this case the pumping test is effective. Figures 3.2.1 show areas of influence for the three cases for the middle well pumping at a frequency of 0.4189 s^{-1}. As it is expected, it is evident how a larger diffusivity affects a larger area. Furthermore, same results for other wells show that the two upper wells and the eastern one have smaller areas of influence because they are located in a low-diffusivity zone of the synthetic field. This should not be a problem because also the absence of signal is an information for the model. Another final observation is that high frequencies have smaller areas while small ones influence larger areas. In particular, frequency 1.2566 s^{-1} produces effected areas that never reach surrounding wells while frequency 0.1256 s^{-1} has areas of influence reaching observation wells, not for the three wells located in the low diffusivity zone.

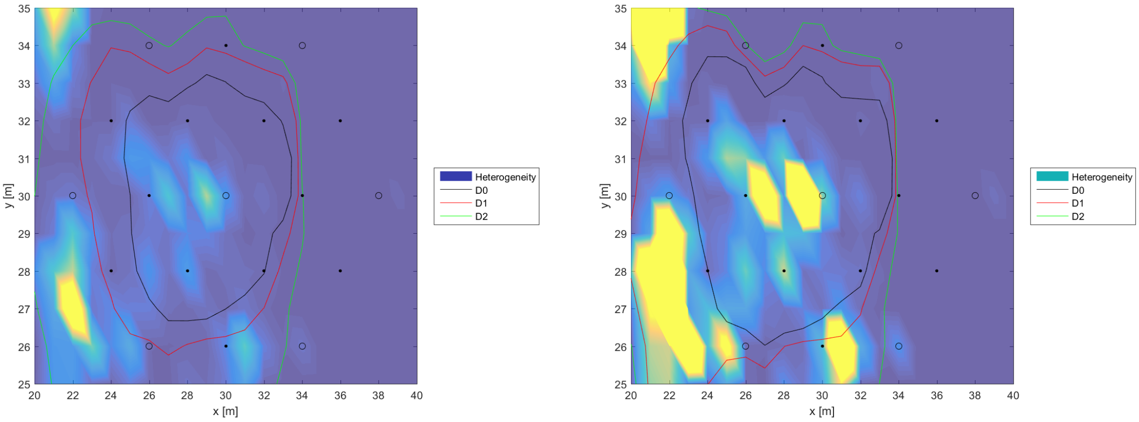

Figure 3.2.1: Influence radius for several diffusivities - frequency 0.4189 s^{-1} - Scenarios 1 and 2

3.3 Pumping test analysis: discharge and frequency

3.3.1 Discharge

In this analysis, the amount of water pumped in by the well is studied. Even if it clearly influences the local head, the discharge should not effects influence area as here defined. Therefore, it should not be relevant in this code because calculated parameters of conductivity and storativity are not correlated to it.
The experiment involves three cases which discharges are respectively 1 l/s, 5 l/s and 10 l/s and their areas of influence are compared. As expected, in figures 3.3.1 only one line can be observed. This is simply due to the fact that areas are perfectly overlapping and therefore the check is verified.

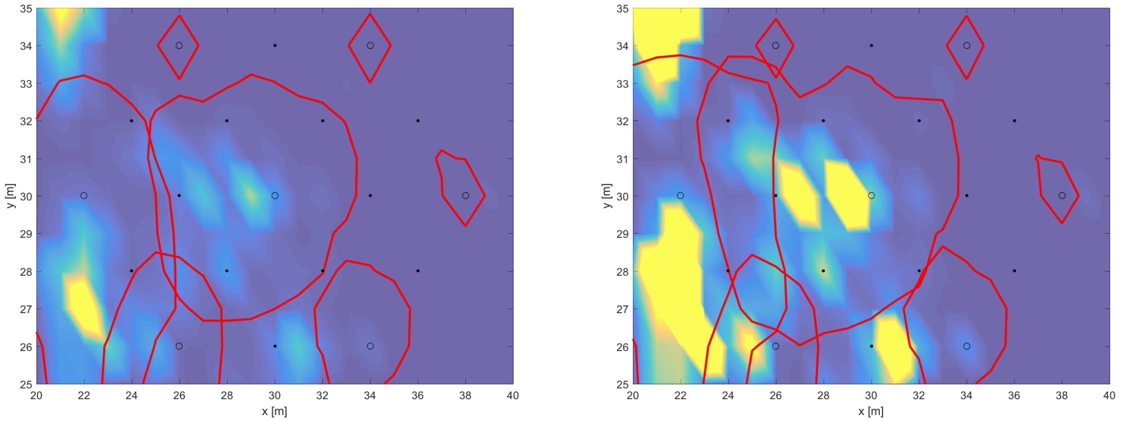

Figure 3.3.1: Influence radius for several discharges - frequency 0.4189 s^{-1} - Scenarios 1 and 2

3.3.2 Frequency

As already mentioned in section 3.2, frequency clearly affects areas of influence. This last analysis aims to enlarge the set of frequencies in order to select the most appropriate ones for following computations. Table 3.4 lists all frequencies analyzed.

Table 3.4: Frequencies used in the sensitivity analysis

Period $[s]$	Frequency $[s^{-1}]$	Radius of influence $[m]$
5	1.2566	3.69
10	0.6283	5.21
15	0.4189	6.38
20	0.3142	7.37
50	0.1256	11.66
100	0.0628	16.49
150	0.0419	20.18
200	0.0314	23.32
300	0.0209	28.58

This analysis simply show the area of influence for each pumping well for each frequency. As for previous analyses, only results of the middle well are reported in this section. Figures 3.3.2 clearly shows how a smaller frequency influence a larger area around the well. The shape is not circular, as would be for an homogeneous field, but it evidently depends on local diffusivity, enlarging in high diffusivity zones. Considering that simulations would be more precise if areas are large enough to

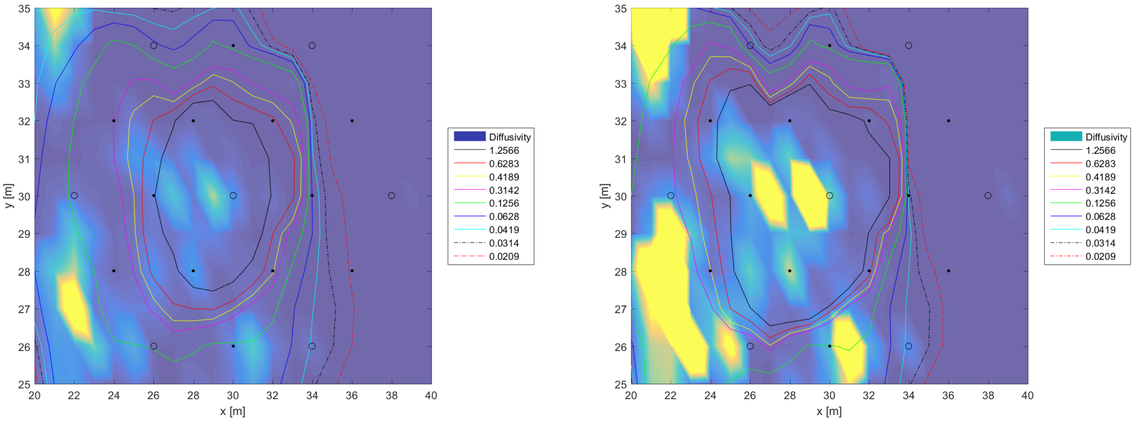

Figure 3.3.2: Influence radius for several frequencies - pumping well (30,30) - Scenarios 1 and 2

include observing wells, there is the possibility to enlarge them with low frequencies and a higher diffusivity. However, it is important to check that areas are not too large to beat boundaries. Ideally, boundaries should not be influenced by the pumping well, but here it is possible to have a slight influence because both synthetic and reproduced fields have the same problem. Following figures 3.3.3 show areas of the middle well for several frequencies in a field with higher diffusivity (test D2 in section 3.2). Figures 3.3.4 represent the most critical well located in (22,30): it influences a very large area that approach boundaries, even if it does not touch them. However, it is important to remember that this area is the 10% of the maximum head, meaning that in reality the area is larger and therefore it probably slightly influences the boundary.

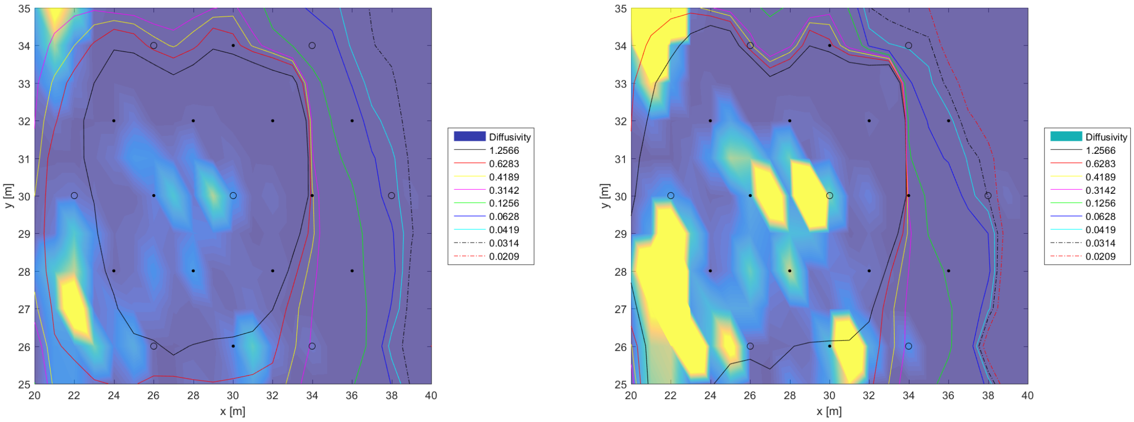

Figure 3.3.3: Influence radius for several frequencies - pumping well (30,30) and D2 - Scenarios 1 and 2

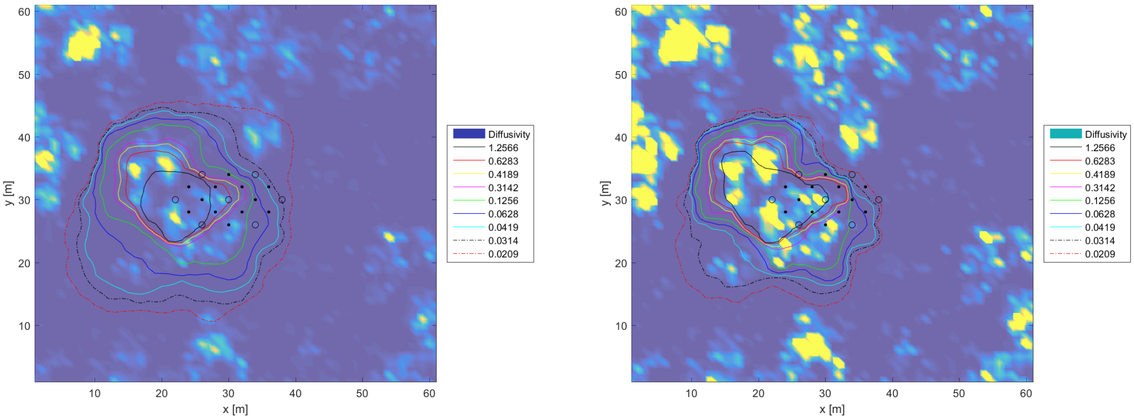

Figure 3.3.4: Influence radius for several frequencies in whole field - pumping well (22,30) and D2 - Scenarios 1 and 2

After having understood the influence of frequencies basing on influence areas, this paragraph wants to determine more rigorously those to use in the test. For this purpose, dimensionless governing equations 2.2.12 and a simple case of an horizontal isotropic layer are considered. The 3^{rd} direction is neglected and in the horizontal plane the only spatial scale is I_{Yh} that it is here assumed equal to 4 m, middle value between the two scenarios' spatial scales. Knowing the distance that test should interest (minimum and maximum distances between wells), approximate radii are calculated using two simple equations, 3.2.1 and first of 2.2.11:

$$\begin{aligned} \omega^* &= \frac{D}{I_{Yh}^2}\omega \\ R^* &= \sqrt{TD} = \sqrt{\frac{2\pi}{\omega^*}D} = \sqrt{\frac{2\pi I_{Yh}^2}{D\omega}D} = \sqrt{\frac{2\pi}{\omega}}I_{Yh} \end{aligned} \tag{3.3.1}$$

In table 3.5 approximate influence radii with correlated frequencies are reported. Looking at this table, it is evident how ideal frequencies depend on field diffusivity, that is unknown a priori. Therefore, in reality technicians should use a large range of frequencies to be sure to include the correct ones.

Table 3.5: Frequencies and correlated influence radii for the three diffusivity cases

R^*	$R = \frac{R^*}{I_{Yh}}$	$\omega = \frac{2\pi}{R^2}$	$\omega^*_{D0} = \frac{D0}{I^2_{Yh}}\omega$	$T_{D0} = \frac{2\pi}{\omega^*_{D0}}$	ω^*_{D1}	T_{D1}	ω^*_{D2}	T_{D2}
$[m]$	$[-]$	$[-]$	$[rad/s]$	$[s]$	$[rad/s]$	$[s]$	$[rad/s]$	$[s]$
4	1.00	6.2800	0.3925	16.00	1.0676	5.88	2.9006	2.17
5	1.25	4.0192	0.2512	25.00	0.6833	9.19	1.8564	3.38
6	1.50	2.7911	0.1744	36.00	0.4745	13.24	1.2891	4.87
7	1.75	2.0506	0.1282	49.00	0.3486	18.01	0.9471	6.63
8	2.00	1.5700	0.0981	64.00	0.2669	23.53	0.7251	8.66
9	2.25	1.2405	0.0775	81.00	0.2109	29.78	0.5730	10.96
10	2.50	1.0048	0.0628	100.00	0.1708	36.76	0.4641	13.53
11	2.75	0.8304	0.0519	121.00	0.1412	44.49	0.3835	16.37
12	3.00	0.6978	0.0436	144.00	0.1186	52.94	0.3223	19.49
13	3.25	0.5946	0.0372	169.00	0.1011	62.13	0.2746	22.87
14	3.50	0.5127	0.0320	196.00	0.0872	72.06	0.2368	26.52
15	3.75	0.4466	0.0279	225.00	0.0759	82.72	0.2063	30.45
16	4.00	0.3925	0.0245	256.00	0.0667	94.12	0.1813	34.64
17	4.25	0.3477	0.0217	289.00	0.0591	106.25	0.1606	39.11
18	4.50	0.3101	0.0194	324.00	0.0527	119.12	0.1432	43.84

3.4 Selected parameters

After having fully understood the code sensibility to some important characteristics, in this section their values are decided for following simulations.
Firstly, the solver order is 1. Indeed, it was previously showed how the second order does not lead to a relevant accuracy improvement for the oscillatory pumping test, while computational times are much longer.

Secondly, the average diffusivity of the field is arbitrarily fixed to 1 m^2/s, scenario D0 in section 3.2. This choice does not lead to any advantage, the only consequence is the need to carefully select appropriate frequencies complying with table 3.5. In all following synthetic cases, frequencies used are those reported in table 3.7.
Finally, last values to fix are statistical parameters and the discharge rate. Statistical parameters are chosen only following the already defined average diffusivity. Regarding the variance, it is selected equal to 4 because it is a high value, usually difficult to reproduce and rarely analyzed with other models. Finally, the discharge rate is arbitrarily fixed to 1 l/s. All selected parameters are reported in table 3.6.

Table 3.6: Selecte parameters

	m_Y $[m]$	σ_Y^2 $[m^2]$	m_Z $[m]$	σ_Z^2 $[m^2]$	I_x $[m]$	I_y $[m]$	Q $[l/s]$
Scenario	-8	4	-8	4	3	5	1

Table 3.7: Selected frequencies

R^* $[m]$	ω_{D0}^* $[rad/s]$	T_{D0} $[s]$
6	0.1744	36.00
8	0.0981	64.00
11	0.0519	121.00
13	0.0372	169.00

In order to check if pumping wells influence observing ones with these selected parameters, following are reported areas of influence of wells for this field and these frequencies. Moreover, in order to have a more precise check, the influence area is here defined as the area in which head variation has a value major or equal to 1% of the maximum change in the pumping well location. Note that areas are in reality still bigger than what it is shown but influence is very little. In figures 3.4.1 is evident how almost all wells are able to reach with their areas other wells. Therefore frequencies used are correct and well selected. Only exceptions are sixth and seventh wells, whose areas are very close to the pumping well and they do not include other observing wells.

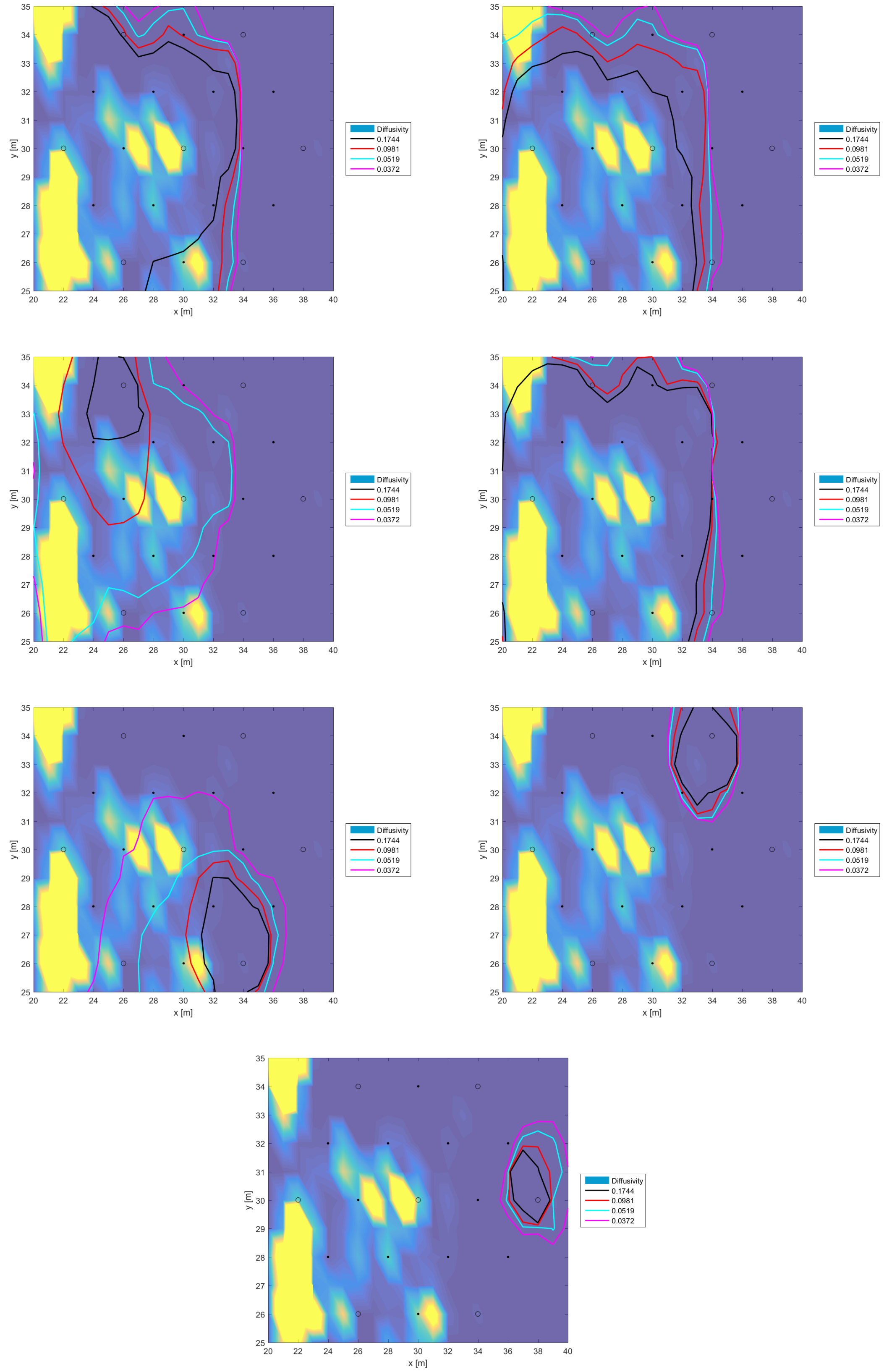

Figure 3.4.1: Influence areas for all selected frequencies in the chosen field, for wells 1 to 7

4

Synthetic cases

The previous sensitivity analysis evaluated the better parameters to use in the code in terms of solver order, diffusivity of the field and frequencies used in the test. This step was very useful to determine some important values used following.
At this point, the code is tested in different simulations in order to observe how it behaves in simpler and more complex fields. The main purpose is to fully understand limitations and strong points of the model.

4.1 Reported results

In order to understand how much the inference is able to correctly reproduce the real field, there are many results that can be shown. It is important to specify here that not all inferred simulations are described in this work, only the most important simulations are here reported: the best simulation with the highest modified Nash-Sutcliffe efficiency and the average of all realizations (following simply called mean simulation). The latter is the simulation often inferred by applied models.

First of all, log-conductivity, log-storativity and diffusivity field maps of best and mean simulations are reported in order to visibly compare their heterogeneity and pattern.
Anyway, maps are only an aspect in assessing the quality of the inversion, other results reported are Y and Z values along transects where wells and pilot points are located. Diagrams report reference values, ones of the best simulation and ones of the average of all realizations.

Then, probability distribution functions (pdfs) of properties for all realizations are reported. Even if this result does not show the local correctness of the simulation, it highlights the goodness of estimated statistical parameters.
Scatter plots of simulated versus observed amplitudes and phases are also inserted. This is a really effective tool, because it directly compares observations of the reference field and of the best simulation. It is easy to read: closer points are to the diagonal line, closer simulated values are to observed ones, and therefore better is the realization.
However, all these results evaluate the goodness of reproduction of field hydraulic properties. Though, the real aim of this model is the reproduction of the contaminant spreading, therefore the most important aspect here is the contaminant transport. In order to evaluate this feature, two typologies of results are listed: breakthrough curves and connectivity functions. The firsts are determined for all realizations by using MODFLOW and MODPATH models. The field is exactly reproduced in geometry and grid spacing, while hydraulic gradient of 0.001 is assigned. The particle tracking is computed in steady state conditions after solving the flow with conductivity fields obtained from the inversion of the oscillatory pumping tests data. An important observation is that in steady state the only relevant property is conductivity, storativity plays a role only in transient conditions. Therefore, this result is an important check only for one property. Connectivity functions indicate connectivity along flow paths with hydraulic conductivity above a given threshold, typically fixed based on the quantiles of the pdf of hydraulic conductivity. For example, connectivity for high quantiles, say 90%, indicates fast pathways.
Given the difficulty to correctly and fully compare simulations with many aspects to take into consideration, it is helpful to have more mathematically based results. Several tables are reported with reference versus simulated statistics, absolute errors, efficiency coefficients for signal, amplitude and phase, and connectivity indexes determined as 1.2.12.

The signal efficiency is determined with the normalized coefficient 2.2.18. While for amplitude and phase efficiencies, the NashSutcliffe efficiency coefficient is used. It is defined as:

$$NS = 1 - \frac{\sum_{obsW=1}^{n_{sample}} (H_{sim} - H_{obs})^2}{\sum_{obsW=1}^{n_{sample}} (H_{obs} - \overline{H_{obs}})^2} \tag{4.1.1}$$

where n_{sample} is the number of observation, H_{sim} are head variations in observing wells locations for tested frequencies, H_{obs} are corresponding observed head variations, $\overline{H_{obs}}$ is the mean of observed values.

4.2 Tests organization

Several tests have been performed in order to explore the capability of oscillatory pumping tests in inferring hydraulic property variations. Tests here reported are: 1) a simple chessboard with block-type variability for both conductivity and storativity (Cb); 2) a heterogeneous storativity and a homogeneous conductivity (K0); 3) a heterogeneous conductivity and homogeneous storativity (S0); and finally 4) hydraulic conductivity and storativity both heterogeneous (SK). In this case, two tests are carried out in order to understand the importance of the extent of ranges set as input in which the PSO algorithm searches possible solutions. We considered also the cases of false negative and false positive as follows: the K-false negative test (KNeg) aims to reproduce heterogeneous reference fields of conductivity and storativity but inversion is performed by assuming correctly that the Z field is heterogeneous and incorrectly that the Y one is homogeneous, while the equivalent S-false negative test (SNeg) assumes a homogeneous storativity field and a heterogeneous conductivity field; the K-false positive test (KPos) aims to infer a homogeneous Y true field and a heterogeneous Z true field by assuming both heterogeneous in the inversion, while the equivalent S-false positive test (SPos) aims to infer a heterogeneous conductivity field and a homogeneous storativity field still assuming both heterogeneous.

Table 4.1: Summary of modeling scenarios considered

Case	True field		Assumed field in inversion	
	Log T	Log S	Log T	Log S
Cb	Chessboard	Chessboard.	Gaussian heterog.	Gaussian heterog.
K0	Homogeneous	Gaussian heterog.	Homogeneous	Gaussian heterog.
S0	Gaussian heterog.	Homogeneous	Gaussian heterog.	Homogeneous
SK	Gaussian heterog.	Gaussian heterog.	Gaussian heterog.	Gaussian heterog.
KNeg	Gaussian heterog.	Gaussian heterog.	Homogeneous	Gaussian heterog.
SNeg	Gaussian heterog.	Gaussian heterog.	Gaussian heterog.	Homogeneous
KPos	Homogeneous	Gaussian heterog.	Gaussian heterog.	Gaussian heterog.
SPos	Gaussian heterog.	Homogeneous	Gaussian heterog.	Gaussian heterog.

4.3 Chessboard

The first test is a simulation of a very simple case, often used in applications where heterogeneity is introduced by identifying blocks in which the hydraulic conductivity can be assumed homogeneous. These blocks, showing contrasting hydraulic properties, are typically identified from geological analysis of the simulation domain. In particular, we consider the case of a domain divided into 4 blocks with homogeneous, but contrasting, hydraulic property variations. In short we call this case chessboard.. The field is divided into four squares and log-conductivity and log-storativity values are assigned to each one in such a way to have different fields but with an average diffusivity of 1 m^2/s. Values of the central zoom are showed in figure 4.3.1.

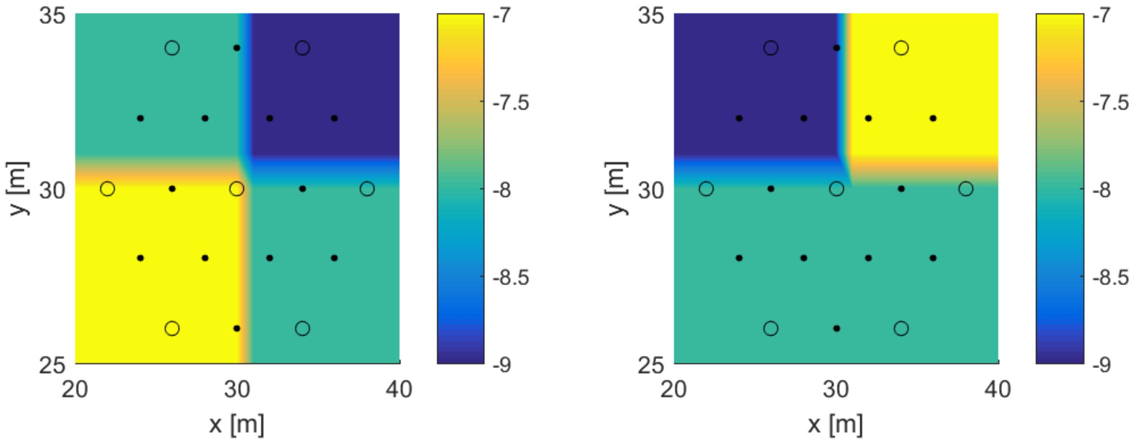

Figure 4.3.1: Reference fields, Chessboard test

4.3.1 Results and observations

Figures 4.3.2 and 4.3.3 show the best hydraulic conductivity and storativity fields (i.e., the fields showing the highest Nash-Shutcliffe coefficient 2.2.18), and the ensemble mean, computed by using forty inversions of the same set of oscillatory pumping tests, respectively. By purpose in the inversion we assumed the wrong model of spatial variability: fields were considered as a SRF with exponential covariance function (2.2.15) despite the real spatial variability is of four blocks with homogeneous, but contrasting hydraulic properties. A first look at the maps suggests that they are not so close to reference fields. Though, observing with more attention it is possible to

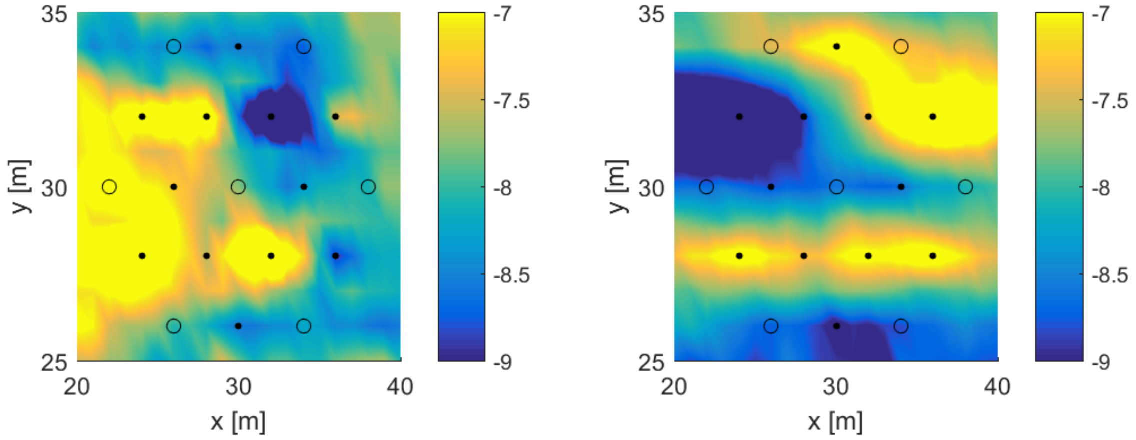

Figure 4.3.2: Best fields, Chessboard test

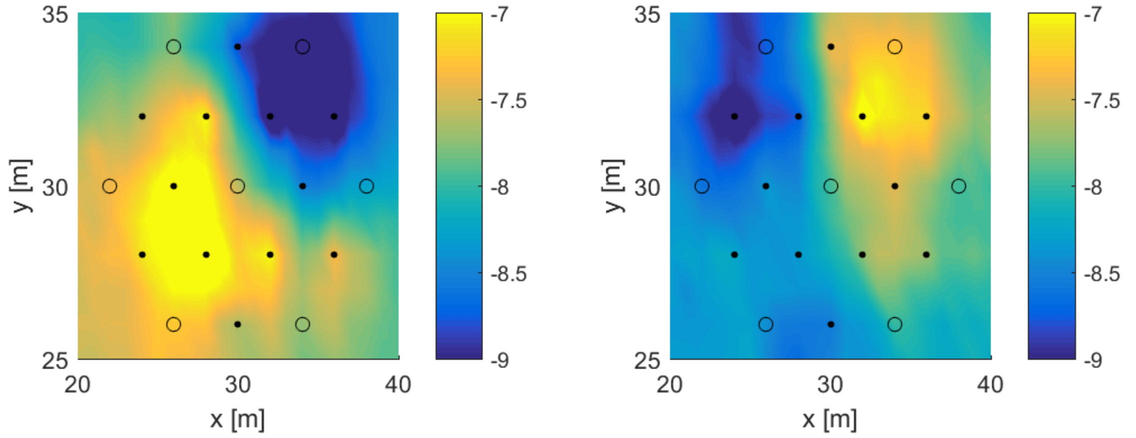

Figure 4.3.3: Mean fields, Chessboard test

notice many similarities.

Indeed, looking at the Y field of the best realization, the area of lower conductivity is correctly located at the northeastern corner while other squares have higher conductivity. In the other three corners a high conductivity area is present, mainly in the southwestern square and it spreads on other two zones. In average, these last two squares are correctly characterized by medium values. The storativity field of the best solution is also quite good: northern corners are very well identified, highest and lowest values are located in the two close squares; while the southern area is not characterized homogeneously by medium values but by two belts off high and low conductivity.

The ensemble average of all inversions shows a better resemblance with the reference field. The lowest log-conductivity values on the northeastern corner, the highest ones on the southwestern one and other two corners are quite well characterized by medium values. Regarding the log-storativity field, similar observations can be done. The model is able to correctly locate highest and lowest values, even if in this case the lowest ones in the northwestern corner are slightly noticeable. The southern area is correctly characterized by medium values.

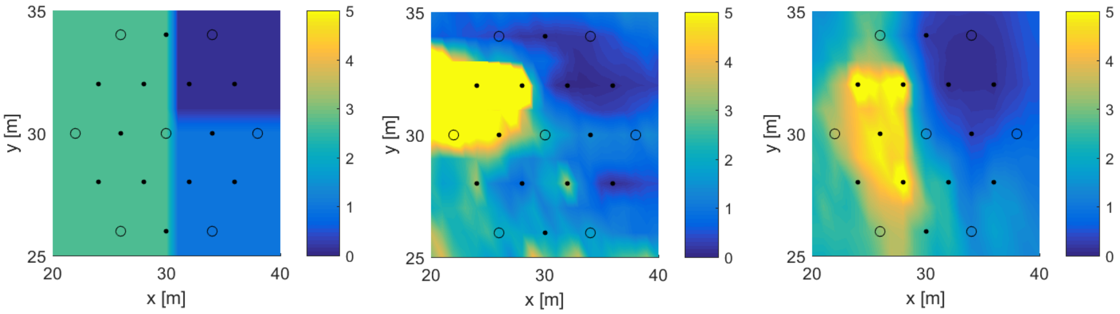

Figure 4.3.4: Diffusivity maps of the reference, best and mean fields, Chessboard test

Figures 4.3.4 show how the diffusivity is reproduced by the model. The best simulation correctly identifies the low values in the eastern zone while it overestimates high values on the West. The mean simulation reproduces much better the diffusivity pattern, identifying the lowest and medium values. Also the western area is characterized by medium-high values. Globally the diffusivity pattern is correctly captured.

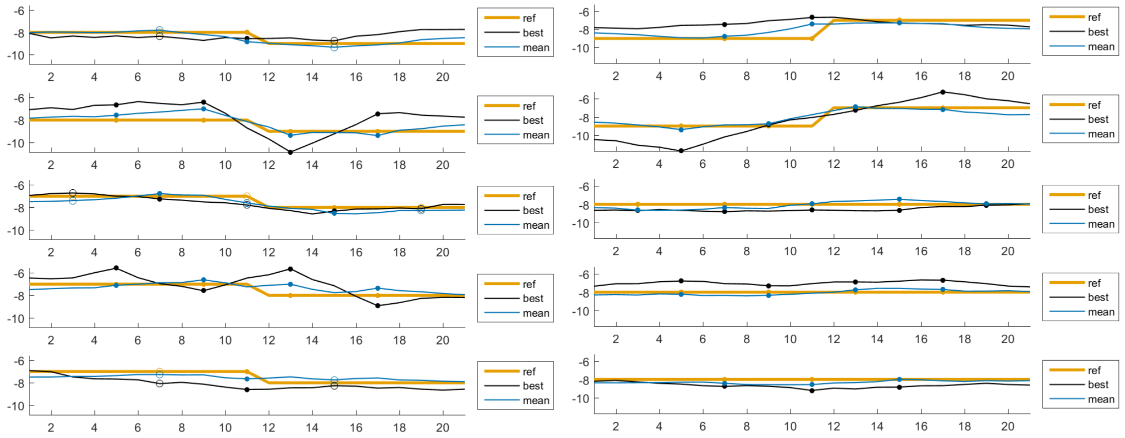

Figure 4.3.5: Transects of reference, best and mean fields in wells and power point rows, Y and Z fields, Chessboard test

Figures 4.3.5 report log-conductivity and log-storativity values along five transects. These transects are located in lines where wells and pilot points are placed, respectively indicated as circles and full points. From this diagrams it is evident how the average of realizations tends to be quite flat and in observation points it is usually close to reference values. On the contrary, the best solution is usually more fluctuating. Values in pilot points are relatively far from reference ones, while in wells locations values are coherent.

Specifically on conductivity diagram, best solution is smooth and similar to the mean in transects where wells are located. The third and fifth transects show similar decreasing trend in x, even if simulations are not able to correctly reproduce the step. In the first transect, not even the decreasing trend is captured by the best solution. In the second and fourthtransects, where only pilot points are present, solutions are more fluctuating, mainly the best one. Looking at the mean solution, it respects the trend and it is also able to reproduce the step. While the best solution has a general decreasing trend but simulation is far from the reference.

Concerning storativity diagram, the last three transects are all flat and close to the reference. The first transect shows the mean simulation that follows the reference seeing also the upward step and the best solution that has a smooth increasing trend. The second one reports a well reproducing mean and a best solution that notices the step but it emphasizes it too much.

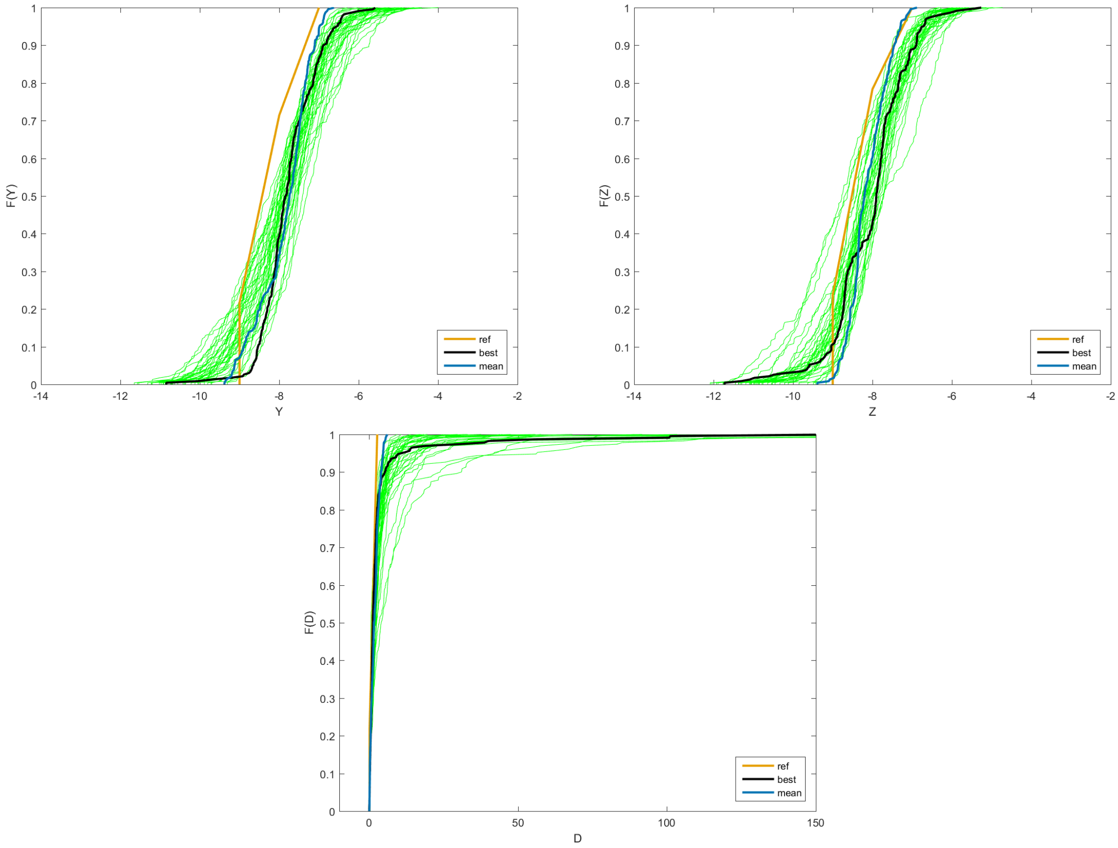

Figure 4.3.6: Empirical cumulative distribution functions of Y, Z and D, Chessboard test

Observing cumulative distribution functions of fields it is possible to notice how reference curves have a completely different shape from others. This is obvious because the chessboard is not a Gaussian field while all realizations are created as Gaussian fields by HYDROGEN. The mean curve tends to be more vertical than the best one, that is because mean tends to remove extreme values and concentrate on medium values, therefore its range is quite small comparing to all single realizations. Also in the diffusivity field it is noticeable how the best curve is far from the reference as well as many realizations, while the mean one is closer to it.

Table 4.2: Statistical parameters comparison, Chessboard test

Field	m_Y	σ_Y^2	m_Z	σ_Z^2	$I_x(Y)$	$I_y(Y)$	$I_x(Z)$	$I_y(Z)$
Reference	-7.93	0.50	-8.02	0.45	6.43	3.37	1.89	2.98
Best	-7.77	0.48	-8.04	0.86	3.35	1.68	5.37	0.98
Mean	-7.86	0.41	-8.12	0.23	5.60	3.06	1.37	4.55

This good reproduction of a non-Gaussian field is confirmed by table 4.2 that reports statistical parameters of best and mean solutions compared to reference parameters evaluated in the central area of interest. It is impressive how close inferred means are to reference ones. Indeed, means have relatives errors lower than 2% while variances have higher relative errors, mainly for the log-storativity where the variance is doubled or halved in simulations. Unfortunately, integral scale are incorrectly reproduced: best solution is worse than main one in this case.

Table 4.3: Mean absolute errors of fields and efficiency coefficients, Chessboard test

(a)

Field	Mean($\|\Delta Y\|$)	Mean($\|\Delta Z\|$)
Best	0.5645	0.6722
Mean	0.3109	0.3777

(b)

	NSEE
signal	0.84
amplitude	0.94
phase	0.97

Next reported results are scatter plots comparing observed values of the reference field and simulated values of the best solution of both amplitude and phase of the signal. To be clear, if a point is exactly in the diagonal line, it means that simulation was able to perfectly evaluate the amplitude or phase of the signal in a specific

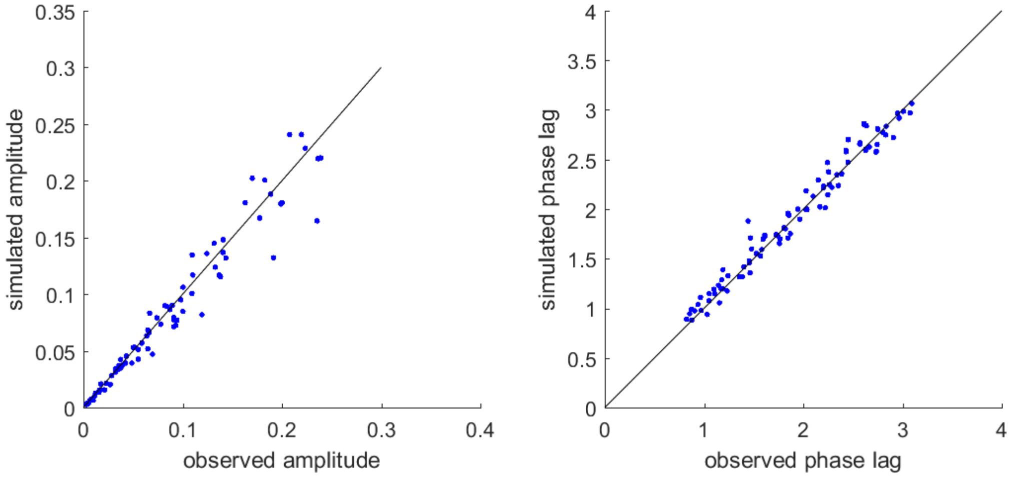

Figure 4.3.7: Scatter plots of simulated versus observed values of amplitude and phase, Chessboard test

observing well for a certain well pumping with a specific frequency. Moreover, in a scatter plot there are as many points as observations. In this work, there are seven wells, one is pumping and six are observing, and each well pumps with four different frequencies. Therefore, observations are $6 \times 7 \times 4 = 168$. Observing plots in figure 4.3.7, results are very good in this chessboard test. Indeed, it is possible to see how close points are to the diagonal line, mainly in the phase scatter plot while they are a little bit farer in the amplitude one. Therefore, the best simulation reproduced a better phase than amplitude of the signal.

Tables 4.3(a) and 4.3(b) are an important indication of the general reproduction efficiency and they confirm what just noticed. First one reports the mean of local absolute errors for the log-conductivity and log-storativity fields. This case, values are all lower than 0.7 and it is definitely a good result. The second table shows how good are the results reporting efficiencies of the signal defined with the modified NashSutcliffe efficiency (equation 2.2.18) and of amplitude and phase defined with the NashSutcliffe efficiency (equation 4.1.1). These three efficiencies are all very high and close to the maximum 1.

As already mentioned describing the reported results in section 4.1, the comparison of breakthrough curves and connectivity functions is the most important result

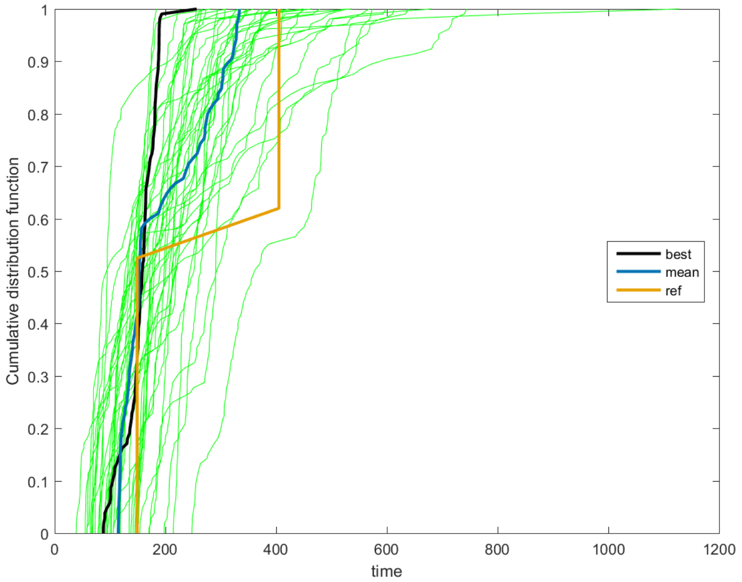

Figure 4.3.8: Breakthrough curves of realizations, Chessboard test

for contaminant transport. In this case, breakthrough curves of simulations have different shape from the reference curve, due to the fact that Gaussian fields demand to reproduce a non-Gaussian one. However, curves are not so far from the reference. Most of particles of simulations are faster than ones of the reference, except for tails that go longer in time. It is noticeable how the best solution is much faster and not well representative: fastest particles arrive at the end of the interested area, 40 m, in about 100 d (150 d for the reference simulation) and slowest in around 200 d (400 d in the reference). On the contrary, the mean of realizations is more capable to capture the shape of the reference breakthrough curve with fastest particles closer to 150 d and lowest ones arriving in more than 300 d.

Connectivity indexes and functions are evaluated as already explained in section 1.2.6. The index, determined with equation 1.2.12, shows how both mean and best simulations underestimate the value of the reference connectivity index, even if they are quite close to it. Looking at the functions for the three defined thresholds (50, 75, 90 percentiles), a similar observation can be done. Simulations underestimate the reference connectivity for all considered lags going well below bounds defined as the reference field plus/minus the standard deviation of all fields for that specific lag. Figures 4.3.9 show that for a low threshold the connectivity is globally high while for the 75 and 90 percentiles the connectivity curve declines similarly for the best and mean, but both faster than in the reference field, suggesting that the inferred fields are less connected than the reference, as it is expected due to the wrong model of spatial variability.

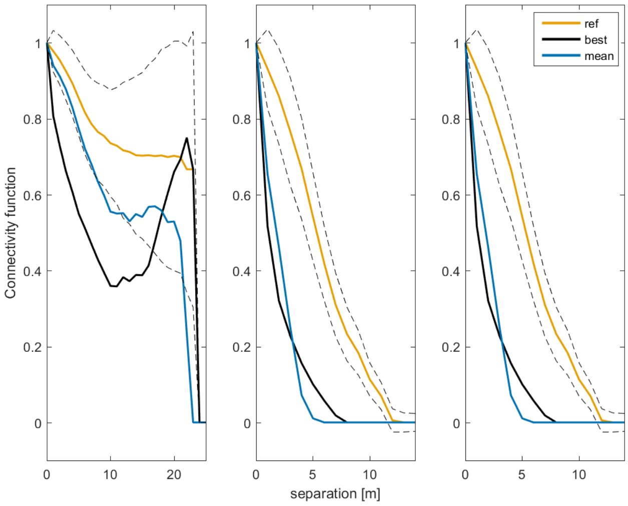

Figure 4.3.9: Connectivity functions (50, 75, 90 percentile), Chessboard test

Table 4.4: Connectivity indicator CI, Chessboard test

	CI
reference	1.78
best	1.56
mean	1.64

Considering all these results it is evident how simulations are able to infer reference properties of a simple non-Gaussian field. In particular, the best solution is good on estimating heterogeneity and on reproducing the observed signal. However, the average of all realizations identifies patterns more precisely and well infer statistics of the true field. Observing results about contaminant transport, breakthrough curves are close to the reference but its trend is not correctly inferred neither by the best nor the mean simulation, connectivity function too is always underestimated by simulations.

4.4 One heterogeneous and one homogeneous fields

A simple test is done assuming that one property of the field is homogeneous, the homogeneity is known a priori but the exact value of the property is not known. Given the knowledge of homogeneity for a certain property, the range in the PSO algorithm of the respective variance is fixed equal to 0, as reported in table 4.5.

Table 4.5: Parameters and ranges for the two homogeneous tests (K0 and S0)

		$mean(Y)$	$var(Y)$	$mean(Z)$	$var(Z)$	I_x	I_y
K0	true parameters	-8	0	-8	4	3	5
	range	-9.50	0.00	-9.50	3.00	1.00	1.00
		-6.50	0.00	-6.50	5.00	7.00	7.00
S0	true parameters	-8	4	-8	0	3	5
	range	-9.50	3.00	-9.50	0.00	1.00	1.00
		-6.50	5.00	-6.50	0.00	7.00	7.00

4.4.1 Results and observations for 'Homogeneous K'

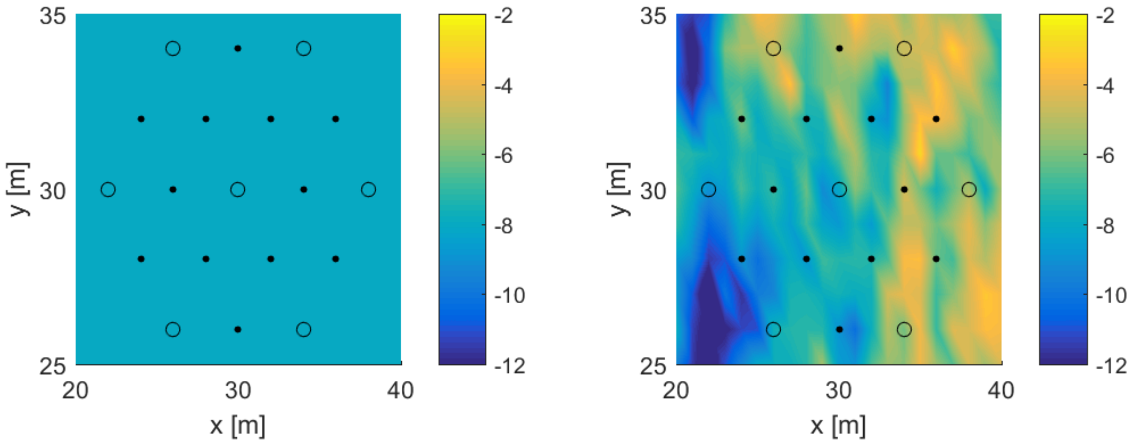

Figure 4.4.1: Reference fields, K-homogeneous test

Figures 4.4.2 and 4.4.3 are the maps of log-conductivity and log-storativity reproduced respectively by best and mean simulations. Observing the best reproduction of the Z field that is more of interest in this specific test, it is noticeable that the inferred field is able to reproduce approximatively the reference field. Indeed, the pattern is identified with a low-values belt on the West and medium-high values on the eastern area. However, the two spots characterized by lowest values on the western corners are only partially reproduced, while a high-values spot is incorrectly inferred on the East. Considering now the mean simulation, the inferred Z field correctly recognizes the pattern of the property but it is not able to see extreme values,

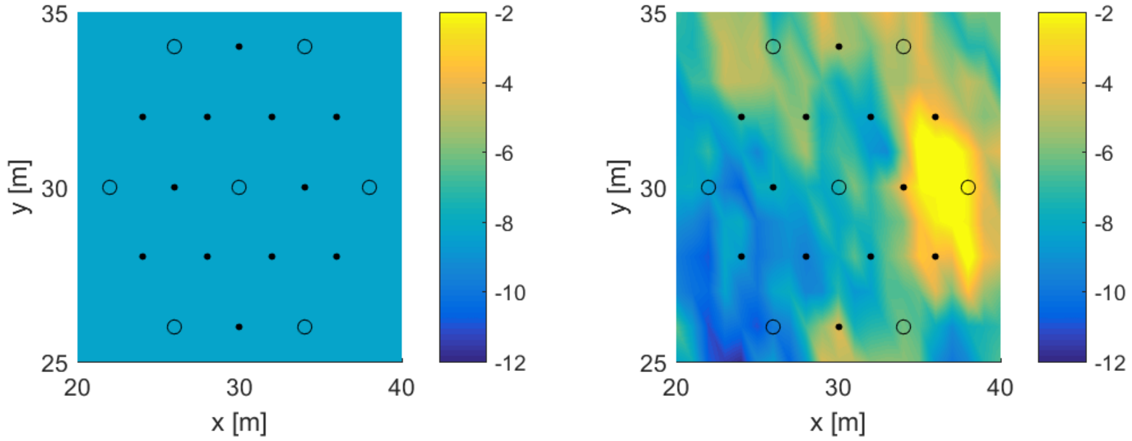

Figure 4.4.2: Best fields, K-homogeneous test

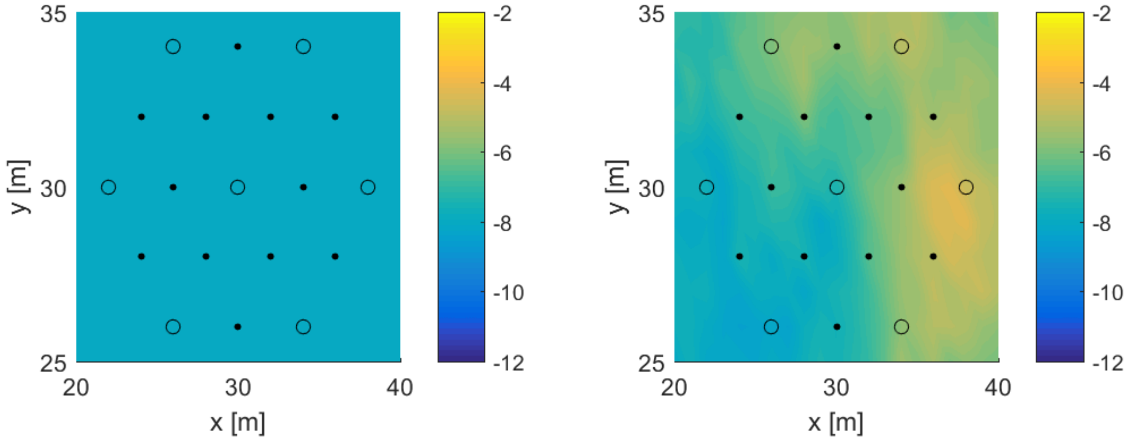

Figure 4.4.3: Mean fields, K-homogeneous test

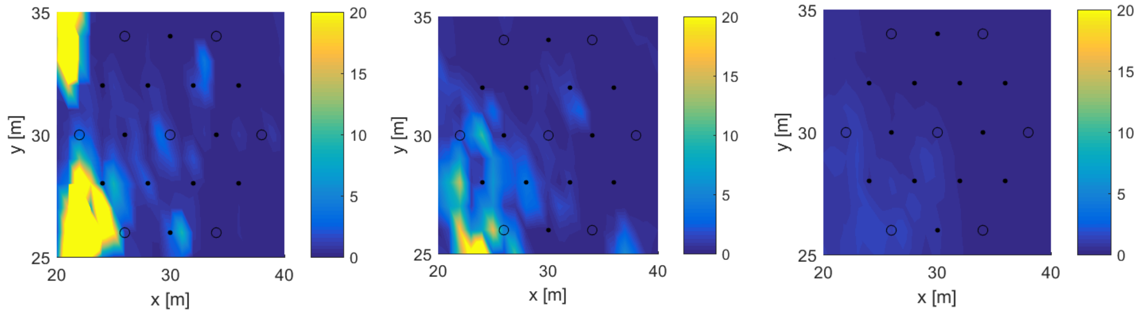

Figure 4.4.4: Diffusivity maps of the reference, best and mean fields, K-homogeneous test

such as lowest ones on the West, and heterogeneity on the eastern zone.

The inference of diffusivity is strongly influenced by the inference of conductivity and storativity, due to the homogeneous Y field it is mainly influenced by the Z field. Figures 4.4.4 clearly show how the best simulation approximatively reproduces the pattern but it misses highest values due to the missed reproduction of lowest values of log-storativity on western corners. The mean simulation infers an almost homogeneous diffusivity field.

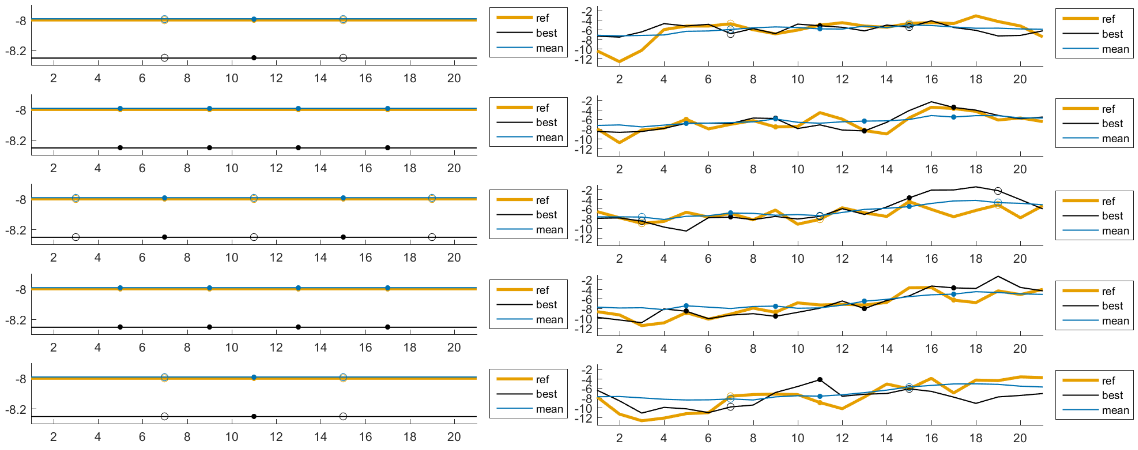

Figure 4.4.5: Transects of reference, best and mean fields in wells and power point rows, Y and Z fields, K-homogeneous test

Transects in figures 4.4.5 give more detailed information on simulations reproduction. Indeed, it is evident how close the average of all realizations reproduces the homogeneous log-conductivity field, while the best simulation is slightly worst. Moving to the more dynamic Z transects, the mean simulation is able to follow the general trend of the reference field but it does not reproduce oscillations and extreme values. On the contrary, the best simulation follows much more precisely the reference line but it incorrectly estimates values mainly on lateral zones. It is noticeable how it tends to overstate values on western belt (see first, second and fifth transects) and on central eastern area (see third and forth transects). Always on the eastern area, the best simulation also understates the reference field on first and last transects.

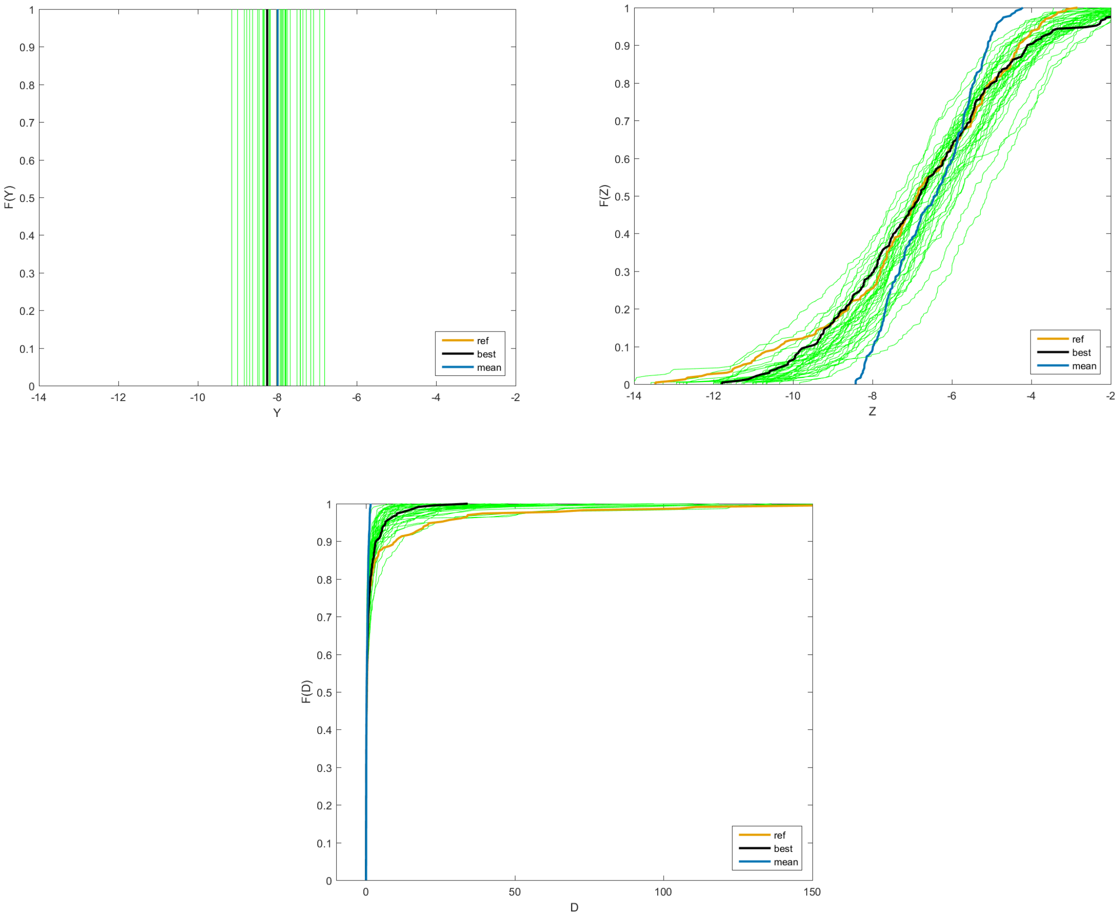

Figure 4.4.6: Empirical cumulative distribution functions of Y, Z and D, K-homogeneous test

Heterogeneity of the field is reproduced by simulations even if they are not so

Table 4.6: Statistical parameters comparison, K-homogeneous test

Field	m_Y	σ_Y^2	m_Z	σ_Z^2	$I_x(Z)$	$I_y(Z)$
Reference	-8.00	0.00	-6.95	4.92	2.42	3.28
Best	-8.25	0.00	-6.75	4.83	2.53	2.84
Mean	-7.99	0.00	-6.50	1.21	3.45	4.73

precise. Therefore, let's compare probability distributions to see how statistics are inferred. Observing the Y pdf it is evident what already mentioned before, the mean simulation exactly infers the reference while the best simulation is close to it but it is not overlapping. On the contrary, the best Z pdf overlaps the reference pdf in medium values while it underestimates probability of extreme values. The mean simulation in this case is too steep because it does not consider extreme values and it is more concentrated on central values. Diffusivity probabilities are not inferred precisely by the best simulation that, as already above-mentioned, does not reproduce highest diffusivity values. The reference field reaches values of 150 m^2/s while the best one does not even reach 50 m^2/s. The diffusivity pdf of the main simulation is almost vertical and it does not see high values at all.

Table 4.6 confirms what just observed about statistical reproduction. Indeed, best simulation infers correctly both means and variances, the Y variance easily for assumption while the Z variance is precisely inferred. Also integral scales of Z are very well identified. The mean simulation returns very precise means, mainly the Y mean, but it strongly undervalues the log-storativity variance.

Scatter plots of simulated observations on wells in figures 4.4.7 show that amplitude of the signal is inferred precisely while phase is sometimes misunderstood. Indeed, points on the amplitude plot are quite concentrated along the diagonal while points of the phase plot are much more scattered in the graph. These outside points are numerous and the majority of them is due to the seventh well, located in a low diffusivity field that obstructs signal transmission. Same observations are done watching at table 4.7(b) where the amplitude efficiency is very high while the phase one is low and it strongly influences the overall signal efficiency.

Results of figure 4.4.8 are very similar to previous conductivity results. Indeed,

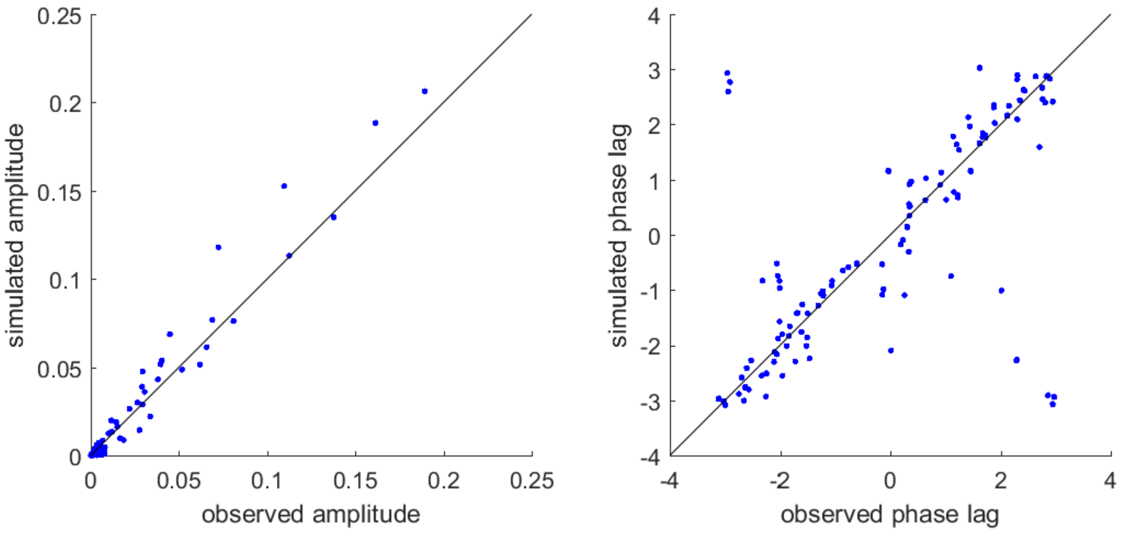

Figure 4.4.7: Scatter plots of simulated versus observed values of amplitude and phase, K-homogeneous test

Table 4.7: Mean absolute errors of fields and efficiency coefficients, K-homogeneous test

(a)

Field	Mean($\|\Delta Y\|$)	Mean($\|\Delta Z\|$)
Best	0.2538	1.5516
Mean	0.0076	1.2712

(b)

	NSEE
signal	0.38
amplitude	0.94
phase	0.36

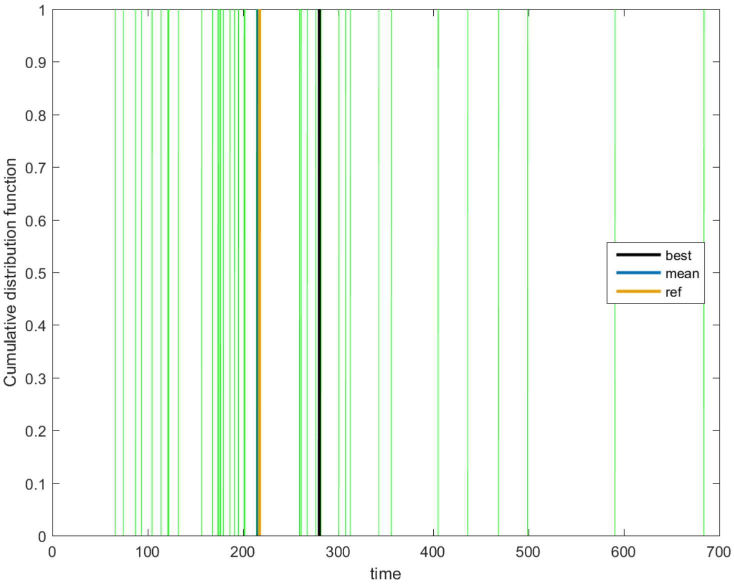

Figure 4.4.8: Breakthrough curves of realizations, K-homogeneous test

considering that contaminant transport is mainly linked to the conductivity field and that it is homogeneous, breakthrough curves are here vertical because particles uniformly reach the border. Moreover, due to the precise reproduction of the conductivity field by the mean simulation, also in this result the mean curve is really close to the reference one, while the best one is less precise.
Same observations can be done for connectivity functions. Indeed, the mean simulation perfectly overlaps the reference one while the best simulation understates it. The connectivity index is the same for all simulations, due to the uprightness of breakthrough curves.

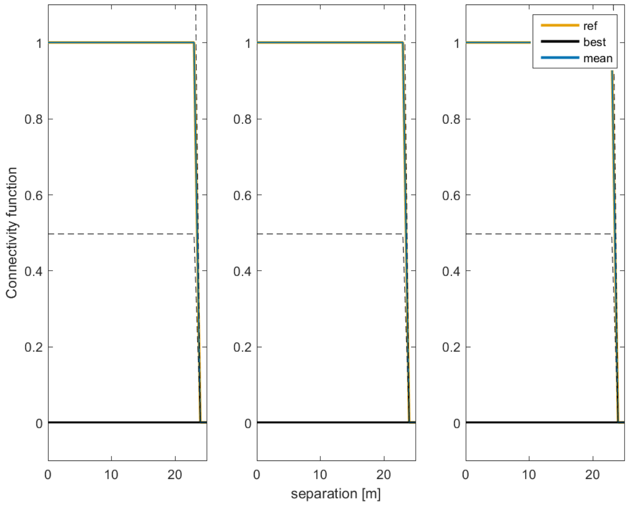

Figure 4.4.9: Connectivity functions (50, 75, 90 percentile), K-homogeneous test

Table 4.8: Connectivity indicator CI, K-homogeneous test

	CI
reference	1.00
best	1.00
mean	1.00

4.4.2 Results and observations for 'Homogeneous S'

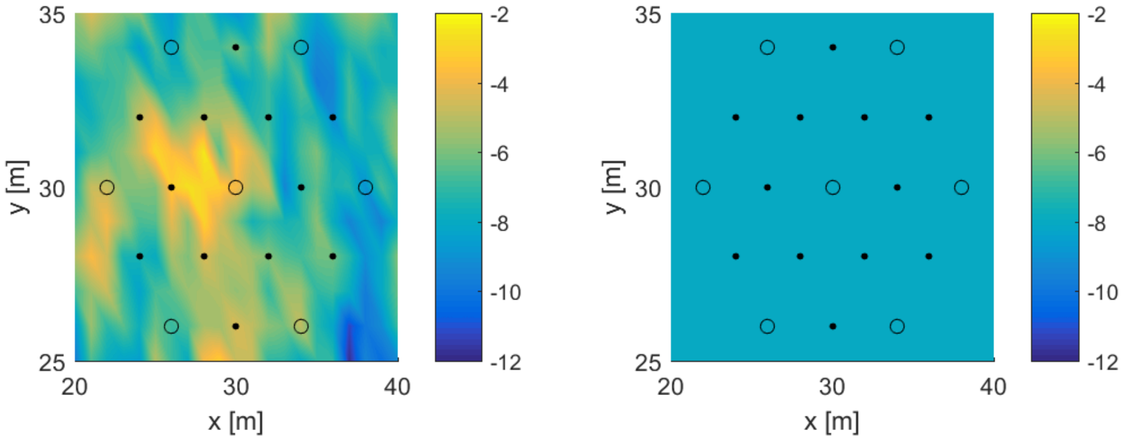

Figure 4.4.10: Reference fields, S-homogeneous test

Figures 4.4.11 and 4.4.12 show how the model is able to infer the reference field visible in figures 4.4.10. Looking first at the best simulation, it is possible to notice how well it reproduces fields. The log-storativity field is obviously homogeneous and close to the reference value, the log-conductivity field is very heterogeneous and

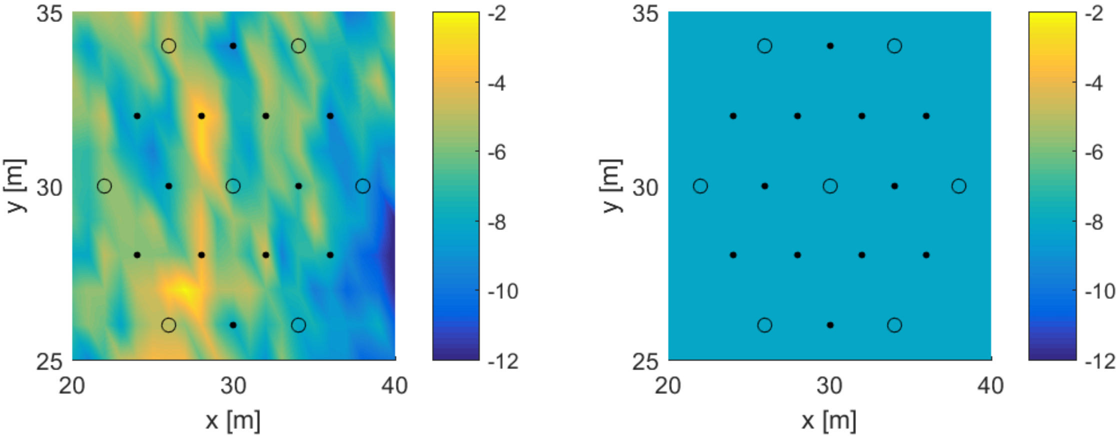

Figure 4.4.11: Best fields, S-homogeneous test

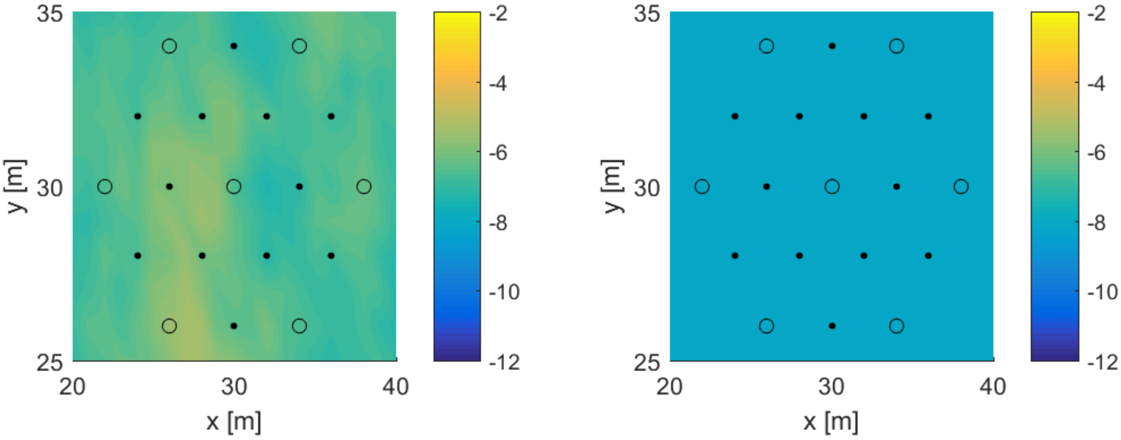

Figure 4.4.12: Mean fields, S-homogeneous test

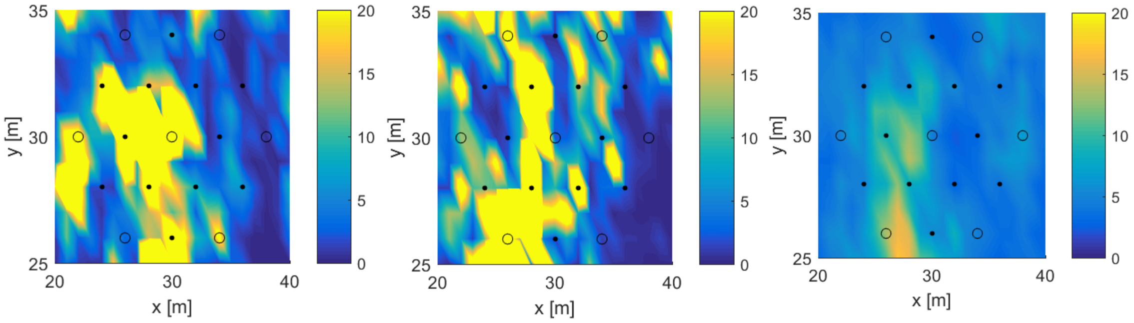

Figure 4.4.13: Diffusivity maps of the reference, best and mean fields, S-homogeneous test

quite similar to the reference pattern. Indeed, it infers the low-conductivity belt on the East, while on the West it recognizes a fragmented medium-conductivity area with high values that are not as compacted as they should be. Looking now at the mean fields, the Z field is of course homogeneous and close to the reference value, while the Y field is too homogeneous to reproduce correctly the reference field. A slightly visible pattern identifies the central high-conductivity zone and the eastern low-conductivity belt. However, in the mean process extreme values are eliminated and values are very close to the average of the reference field.

The diffusivity field depends on conductivity and storativity fields following its definition 1.2.10. Considering that the log-storativity is homogeneous in the field, the diffusivity patter will depend mainly on the log-conductivity. Indeed, observations done in the Y field can be here repeated. The best field well recognizes the eastern belt and the central high values zone, even if it is more scattered than the reference. While the mean field does not show any extreme value but it identifies the pattern.

Figures 4.4.14 illustrates properties variations along the five transects where wells and pilot points are placed. Regarding log-conductivity, it is evident how the mean is almost flat, it slightly follows the reference trend. On the contrary, the best simulation has a more oscillating tendency. Usually, it is able to follow the reference trend overstating some points. Only in the third transect there is a relatively long section where the best conductivity is below the reference and this is the central

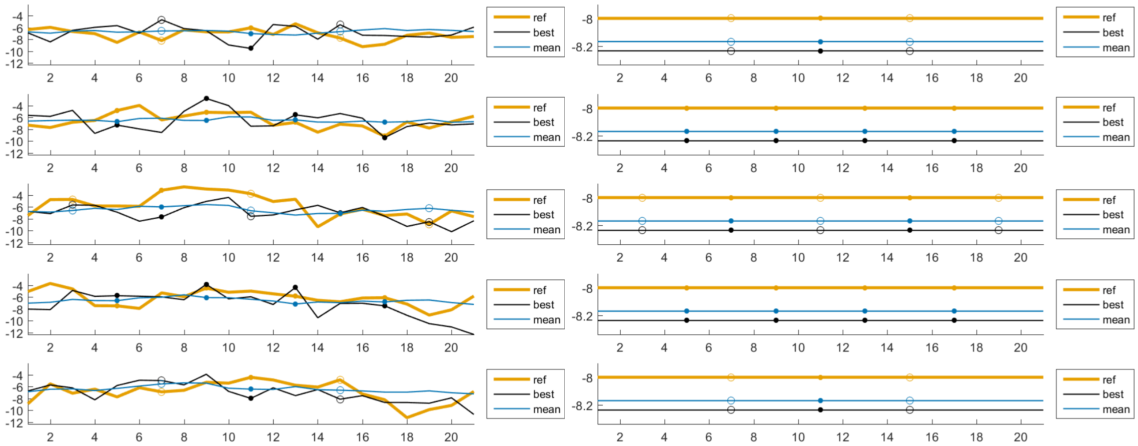

Figure 4.4.14: Transects of reference, best and mean fields in wells and power point rows, Y and Z fields, S-homogeneous test

high-values zone that was already noticed as missed in the maps. Regarding log-storativity, transects are all completely flat due to the assumption of homogeneity. Both mean and best simulations underestimate the value but they are very close to it.

Observing in figures 4.4.15 the probability distributions of the three properties (log-conductivity, log-storativity and diffusivity), it is evident how the best simulation well captures the statistics of the reference field. Indeed, in the Y distribution function, all simulations are close to the reference, the best pdf almost overlaps the reference one on highest values while it overstates low-medium values. On the contrary, the mean pdf completely misunderstands the distribution because it excludes extreme values and concentrates on medium values, as already noticed in maps 4.4.12. Looking at Z pdfs, the first observation is that they are all vertical due to the assumption of homogeneity and they are all close to the reference, indeed the range of variation for Z in the PSO algorithm, in which particles can move, is $(-9.5; -6.5)$, but due to the PSO optimization the final values are much closer to the reference value, approximatively from -8.6 to -7.8. Finally, the diffusivity pdfs are also quite close to the reference, understandable considering that log-conductivity and log-storativity statistics are both well inferred. The best simulation almost overlaps the reference along all diffusivity values, mostly in highest values while best pdf overstates low-medium values, as for the conductivity pdf. Also in this property, the mean pdf misunderstands the reference due to the exclusion of extreme values.

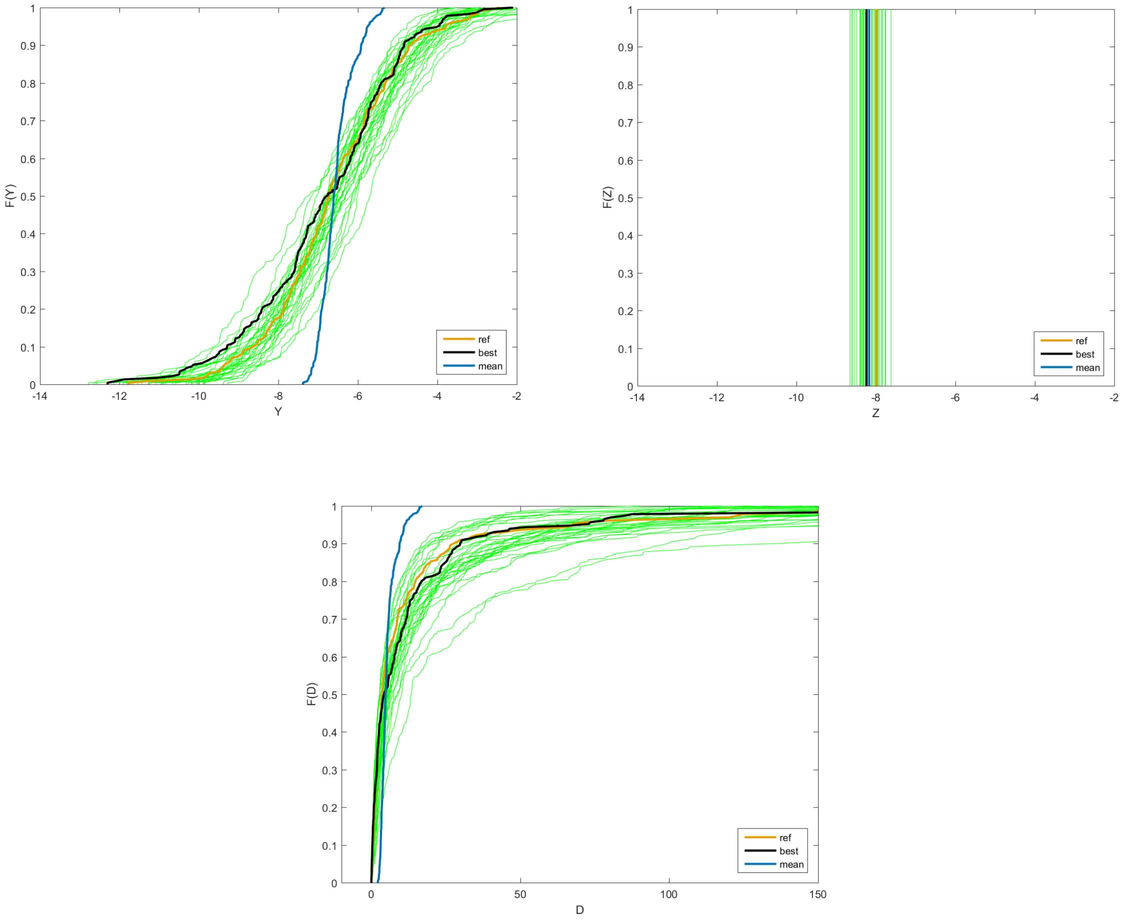

Figure 4.4.15: Empirical cumulative distribution functions of Y, Z and D, S-homogeneous test

Table 4.9: Statistical parameters comparison, S-homogeneous test

Field	m_Y	σ_Y^2	m_Z	σ_Z^2	$I_x(Y)$	$I_y(Y)$
Reference	-6.63	2.61	-8.00	0.00	0.79	2.18
Best	-6.85	3.25	-8.24	0.00	0.76	2.77
Mean	-6.54	0.17	-8.17	0.00	0.44	3.52

Table 4.9 states what was already noticed from previous results. Both best and mean simulations well infer means of the properties, but the mean simulation strongly underestimates the variance of the log-conductivity creating a smoothed field, while the best simulation recognizes a high variance, even if it overestimates it. Moreover, it is observable that integral scales of the log-conductivity are precisely inferred by the best simulation. Of course, there are no integral scales for the log-storativity given that it is homogeneously distributed.

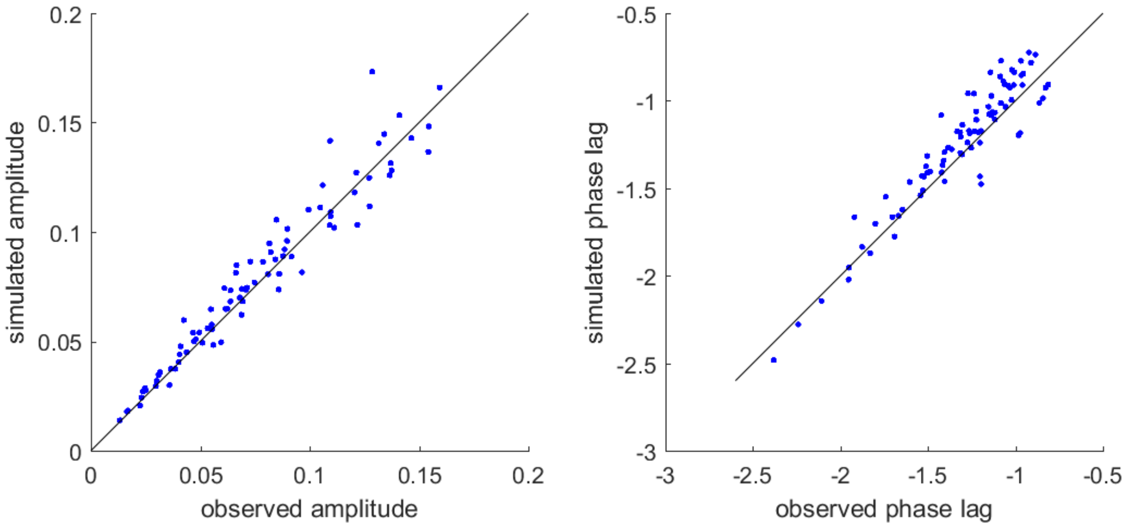

Figure 4.4.16: Scatter plots of simulated versus observed values of amplitude and phase, S-homogeneous test

Scatter plots of amplitude and phase of the signal show how the observed signal in the observing wells is globally well inferred by the best simulation. The phase points concentrate along the diagonal while the amplitude ones seem slightly more diffused but it is only for few points.

Table 4.10(a) reports absolute mean errors for the log-conductivity that are quite high, while they are very low for the homogeneous log-storativity. While table 4.10(b) reports efficiency coefficients for the signal, the amplitude and the phase. It is noticeable that they are very high, the amplitude efficiency reaches almost the maximum while the phase is lower and it is the component that lowers the overall signal efficiency.

Table 4.10: Mean absolute errors of fields and efficiency coefficients, S-homogeneous test

(a)

Field	Mean($\|\Delta Y\|$)	Mean($\|\Delta Z\|$)
Best	1.6597	0.2353
Mean	1.2062	0.1682

(b)

	NSEE
signal	0.84
amplitude	0.93
phase	0.82

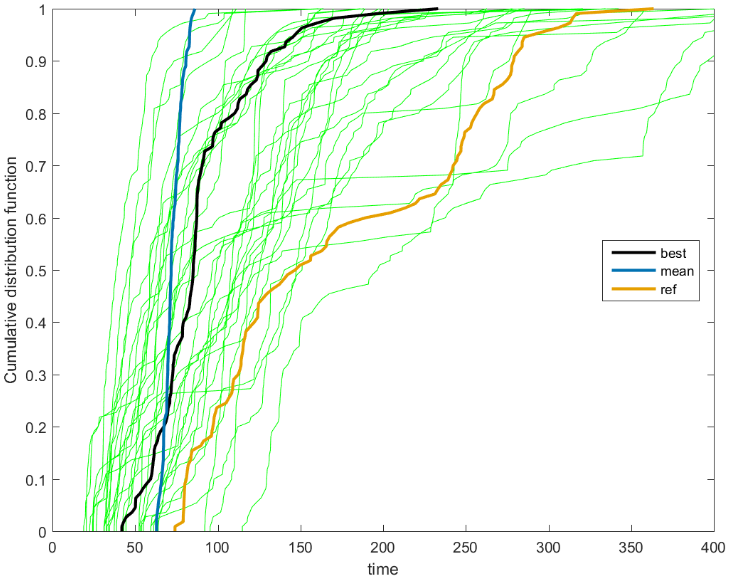

Figure 4.4.17: Breakthrough curves of realizations, S-homogeneous test

Observing simulated breakthrough curves, the majority of the curves are characterized by a steep foot and a long tail, registering a clear distinction between faster and slower particles. Particles are relatively slow on the reference field, considering that almost all inferred curves are faster. Faster particles reach the end of the central area of interest in about 75 d, while slower ones take more than 300 d. The best simulation has a faster breakthrough curve, in which faster particles take less than 50 d, while slower ones more than 150 d. On the contrary, the mean curve has a different trend that does not clearly differentiate faster and slower particles due to the smoothed field.

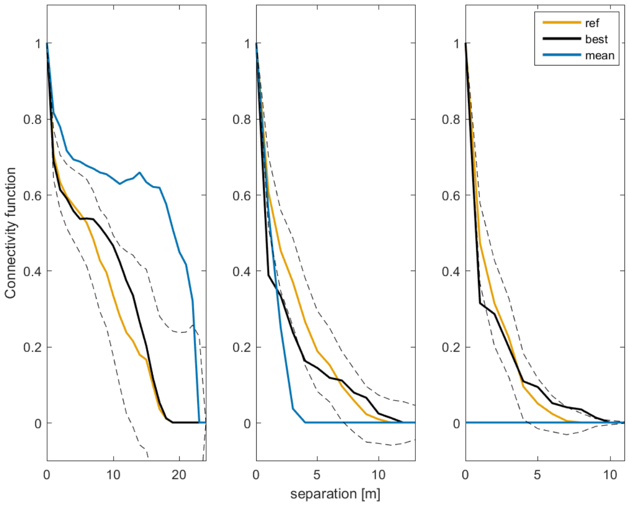

Figure 4.4.18: Connectivity functions (50, 75, 90 percentile), S-homogeneous test

Table 4.11: Connectivity indicator CI, S-homogeneous test

	CI
reference	1.47
best	1.33
mean	1.65

The final results consider the connectivity of the field. Figure 4.4.18 reports the connectivity function versus the space lag between two pixels. In this test, the best curve is very close to the reference one in all the thresholds used. This is very promising because it means that, even if the field's heterogeneity and the breakthrough curve are not fully captured, the connectivity field is understood for both high and medium values. On the contrary, the mean curve overestimates the connectivity for the 50 percentile, underestimates for the 75 percentile and it does not see anything for the 90 one. Looking at the connectivity index reported in table 4.11, both best and mean simulations result in indexes close to the reference value, but this is a much more limited comparison respect to the connectivity function.

4.4.3 Observations

In this section, the model is facilitated in its work considering that one statistical parameter, the variance of one property, is known and therefore not variable in the inversion and optimization procedures.

In the first K-homogeneous test, the reference log-conductivity field is homogeneous and it is known. The reproduction by the best simulation of fields is quite precise even if it misses lowest values and it exaggerates on highest ones. However, pdfs and statistics are well reproduced. As for the Y field, results about contaminant transport are nice but not precise, the mean simulation gets a better reproduction. The overall efficiency is 0.38.

The latter test, S-homogeneous test, results in a good reproduction of the heterogeneous log-conductivity field, the homogeneous log-storativity one, pdfs and statistics. Indeed, it has a overall efficiency of 0.84. Breakthrough curves are close to the reference but the best solution is not able to reproduce it, while connectivity functions are very close to the reference one mainly for higher thresholds.

Summarizing all results, it is possible to conclude that in both tests the average of all realizations is able to identify the heterogeneity pattern but it does not reproduces extreme values, misunderstanding pdfs and statistics. Moreover, it is evident how the S-homogeneous test reaches a much higher overall efficiency than the K-homogeneous one and it is mainly linked to the much higher phase efficiency. This large difference is not so evident by properties maps, however transects and pdfs are reproduced better in the latter test.

4.5 Heterogeneous fields with small and large PSO ranges

The model here tested uses the PSO algorithm to optimize solutions, as explained in section 2.2. In order to use this algorithm, it is necessary to give as input ranges of possible variation for the parameters, so that PSO has a delimited multidimensional search space. These ranges are unknown a priori and therefore they have to be large enough to be sure to include real field parameters. However, if some information is available thanks to preliminary tests or knowledge, it is possible to considerably reduce ranges. This way, the computational time decreases.

This section aims to compare tests of the same random field obtained using first small ranges, then larger ones. The reference field is characterized by the previously selected parameters, reported in table 3.6. Fields obtained are showed following while ranges are reported in table 4.12.

Table 4.12: Ranges used for the PSO algorithm in Small and Large ranges scenarios (SK_S and SK_L)

	$mean(Y)$	$var(Y)$	$mean(Z)$	$var(Z)$	I_x	I_y
Small	-9.50	3.00	-9.50	3.00	1.00	1.00
	-6.50	5.00	-6.50	5.00	7.00	7.00
Large	-12.00	0.01	-12.00	0.01	1.00	1.00
	-4.00	6.00	-4.00	6.00	10.00	10.00

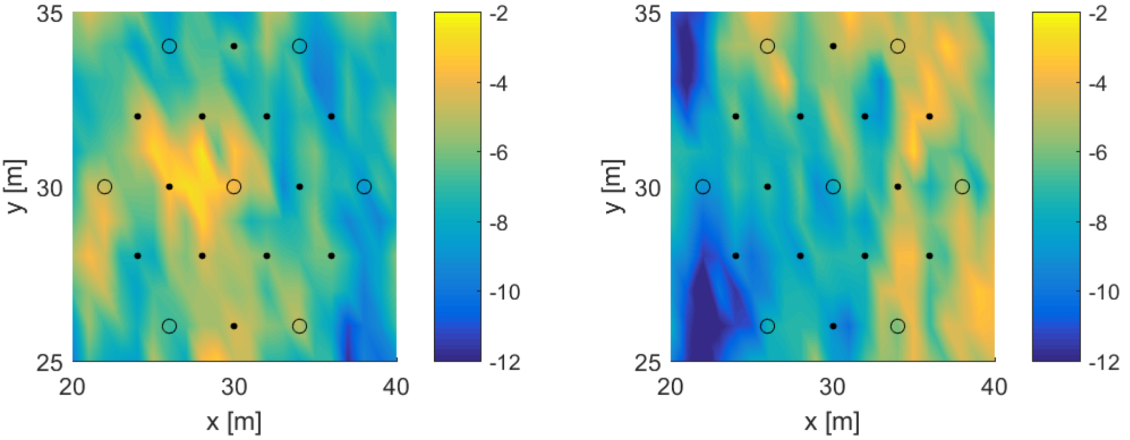

Figure 4.5.1: Reference fields, PSO ranges test

4.5.1 Results and observations for 'Small ranges'

Figures 4.5.2 and 4.5.3 are fields inferred by the best simulation and the average of all simulations. Ideally they should be similar to reference fields 4.5.1. In this test, fields seem to get wrong inference. Indeed, the best simulation results in a log-conductivity field that recognize the central medium-high Y spot, but it strongly underestimates the western area and the northeastern corner. Looking at the log-storativity field, the best simulation identifies the low conductivity zone but it wrongly places lowest values. Moreover it underestimates the south-eastern corner and overestimates the north-eastern one. On the contrary, the mean simulation identifies correctly the pattern of the field but, as in previous cases, it does not consider extreme values, so it is very smooth.

Diffusivity field is quite well inferred by best and mean simulations. Highest values are correctly located on the western area, even if the best solution misses some

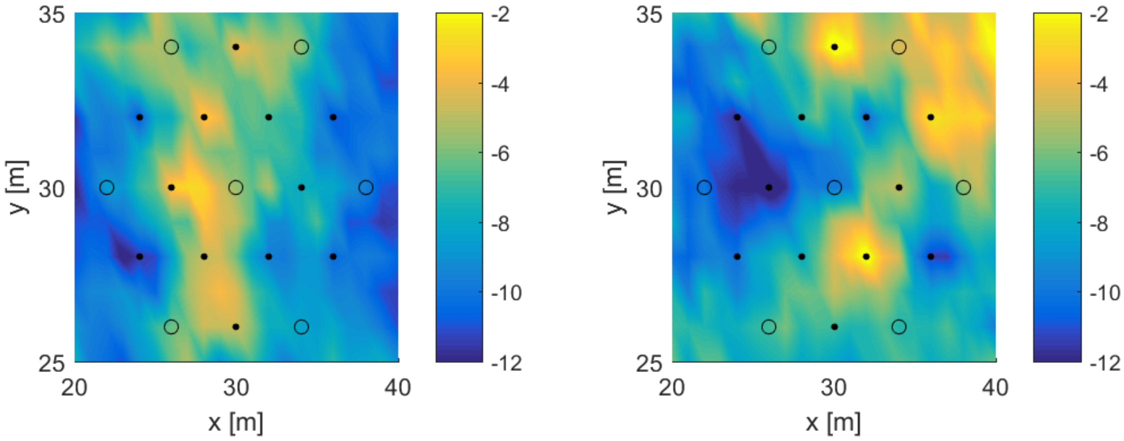

Figure 4.5.2: Best fields, Small PSO ranges test

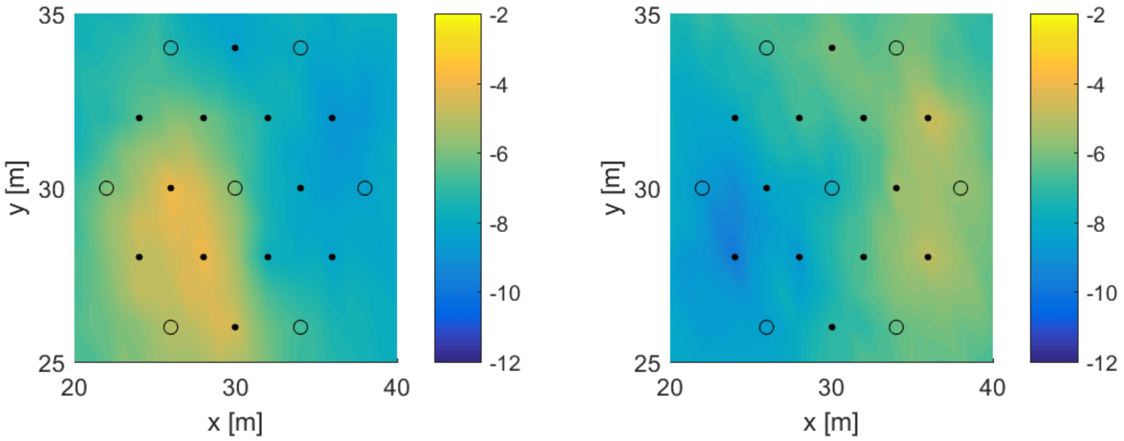

Figure 4.5.3: Mean fields, Small PSO ranges test

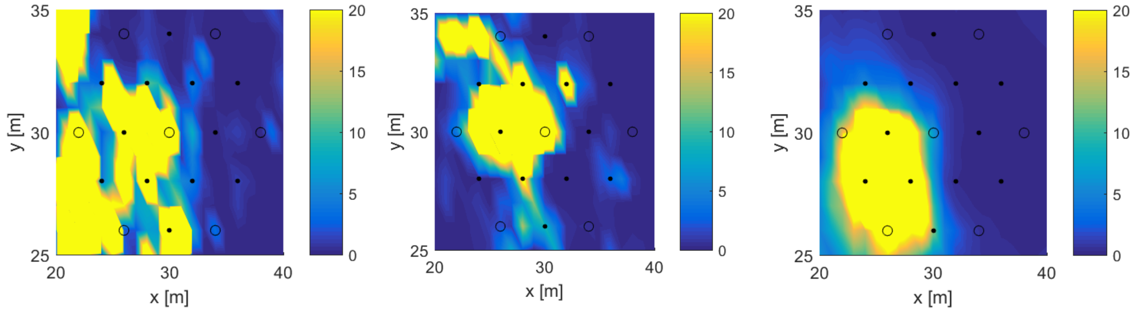

Figure 4.5.4: Diffusivity maps of the reference, best and mean fields, Small PSO ranges test

high values on the South while the mean simulation concentrates too much the high values on the southwestern corner.

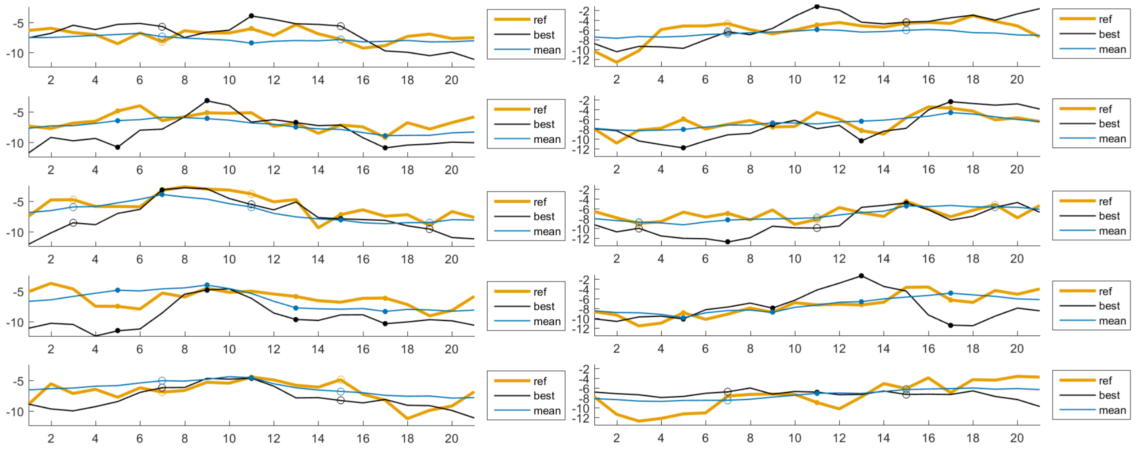

Figure 4.5.5: Transects of reference, best and mean fields in wells and power point rows, Y and Z fields, Small PSO ranges test

Figures 4.5.5 report log-properties' values along five transects. Observing the log-conductivity graph, the mean simulation does not oscillate but it goes up and down following the general trend of the reference. This is particularly evident on the second and third transects. The best simulation fluctuates more but it still follows the reference. However, it is noticeable how it well capture the central zone, while it always underestimates lateral conductivities. This underestimation was already observed on the maps. Moving to log-storativity transects, similar observations can be done for trends of best and mean simulations. There is not the lateral underestimation but it is noticeable a general undervaluation on the western part of the first three transects an on the eastern part of the last two, other sectors are reproduced correctly.

Concerning the probability distributions of properties, the log-conductivity pdf clearly does not find the reference curve, it strongly increases the probability of low-medium values. On the contrary the log-storativity pdf is very close to the reference one and it perfectly overlaps in lowest values. The diffusivity pdf is consequently wrong depending on the other two properties, this case the wrong log-conductivity pdf causes a wrong diffusivity pdf along all values. As in previous cases, the mean

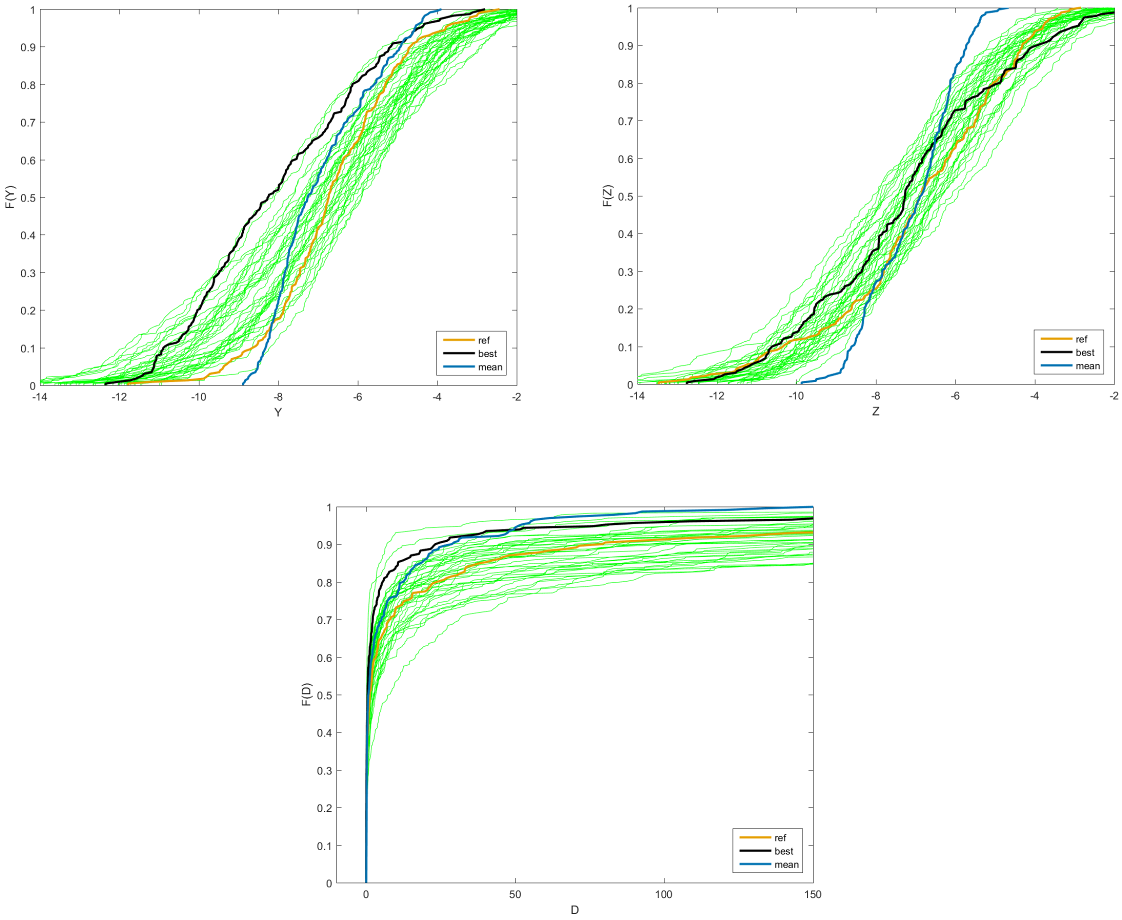

Figure 4.5.6: Empirical cumulative distribution functions of Y, Z and D, Small PSO ranges test

Table 4.13: Statistical parameters comparison, Small PSO ranges test

Field	m_Y	σ_Y^2	m_Z	σ_Z^2	$I_x(Y)$	$I_y(Y)$	$I_x(Z)$	$I_y(Z)$
Reference	-6.63	2.61	-6.95	4.92	0.79	2.18	2.42	3.28
Best	-8.02	4.54	-7.22	5.85	-0.22	4.18	3.05	2.13
Mean	-6.91	1.61	-7.07	1.20	3.17	4.20	2.80	5.03

simulation does not correctly infer the reference statistics mainly due to the exclusion of extreme values.

Table 4.9 clearly demonstrates that the best solution does not complies with the statistics of the central zone of interest. Indeed, mean and variance of the log-conductivity is strongly overvalued, while ones of the log-storativity are slightly overestimated. Regarding the integral scales, ones of Y are much greater than the reference except for an unusual negative value of the best reference that expresses the difficulty of the best simulation to infer the log-conductivity field. On the contrary, the inferred log-storativity integral scales are quite good.

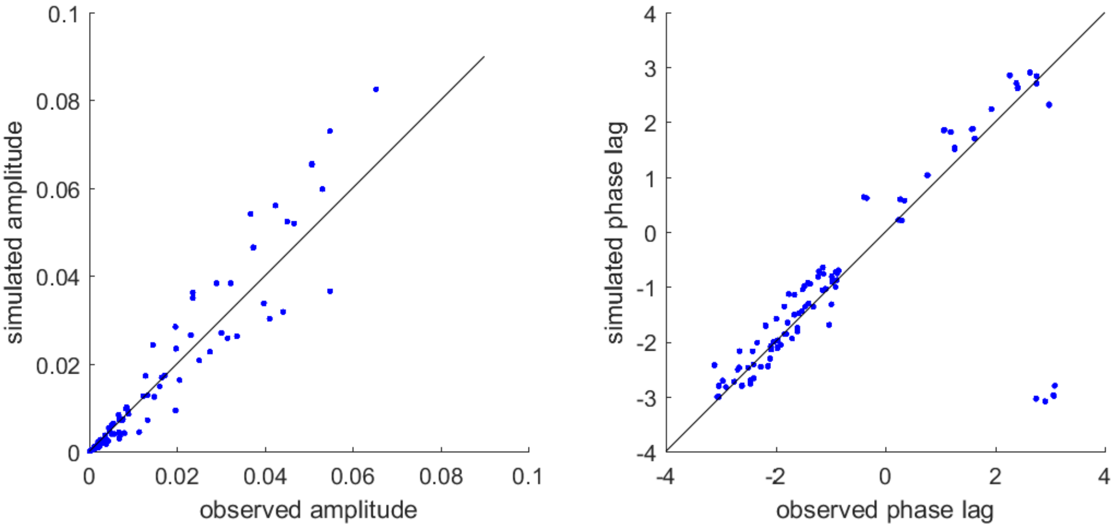

Figure 4.5.7: Scatter plots of simulated versus observed values of amplitude and phase, Small PSO ranges test

Scatter plots of the amplitude and phase of the signal show how much the simulation was able to infer observations. In figures 4.5.7, amplitude points are close to the diagonal but quite spread. On the contrary, phase points are more concentrated around the diagonal except for points completely out in which best simulation strongly underestimates observations. They will probably reduce drastically the efficiency. These extern points are all tests of the seventh well (the eastern one), this well is located in a low diffusivity area, therefore the signal does not travel easily and probably it is the cause of this error.

Table 4.14: Mean absolute errors of fields and efficiency coefficients, Small PSO ranges test

(a)

Field	Mean($\|\Delta Y\|$)	Mean($\|\Delta Z\|$)
Best	2.1255	2.1231
Mean	1.1808	1.3124

(b)

	NSEE
signal	0.59
amplitude	0.86
phase	0.44

Table 4.14(b) shows that the overall signal efficiency is quite good, with a modified NashSutcliffe efficiency of 0.59. The efficiency of the signal amplitude is high, but the phase one is low. This is probably due to the seventh well, considering the four outside points that were noticed in figure 4.5.7.

Figure 4.5.8 reports breakthrough curves of simulations. It is evident how the best simulation gets a wrong curve which particles are much slower than the reference field. Indeed, faster particles arrive at 40 m in 400 d, when also the slowest particles of the reference already arrived. In this case, the mean simulation is much closer to the reference even if it is much more uniform on arrival times and it is almost vertical, as usual.

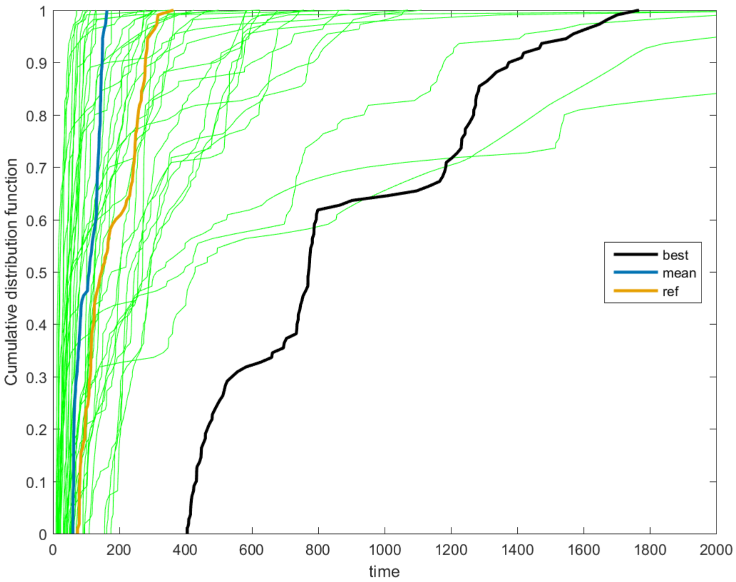

Figure 4.5.8: Breakthrough curves of realizations, Small PSO ranges test

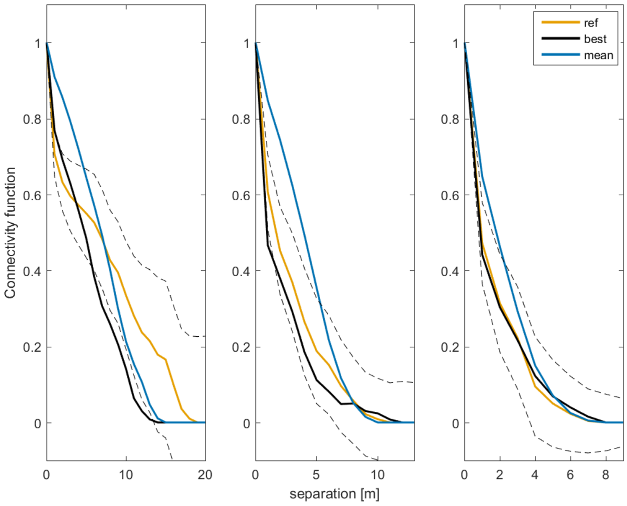

Figure 4.5.9: Connectivity functions (50, 75, 90 percentile), Small PSO ranges test

Table 4.15: Connectivity indicator CI, Small PSO ranges test

	CI
reference	2.19
best	2.09
mean	1.74

Results of the connectivity suggest that the best simulation is able to well infer this property. Indeed, in figure 4.5.9, the curve of the best simulation underestimates the connectivity for the lowest threshold of 50 percentile, but it is not so far from the reference curve. For higher thresholds the fit improves reaching almost the complete overlapping for the 90 percentile. Also the mean curve improves with higher thresholds but it is not as good as the best one. This good fit is confirmed by the connectivity indexes reported in table 4.15, where the best simulation has an index very close to the reference one while the mean is lower but still quite close.

4.5.2 Results and observations for 'Large ranges'

At a first sight, maps obtained in Large PSO ranges tests are better than ones in Small PSO ranges one. Indeed, observing the best log-conductivity field in figure 4.5.10, the low-conductivity belt is recognized in the eastern side, while medium values are inferred on the West. The central high Y zone is localized, even if it wrongly spreads on the South. The log-storativity field is even better, it recognizes the western low Z belt and the eastern medium values, where two low-values spots are correctly identified. As usual, mean fields well identifies patterns but they are not able to see extreme values.

Diffusivity fields of both the best and the mean simulations correctly infer a high-value zone in the southwestern corner even if in the reference field this area is less compacted and more scattered. Moreover, the best simulation wrongly identifies a high-diffusivity spot on the North. However, inferences are globally good.

Figures 4.5.13 report the log-properties values along five transects. As for previous tests, the mean simulation has smooth trends that follow reference ones. On the contrary, the best simulation is more oscillating. Concentrating on the log-conductivity graph, the best solution approximatively follows the reference trend. In the first transect it almost overlaps the reference simulation, while in others it is less precise. Mainly in the last two transects there is a relevant overestimation between 4 m and 8 m that coincides with the overvalued southern area noticed in figures 4.5.10. Observing log-storativity transects, the best solution follows the reference trend, usually recognizing ups and downs but exaggerating them sometimes.

Figures 4.5.14 display good inferences of the probability distributions of the properties. Looking at the log-conductivity function, the best simulation perfectly over-

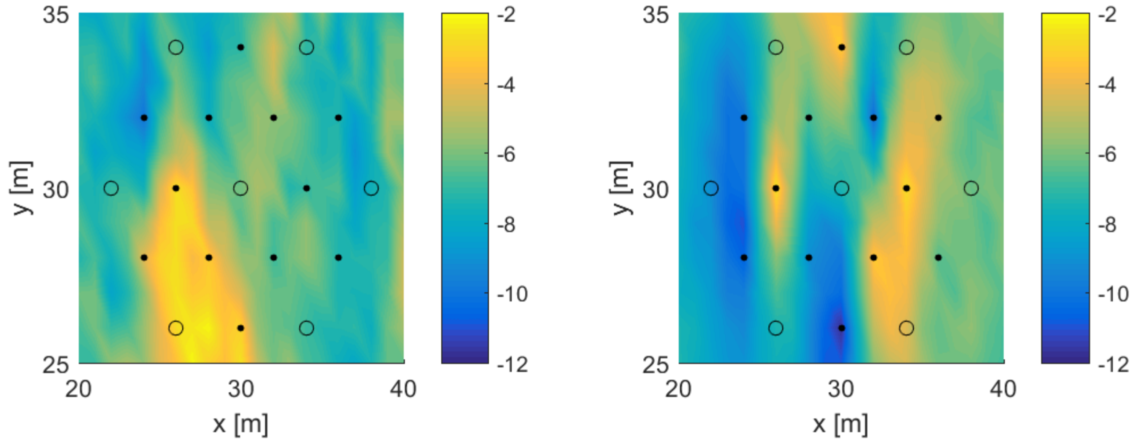

Figure 4.5.10: Best fields, Large PSO ranges test

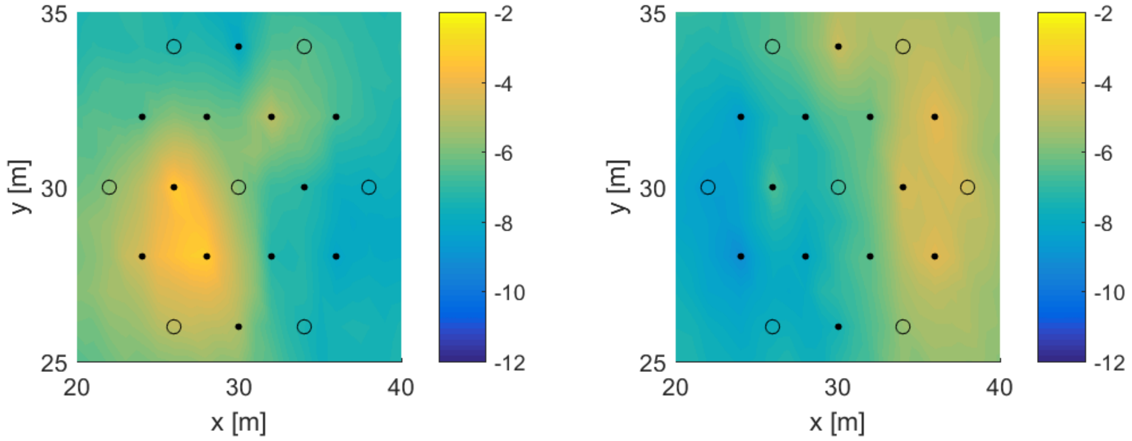

Figure 4.5.11: Mean fields, Large PSO ranges test

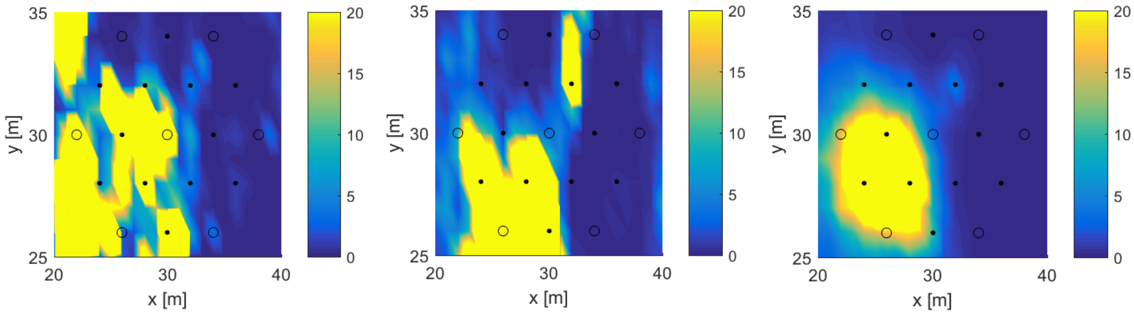

Figure 4.5.12: Diffusivity maps of the reference, best and mean fields, Large PSO ranges test

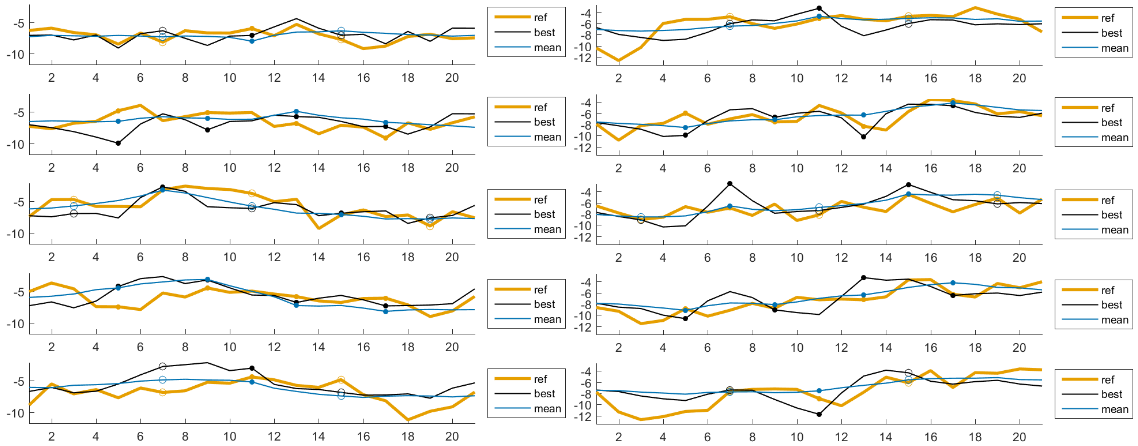

Figure 4.5.13: Transects of reference, best and mean fields in wells and power point rows, Y and Z fields, Large PSO ranges test

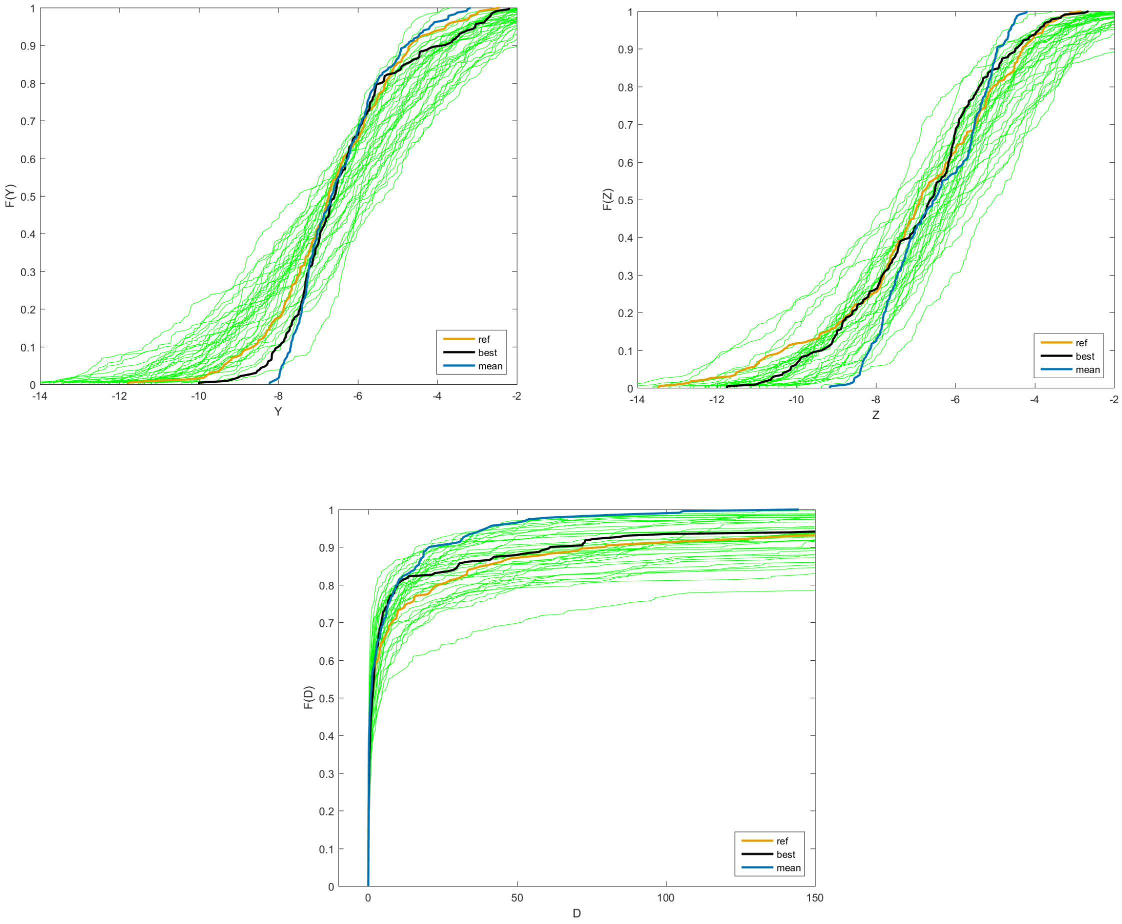

Figure 4.5.14: Empirical cumulative distribution functions of Y, Z and D, Large PSO ranges test

Table 4.16: Statistical parameters comparison, Large PSO ranges test

Field	m_Y	σ_Y^2	m_Z	σ_Z^2	$I_x(Y)$	$I_y(Y)$	$I_x(Z)$	$I_y(Z)$
Reference	-6.63	2.61	-6.95	4.92	0.79	2.18	2.42	3.28
Best	-6.33	2.29	-6.85	3.38	1.34	3.04	1.28	4.09
Mean	-6.42	1.31	-6.48	1.58	2.39	3.29	3.06	5.46

laps the reference pdf for medium values while it underestimates the probability of highest and lowest values. Moreover, for this property, the mean pdf is also good, inasmuch it overlaps the reference pdf for medium and high values, even if it gets a completely wrong estimation of lowest values. Even in the log-storativity distributions, the best pdf overlaps the reference one for medium-high values while it still underestimates lowest values. The mean pdf is more steep than others pdfs and it is not able to infer correctly extreme values, mainly lowest ones. Due to the good best inferences of Y and Z, the diffusivity distribution of the best simulation is very close to the reference one, the bigger error is for lowest diffusivity values.

Table 4.16 confirms that the best simulation well infers reference statistics, as already noticed in the distributions figures 4.5.14. In this table it is possible to notice how both means are close to the reference one and variances are much closer than previous tests. Also the integral scales are slightly better than before, even if there are relative errors around 50%. For the mean simulation, same observations can be done, except for the variances which are strongly underestimated.

Scatter plots of the amplitude and phase of the signal in figures 4.5.15 show how the best simulation generally underestimates the amplitude of observations while it correctly estimates the phase except for two points far from the diagonal due to the sixth well, located in a low diffusivity zone that causes a poor signal transmission.

Table 4.17(a) reports mean absolute errors that are smaller than errors in the Small PSO ranges test even if they are not small. Furthermore, the overall efficiency is quite good, with a modified NS efficiency of 0.69. Both amplitude and phase NS efficiencies are higher, noticing that the phase efficiency decreases mainly due to the sixth well, as mentioned in the scatter plots comment.

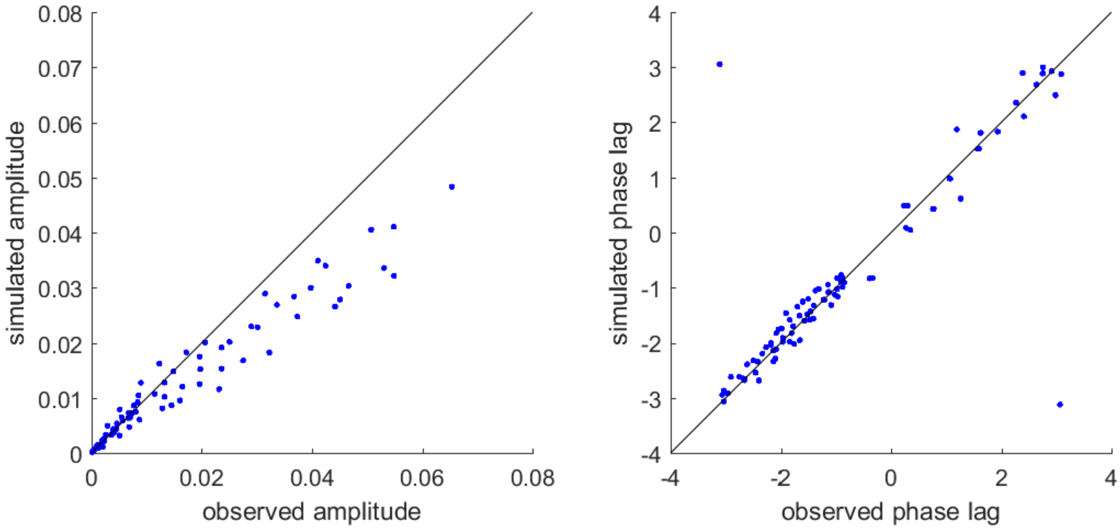

Figure 4.5.15: Scatter plots of simulated versus observed values of amplitude and phase, Large PSO ranges test

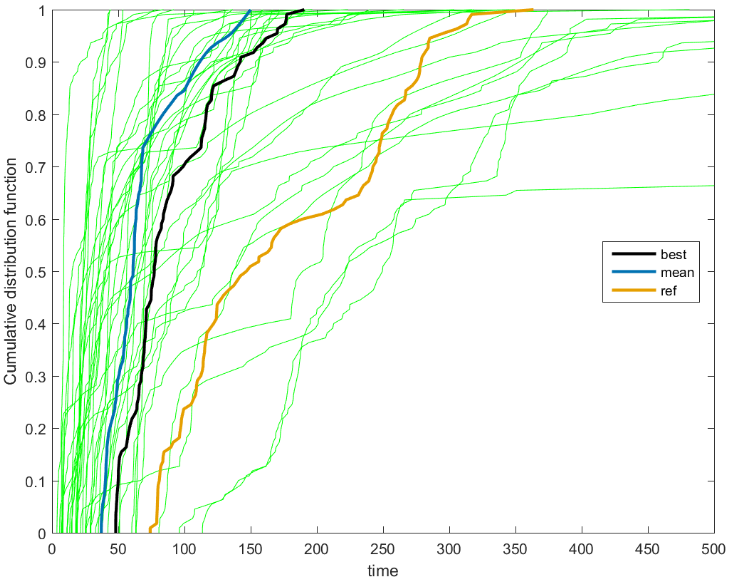

Figure 4.5.16: Breakthrough curves of realizations, Large PSO ranges test

Table 4.17: Mean absolute errors of fields and efficiency coefficients, Large PSO ranges test

(a)

Field	Mean($\|\Delta Y\|$)	Mean($\|\Delta Z\|$)
Best	1.505565	1.661928
Mean	1.17339	1.2663

(b)

	NSEE
signal	0.69
amplitude	0.83
phase	0.70

Breakthrough curves of the several simulations reported in figure 4.5.16 are scattered along a big range of time. The majority is steep at first and then it has a long tail due to slowest particles. The curve of the best simulation has a trend similar to the reference one but it is much faster, indeed fastest particles arrive at the end of the central zoom of interest after 50 d (75 d for the reference ones) and slowest after almost 200 d (more than 350 d for the reference ones). The curve of the average of all simulations has the same trend but it is even faster. Even if previous results of this test are good, the breakthrough curve is not yet fully captured.

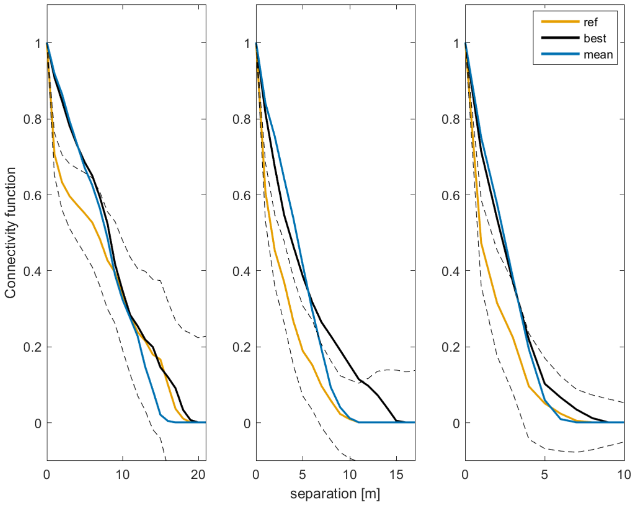

Figure 4.5.17: Connectivity functions (50, 75, 90 percentile), Large PSO ranges test

Table 4.18: Connectivity indicator CI, Large PSO ranges test

	CI
reference	2.19
best	1.82
mean	1.77

Finally looking at the connectivity parameter, the best simulation is able to reproduce the connectivity function of the reference fields. Indeed, best connectivity function is very close to the reference curve for lowest thresholds while the fit is slightly worst for the highest threshold (90 percentile). The mean simulation approximatively fits the reference curve even if it tends to overestimate it. This approximatively good fit of the best simulation is confirmed by the connectivity index in table 4.18, while the index of the mean simulation is slightly lower and farer from the reference index.

4.5.3 Observations

The reproduction of a heterogeneous field for both log-conductivity and log-storativity was here done in two tests characterized by different ranges where particles can search on the PSO algorithm.

The first one is characterized by small search ranges around the means of the reference. This case could apply in reality if there is previous information about the field. Looking at obtained results in general, it is evident how this test with a decent efficiency of 0.59 obtains good but not precise maps of properties' heterogeneity and an incorrect log-conductivity pdf and statistics. On the contrary, pdf and statistics of the log-storativity field are reproduced relatively well. Regarding the contaminant transport results, the reference breakthrough curve is completely misunderstood by the best simulation, while the connectivity function of the best solution is quite close to the reference one, especially for higher thresholds.
The latter test is characterized by large ranges of the PSO optimization. Observing all results, there is a general improvement in the reproduction. First of all, heterogeneity maps of the best simulation are very close to the reference ones looking both at values and at patterns. Secondly, probability distributions are well inferred except for lowest values which probability is underestimated for both properties. Still, the transport results are the worst: breakthrough curves are not yet reproduced correctly even if they are better than ones of the previous test, while the connectivity function of the best solution is close but not overlapping the reference one.
As in the previous section of homogeneous properties, the best simulation returns better results than the average of all realizations. Indeed, it is better on identifying extreme values, therefore heterogeneity, pdfs and statistics.

It is possible to conclude from this section that the time-consuming Large PSO ranges test returns better results on heterogeneity inference even if it is still not able to capture the particle transport behavior of the reference field. This better solution of the second test was unexpected but, looking for the cause, it was noticed the best field has high correct variances of the central area of interest but low incorrect variances of the whole field. This probably means that PSO ranges for variances has to be set large enough to allow the inferred field to adapt to observations in the central area, completely misunderstanding values in external areas not of interest.

4.6 False positives and negatives

The aim of this section is to assess the accuracy and the effectiveness of the methodology in case distributions of log-conductivity or log-storativity are incorrectly assumed.

First case considers both Y and Z true fields characterized by a Gaussian distribution, however in the inversion it is assumed only the latter as a stochastic process while the former is assumed spatially constant (KNeg). The second case is similar to the previous one, but Z is the property considered homogeneous in the entire field (SNeg). The third case considers a true field where Y is a stochastic process while Z is homogeneous, however inversion assumes both properties as stochastic processes (KPos). The last case is similar to the previous one, but the Y field is wrongly assumed: the reference field is homogeneous, while it is inverted as heterogeneous (SPos). A summary table is reported following to show true parameters and ranges used.

4.6.1 Results and observations for 'K false negative'

Figures 4.6.1 and 4.6.2 show the inferred fields of best and mean simulations. The reference fields are not here reinserted because they were already reported in figures 4.5.1, the difference is that here the conductivity field is assumed homogeneous and therefore it does not show any similarity on patterns and the value is close to the average of the field. Neglecting this field that is obviously wrong passing from a reference field with a high variance of 4 to a homogeneous field (variance equal to 0), the focus of the discussion in this test is the inference of the log-storativity field.

Table 4.19: Parameters and ranges for false positive and negative tests

		$mean(Y)$	$var(Y)$	$mean(Z)$	$var(Z)$	I_x	I_y
KNeg	true parameters	-8	4	-8	4	3	5
	range	-9.50	0.00	-9.50	3.00	1.00	1.00
		-6.50	0.00	-6.50	5.00	7.00	7.00
SNeg	true parameters	-8	4	-8	4	3	5
	range	-9.50	3.00	-9.50	0.00	1.00	1.00
		-6.50	5.00	-6.50	0.00	7.00	7.00
KPos	true parameters	-8	0	-8	4	3	5
	range	-9.50	0.00	-9.50	3.00	1.00	1.00
		-6.50	6.00	-6.50	5.00	7.00	7.00
SPos	true parameters	-8	4	-8	0	3	5
	range	-9.50	3.00	-9.50	0.00	1.00	1.00
		-6.50	5.00	-6.50	6.00	7.00	7.00

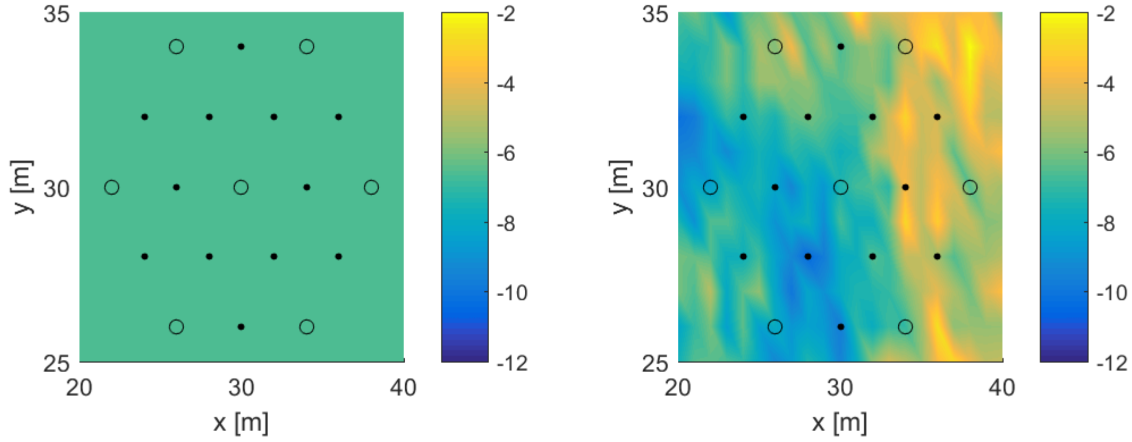

Figure 4.6.1: Best fields, K-false negative test

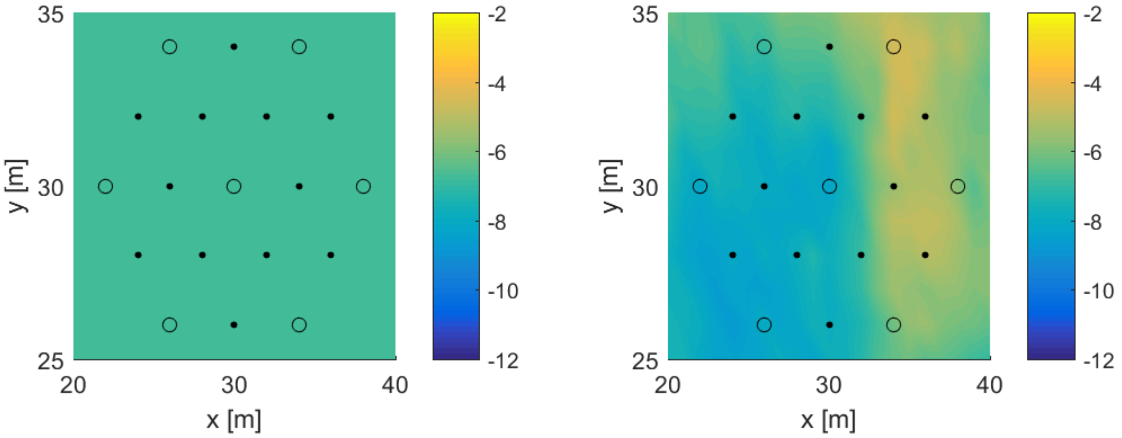

Figure 4.6.2: Mean fields, K-false negative test

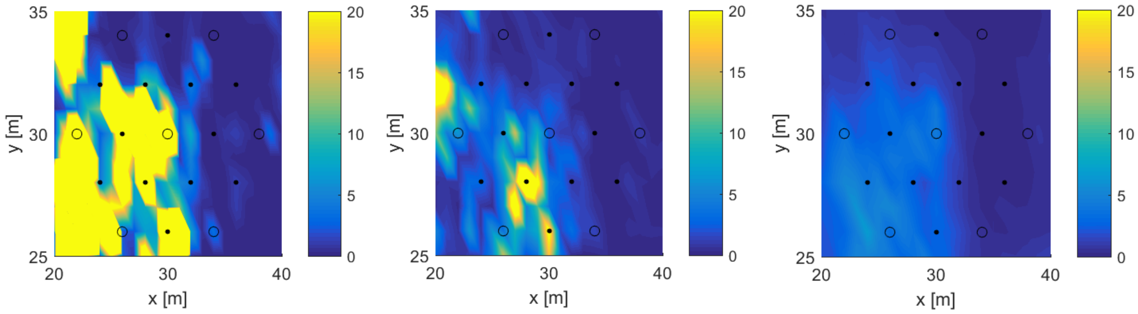

Figure 4.6.3: Diffusivity maps of the reference, best and mean fields, K-false negative test

The field is very similar to the reference field, the eastern area is correctly identified as a medium-high storativity area with a branch on the North. However, the best solution misses to identifies the two corners with lowest storativity. As for previous tests, the mean simulation succeeds on the pattern reproduction but it is not able to recognize extreme values.

Figures 4.6.3 show how best and mean simulations are able to differentiate low and medium diffusivity areas. However, they completely miss to reproduce highest values of diffusivity. This of course is due to the wrong assumption of a homogeneous conductivity fields that leads to not consider high values on the western area.

Observing transects in figure 4.6.4, clearly the log-conductivity figure shows the

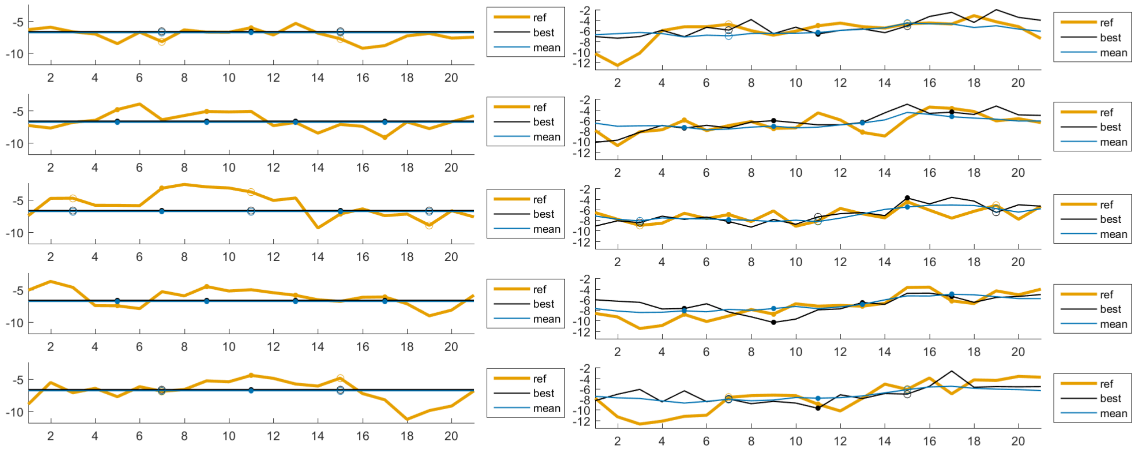

Figure 4.6.4: Transects of reference, best and mean fields in wells and power point rows, Y and Z fields, K-false negative simulation

homogeneous inferred simulations that do not reproduce the real heterogeneity. In the log-storativity figure, instead, the best simulation well follows the reference field in its oscillating tendency. However, in this figure it is noticeable what was already evident on the map 4.6.1, on the left of first and last two transects there is s strong overestimation of Z values. This way two spots of low values are completely missed. Concerning the mean simulation, it follows quite precisely the reference trend but it is flattered and it does not reproduces smaller oscillations.

The probability distributions of the properties in figures 4.6.5 display how wrong is the inference of the log-conductivity due to the incorrect original assumption. On the contrary, it reports quite a good reproduction of the log-storativity pdf, even if it strongly understates the probability of lowest values. The mean curve overlaps to the reference pdf for medium values but it misses extreme storativities. Mainly due to the incorrect inference of the conductivity, the diffusivity distribution is also incorrect. It does not consider highest values: the reference diffusivity field reaches values of about 300 m^2/s while the best simulation does not even reach 50 m^2/s. The main diffusivity field is very flattered and consequently the curve is almost vertical without upper tail.

Table 4.20 reports all inferred statistics. It is evident from this table that the best simulation succeeds on estimating the log-conductivity average but, clearly, it eval-

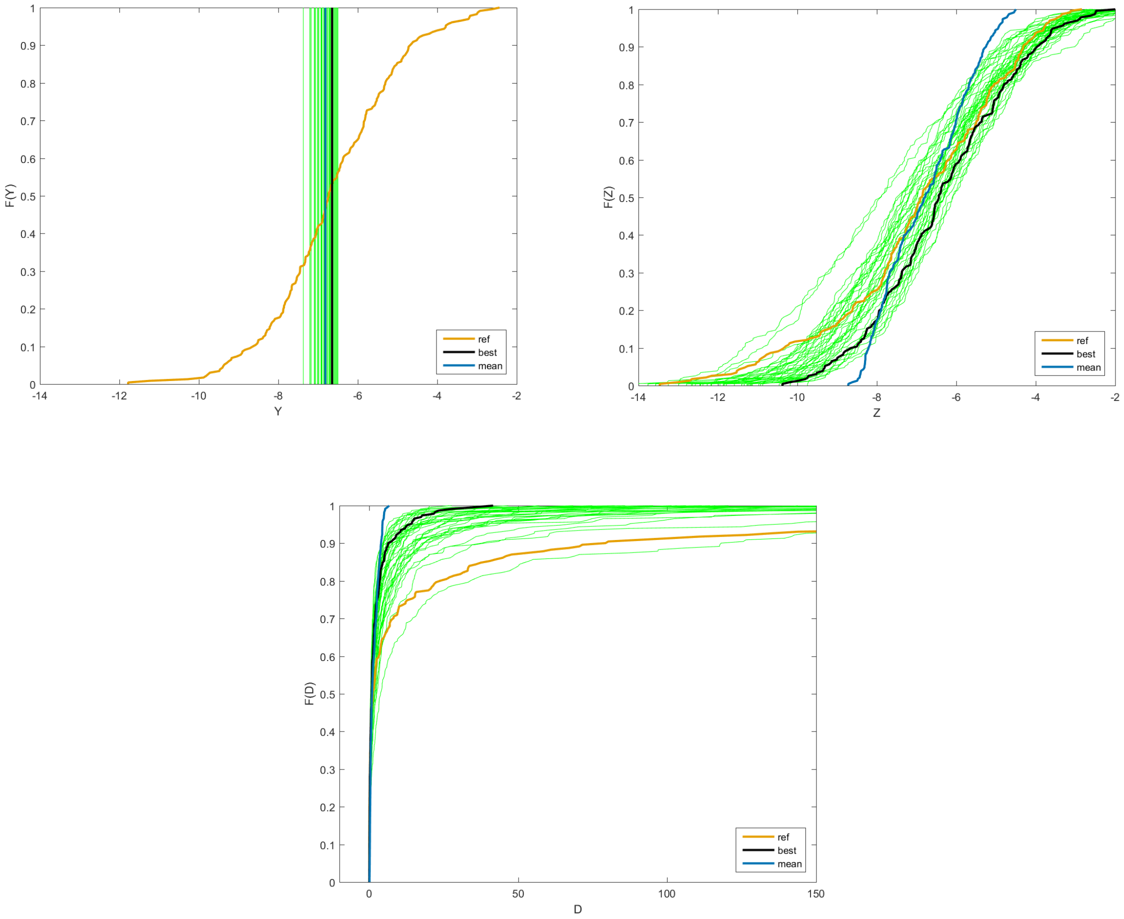

Figure 4.6.5: Empirical cumulative distribution functions of Y, Z and D, K-false negative test

Table 4.20: Statistical parameters comparison, K-false negative test

Field	m_Y	σ_Y^2	m_Z	σ_Z^2	$I_x(Y)$	$I_y(Y)$	$I_x(Z)$	$I_y(Z)$
Reference	-6.63	2.61	-6.95	4.92	0.79	2.18	2.42	3.28
Best	-6.65	0.00	-6.33	3.09			2.46	3.59
Mean	-6.82	0.00	-6.76	1.23			3.23	4.84

uates a wrong variance equal to 0. On the other hand, looking at the log-storativity statistics, the mean of the best simulation is enough close to the reference mean of the central area and also the variance has a decent inference. Instead, surprising are the simulated integral scales that in previous tests were not well inferred while here they are very close to the reference ones. Observing the statistics of the mean simulation, means are well inferred but variances are underestimated, the Y variance for assumption and the Z variance for a faulty inference.

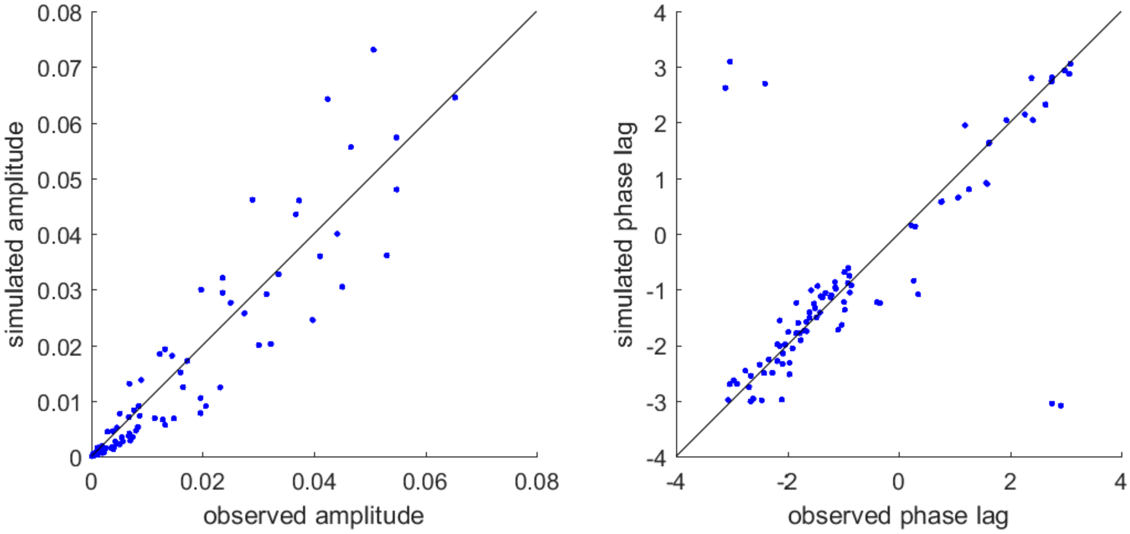

Figure 4.6.6: Scatter plots of simulated versus observed values of amplitude and phase, K-false negative test

Scatter plots in figures 4.6.6 illustrate how the periodic signal is simulated by the best solution considering its two components: amplitude and phase. The amplitude is rather diffused with tests both underestimated and overestimated, while the phase is more concentrate along the diagonal except for some outsiders due to sixth and seventh wells. These points probably cause the low phase efficiency (reported in table 4.21(b)) that strongly reduces the overall signal efficiency.

Table 4.21(a) lists mean absolute errors that are rather high, especially for the log-storativity field. Indeed, the log-conductivity field is inferred as a homogeneous field with an average value of the reference field, therefore errors are lower in average. While for the heterogeneous Z field it is easier to have big discrepancies.

Table 4.21: Mean absolute errors of fields and efficiency coefficients, K-false negative test

(a)

Field	Mean($\|\Delta Y\|$)	Mean($\|\Delta Z\|$)
Best	1.2669	1.6093
Mean	1.2672	1.2922

(b)

	NSEE
signal	0.50
amplitude	0.83
phase	0.34

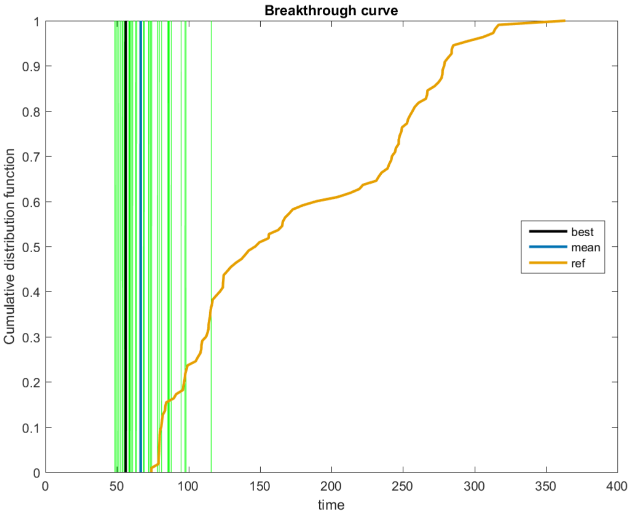

Figure 4.6.7: Breakthrough curves of realizations, K-false negative test

In this test, breakthrough curves are not a useful result because they strongly depend on the conductivity field. Therefore, incorrectly assuming a homogeneous field leads to have a completely different type of curves, as illustrated in figure 4.6.7. Indeed, homogeneous fields result in a uniform movement of particles that arrive all at the same time. On the contrary, a heterogeneous field is characterized by preferential pathways where particles are much faster than outside this routes.

As for breakthrough curves, connectivity depends mainly on the conductivity field, that here is incorrectly assumed. Indeed, looking at the connectivity function in figure 4.6.8, both best and mean simulations are not able to reproduce the reference connectivity function. This incorrect reproduction is evident also in indexes reported in table 4.22.

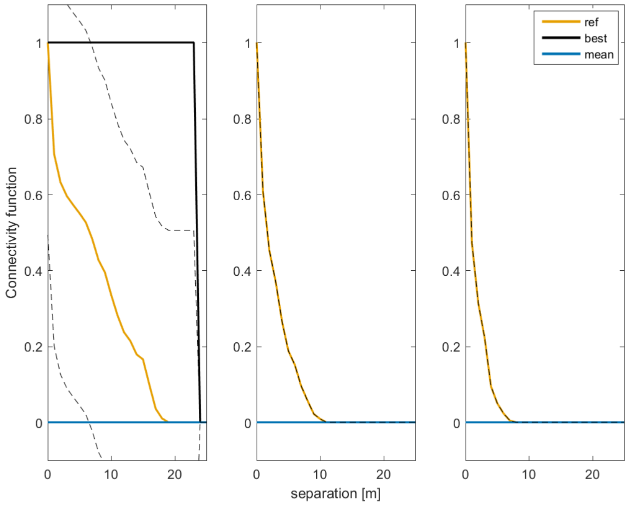

Figure 4.6.8: Connectivity functions (50, 75, 90 percentile), K-false negative test

Table 4.22: Connectivity indicator CI, K-false negative test

	CI
reference	2.19
best	1.00
mean	1.00

4.6.2 Results and observations for 'S false negative'

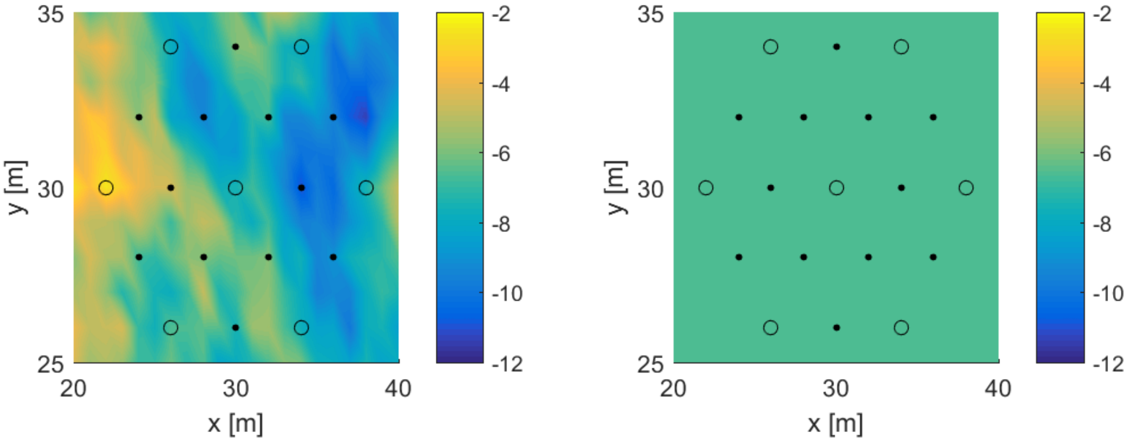

Figure 4.6.9: Best fields, S-false negative test

Figures 4.6.9 and 4.6.10 display the inference of reference fields 4.5.1 obtained by

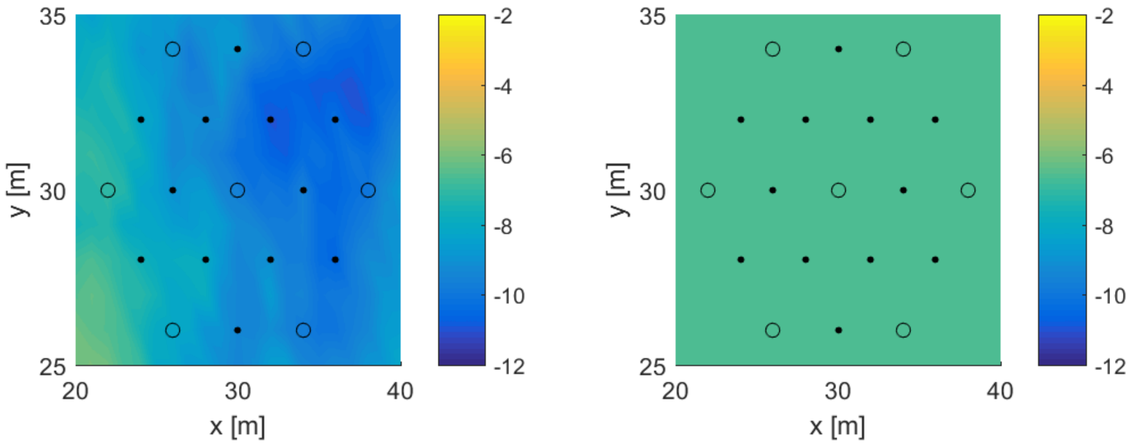

Figure 4.6.10: Mean fields, S-false negative test

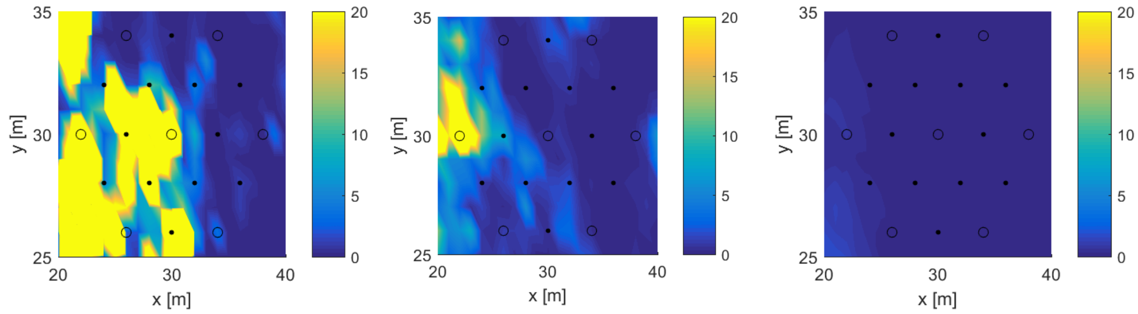

Figure 4.6.11: Diffusivity maps of the reference, best and mean fields, S-false negative test

best and mean simulations incorrectly assuming that the log-storativity field is homogeneous. The best simulation is able to approximatively infer the log-conductivity field, indeed it is reproduced the eastern low-values belt and the medium-high conductivity area on the West. The main incorrectness is the movement to the left of the spot with highest values. On the contrary, the mean simulation barely notices the western area with higher values.

The reference diffusivity field in figure 4.6.11 is reproduced rather well by the best simulation that correctly infers the pattern but it is not able to see highest values. This is due to the assumed homogeneous field of the log-storativity field that does not allow to see areas with lowest values. Furthermore, looking at the field inferred by the mean simulation, it does not even infer the diffusivity pattern.

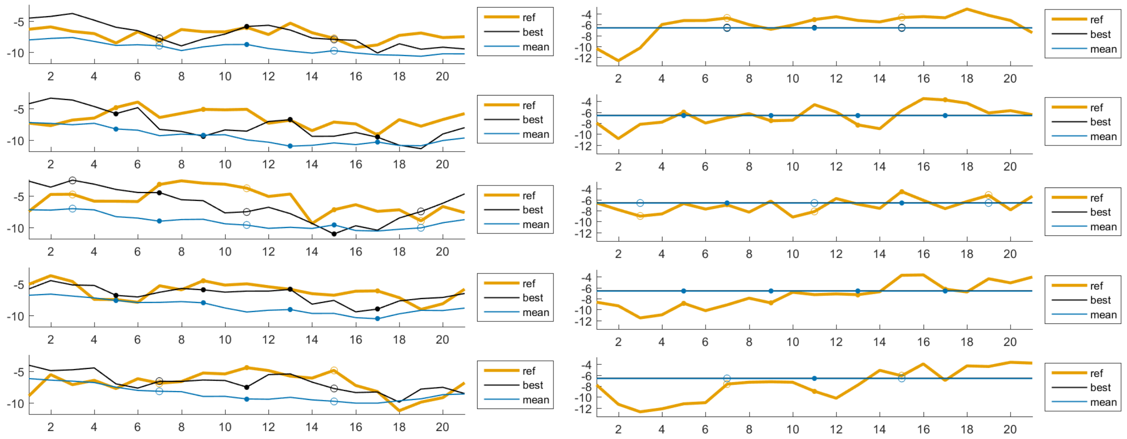

Figure 4.6.12: Transects of reference, best and mean fields in wells and power point rows, Y and Z fields, S-false negative test

Observing the five transects of the log-conductivity in figure 4.6.12, it is evident how the best simulation is partially able to follow the reference curve. Mainly the second and third transects are incorrectly inferred. Moreover, it can be noticed in these two transects how the area of highest values between 6 m and 12 m is moved to the left. The mean simulation is flattered, as in all tests, but in this case it is always underestimating reference values. Regarding log-storativity transects, best and mean simulations are perfectly overlapping and flat due to the initial assumption.

From figures 4.6.13, it is evident how few are the simulations respect to other tests. This is due to the fact that in the post-processing only simulations with positive efficiency are considered, in this test only 14 out of 40 simulations are positive. The log-conductivity pdf of the best simulation is actually the closest pdf to the reference one. They overlap for highest values, while the best simulation overestimates the probability of lowest values. The pdf of the mean simulation is incorrect, it overstates probability of all values and it does not consider highest values. Observing the log-storativity pdf, same observation done for the Y pdfs of the K-false negative observations can be here replied. Indeed, the statistics of the Z field are completely wrong due to the incorrect initial assumption. Consequently, the diffusivity pdf is also incorrect because it strongly depends on both conductivity and storativity.

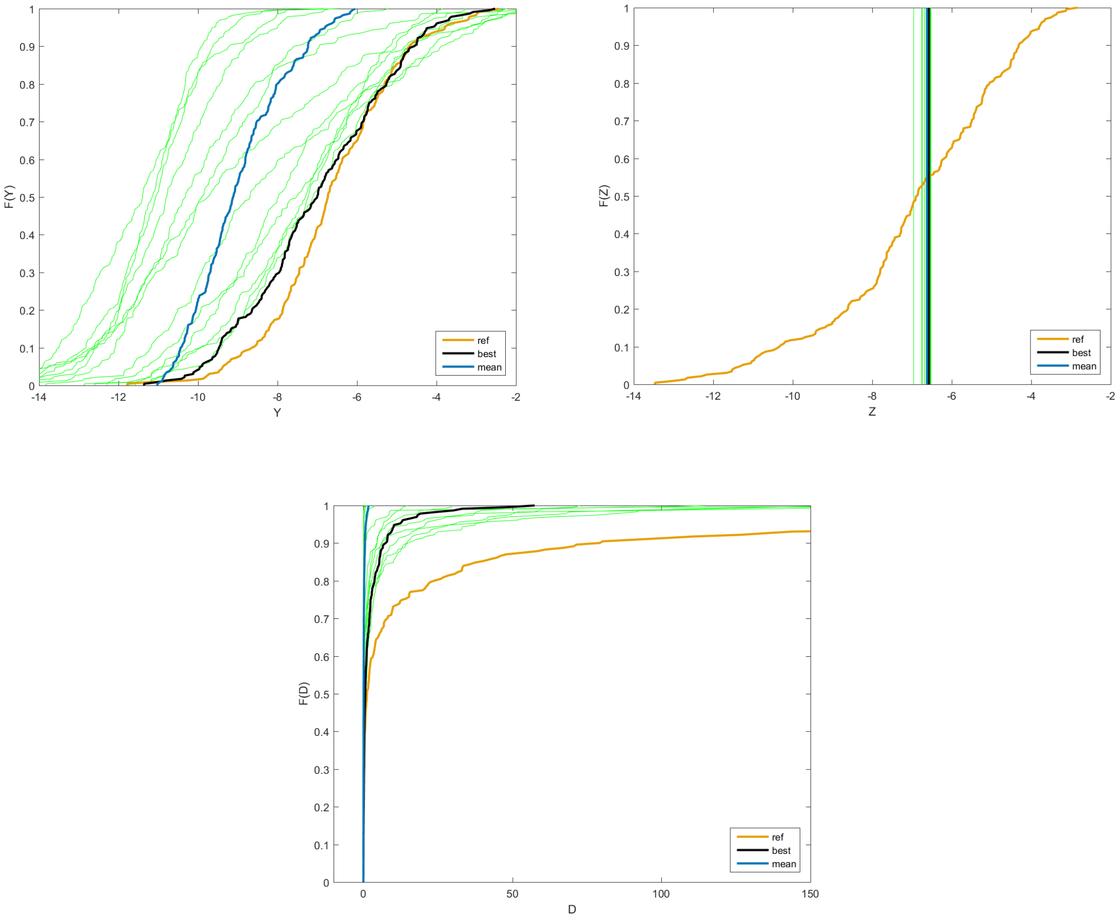

Figure 4.6.13: Empirical cumulative distribution functions of Y, Z and D, S-false negative test

Table 4.23: Statistical parameters comparison, S-false negative test

Field	m_Y	σ_Y^2	m_Z	σ_Z^2	$I_x(Y)$	$I_y(Y)$	$I_x(Z)$	$I_y(Z)$
Reference	-6.63	2.61	-6.95	4.92	0.79	2.18	2.42	3.28
Best	-6.99	3.32	-6.58	0.00	1.99	3.83		
Mean	-8.97	1.37	-6.62	0.00	2.99	4.95		

Table of statistic parameters 4.23 shows how much incorrect are statistics of the log-storativity field, obviously the variance but also the mean is not accurate for both best and mean simulations. For what regards the log-conductivity field, the best simulation approximate the mean with a relative error of 5% and the variance with a relative error of 30%. The mean simulation obtains still worst statistics completely misunderstanding the mean and underestimating the variance. Also the integral scales are not correctly inferred, they are strongly overestimated.

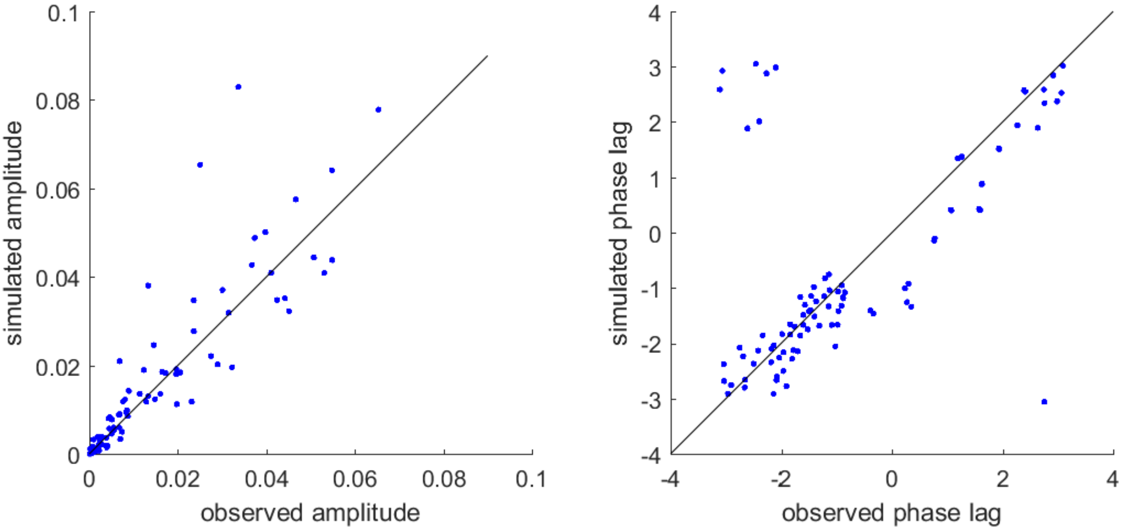

Figure 4.6.14: Scatter plots of simulated versus observed values of amplitude and phase, S-false negative test

As for the previous test, scatter plots in figure 4.6.14 display a scattered amplitude scatter plot, while the phase one is generally more concentrate along the diagonal with some external points mainly corresponding to sixth and seventh wells that are

Table 4.24: Mean absolute errors of fields and efficiency coefficients, S-false negative test

(a)

Field	Mean($\|\Delta Y\|$)	Mean($\|\Delta Z\|$)
Best	1.7083	1.7755
Mean	2.5619	1.7712

(b)

	NSEE
signal	0.31
amplitude	0.68
phase	0.09

located in low diffusivity areas.

Table 4.24(a) reports high mean absolute errors for both fields. Moreover, it is confirmed that the mean field is completely wrong on the Y field. The efficiencies table 4.24(b) shows a low overall signal efficiency of 0.31 mostly due to the very low phase efficiency.

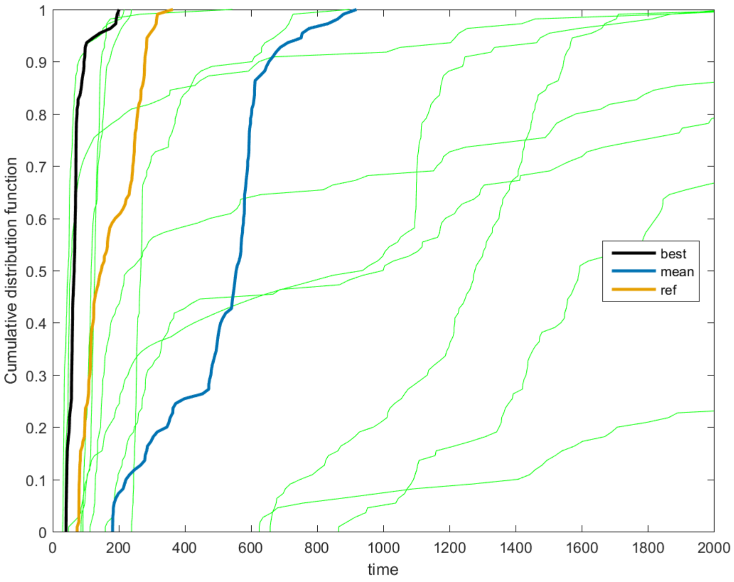

Figure 4.6.15: Breakthrough curves of realizations, S-false negative test

The majority of breakthrough curves inferred in this test are characterized by a very long tail but the best simulation has a curve close to the reference even if it does not follows its trend. The mean simulation results in a breakthrough curve characterized by slower particles. Slowest particles of the mean simulation arrive after about 900 d while the slowest ones of the reference field arrive after not even 400 d.

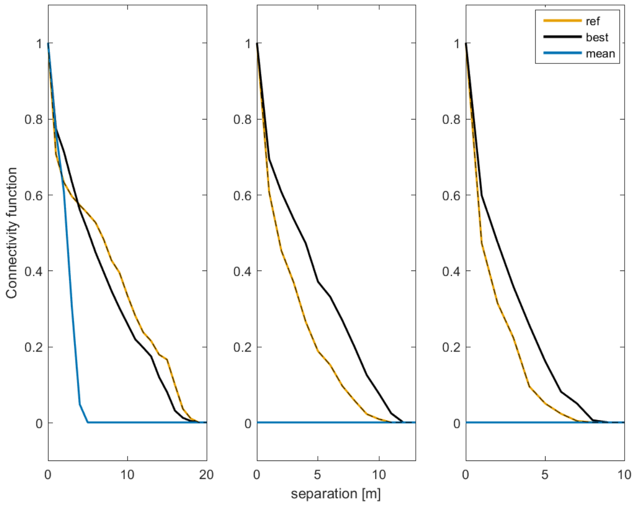

Figure 4.6.16: Connectivity functions (50, 75, 90 percentile), S-false negative test

Table 4.25: Connectivity indicator CI, S-false negative test

	CI
reference	2.19
best	1.71
mean	2.73

As for the breakthrough curves, results of the connectivity confirm the goodness of the best simulation to recognize fast pathways and the lacking reproduction of the mean simulation. This is clearly visible in figure 4.6.16 where the mean always underestimates the connectivity due to the general underestimation of the log-conductivity.

4.6.3 Results and observations for 'K false positive'

The last two tests aim to understand how the model works inferring a homogeneous field that is assumed heterogeneous in the PSO algorithm. In the first test, the log-conductivity field is homogeneous, as visible in figures 4.6.17. The best simulation in figures 4.6.18 imprecisely inferred fields. Indeed, the Y field is basically homogeneous with two spots of low values and two of high values. These spots surely influence statistical parameters, such as the variance. Observing the log-storativity field, it seems quite different because the two western corners are not identified as low values, while the southern spot is shifted more centrally. Moreover, the best

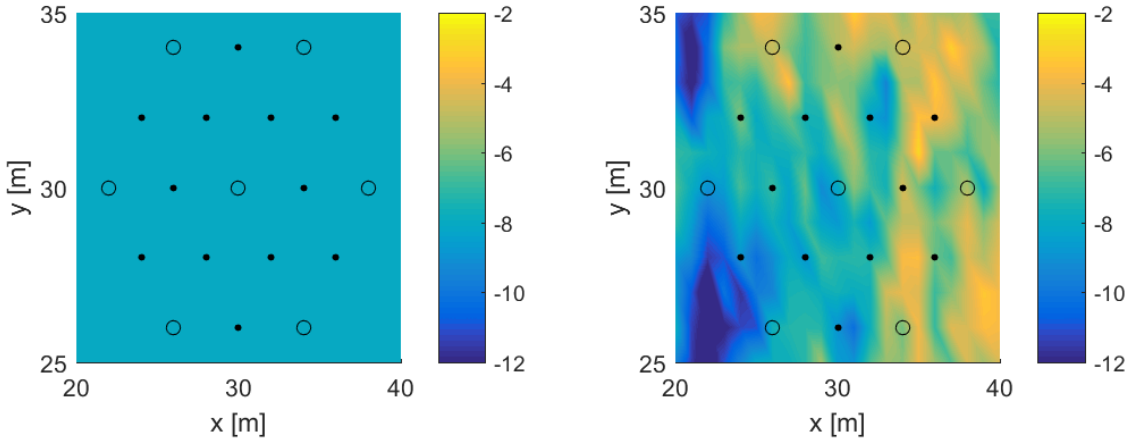

Figure 4.6.17: Reference fields, K-false positive test

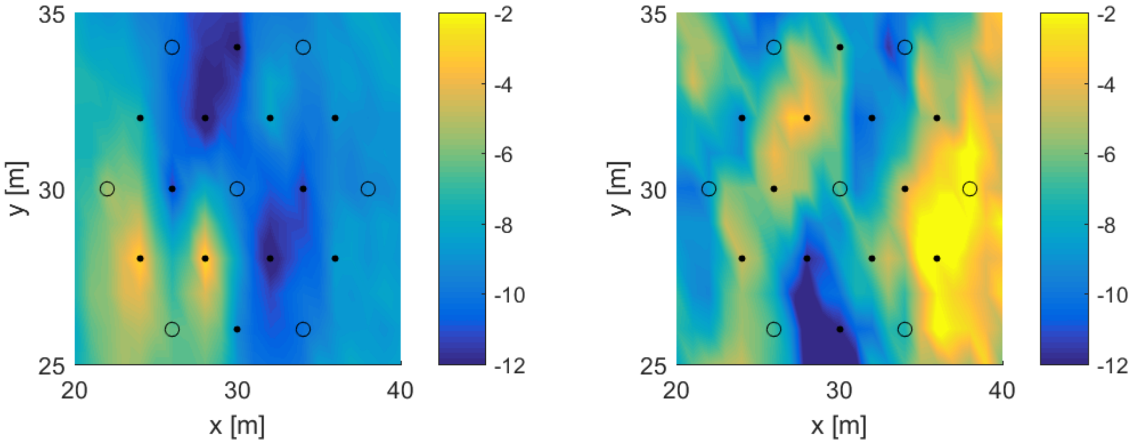

Figure 4.6.18: Best fields, K-false positive test

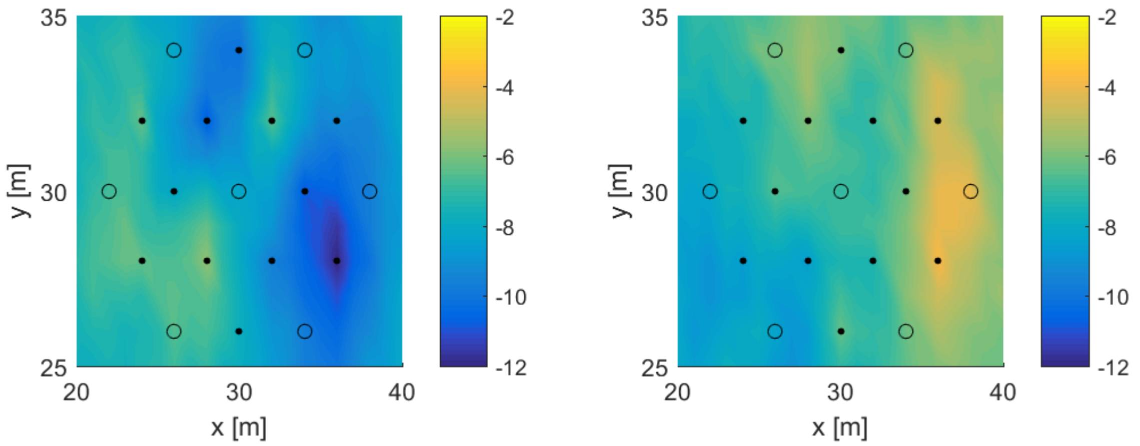

Figure 4.6.19: Mean fields, K-false positive test

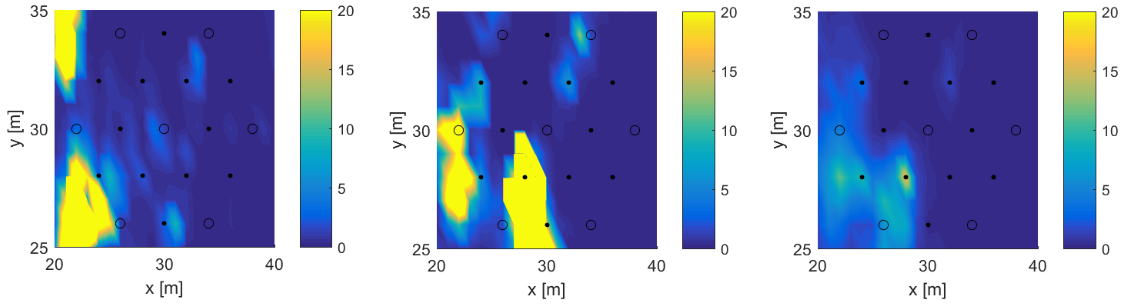

Figure 4.6.20: Diffusivity maps of the reference, best and mean fields, K-false positive test

simulation infers a zone of high values that is not present in the reference field. The mean simulation captures quite well the homogeneous Y field, except for an inferred low-values spot, and it is able to correctly infer the Z pattern even if it does not see lowest values.

The pattern of the reference diffusivity field is inferred by best and mean simulations but highest values are wrongly located by the best simulation or they are not placed at all by the mean simulation. Regarding the incorrect location of highest values for the best simulation, it is mainly due to the wrong placement of lowest values on the log-storativity field. While the missed identification of highest values for the mean simulation is still mainly due to the leaking inference of extreme low values in the Z field.

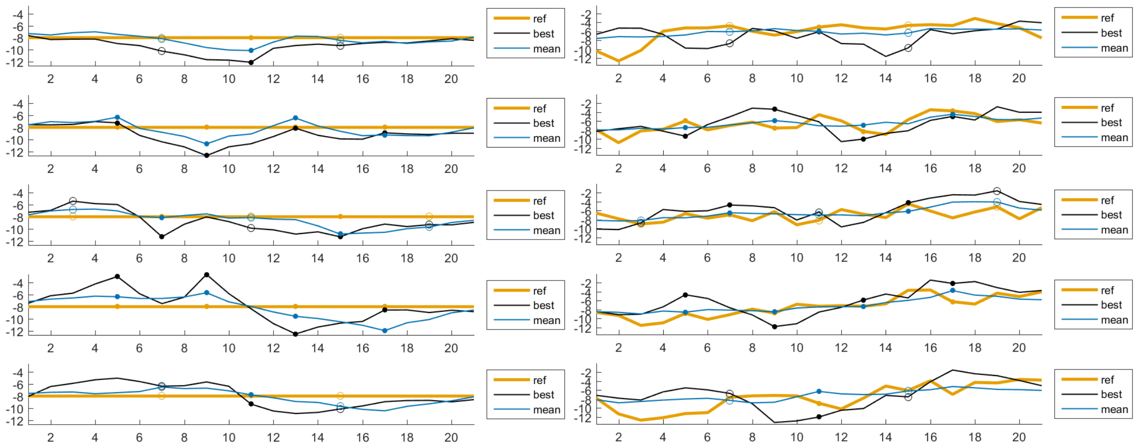

Figure 4.6.21: Transects of reference, best and mean fields in wells and power point rows, Y and Z fields, K-false positive test

From figures 4.6.21 reporting properties values along five transects, it is evident how simulations are not able to reproduce correctly the homogeneous log-conductivity field. Indeed, best simulation is really oscillating (see the fourth transect) and it understates Y values in the first three transects. The mean simulation is smoother than the best one but however it still reproduces the field as heterogeneous. Observing log-storativity transects, it is noticeable how in central transects best and mean simulations follow the trend of the reference field. On the contrary, in the first transect the best simulation always understates the reference curve except for the low-values spot on the left, while the mean simulation is almost flat. The last transect is incorrectly inferred by the best simulation that trades places of low and high-values spots, while the mean simulation does not even see these spots.

In order to understand the statistical reproduction of simulation, probability distributions are now compared. As expected, simulations are not able to correctly reproduce Y pdf because the reference field is homogeneous and therefore its function is vertical, while simulations are Gaussian fields with higher variances and therefore characterized also by extreme values. On the contrary, the Z pdf is approximatively reproduced by simulations. Indeed, it is possible to see how the pdf of the best simulation is quite close to the reference one. They overlap for lowest values while the best simulation overvalues probability of medium values and undervalues one of highest values. Moreover, it considers higher values than ones of the reference

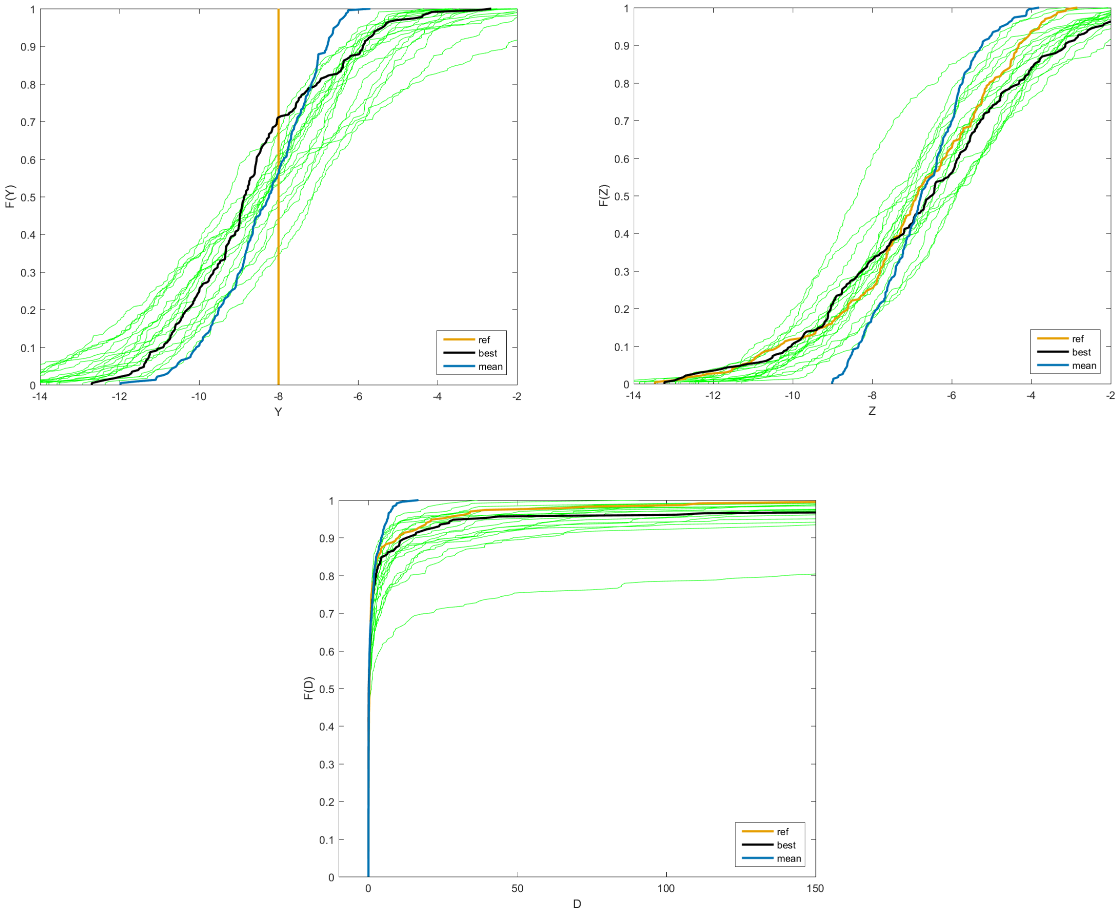

Figure 4.6.22: Empirical cumulative distribution functions of Y, Z and D, K-false positive test

Table 4.26: Statistical parameters comparison, K-false positive test

Field	m_Y	σ_Y^2	m_Z	σ_Z^2	$I_x(Y)$	$I_y(Y)$	$I_x(Z)$	$I_y(Z)$
Reference	-8.00	0.00	-6.95	4.92			2.42	3.28
Best	-8.63	3.54	-6.66	7.03	1.63	3.18	0.85	2.99
Mean	-8.34	1.53	-6.72	1.40	1.17	3.96	3.29	4.80

field. The mean simulation captures central values but it does not infer extreme values, therefore it is more steep. Despite these incorrectnesses for log-conductivity and log-storativity pdfs, the reference diffusivity pdf is inferred rather precisely by the best simulation, while the mean one does not reproduce diffusivity values higher than 20 m^2/s, while the reference field has diffusivity values reaching 150 m^2/s.

Table 4.26 confirm what just noticed on statistics. Indeed, variances of the Y field are much higher than 0, in particular the best simulation infers a field much more heterogeneous. On the contrary means are quite close to the reference value. Moving to the Z field, means are inferred precisely while variances are incorrect: the best simulation infers a very heterogeneous field while the mean simulation infers a smoothed field. Integral scales are also incorrect.

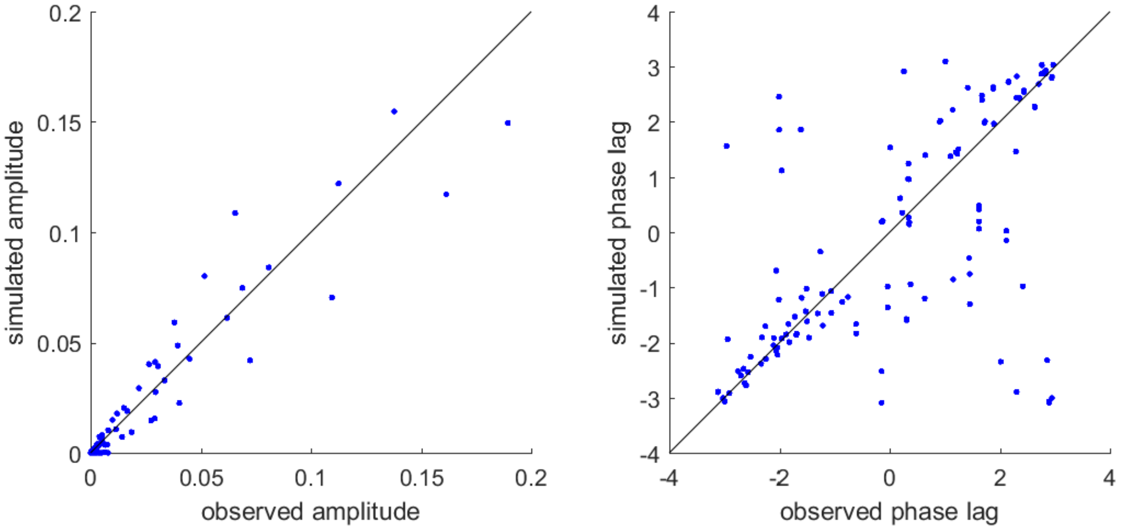

Figure 4.6.23: Scatter plots of simulated versus observed values of amplitude and phase, K-false positive test

Looking scatter plots of amplitude and phase of signal observed in figures 4.6.23, it is evident how the amplitude plot is characterized by points close to the diagonal while the phase one has almost half of points far from the diagonal and the scatter plot results very scattered. As noticed in previous tests, a bad phase scatter plot and therefore a low phase efficiency results in a low overall signal efficiency. This is confirmed by table 4.27(b) where the high amplitude efficiency is not so helpful in the overall efficiency.

Table 4.27: Mean absolute errors of fields and efficiency coefficients, K-false positive test

(a)

Field	Mean($\|\Delta Y\|$)	Mean($\|\Delta Z\|$)
Best	1.6444	2.5411
Mean	1.0394	1.3517

(b)

	NSEE
signal	0.34
amplitude	0.90
phase	0.39

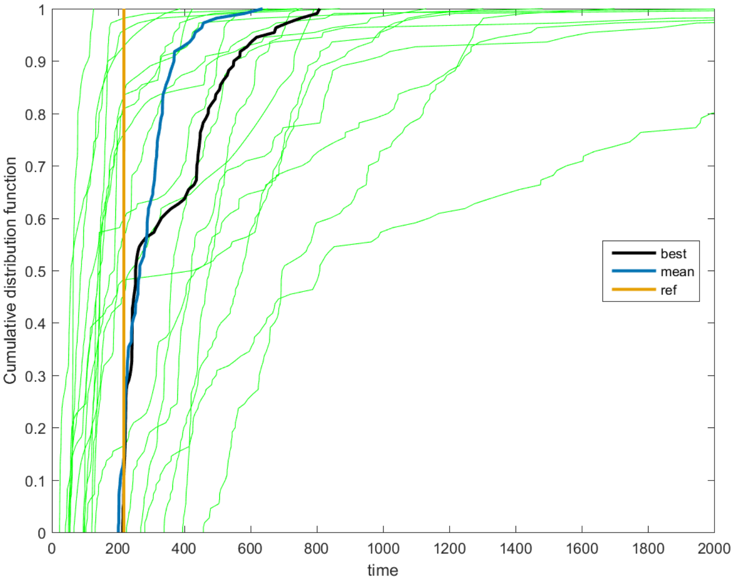

Figure 4.6.24: Breakthrough curves of realizations, K-false positive test

Breakthrough curves of this test in figure 4.6.24 are close to the reference curve but the majority of them has a completely different shape. Indeed, the reference curve is vertical due to the homogeneous log-conductivity field, while other curves are mostly characterized by a long tail for slowest particles. This results in best and mean simulations with curves that are very precise on the arrival time of fastest particles while they overstate arrival time of slowest ones, mainly the best simulation.

Also connectivity functions have completely different shapes. Indeed, the reference field has a constant connectivity function for all the field length while it changes for all other Gaussian simulations that strongly underestimate values. Important to specify that there is only one graph because in this case thresholds are percentiles of a constant field, therefore they are equal. The incorrect inference of connectivity is also noticeable from connectivity indexes in table 4.28.

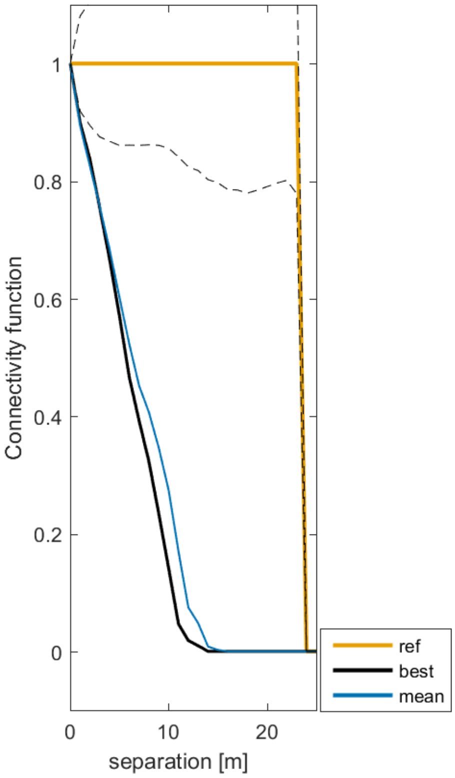

Figure 4.6.25: Connectivity functions (50, 75, 90 percentile), K-false positive test

Table 4.28: Connectivity indicator CI, K-false positive test

	CI
reference	1.00
best	1.60
mean	1.41

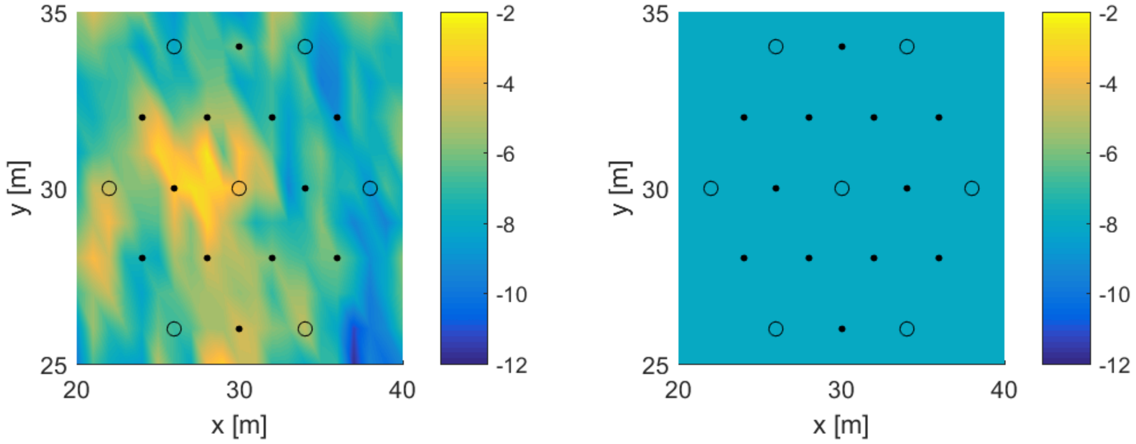

Figure 4.6.26: Reference fields, S-false positive test

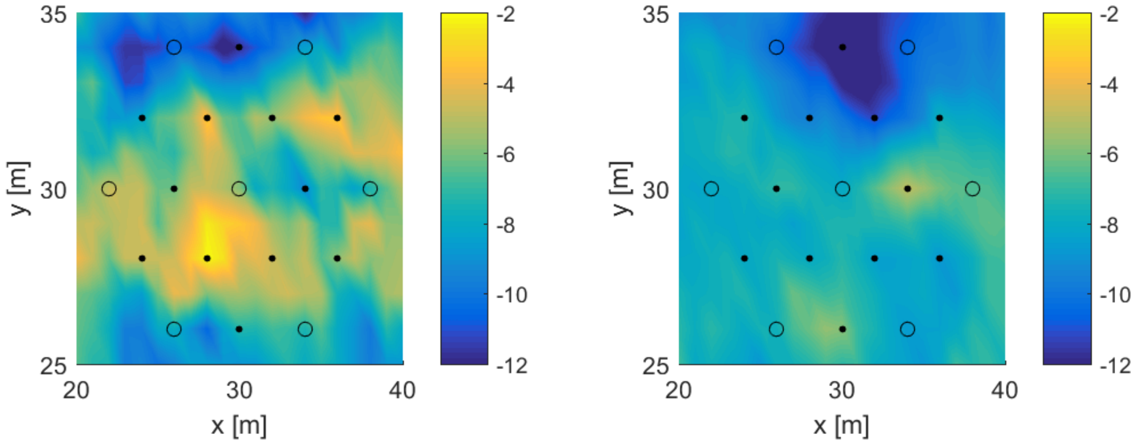

Figure 4.6.27: Best fields, S-false positive test

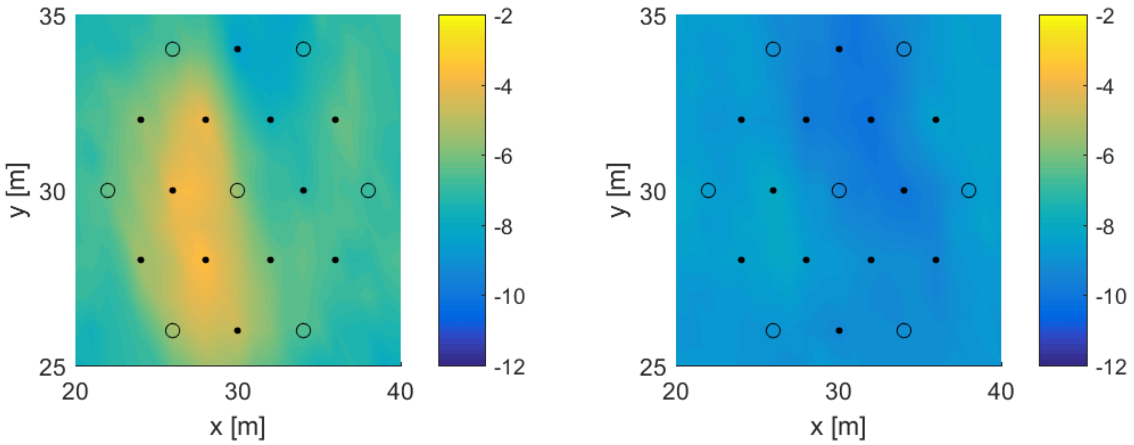

Figure 4.6.28: Mean fields, S-false positive test

4.6.4 Results and observations for 'S false positive'

Figures 4.6.26 show reference fields for this test in which the log-storativity field is homogeneous but it is considered heterogeneous for assumption. The homogeneous field is inferred by the best simulation, figures 4.6.27 as slightly heterogeneous except for the relevant low-values spot on the North. This is an error that will lead to a lower efficiency and a relevant variance of the simulated field. On the contrary the mean Z field shows in figures 4.6.28 some heterogeneity but it is very light and it can be considered homogeneous. Regarding the log-conductivity field, it is approximatively inferred by the best simulation. The field recognizes the high-values spot in the center and the majority of middle values but it infers two horizontal belts characterized by low values. The southern one is partially correct because the southern corners of the reference Y fields have also low values, but the northern one with very low values is not linked to the reference field that in that area has medium values. As usual, the mean simulation succeeds on identifying the pattern but it misses extreme values.

Figures 4.6.29 show how badly the diffusivity field is inferred by simulations. Indeed, the best simulation wrongly identifies northeastern areas with high diffusivity and it does not see the low-values belt on the East. While the mean simulation sees the high diffusivity area on the South-West but it is too compact and large.

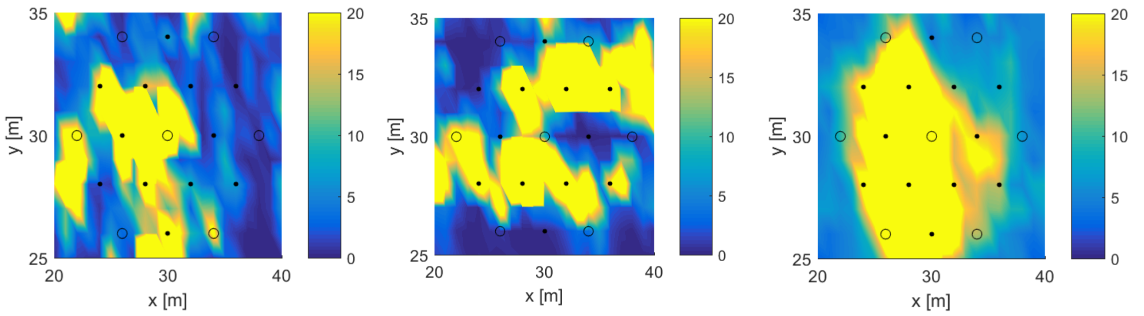

Figure 4.6.29: Diffusivity maps of the reference, best and mean fields, S-false positive test

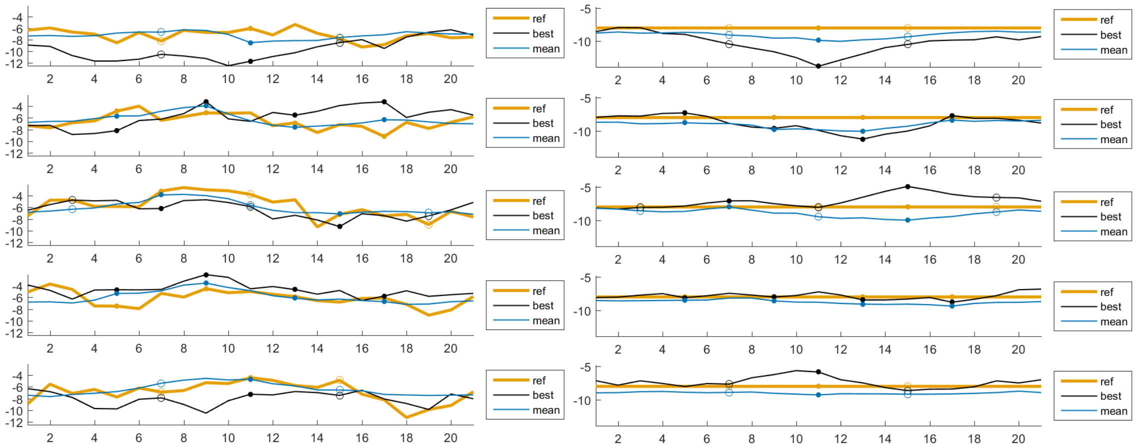

Figure 4.6.30: Transects of reference, best and mean fields in wells and power point rows, Y and Z fields, S-false positive test

Transects in figure 4.6.30 show how properties are inferred in fivetransects. From the log-storativity transects it is visible how the mean simulation is able to reproduce this homogeneous field. It is almost flat but always underestimating the reference. The best simulation also reproduces quite precisely the reference but it overstates some points, mainly on the northern zone. Concerning log-conductivity transects, the mean simulation is still flatter than the best one but it follows anyway the general trend of the reference field. The best simulation better follows the trend but in first and lasttransects the inference strongly undervalues the reference field in almost all the transect length.

Comparing probability distributions is possible to understand how statistics of

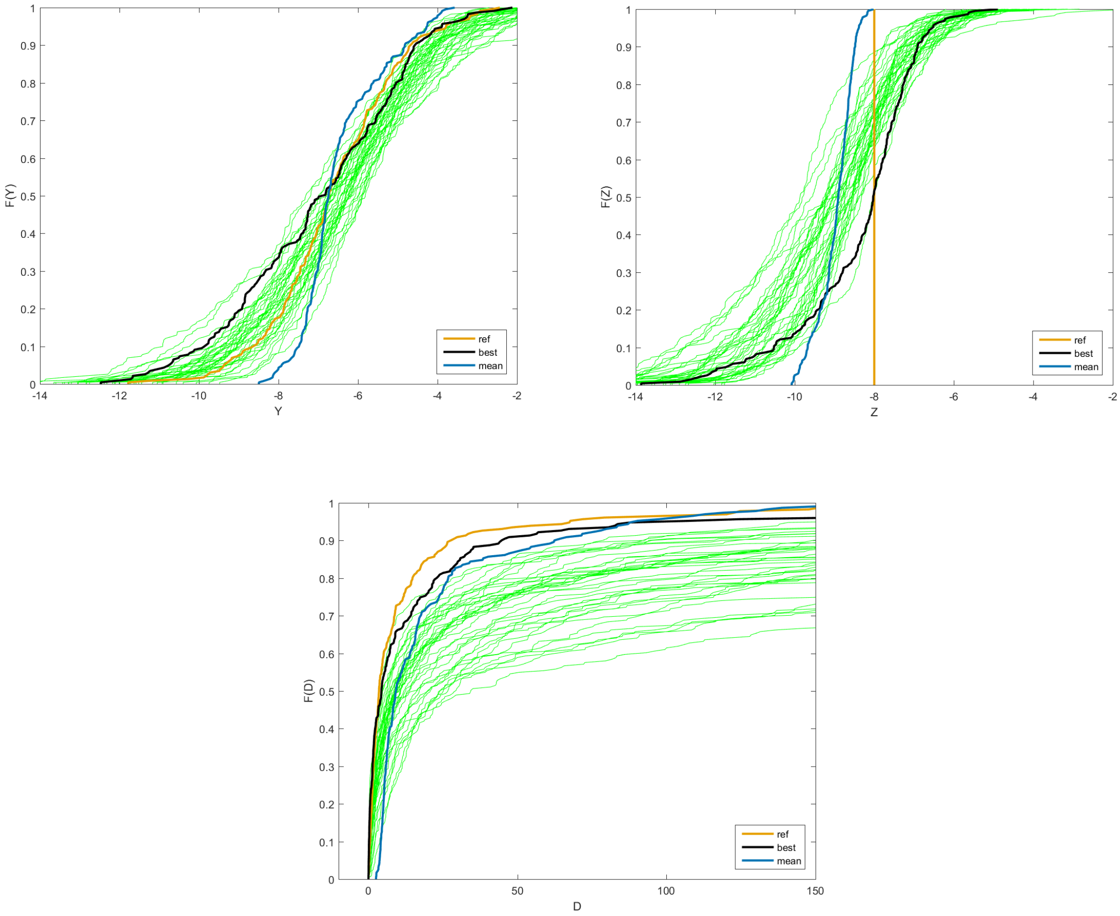

Figure 4.6.31: Empirical cumulative distribution functions of Y, Z and D, S-false positive test

Table 4.29: Statistical parameters comparison, S-false positive test

Field	m_Y	σ_Y^2	m_Z	σ_Z^2	$I_x(Y)$	$I_y(Y)$	$I_x(Z)$	$I_y(Z)$
Reference	-6.63	2.61	-8.00	0.00	0.79	2.18		
Best	-7.02	4.58	-8.37	2.25	7.39	-0.08	7.21	0.98
Mean	-6.46	1.08	-8.97	0.19	1.99	3.70	-0.77	3.80

the field are inferred. Regarding the Z pdfs, it is evident how simulations are not able to infer the reference curve because this describes a homogeneous field while they describe heterogeneous fields with variable variances. The best simulation is a typical Gaussian distribution with a long tail for lowest values. The mean simulation is much more steep and has a concentrated range of values, however it is shifted to lower values and it does not center the mean. Looking at the Y pdf, the function of the best simulation overlaps the reference pdf for medium-high values but it overvalues the probability of lowest values. The mean simulation strongly underestimates the probability of lowest values but it follows the reference pdf for highest values even if these two curves do not overlap. Finally the probability distribution function of the diffusivity field is reproduced quite well by best and mean simulations. The pdf of the best simulation is the closest pdf to the reference even if it undervalues the probability of all diffusivity values. The mean simulation gets also a nice function especially because it overlaps the reference one for highest values.

Table 4.29 shows what already said previously. Concerning the homogeneous Z field, the best simulation infers a mean close to the reference one but it is wrong on the variance estimation. On the contrary, the mean simulation considerably undervalues the mean but it nicely infers a low variance recognizing the homogeneous field. The Y field is inferred by both best and mean simulations precisely for what concerns the mean while the reference variance is overvalued by the best simulation and it is undervalued by the mean one. Integral scales are absolutely incorrect.

Even if statistical parameters seem to be incorrectly inferred, scatter plots show a decent reproduction of observations in wells' locations, both for the amplitude and the phase of the signal. This is also noticeable in table 4.30(b) where both amplitude and phase efficiencies are higher than 0.90 and consequently the overall signal efficiency is also high.

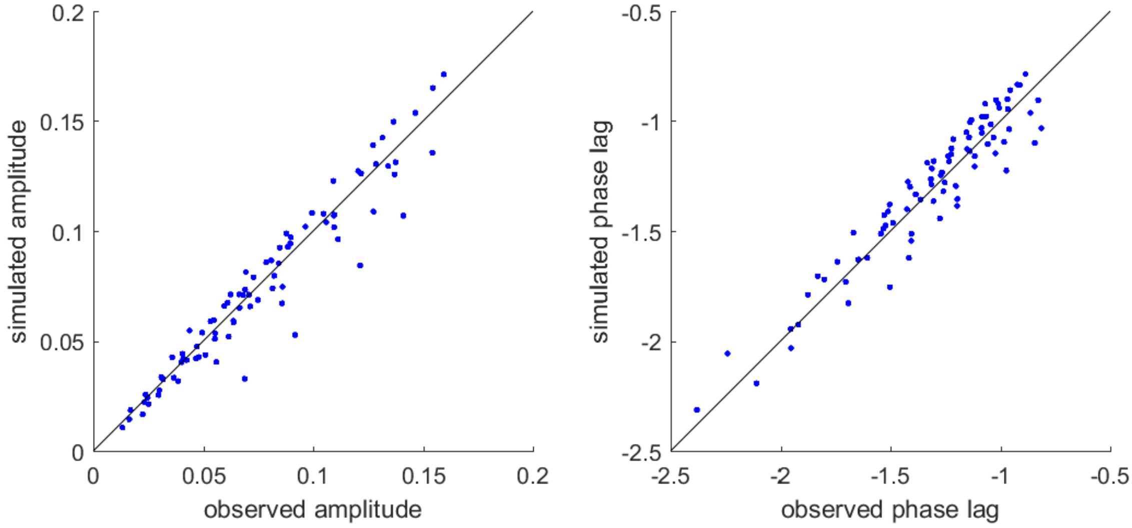

Figure 4.6.32: Scatter plots of simulated versus observed values of amplitude and phase, S-false positive test

Table 4.30: Mean absolute errors of fields and efficiency coefficients, S-false positive test

(a)

Field	Mean($\|\Delta Y\|$)	Mean($\|\Delta Z\|$)
Best	1.9439	1.0907
Mean	1.1233	0.9719

(b)

	NSEE
signal	0.86
amplitude	0.92
phase	0.90

In order to analyze the reproduction of the behavior of the reference field in situations of contaminant transport, breakthrough curves are an important result. In this case, curves of simulations are quite different from the reference because they tend to have longer tails, therefore longer arrival times of slowest particles. Indeed, it is evident how the best simulation produces a curve that has very short arrival times for faster particles and very long times for slowest one, this is confirmed by the huge connectivity index in table 4.31. On the contrary, the mean simulation results in a curve that is almost vertical, with first particles as fast as reference ones and last particles much faster than reference ones. However, it has a better reproduction of curves respect to the best simulation.

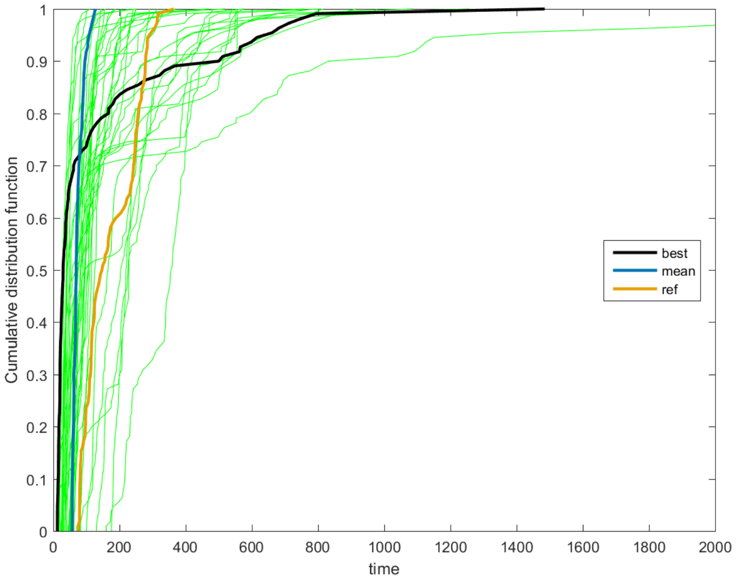

Figure 4.6.33: Breakthrough curves of realizations, S-false positive test

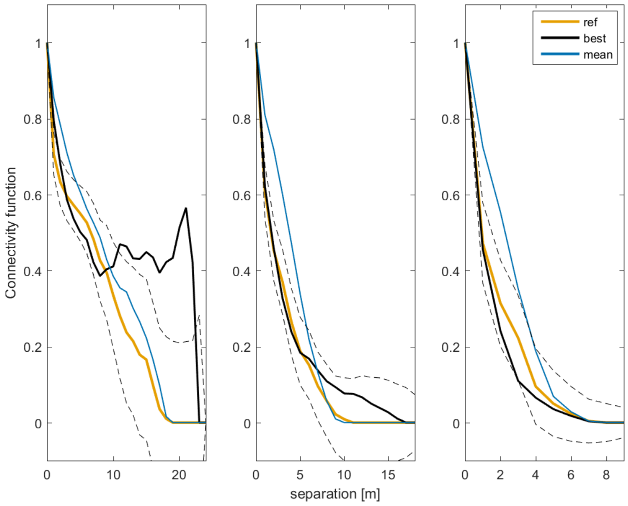

Figure 4.6.34: Connectivity functions (50, 75, 90 percentile), S-false positive test

Table 4.31: Connectivity indicator CI, S-false positive test

	CI
reference	2.19
best	10.01
mean	1.30

Looking instead at connectivity functions in figure 4.6.34, mean simulation infers the reference function better than the best simulation only for the lowest threshold, while for higher one the best simulation has a very good inference.

4.6.5 Observations

The aim of this section is to understand the robustness of the model to incorrect initial assumptions that inevitably lead to wrong reproduction. This is done by two type of tests: false negative where an highly heterogeneous property is assumed homogeneous, false positive where a homogeneous property is assumed as heterogeneous.

The first test, K-false negative, obviously does not reproduce correctly the log-conductivity field but it reproduces quite well log-storativity heterogeneity, pdfs and statistics. The main incorrectness is the lacking identification of the two low-values spots. In any case, the total efficiency is quite high, 0.50. Due to the incorrect assumption of conductivity field, both breakthrough curve and connectivity function are wrongly reproduced.
The S-false negative test has a lower overall efficiency of 0.31. Indeed, obviously the log-storativity field is incorrect and the log-conductivity field is reproduced quite well even if there is a movement of the high-values spot. The log-conductivity pdf overlaps to the reference one for highest values but it overvalues the probability of low-medium values. The efficiency results quite low because the observations are not reproduced correctly how is noticeable from the two strewn scatter plots, mainly the phase one. Results about contaminant transport are relatively good.
The false positive test for the log-conductivity aims to reproduce a homogeneous reference field not knowing this a priori, therefore the variance can be searched by the PSO algorithm in a large range from 0 to 6. The overall efficiency of this test is only 0.34. Indeed, maps of the best simulation show a too heterogeneous Y field and a Z field with shifted low-values spots and with a high-values spot that is absent in the reference. However, the probability distribution functions are nice, except for the Y pdf. Inferred breakthrough curves and connectivity functions are far from the reference ones due to the missed reproduction of the homogeneous log-conductivity field.
The S-false positive test results in a general efficiency much higher than the other false positive test, 0.86. The best simulation is able to reproduce an almost homogeneous Z field, except for a low-values spot, and also the Y field is nice even if there

are some incorrectnesses. As in the previous test, inferred pdfs well reproduce the reference one, except for the Z field. Breakthrough curves are still far from the reference one, while connectivity functions for highest thresholds are precisely inferred.

From these false tests it is possible to conclude that doing a wrong assumption is obviously damaging for the overall simulation, except for the false positive of the log-storativity. Indeed, it is expected that false positives result in better simulations due to the fact that the possible range of the variance is large enough to include homogeneity, variance equal to 0. On the contrary, in false negatives the wrong assumption does not allow any adaptation. However, only the false positive of Z has a high signal efficiency, even if single results seem not to be so nice, mainly inferred statistical parameters.

5

Conclusions

In this study we proposed to use the steady-periodic model for oscillatory pumping test in a Monte Carlo framework in order to characterize hydrogeological formations. Using the model with governing equations in the Fourier space, we were able to overcome limitations of existing models for aquifers understanding, such as computational time, consideration of heterogeneity and comprehension of all possible realizations. Below are reported in detail different contributions of this study.

First, we discuss the choice of reproduction of the only non-Gaussian field: the Chessboard. This test was at first introduced to try out the model in a simple field, characterized by a block-type variability. However, the good inference showed to be not only a positive starting point for generally heterogeneous fields, but it showed how Gaussian realizations of this test can nicely reproduce a geometric field. Indeed, their easy adaptability to observational data make them succeed in simulating a field that is really different from a Gaussian one.

By looking at results of all tests on random fields, where at least one of the properties' fields is heterogeneous, another observation can be made on the comparison between realizations of the ensemble average of all inversions and of the best realization of the Monte Carlo framework. The former returns inferred properties' maps where the heterogeneity pattern is correctly identified and reproduced but extreme values are always absent. Therefore, statistics of the properties are incorrect, mainly because variances are too low. On the contrary, the best simulation returns inferred fields whose heterogeneity is close to the one taken as reference, despite it is not identical. The best realization notices and places extreme values even if they are sometimes shifted in the map. However, statistical parameters are usually closer to the reference field as for probability density functions. This is an important observation because many models actually used invert observational data in order to obtain the average of all simulations, therefore to evaluate deterministically the field. However, it is mainly the uncertainty to allow the reproduction of extreme values and that makes maps of the best simulation closer to reference fields.

Thirdly, realizations of the model are generally able to infer accurately statistical parameters and probability distribution functions of properties. Local properties in the field are approximatively reproduced giving an indicative idea of their heterogeneity. Moreover, for situations of pollutant spreading, contaminant transport on the field is crucial but difficult to infer even with the model proposed in this study. Indeed, breakthrough curves have never been correctly reproduced by simulated fields. Inferred fields produce curves that surround the reference one but best and mean simulations never overlap it, mainly arrival times of slowest particles are different. On the other hand, the connectivity function is usually reproduced quite precisely by the best simulation. It is a good result because it is a relevant indication of the presence of fast pathways where contaminant particles can move. Even if arrival times of particles are wrongly inferred, the connectivity function gives an idea about the possibility for particles to move fast and therefore it is an indication of contaminant spreading risk. This approximate information could be very useful for sensible areas.

We discuss now the comparison between Small and Large PSO ranges tests. This paper applies a model that uses as input ranges of data where the PSO algorithm searches statistical parameters of the field. If prior information of the field is available, it is possible to narrow these ranges in order to speed the model and reduce the computational time. Moreover, inferred fields were expected to have higher efficiencies than those from Large PSO ranges test because parameters are closer to the real value. Instead, the latter test returns better results both in terms of heterogeneity maps and statistics. Large ranges on the PSO algorithm allow the inferred field to adapt to the available observations about the area of interest, even if the inference of the external area is completely misunderstood in its values and statistics. Our comment is mainly directed to ranges of variances, therefore a good compromise could be to have small ranges for means and large ones for variances. By applying this arrangement the adaptation of the inferred field is allowed but computational time decreases.

Final tests, where the spatial variability of the reference field was wrongly assumed, make clear how obviously damaging is to do an incorrect initial assumption. However, looking at heterogeneity maps of the best simulation, it seems that log-storativity is more adaptable to observations than log-conductivity. For example, in performing the K-false negative test it is evident how well the storativity is able to follow the reference trend, while the same adaptation does not happen for the conductivity field when performing the S-false negative test. Finally, false positive tests were carried out: the reference field of one property is homogeneous and it is assumed heterogeneous in the inversion with the possibility to be inferred with very low variance. Also in this case, log-storativity results more adaptable because in the S-false positive test the Z field results almost homogeneous, while the Y field is more heterogeneous in the K-false positive test.

Following the previous observation on adaptation and looking at Y-Z error tables of all tests, it is possible to conclude that the applied model is more sensible to storativity than conductivity. The general rule noticed is that tests with a low efficiency of the best simulation are all characterized also by a low efficiency of the phase and by high mean absolute errors of the log-storativity field, while the error of the log-conductivity field seems not to be so relevant. Indeed, for each group of tests (one homogeneous field, both heterogeneous fields, false negatives and positives) the test with best reproduction and highest efficiency is always the one with lowest mean absolute error of Z. This is understandable given that the model is applied to an oscillatory pumping test and therefore it works in transient conditions. Indeed, it is widely known that in steady conditions conductivity is the relevant parameter, while in transient conditions storativity turns into the main parameter even if conductivity still influences. However, motivations of this observation are focused on results of the best simulation that is identified by maximizing the modified NashSutcliffe efficiency. Could be that using another formula to maximize on the inversion process leads to different best simulations. Even so, it is reasonably conceivable that different formulas lead to slightly different best realizations but there are not significant changes. Therefore, it could be recommendable to combine this model which accurately infers storativity of the field with an alternative model for field's connectivity estimation.

In conclusion, we summarize the main point highlighted by this study. First, the best simulation is usually able to accurately reproduce the field while ensemble average of all inversions identifies the heterogeneity pattern, but it lacks in estimation of extreme values therefore misunderstanding statistics of the field. Second, contaminant transport remains difficult to infer accurately: breakthrough curves are never correctly reproduced, while connectivity functions are approximatively inferred by the best simulation. Third, when using the PSO algorithm, large ranges of variances should be set as input in order to let the inference adapt to observations. Finally, the model better reproduces storativity than conductivity, due to working conditions on the field.

Glossary of Notations

α	Geoformation compressibility	$[T^2L/M]$
β	Water compressibility	$[T^2L/M]$
γ	Liquid unit weight	$[M/(L^2T^2)]$
Γ_d	Dirichlet boundary	
Γ_n	Neumann boundary	
Γ_w	Linearized water table boundary	
θ	Unknown vector with structural parameters of the geostatistical model of spatial variability, on the inversion procedure	
μ	Fluid viscosity	$[M/(LT)]$
ρ	Fluid density	$[M/L^3]$
σ_Y^2	Variance of the log-conductivity	[-]
σ_Z^2	Variance of the log-storativity	[-]
τ	Connectivity function	[-]
Φ_ω	Phasor, a complex-valued field variable	[-]
ω	Frequency	$[rad/T]$
$\omega_{inertia}$	Inertia weight of the PSO algorithm	[-]
Ω	Domain of interest	
$\Im$	Forward nonlinear operator	
$\mathbf{a}$	Unknown vector of Y and Z values, on the inversion procedure	
$\bar{\mathbf{a}}$	Vector of prior estimates of $\mathbf{a}$, on the inversion procedure	
A	non-Hermitian symmetric matrix, on the COCR solver	[-]
b	Saturated thickness of the aquifer	$[L]$

c_1	Constant cognitive learning factor, on the PSO algorithm	[-]
c_2	Constant social learning factor, on the PSO algorithm	[-]
C_a	Covariances function of $\mathbf{a}$, on the inversion procedure	[-]
C_v	Covariances function of $\mathbf{v}$, on the inversion procedure	[-]
C_Y	Axisymmetric exponential covariance function, on the inversion procedure	[-]
CI	Connectivity index	[-]
d	Distance	$[L]$
dx_i	Distance in i-th direction	$[L]$
D	Diffusivity	$[L^2/T]$
e	Anisotropy ratio	[-]
g	Gravitational acceleration	$[L/T^2]$
g_{best}^k	Global best position, index k refers to the iteration, on the PSO algorithm	$[L]$
h	Head change from an initial steady condition	$[L]$
H	Hydraulic head	$[L]$
H_{obs}	Observed head variations in observing wells locations	$[L]$
$\overline{H_{obs}}$	Mean of observed head variations in observing wells locations	$[L]$
H_{sim}	Simulated head variations in observing wells locations	$[mL]$
I_{Yh}	Horizontal spatial scale of conductivity	$[L]$
I_{Yv}	Vertical spatial scale of conductivity	$[L]$
k	Intrinsic permeability of the soil	$[L^2]$
K	Conductivity	$[L/T]$
K_{ef}	Effective hydraulic conductivity	$[L/T]$
K_G	Geometric conductivity	$[L/T]$
K_H	Harmonic conductivity	$[L/T]$
m_Y	Mean of the log-conductivity	[-]
m_Z	Mean of the log-storativity	[-]
n	Fraction of void space	[-]
$\mathbf{n}$	Outward normal of a certain boundary	[-]
n_e	Effective porosity	[-]
N_ω	Number of frequencies used	[-]
N_θ	Number of parameters of the geostatistical model of spatial variability	[-]

N_k	Number of measurements	[-]
N_{pp}	Number of pilot points	[-]
N_{tests}	Number of tests	[-]
N_{ws}	Number of wells	[-]
$ntime$	Number of time steps	[-]
$p^k_{best,i}$	Partial best position, indexes i and k refer respectively to the particle and the iteration, on the PSO algorithm	$[L]$
$\mathbf{p}_n$	Search direction, on the COCR solver	$[L]$
P	Gauge pressure	$[M/(LT^2)]$
Pr	Property, log-conductivity or log-storativity	[-]
Pr_k	Log measurement	[-]
Pr_{pp}	Pilot point value of the property	[-]
Pr^c	Conditioned field	[-]
$Pr^{c,p}$	Prior log field	[-]
q	Volumetric water source	$[(L^3/T)/L^3]$
q_i	Specific discharge in i-th direction, called Darcy's velocity	$[L/T]$
Q	Peak flow rate	$[(L^3/T)/L^3]$
$\mathbf{r}$	Distance vector between two points	$[L]$
$\mathbf{r}_n$	Residual vector, on the COCR solver	$[L]$
R	Radius of influence of a pumping well	$[L]$
s	Drawdown	$[L]$
S	Storativity	[-]
S_r	Specific retention	[-]
S_s	Specific storage	$[L^{-1}]$
S_y	Specific yield	[-]
t	Time	$[T]$
t_{av}	Average arrival time	$[T]$
$t_{5\%}$	Arrival time of the faster 5% of particles	$[T]$
T	Period	$[T]$
T	Transmissivity	$[L^2/T]$
$\mathbf{v}$	Vector of the measurement errors, on the inversion procedure	$[m]$
v^k_i	Spatially constant drift, indexes i and k refer respectively to the particle and the iteration, on the PSO algorithm	$[L]$
V_t	Total volume	$[L^3]$

V_v	Void volume	$[L^3]$
$\mathbf{x}$	Vector of spatial coordinates	$[L]$
$\mathbf{x}_n$	n-th approximate solution, on the COCR solver	$[L]$
Y	Log-conductivity	[-]
z	Elevation at the piezometer bottom	$[L]$
$\mathbf{z}$	Vector of head measurements	$[m]$
Z	Log-storativity	[-]

COCR	Conjugate A-Orthogonal Conjugate Residual
COCG	Conjugate Orthogonal Conjugate Gradient
M-OHI	Multifrequency Oscillatory Hydraulic Imaging
NS	Nash-Sutcliffe efficiency
NSEE	Modified Nash-Sutcliffe efficiency
QMR	Quasi-Minimal Residual method
RSF	Random Space Function
pdf	Probability distribution function
PSO	Particle Swarm Optimizer
D0	Scenario with average diffusivity equal to 1 m^2/s
D1	Scenario with average diffusivity equal to 2.72 m^2/s
D2	Scenario with average diffusivity equal to 7.39 m^2/s
Cb	Chessboard test
K0	K-Homogeneous test
S0	S-Homogeneous test
SK	Heterogeneous properties test
KNeg	K-false negative test
SNeg	S-false negative test
KPos	K-false positive test
SPos	S-false positive test

Bibliography

Vedat Batu. *Aquifer hydraulics: a comprehensive guide to hydrogeologic data analysis*, volume 1. John Wiley & Sons, 1998.

Alberto Bellin and Yoram Rubin. HYDRO_GEN: A spatially distributed random field generator for correlated properties. *Stochastic Hydrology and Hydraulics*, 10 (4):253–278, nov 1996.

Marco Bianchi, Chunmiao Zheng, Crystal Wilson, Geoffrey R. Tick, Gaisheng Liu, and Steven M. Gorelick. Spatial connectivity in a highly heterogeneous aquifer: From cores to preferential flow paths. *Water Resources Research*, 47:W05524, 2011.

Geoffry C. Bohling and James J. Butler. Inherent limitations of hydraulic tomography. *Groundwater*, 2010.

Michael Cardiff. A field proofofconcept of aquifer imaging using 3D transient hydraulic tomography with modular, temporarilyemplaced equipment. *Water Resources Research*, 2012.

Michael Cardiff and Warren Barrash. 3-D transient hydraulic tomography in unconfined aquifers with fast drainage response. *Water Resources Research*, 47(12), dec 2011.

Michael Cardiff, Tania Bakhos, Peter K. Kitanidis, and Warren Barrash. Aquifer heterogeneity characterization with oscillatory pumping: Sensitivity analysis and imaging potential. *Water Resources Research*, 49(9):5395–5410, 2013a.

Michael Cardiff, Warren Barrash, and Peter K. Kitanidis. Hydraulic conductivity imaging from 3-D transient hydraulic tomography at several pumping/observation densities. *Water Resources Research*, 49(11):7311–7326, 2013b.

Marta Castagna and Alberto Bellin. A Bayesian approach for inversion of hydraulic tomographic data. *Water Resources Research*, 45(4), 2009.

Marta Castagna, Matthew W. Becker, and Alberto Bellin. Joint estimation of transmissivity and storativity in a bedrock fracture. *Water Resources Research*, 47(9), sep 2011.

Marta Castagna, Alberto Bellin, and Gabriele Chiogna. Uncertainty Estimation and Evaluation of Shallow Aquifers' Exploitability: The Case Study of the Adige Valley Aquifer (Italy). *Water*, 7(7):3367–3395, jun 2015.

Catherine Certes and Ghislain de Marsily. Application of the pilot point method to the identification of aquifer transmissivities. *Advances in Water Resources*, 14 (5):284–300, oct 1991.

Roger B. Clapp and George M. Hornberger. Empirical equations for some soil hydraulic properties. *Water resources research*, 14(4):601–604, 1978.

edited by Jaques Deller. *The handbook of groundwater engineering.* 1998.

R. Allan Freeze and John A. Cherry. *Groundwater.* Number 556.3 FRE. 1979.

Monroe A. Hartman, Arlin D. Nicks, and Edd D. Rhoades. Field experiment on washita river. 1969.

Charles Harvey and Steven M. Gorelick. / Mississippi. 36(3):637–650, 2000.

Kevin M. Hiscock and Victor F. Bense. *Hydrogeology: principles and practice - Second edition.* 2014.

Dennis McLaughlin and Lloyd R. Townley. A reassessment of the groundwater inverse problem. *Water Resources Research*, 1996.

Sebastiano Piccolroaz. Prediction of lake surface temperature using the air2water model: guidelines, challenges, and future perspectives. *Advances in Oceanography and Limnology*, 7(1):36–50, 2016.

Yoram Rubin. *Applied stochastic hydrogeology.* 2003.

Yoram Rubin and Susan S. Hubbard. *Hydrogeophysics.* 2005.

Tomohiro Sogabe and Shao-Liang Zhang. A COCR method for solving complex symmetric linear systems. *Journal of Computational and Applied Mathematics*, 199(2):297–303, 2007.

Tatiana D. Streltsova. *Well Testing in Heterogeneous Formations.* 1988.

Andrew W. Western, Gunter Bloschl, and Rodger B. Grayson. Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resources Research*, 37(1):83–97, 2001.

Xiaoyong Zhan, Geoffry C. Bohling, James J. Butler, and Li Zheng. Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities. 2002.

Chunmiao Zheng, Marco Bianchi, and Steven M. Gorelick. Lessons learned from 25 years of research at the MADE site. *Ground Water*, 49(5):649–662, 2011.

Acknowledgements

I would like to express my gratitude to my supervisor Prof. Alberto Bellin for the useful guide he provided in the development of this thesis. His comments, remarks and engagement have always supported my learning process. I am thankful to Ing. Sebastiano Piccolroaz and Ing. Diego Avesani for having discussed with me doubts and problems helping me to understand and reflect. I would like also to thank Prof. Luis Filipe Tavares Ribeiro for the continuous support despite the long distance. Furthermore, I thank Ivano Cristofolini that provided me with very good advices any time I was facing problems with the cluster.

In this final work of my master, I would like to thank also people who helped me along my double degree experience in Lisbon: Prof. Marco Toffolon, Prof. Rodrigo Proena de Oliveira. A particular thanks goes to Virna Eccli that helped me navigating through the convolutet bureaucratical processes.

However, not less important is the huge support I had from my friends. The wonderful girls of MC2005, the very large group of university colleagues who walked this entire educational trip with me and who joyfully welcomed me when I came back from Lisbon, and finally, Veronica, my crazy friend. I would like to thank also all friends who made my Portuguese experience special, above all a warm embrace to Mere and Vanessa. Lastly, Matteo whom I will never thank enough for his constant support and patience even in the worst moments.

I thank my sister who always helped me whenever I needed, regardless the reason or the everlasting phisical long distance.
Last but not least, I want to thank my relatives, in Italian..

Ringrazio le mie nonne, cosí diverse e cosí speciali, che mi hanno cresciuta insegnandomi a stare coi piedi per terra, senza mai perdere la grande voglia di sperimentare. Ringrazio infine ed infinitamente i miei genitori per avermi sempre sostenuto in ogni mia scelta, ma devo loro innanzituto grande gratitudine per avermi cresciuto con la curiositá di sapere e la voglia di scoprire. Solo grazie a voi sono arrivata fino a qui oggi.