Conditional stochastic mapping of transport connectivity

D. Fernàndez-Garcia,[1] P. Trinchero,[2] and X. Sanchez-Vila[1]

Received 18 August 2009; revised 13 January 2010; accepted 1 February 2010; published 9 October 2010.

[1] We present a method for the stochastic simulation of point-to-point transport connectivity honoring data from three types of information: (1) travel time estimates obtained from field tracer tests; (2) estimates of flow connectivity indicators obtained from the relatively fast or slow flow response that is observed at a point location given the flow impulse at another location, and (3) measurements of transmissivity at a local scale. The method thus efficiently integrates data obtained from different hydraulic tests, each sampling different areas within the aquifer. To achieve this, we first extend the concept of point-to-point flow connectivity and transport connectivity, mathematically formulated by Trinchero et al. (2008) for pumping conditions, to support a more general flow configuration. Interestingly, point-to-point flow connectivity can be generally seen as a weighted integral of transmissivity over the entire domain, the weighting function being proportional to the sensitivity of heads with respect to the natural log of transmissivity per unit of aquifer volume. On the contrary, point-to-point transport connectivity is a weighted integral along the particle path of the solute mass that involves two variables: transmissivity and flow connectivity. Each variable has its own distinct weighting function. The weighting function of transmissivity is inversely proportional to both the homogeneous travel time and the point velocity sampled along the travel path. On this basis, we show how to generate conditional point-to-point transport connectivity maps. The method avoids the inference of cross-covariance functions between variables measured over different scales and sampled areas (which cannot be otherwise estimated with a few data measurements) by expressing them as a function of the local transmissivity covariance function. An example of the method is provided to evaluate the worth of including tracer data to delineate capture zones of abstraction wells originally defined from local transmissivity measurements. Monte Carlo simulations reveal that the impact of including tracer data is a maximum when the travel time data are obtained at a location different than that of transmissivity measurements. The reason is that weighting functions give larger weights to the injection location, so introducing tracer test data at points where transmissivity is already known is somewhat redundant.

Citation: Fernàndez-Garcia, D., P. Trinchero, and X. Sanchez-Vila (2010), Conditional stochastic mapping of transport connectivity, *Water Resour. Res.*, *46*, W10515, doi:10.1029/2009WR008533.

1. Introduction

[2] The concept of hydraulic connectivity lacks a rigorous mathematical definition in porous media. Still, connectivity is an easily understandable subject. Loosely speaking, two points of an aquifer are well connected if an action taking place at one of them has a fast and observable response at the other one. Yet this response at a given point should be compared with some average or characteristic behavior of the aquifer, leading to a somewhat ill-defined property.

[3] Initial work on connectivity focused on the characteristic response associated with the domain as a whole. From a qualitative standpoint, the interconnection of high hydraulic conductivity values (K) increases the amount of water than can flow with respect to aquifers where high-K bodies are not connected. This is obviously the case for a fully stratified medium, where it is well known that flow takes place preferentially parallel to stratification. In more realistic scenarios, such as sand bodies embedded in a less conductive matrix, the flow system in the aquifer is mainly controlled by the continuity and interconnectedness of the sand rather than by its actual local hydraulic conductivity values, which was reported first by *Fogg* [1986] and later extended to transport (actually travel time estimates) by *Poeter and Townsend* [1994].

[4] These findings were numerically confirmed by a number of authors who used different approaches to generate fields with an enhanced connectivity of the points displaying local high K values. As an example, *Sanchez-Vila et al.* [1996] used a Boolean approach to embed interconnected elongated features over a realization of a multi-Gaussian field. Their main conclusion was that the presence of such structures yielded effective transmissivity

[1]Department of Geotechnical Engineering and Geosciences, Technical University of Catalonia, Barcelona, Spain.
[2]Amphos XXI Consulting, Barcelona, Spain.

0043-1397/10/2009WR008533

values, T_{eff}, larger than those corresponding to a multi-Gaussian field. About the same time, *Gómez-Hernández and Wen* [1998] generated connected fields by increasing the correlation lengths of extremely high or low K values in a multi-indicator model. They found that travel times estimated from multi-Gaussian fields are not conservative, meaning that pollutants can travel faster than predicted from such random field models. These findings regarding T_{eff} and fast travel times were later confirmed by *Zinn and Harvey* [2003], who developed their own ad hoc connected fields. *LaBolle and Fogg* [2001] showed that connectivity of low-permeability hydrofacies plays an important role in the success of remediation actions.

[5] Thus, the concept of aquifer connectivity is directly related to the presence of conductive features that display some spatial continuity. As such, geostatistical methods based on two-point statistics (such as variograms) are not able to capture most of the important connectivity features occurring in natural formations (i.e., preferential flow channels [*Western et al.*, 2001; *Kerrou et al.*, 2008]). In past years a significant number of connected fields have been used because, contrary to multi-Gaussian fields, they produce breakthrough curves displaying power-law tailing [e.g., *Zheng and Gorelick*, 2003; *Willmann et al.*, 2008; *Fernàndez-Garcia et al.*, 2009], similar to that observed in real aquifers. These fields can be constructed using models such as multiple-point geostatistics [e.g., *Strebelle*, 2002].

[6] The picture of aquifer connectivity is completed by discussing point-to-point connectivity [e.g., *Trinchero et al.*, 2008]. This concept is more intuitive because it relies on the loose definition specified earlier. For example, two points can be said to be well connected whenever a pumping operation taking place at one of them is noticeably observed at the other one in a short period of time. If this pumping is used for parameter identification using, for example, the Cooper-Jacob method, a fast response would lead to an estimation of the storage coefficient, S_{est}, smaller than the actual value [*Meier et al.*, 1998]. Then S_{est} becomes an apparent value [*Sanchez-Vila et al.*, 2006], and the ratio between S_{est} and the real S value (provided it exists) can be regarded as an indicator of point-to-point connectivity [*Sanchez-Vila et al.*, 1999; *Knudby and Carrera*, 2005]. Another indicator of flow connectivity was suggested by *Knudby and Carrera* [2006] to be the apparent diffusivity D_r (estimated transmissivity divided by estimated storage coefficient). The reliability of this indicator was assessed using a Monte Carlo analysis. They found that D_r has a certain degree of correlation with indicators of both transport and flow global connectivity. *Frippiat et al.* [2009] suggested that the head and flow variances can be used to identified preferential paths and flow barriers. Flow distributions were also used for identifying the presence of channels in fractured media [*de Dreuzy et al.*, 2001].

[7] Alternative definitions of connectivity indicators can be considered when the outcome of tracer tests is considered. We could also state that two points are well connected if pumping at one of them and injecting a conservative tracer in the other leads to faster tracer travel time than that expected in a homogeneous medium. This interpretation of the fast arrival of tracers was used by *Sanchez-Vila and Carrera* [1997] and *Fernàndez-Garcia et al.* [2002] to characterize the presence of preferential flow paths occurring between the pumping well and the injection location.

[8] Despite the existence of different indicators of connectivity, it is reasonable that all of them should be somehow related. *Trinchero et al.* [2008] analyzed the relationship between two of these indicators: the storage coefficient estimated using the Cooper-Jacob method, S_{est} (indicator of flow connectivity), and the relative advective travel time, t_a (indicator of transport connectivity). These authors assessed this relationship in multi-Gaussian random fields and developed a framework for the delineation of capture zones around an abstraction well used for the production of drinking water. Nevertheless, in its current formulation the analytical solution is limited to convergent flow conditions induced by a single pumping well and it is not directly applicable to real applications because it requires a full knowledge of the aquifer (spatial distribution of transmissivity) that is usually not available and is economically unfeasible to obtain in standard field campaigns.

[9] We start by generalizing the relationship between flow and transport connectivity indicators to incorporate a general flow configuration. Then we present a stochastic framework for the delineation of connectivity patterns using a limited and sparse number of measurements. The lack of complete knowledge of the variables involved in the problem is overcome by treating them as regionalized variables or random functions. The methodology allows for conditioning the results to three types of data obtained from different field tests and measurement scales, namely (1) travel time estimates obtained from field tracer tests, (2) estimates of flow connectivity indicators, and (3) measurements of transmissivity at a local scale. The fact that the tracer and flow connectivity data depend on the separation distance of the field tests renders the method the capability of efficiently integrating data obtained from different hydraulic tests, each one sampling different areas within the aquifer. The ability of the methodology to properly delineate capture zones is assessed through stochastic estimation (i.e., ordinary co-kriging) and sequential Gaussian simulations based on different sets of measurements.

2. General Theory: Point-to-Point Connectivity

2.1. Flow Connectivity Between Two Points

[10] We consider a two-dimensional heterogeneous confined aquifer under quasi steady state flow conditions. The heterogeneous structure is characterized by a spatially varying natural log of transmissivity, $\mathbf{Y}(\mathbf{x}) = \ln \mathbf{T}(\mathbf{x})$, where $\mathbf{T}$ is the local transmissivity tensor. All other properties are assumed homogeneous throughout the domain. Our goal is to examine the time response of the pressure head at a given observation location, $\mathbf{x}_i$, given a (constant-in-time) flow impulse at another location, $\mathbf{x}_p$. We only consider the case in which these two points are actually connected through a flow path, so there is a specific streamline that passes through them. This setup includes, for example, convergent or dipole flows in the presence of a regional trend or not.

[11] We start by decomposing $\mathbf{Y}(\mathbf{x})$ around an equivalent homogeneous transmissivity tensor, $\mathbf{Y}_0 = \ln \mathbf{T}_0$, that is obtained by replacing the heterogeneous system by a homogeneous one based on some criteria (e.g., the reproduction of the total net flow discharge throughout the domain). In general, even when transmissivity is locally isotropic, $\mathbf{Y}(\mathbf{x}) = Y(\mathbf{x})\mathbf{I}_d$ ($\mathbf{I}_d$ is the identity matrix), the equivalent

transmissivity of a heterogeneous medium is better described as a full tensor, so we can formally write

$$\mathbf{Y}(\mathbf{x}) = \mathbf{Y}_0 + \mathbf{Y}'(\mathbf{x}), \tag{1}$$

where $\mathbf{Y}'(\mathbf{x})$ is the deviation from the equivalent homogeneous tensor. By replacing $\mathbf{Y}(\mathbf{x})$ by $\mathbf{Y}_0$ the corresponding flow problem is written as

$$\nabla \cdot \left(e^{\mathbf{Y}_0} \nabla h_0\right) + f^w(\mathbf{x}_p) = S\frac{\partial h_0}{\partial t}, \tag{2}$$

where f^w is the source/sink term, which includes the constant-in-time impulse at $\mathbf{x}_p$, and S is the aquifer storage coefficient. Solving (2) with appropriate boundary conditions and assuming quasi steady state flow, one obtains the head field $h_0(\mathbf{x})$ of the equivalent homogeneous medium, which allow us to decompose $h(\mathbf{x})$ into

$$h(\mathbf{x}) = h_0(\mathbf{x}) + h'(\mathbf{x}), \tag{3}$$

where h' is the deviation from the homogeneous solution. Here we have avoided writing the time dependence of groundwater heads to indicate that h, h_0, and h' are referred to as pseudo steady state conditions (large times, i.e., $t \to \infty$).

[12] In a general anisotropic medium the flow vector and the hydraulic gradient are not necessarily collinear. In this case, it is convenient to express Darcy's law in terms of directional transmissivity, $T(\mathbf{x})$, which is defined as the ratio between the specific discharge at a given point $\mathbf{x}$ and the component of the hydraulic gradient in the direction of flow,

$$T(\mathbf{x}) = \frac{\|\mathbf{T}(\mathbf{x})\mathbf{J}(\mathbf{x})\|^2}{[\mathbf{T}(\mathbf{x})\mathbf{J}(\mathbf{x})] \cdot \mathbf{J}(\mathbf{x})}, \tag{4}$$

where $\mathbf{J}(\mathbf{x})$ is the hydraulic gradient at the $\mathbf{x}$ point. Note that we used the convention that scalar attributes refer to the corresponding directional property. Because of this convention, the hydraulic head at a given location of the aquifer is a function of the directional transmissivity field, $T(\mathbf{x})$. On the basis of this condition and denoting $Y(\mathbf{x}) = \ln T(\mathbf{x})$, we can formally represent the head at a given $\mathbf{x}_i$ location as

$$h(\mathbf{x}_i) = F_i(Y), \tag{5}$$

where F_i is a functional that relates a given (directional) transmissivity field, $Y(\mathbf{x})$, to the head value at $\mathbf{x}_i$. We write $h(\mathbf{x}_i)$ in terms of the expansion

$$h(\mathbf{x}_i) = F_i(Y_0) + F_i^{(1)}(Y_0)\delta Y + \frac{1}{2!}F_i^{(2)}(Y_0)\delta Y^2 + \cdots, \tag{6}$$

where the operator $F_i^{(n)}(Y_0)$ is known as the nth-order Fréchet derivative of $F_i(Y)$ [e.g., *Milne*, 1980]. The first-order derivative is referred to simply as the Fréchet derivative. Realizing that $F_i(Y_0) = h_0$ in (6), the perturbation h' can be written as

$$h'(\mathbf{x}_i) = F_i^{(1)}(Y_0)\delta Y + O(\|\delta Y\|^2). \tag{7}$$

[13] If higher-order terms represented by $O(\|\delta Y\|^2)$ are neglected, the problem is linear and can be expressed as a convolution integral

$$h'(\mathbf{x}_i) = \int_{\Re^2} K_i(\mathbf{x}_i - \mathbf{x}, Y_0)\delta Y(\mathbf{x})d\mathbf{x}, \tag{8}$$

where $K_i(\mathbf{x}, Y_0)$ is the Fréchet kernel associated with the $h(\mathbf{x}_i)$ value. This kernel establishes the relationship between a small (first-order) perturbation in the transmissivity field, $Y(\mathbf{x})$, and the corresponding change in the head field.

[14] In a mildly heterogeneous aquifer (small variance of the natural log of transmissivity, σ_Y^2), we have that $Y' \approx \delta Y$ (i.e., the perturbation around some mean value), and a first-order approximation of h' is written as

$$h'(\mathbf{x}_i) = \int_{\Re^2} K_i(\mathbf{x}_i - \mathbf{x}, Y_0)Y'(\mathbf{x})d\mathbf{x}. \tag{9}$$

[15] Given the observation of the head response, $h(\mathbf{x}_i)$, at a given location, $\mathbf{x}_i$, two points are well connected in terms of flow when the head response is faster than expected from the equivalent homogeneous flow problem given by (2). That is to say, at pseudo steady state conditions (large times), one observes that $h(\mathbf{x}_i) < h_0(\mathbf{x}_i)$ or $h'(\mathbf{x}_i) < 0$. On the basis of this observation, we can consider the following indicator of flow connectivity between two points of the aquifer:

$$\omega'(\mathbf{x}_i, \mathbf{x}_p) = \frac{h'(\mathbf{x}_i)}{m_0}, \tag{10}$$

where negative ω' values denote strong point-to-point flow connectivity, and m_0 is a constant positive parameter that normalizes h'. Thus, the indicator of flow connectivity between two points can generally be written as

$$\omega'(\mathbf{x}_i, \mathbf{x}_p) = -\int_{\Re^2} U(\mathbf{x}_i, \mathbf{x}_p, \mathbf{x})Y'(\mathbf{x})d\mathbf{x}, \tag{11}$$

where

$$U(\mathbf{x}_i, \mathbf{x}_p, \mathbf{x}) = -m_0^{-1}K_i(\mathbf{x}_i - \mathbf{x}, Y_0) \tag{12}$$

is a weighting function of the natural log of transmissivity that establishes the relationship between flow connectivity and the departure of $Y(\mathbf{x})$ from the equivalent homogeneous solution.

[16] Analytical expressions of the Fréchet kernels can be derived using either the adjoint Green's function approach or the series expansion approach. Assuming radial flow conditions induced by a pumping well, *Sanchez-Vila et al.* [1999] derived the corresponding Fréchet kernels valid at large times. In this flow configuration, the indicator of point-to-point flow connectivity, ω', has an interesting meaning. By normalizing the pressure head in (10) by

$$m_0 = \frac{Q_w}{4\pi T_0}, \tag{13}$$

they found that ω' expresses the ratio between the estimated storage coefficient obtained by interpreting the late-time

behavior of drawdown data using Jacob's method and the true aquifer storage coefficient [*Sanchez-Vila et al.*, 1999],

$$\omega' = \ln\left(\frac{S_{\text{est}}}{S}\right), \tag{14}$$

and the actual expression in terms of Fréchet kernels is

$$\omega'(r_i) = -\int_{\Re^2} U(r,\theta)Y'(r,\theta)dV, \tag{15}$$

$$U(r,\theta) = \frac{r - r_i\cos\theta}{r(r^2 + r_i^2 - 2rr_i\cos\theta)}, \tag{16}$$

where (r, θ) are the cylindrical coordinates centered at the pumping well, and r_i is the radial distance between the pumping and the observation wells. For transient flow conditions, *Knight and Kluitenberg* [2005] derived explicit analytical expressions of the Fréchet kernels for both pumping and slug tests.

[17] For problems where it is not possible to derive an explicit expression of the Fréchet kernels, one can also parameterize the $Y(\mathbf{x})$ field to numerically solve for the sensitivities using (1) a first-order perturbation approach, (2) the sensitivity equation, or (3) the adjoint-equation method. In these cases, it is convenient to know the relationship between Fréchet kernels and the traditional sensitivity coefficients. By representing the continuous random field $Y(\mathbf{x})$ by a finite set of attributes associated with a support volume Ω, $\{Y_i,\ i = 1,\ \ldots,\ M\}$, given, for example, by the elements of a numerical mesh, equation (9) is reduced to

$$h'(\mathbf{x}_i) = \sum_{k=1}^{M}\left[\int_{\Omega_i} K_i(\mathbf{x}_i - \mathbf{x}, Y_0)d\mathbf{x}\right]Y_k'. \tag{17}$$

[18] Since the Taylor series expansion of h can also be written as

$$h'(\mathbf{x}_i) = \sum_{k=1}^{M}\frac{\partial F_i(Y_0)}{\partial Y_k}Y_k', \tag{18}$$

it follows that

$$\frac{\partial F_i(Y_0)}{\partial Y_k} = \int_{\Omega_i} K_i(\mathbf{x}_i - \mathbf{x}, Y_0)d\mathbf{x}. \tag{19}$$

[19] If the support volume Ω_i associated with Y_i is sufficiently small so that the integral depends strongly on $K_i(\mathbf{x}_i)$ within Ω_i, then

$$K_i(\mathbf{x}_i, Y_0) \approx \frac{1}{\Omega_i}\frac{\partial F_i(Y_0)}{\partial Y_k}, \tag{20}$$

which expresses that U is proportional to the sensitivity of heads with respect to Y per unit of aquifer volume.

2.2. Transport Connectivity Between Two Points

[20] We aim at quantifying the travel time, t_a, between two points $\mathbf{x}_i$ and $\mathbf{x}_p$ that are actually connected through a streamline φ. The streamline located between $\mathbf{x}_i$ and $\mathbf{x}_p$ is denoted as φ_{ip} and is defined parametrically by $\mathbf{x} = \mathbf{x}(s)$, where s is the arc length of the streamline. By convention, we denote $\mathbf{x}(s)$ simply as $\mathbf{x}_s$, and the value of an attribute or variable at a given streamline location, $\mathbf{x}_s$, is denoted as $a(\mathbf{x}_s) = a(s)$. On the basis of this notation, the characteristic travel time between $\mathbf{x}_i$ and $\mathbf{x}_p$ is

$$t_a = \int_{\varphi_{ip}}\frac{1}{v(s)}ds, \tag{21}$$

where $v(s) = \|\mathbf{v}(\mathbf{x}_s)\|$ is the flow velocity at the $\mathbf{x}_s$ streamline position. By Darcy's law, the magnitude of the groundwater velocity in the direction of flow is

$$v(s) = \frac{T(s)}{b\phi}J(s), \tag{22}$$

where ϕ is the aquifer porosity, b is the aquifer thickness, $J(s) = -dh/ds$ is the hydraulic gradient in the direction of flow, and $T(s)$ is the directional transmissivity.

[21] We start by decomposing the velocity field, $v(s)$, around its equivalent homogeneous velocity field, $v_0(s)$, so that

$$v(s) = v_0(s) + v'(s), \tag{23}$$

where

$$v_0(s) = \frac{T_0(s)}{b\phi}J_0(s), \tag{24}$$

$$v'(s) = v_0(s)\left(e^{Y'(s)}\frac{J(s)}{J_0(s)} - 1\right). \tag{25}$$

[22] Taylor series expansion of the travel time around $v_0(s)$ gives

$$t_a = \int_{\varphi_{ip}}\frac{1}{v_0(s)}\left[1 - \frac{v'(s)}{v_0(s)} + \left(\frac{v'(s)}{v_0(s)}\right)^2 + \cdots\right]ds. \tag{26}$$

[23] Introducing (24) and (25) into (26), and expanding $e^{Y'(s)} = 1 + Y'(s) + 1/2!Y'^2(s) + \cdots$, we have

$$t_a = \int_{\varphi_{ip}}\frac{1}{v_0(s)}\left(1 - Y'(s) - \frac{J'(s)}{J_0(s)} + \cdots\right)ds. \tag{27}$$

[24] In practice, t_a is ill defined because it depends on the true streamline, which is largely uncertain owing to the scarce knowledge of the aquifer. To overcome this limitation, we approximate the true streamline curve, φ, by its homogeneous counterpart, ψ, defined by the parameterization $\mathbf{x} = \mathbf{x}_0(s_0)$, where s_0 is the arc length of ψ. This streamline is known from the solution of the equivalent homogeneous flow problem given by (2). Thus,

$$Y'(s) = Y'(s_0) + \frac{dY'(s_0)}{ds}s' + \cdots, \tag{28}$$

$$\frac{1}{v_0(s)} = \frac{1}{v_0(s_0)} - \frac{1}{v_0^2(s_0)}\frac{dv_0(s_0)}{ds}s' + \cdots, \tag{29}$$

$$\frac{J'(s)}{J_0(s)} = \frac{J'(s_0)}{J_0(s_0)} + \frac{1}{J_0(s)}\frac{dJ'(s_0)}{ds}s' - \frac{J'(s_0)}{J_0^2(s)}\frac{dJ_0(s_0)}{ds}s' + \cdots, \tag{30}$$

and we obtain the following first-order approximation of the travel time response observed at the $\mathbf{x}_p$ location:

$$t_a \approx \int_{\psi_{ip}} \frac{1}{v_0(s_0)} \left(1 - Y'(s_0) - \frac{J'(s_0)}{J_0(s_0)}\right) ds_0. \tag{31}$$

[25] Second-order expansions accounting for uncertainty in Y under uniform and convergent flow conditions were presented by *Rajaram* [1997] and *Riva et al.* [2006b], respectively. The first term in (31) is the zero-order approximation of the travel time, defined as

$$t_{a_0} = \int_{\psi_{ip}} \frac{1}{v_0(s_0)} ds_0. \tag{32}$$

[26] Integrating the third term in (31) by parts, we can express the travel time as a function of the normalized head response $\omega'(s_0)$ (flow connectivity) and $Y'(s_0)$,

$$t_a \approx t_{a_0} \left(1 - \int_{\psi_{ip}} W_0(s_0) Y'(s_0) ds_0 + \int_{\psi_{ip}} W_1(s_0) \omega'(s_0) ds_0 - W_2(s_i)\omega'(s_i) + W_2(s_p)\omega'(s_p)\right), \tag{33}$$

where W_0 is a weighting function of the fluctuations of the natural log of transmissivity (deviations from the equivalent homogeneous value) and W_1 and W_2 are the weighting functions associated with flow connectivity,

$$W_0(s_0) = \frac{1}{t_{a_0} v_0(s_0)}, \tag{34}$$

$$W_1(s_0) = -\frac{1}{t_{a_0}} \frac{d}{ds_0} \left(\frac{m_0}{v_0 J_0}\right), \tag{35}$$

$$W_2(s_0) = \frac{m_0}{t_{a_0} v_0(s_0) J_0(s_0)}. \tag{36}$$

[27] Equation (33) expresses that the travel time response observed at a given location deviates from its expected equivalent homogeneous value because of the interplay of the fluctuations of the natural log of transmissivity and flow connectivity sampled along the solute particle path. The weighting functions W_1 and W_2 are related through

$$\frac{dW_2}{ds_0} = -W_1. \tag{37}$$

[28] The series expansion in (33) is truncated at the first order in σ_Y. In an attempt to overcome the latter truncation, we view (33) as the first two terms of an exponential expansion and we generalize the final solution as

$$t_a \approx t_{a_0} \exp\left(- \int_{\psi_{ip}} W_0(s_0) Y'(s_0) ds_0 + \int_{\psi_{ip}} W_1(s_0)\omega'(s_0) ds_0 - W_2(s_i)\omega'(s_i) + W_2(s_p)\omega'(s_p)\right). \tag{38}$$

[29] This exponentiation is a widely used procedure in stochastic subsurface hydrology [e.g., *Gelhar and Axness*, 1983]. Otherwise, the estimation of t_a through (33) may yield negative values, which are not physically possible. Because the travel time t_{a_0} depends on the initial and final position of the streamline ψ_{ip}, it renders the weighting functions $\{W_0, W_1, W_2\}$ to be not only a function of the position along the particle path $\mathbf{x}_{s_0} = \mathbf{x}(s_0)$ but also a function of the initial and final positions of the streamline, ψ_{ip}. For clarity hereinafter, this dependency was included in the expressions of the weighting functions, so $W_i(s_0) \equiv W_i(\mathbf{x}_i, \mathbf{x}_p, \mathbf{x}_{s_0})$ $(i = 0, 1, 2)$.

[30] Two points of the aquifer, $\mathbf{x}_i$ and $\mathbf{x}_p$, are well connected in terms of transport when the observed travel time response is more rapid than that expected from the equivalent homogeneous problem. On the basis of this conceptualization of transport connectivity, the following indicator of point-to-point transport connectivity between $\mathbf{x}_i$ and $\mathbf{x}_p$ can be defined:

$$\tau'(\mathbf{x}_i, \mathbf{x}_p) = \ln\left(\frac{t_a}{t_{a_0}}\right), \tag{39}$$

where negative τ' values denote strong transport connectivity between $\mathbf{x}_i$ and $\mathbf{x}_p$ and vice versa. Substituting (38) into (39), we obtain

$$\tau'(\mathbf{x}_i, \mathbf{x}_p) \approx - \int_{\psi_{ip}} W_0(\mathbf{x}_i, \mathbf{x}_p, \mathbf{x}_{s_0}) Y'(\mathbf{x}_{s_0}) ds_0 + \int_{\psi_{ip}} W_1(\mathbf{x}_i, \mathbf{x}_p, \mathbf{x}_{s_0}) \omega'(\mathbf{x}_{s_0}, \mathbf{x}_p) ds_0 - W_2(\mathbf{x}_i, \mathbf{x}_p, \mathbf{x}_i)\omega'(\mathbf{x}_i, \mathbf{x}_p) + W_2(\mathbf{x}_i, \mathbf{x}_p, \mathbf{x}_p)\omega'(\mathbf{x}_p, \mathbf{x}_p). \tag{40}$$

[31] From (40) we see that transport connectivity tends to increase (leading to negative τ' values) because of two joint effects: (1) transmissivity values along the particle path are larger than T_0 (the equivalent field transmissivity), so $Y' > 0$; and (2) the hydraulic response along the particle path is faster (small ω') than that attributed to the homogeneous solution. The last two terms in (39) reflect the boundary conditions of flow connectivity at the injection and at the observation location, $\mathbf{x}_i$ and $\mathbf{x}_p$.

3. Analytical Solutions of Transport Weighting Functions

[32] Analytical solutions of the weighting functions W_i can easily be obtained for relatively simple flow systems. Let us consider an infinite confined aquifer. The aquifer is heterogeneous and characterized by an equivalent isotropic transmissivity tensor written as $\mathbf{T}_0 = T_0\mathbf{I}_d$. The flow system is at steady state and determined by the existence of a single pumping well under the influence of a background natural gradient, J_n. The natural gradient is oriented parallel to the x direction such that $\partial h/\partial x > 0$. The well has a small radius ($r_w \to 0$), is located at the origin of coordinates ($\mathbf{x}_p = \mathbf{0}$), and pumps water at a constant rate Q_w. Under these conditions and by the principle of superposition, the velocity field and the travel times associated

with an equivalent homogeneous medium are given by *Bear and Jacobs* [1965]:

$$v_0(x,y) = \sqrt{v_{x_0}^2(x,y) + v_{y_0}^2(x,y)} \tag{41}$$

$$v_{x_0}(x,y) = \frac{q_0}{\phi} + \frac{Q_w}{2\pi b\phi}\frac{x}{x^2+y^2}, \tag{42}$$

$$v_{y_0}(x,y) = \frac{Q_w}{2\pi b\phi}\frac{y}{x^2+y^2}, \tag{43}$$

$$J_0(x,y) = \sqrt{J_{x_0}^2(x,y) + J_{y_0}^2(x,y)}, \tag{44}$$

$$J_{x_0}(x,y) = J_n + \frac{Q_w}{2\pi T_0}\frac{x}{x^2+y^2}, \tag{45}$$

$$J_{y_0}(x,y) = \frac{Q_w}{2\pi T_0}\frac{y}{x^2+y^2}, \tag{46}$$

$$t_{a_0}(x_i,y_i) = \frac{\phi}{q_0}x_i + \frac{\phi Q_w}{2\pi b q_0^2}\ln\left(\frac{Q_w\sin\theta}{Q_w\sin\theta + 2\pi b q_0 y_i}\right), \tag{47}$$

where $v_0(x, y)$ and $J_0(x, y)$ are the velocity and hydraulic gradient at the (x, y) location, respectively; q_0 is the specific discharge defined as $q_0 = T_0J_0/b$; (x_i, y_i) is the solute injection location; and θ is the angle between the injection line and the x axis, that is, $\theta = \tan^{-1}(y_i/x_i)$. By substituting (41)–(47) into (34), the corresponding analytical expressions of the weighting functions associated with the transport connectivity indicator are written as

$$W_0(x,y) = \frac{1}{v_0(x,y)t_{a_0}(x_i,y_i)}, \tag{48}$$

$$W_1(x,y) = \frac{C_1Q_w}{t_{a_0}(x_i,y_i)v_0^3(x,y)(x^2+y^2)}, \tag{49}$$

$$W_2(x,y) = \frac{C_2}{v_0(x,y)J_0(x,y)t_{a_0}(x_i,y_i)}, \tag{50}$$

where C_1 and C_2 are constant values defined as $C_1 = m_0T_0/\pi\phi^2b^2$ and $C_2 = m_0$. Two limiting cases are worth noticing. Whenever the natural gradient J_n is negligible or the injection is near the well, we approach radial flow conditions and the weighting functions can be simply written as

$$W_0(r) = W_1(r) = C_3\frac{2r}{r_i^2}, \tag{51}$$

$$W_2(r) = C_3\frac{r^2}{r_i^2}, \tag{52}$$

where $C_3 = 4\pi T_0m_0/Q_w$. These solutions agree with the work of *Trinchero et al.* [2008], who already showed that under pumping conditions the weighting functions of the transport connectivity indicators are linearly increasing with the radial distance from the well.

[33] On the contrary, whenever the pumping rate Q_w is small or the injection is far away from the well, we get close to uniform flow conditions and, in the limit when $Q_w \to 0$, we obtain

$$W_0(x) = \frac{1}{x_i}, \quad W_1(x) = 0, \quad W_2(x) = \frac{C_2}{J_nx_i}. \tag{53}$$

[34] Here it is important to note the large impact that the flow system has on the transport connectivity indicators. By comparing (53) and (51) we see that, whereas in radial flow conditions the weighting functions W_0 and W_1 increase linearly with the radial distance from the pumping well (thus giving more weight to the transmissivity values far from the pumping well), in uniform flow conditions, all the transmissivity values along the streamline are equally important.

4. Conditional Expectation of Transport Connectivity Indicators

[35] From (40) an estimate of τ' would imply perfect knowledge of the spatial distribution of Y and ω'. Obviously, this is not possible, so we need to estimate τ' from an incomplete knowledge of the system. We wish to get the "best" estimate (and corresponding uncertainty) of τ' ($\mathbf{x}_0$, $\mathbf{x}_{p_0}$) from nearby measurements of (1) transmissivity, $Y(\mathbf{x}_i)$ ($i = 1, \ldots, N_Y$); (2) point-to-point transport connectivity, τ' ($\mathbf{x}_i$, $\mathbf{x}_{p_i}$) ($i = 1, \ldots, N_\tau$); and (3) point-to-point flow connectivity, ω' ($\mathbf{x}_i$, $\mathbf{x}_{p_i}$) ($i = 1, \ldots, N_s$). The data measurements of the indicators of flow and transport connectivity can be representative of different streamlines; for example, the estimates of ω' and τ' can be obtained from pumping and tracer tests performed at different wells. To include these measurements, the $\mathbf{x}_{p_i}$ location of a given streamline denotes the final location of the streamline starting at the $\mathbf{x}_i$ location.

[36] Essentially, we want to obtain the best unbiased linear estimator of τ' ($\mathbf{x}_0$, $\mathbf{x}_{p_0}$) of the form

$$\tau'_{\mathrm{CK}}(\mathbf{x}_0,\mathbf{x}_{p_0}) = \sum_i \lambda_i^Y Y_i + \sum_i \lambda_i^\tau \tau_i' + \sum_i \lambda_i^\omega \omega_i', \tag{54}$$

where τ'_{CK} ($\mathbf{x}_0$, $\mathbf{x}_{p_0}$) is the estimator of transport connectivity between $\mathbf{x}_0$ and $\mathbf{x}_{p_0}$, and by convention we denote $\tau_i' = \tau'(\mathbf{x}_i, \mathbf{x}_{p_i})$ and $\omega_i' = \omega'$ ($\mathbf{x}_i$, $\mathbf{x}_{p_i}$) as measurement values.

[37] The relative contribution of each observation is based on the spatial dependency of the attributes. To this end, we consider $Y(\mathbf{x}) = \ln T(\mathbf{x})$ to describe a correlated random function quantified by its expectation m_Y and the two-point covariance function, C^{yy} ($\mathbf{x}_i$, $\mathbf{x}_j$). The attributes ω' and τ' are linearly dependent on Y' and, therefore, can also be described as two correlated random functions with zero mean (i.e., $m_\tau = 0$ and $m_\omega = 0$).

[38] The weight coefficients of the predictor are obtained by requiring the unbiasedness condition while minimizing the variance of the estimator error, similar to the widely known cokriging interpolation method. Taking the expected value of (54), we get the unbiasedness constraint

$$\sum_i \lambda_i^Y m_Y = 0. \tag{55}$$

[39] The minimization of the variance of the estimator error, $\sigma^2_{CK} = E[(\tau'_{CK} - \tau')^2]$, under the unbiasedness constraint requires the minimization of the objective function L involving one Lagrangian parameter μ:

$$\frac{\partial L}{\partial \lambda_i^Y} = 0, \quad \frac{\partial L}{\partial \lambda_i^\tau} = 0, \quad \frac{\partial L}{\partial \lambda_i^\omega} = 0, \quad \frac{\partial L}{\partial \mu} = 0, \tag{56}$$

where

$$L(\lambda_i^Y, \lambda_i^\tau, \lambda_i^\omega, \mu) = \frac{1}{2} E\left[(\tau'_{CK} - \tau')^2\right] - \mu\left(\sum_i \lambda_i^Y m_Y\right). \tag{57}$$

[40] On the basis of this equation, the weight coefficients are determined by solving the following linear system of equations:

$$\begin{cases} \sum_\beta \sum_j \lambda_j^\beta E(Z'_{\alpha i} Z'_{\beta j}) - \mu m_\alpha = E(Z'_{\alpha i} \tau'_0) & \alpha = Y, \tau, \omega, \quad i = 1, \ldots, n \\ \sum_i \lambda_i^Y m_Y = 0, \end{cases}$$

where $\tau'_0 = \tau'_{CK}(\mathbf{x}_0, \mathbf{x}_{p_0})$, and the random variable $Z'_{\alpha i}$ is defined as

$$Z'_{\alpha i} = \begin{cases} Y'_i = Y_i - m_Y & \text{if } \alpha = Y, \\ \tau'_i = \tau'_i & \text{if } \alpha = \tau, \\ \omega'_i = \omega'_i & \text{if } \alpha = \omega. \end{cases} \tag{58}$$

[41] The minimized estimation variance of $\tau'_{CK}(\mathbf{x}_0, \mathbf{x}_{p_0})$ can then be written as

$$\sigma^2_{CK} = E(\tau'^2) - \sum_\alpha \sum_i \lambda_i^\alpha E(Z'_{\alpha i} \tau'_0), \tag{59}$$

where

$$E(Z'_{\alpha i} Z'_{\beta j}) = \begin{cases} E(Y'_i Y'_j) = C^{yy}(\mathbf{x}_i, \mathbf{x}_j) & \text{if } (\alpha = Y), \quad (\beta = Y), \\ E(Y'_i \tau'_j) = C^{y\tau}(\mathbf{x}_i, \mathbf{x}_j) & \text{if } (\alpha = Y), \quad (\beta = \tau'), \\ E(Y'_i \omega'_j) = C^{y\omega}(\mathbf{x}_i, \mathbf{x}_j) & \text{if } (\alpha = Y), \quad (\beta = \omega'), \\ E(\tau'_i \tau'_j) = C^{\tau\tau}(\mathbf{x}_i, \mathbf{x}_j) & \text{if } (\alpha = \tau'), \quad (\beta = \tau'), \\ E(\tau'_i \omega'_j) = C^{\tau\omega}(\mathbf{x}_i, \mathbf{x}_j) & \text{if } (\alpha = \tau'), \quad (\beta = \omega'), \\ E(\omega'_i \omega'_j) = C^{\omega\omega}(\mathbf{x}_i, \mathbf{x}_j) & \text{if } (\alpha = \omega'), \quad (\beta = \omega'). \end{cases} \tag{60}$$

[42] Typically, we have only a very limited amount of data. Thus, although crucial for the delineation of connectivity patterns, we cannot, in general, estimate the covariance functions between $\{Y', \tau'\}$, $\{Y', \omega'\}$, $\{\tau', \tau'\}$, $\{\tau', \omega'\}$, and $\{\omega', \omega'\}$ because too few pumping and tracer tests are available at a given site. Here we overcome this problem by employing the approximate analytical solution of τ' to express all the required covariance functions as a weighting function of C^{yy}. Explicit expression of the covariance functions are given in Appendix A.

5. Stochastic Simulation of Transport Connectivity Indicators

[43] On the basis of the estimates of point-to-point transport connectivity, one can easily generate alternative, equally probable realizations of τ' while honoring data values of (Y, τ', ω') at different locations. For simplicity, we only consider the sequential Gaussian simulation algorithm, but it could be extended to other simulation techniques such as simulating annealing or multiple-point geostatistics. For more details on the theory of the sequential Gaussian simulation of a random field we refer to the work of *Gómez-Hernández and Journel* [1993].

[44] This approach consists of drawing sequentially the value of a variable (τ') from its conditional probability density function. The attribute τ' is considered a multi-Gaussian random variable and thereby its conditional cumulative density function (CCDF) follows a Gaussian distribution. The mean and variance of the distribution are derived from the previous cokriging system with the only consideration that the conditioning must include not only all available data values of the attributes (Y, τ', ω') but also the previously simulated τ' values.

[45] In short, the algorithm proceeds as follows:

[46] 1. Define a random path that visits each node of the grid map with centroid location $\mathbf{u}$. At each node, retain neighboring information about (Y, τ', ω').

[47] 2. Use the cokriging system defined in section 4 to estimate the mean, $m_\tau(\mathbf{u})$, and variance, $\sigma^2_\tau(\mathbf{u})$, of the CCDF of τ' at the visiting location $\mathbf{u}$.

[48] 3. Draw a random number p from a uniform distribution $U(0, 1)$.

[49] 4. Calculate the simulated value as $\tau'(\mathbf{u}) = G^{-1}(p)\sigma_\tau(\mathbf{u}) + m_\tau(\mathbf{u})$, where $G^{-1}(p)$ is the inverse function of the standard Gaussian CDF.

[50] 5. Add the simulated value $\tau'(\mathbf{u})$ to the data set and go back to step 2 until the path has been completed.

6. Simplifying Approaches

[51] In problems requiring a detailed site characterization, such as in the operation of an underground radioactive waste repository, the covariance and cross-covariance functions of all variables can be directly inferred from data. However, in most practical situations, only a few hydraulic tests are available at a given site, so the approximation of the covariance functions given in Appendix A are needed. In those cases, the large number of space integrals involved in the approximation causes the computation of the covariance functions related to τ' and ω' to be highly computationally demanding.

[52] In an attempt to avoid this burden, we consider two simplifications of the transport connectivity indicators. These simplified expressions are also often needed in tackling many engineering problems, such as in the risk analysis of groundwater remediation strategies [*Bolster et al.*, 2009]. By disregarding the influence of the hydraulic response to transport connectivity, the following approximation can be used:

$$\tau'_I(\mathbf{x}_i, \mathbf{x}_p) = -\int_{\psi_{ip}} W_0(\mathbf{x}_i, \mathbf{x}_p, \mathbf{x}_{s_0}) Y'(\mathbf{x}_{s_0}) ds_0, \tag{61}$$

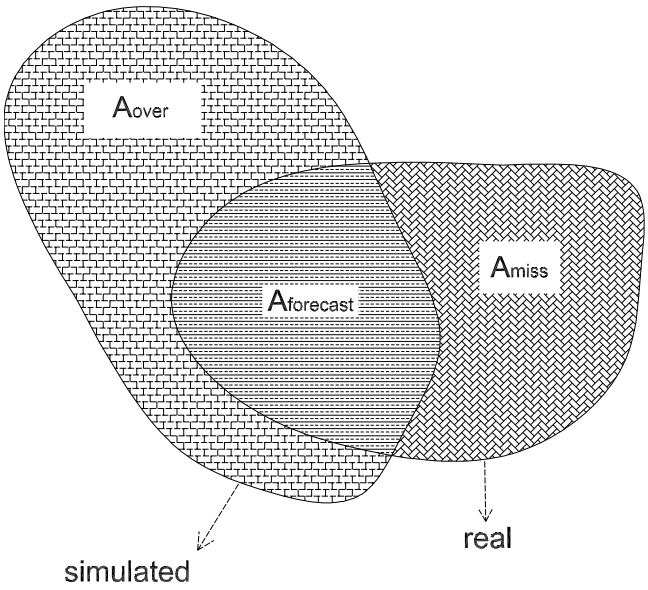

Figure 1. Definition of the captured areas used to determine the performance indices (see equation (64)).

which greatly simplifies equation (40). This simplification was successfully used by *Trinchero et al.* [2008] to delineate time-related capture zones of pumping wells. These authors showed that this approximation causes only a slight overestimation of the low-connectivity patterns, leading to smaller capture zones than desirable in areas of low connectivity. In this context, the following overconservative simplification can be used:

$$\tau'_{cz}(\mathbf{x}_i, \mathbf{x}_p) = \begin{cases} \tau'_I(\mathbf{x}_i, \mathbf{x}_p), & \tau' > 0, \\ 0, & \tau' \le 0. \end{cases} \tag{62}$$

[53] By a conservative approach, we mean that the approximation tends to only slightly underestimate transport connectivity. This tendency is shown in section 7 for pumping conditions, where we compare the captured area given by (61) and (62) with the numerical solution obtained from particle tracking simulations in a Monte Carlo framework.

7. Example of Application

[54] This section presents an illustrative example of the method developed to generate transport connectivity maps. In particular, the example deals with a day-to-day application of the method in which scarce knowledge of the aquifer calls for the combined use of the simplifying approaches and the approximation of the covariance functions of τ' and ω'. The focus here is on the delineation of point-to-point transport connectivity patterns induced by a single pumping well.

[55] We consider a synthetic heterogeneous aquifer whose domain is 101 × 101 units in length. The well is located at the center of the domain and pumps water out at a constant rate (i.e., $Q_w = 30\ L^3/T$). Flow is at steady state and a constant head is fixed at all outer boundaries. Heterogeneity is described by a spatially varying transmissivity, which is modeled as a correlated random function. All other aquifer properties are assumed constant. The distribution of the natural log of transmissivity in the aquifer describes a multi-Gaussian random function with zero mean and a variance varying between $\sigma_Y^2 = 1$ and $\sigma_Y^2 = 2$. The correlation structure is characterized by an isotropic spherical variogram model with an integral scale of 20 units length.

[56] The performance of the method is evaluated using two different approaches: by visual inspection of transport connectivity maps (i.e., comparison of the spatial distribution of τ'_I with its true solution) and by analyzing the behavior of appropriate performance indices. For any given realization of Y, the true solution of point-to-point transport connectivity was numerically calculated via backward particle tracking simulations. To perform the simulations, the simulated steady state flow velocities obtained using a finite difference code, MODFLOW-2000 [*Harbaugh et al.*, 2000], were used in a random walk code, RW3D [*Fernàndez-Garcia et al.*, 2005; *Salamon et al.*, 2006], to simulate backward travel times of particles initially released at the pumping well using a purely advective model. Transport connectivity was then related to the travel time t_a through

$$t_a(\mathbf{x}_i, \mathbf{x}_w) = \frac{\pi r_i^2 b \phi}{Q_w} \exp \tau'(\mathbf{x}_i, \mathbf{x}_w). \tag{63}$$

[57] The chosen performance indices quantify the error produced when forecasting well capture zones obtained from (63). Thus, the analysis evaluates the forecast of the time-related captured zone of 50 days, which is a widely used travel time for designing safeguard zones of an abstraction well. The performance indices are defined as

$$e_{\text{miss}} = \frac{A_{\text{fore}}}{A_{\text{real}}}, \qquad e_{\text{over}} = \frac{A_{\text{over}}}{A_{\text{sim}}}, \tag{64}$$

where A_{fore} is the area of the real capture zone that is correctly predicted, A_{real} is the real captured area, A_{sim} is the area simulated by the model, and A_{over} is the nonreal area that is predicted by the model. Figure 1 illustrates the definition of the different capture areas used to determine e_{miss} and e_{over}. In Figure 1, A_{miss} is the area of the real capture zone that is not identified by the forecast, so $A_{\text{real}} = A_{\text{miss}} + A_{\text{fore}}$. Thus, the first index e_{miss} is the percentage of the real capture zone that has not been identified, whereas e_{over} is the percentage of the forecasted area that is unnecessarily protected. Similar performance indices were used by *Kerrou et al.* [2008].

[58] The example fulfills two main objectives. On the one hand, in section 7.1, we use the synthetic aquifer to evaluate the performance of the simplified expressions of the transport connectivity indicators; on the other hand, the synthetic aquifer is employed in section 7.2 to illustrate the generation of conditional point-to-point transport connectivity maps honoring scarce data obtained from different sources.

7.1. Assessment of τ'_I and τ'_{cz}

[59] The performance of the simplified expression of the connectivity indicators given in section 6, τ'_I and τ'_{cz}, are evaluated using Monte Carlo simulations. We consider 200 realizations of the transmissivity field generated through a sequential Gaussian simulation program, GCOSIM3D [*Gómez-Hernández and Journel*, 1993]. For any given transmissivity field, the estimation of τ'_I and τ'_{cz} was performed numerically using one quadrature of the integrals defined in (61) and (62). The estimated connectivity values

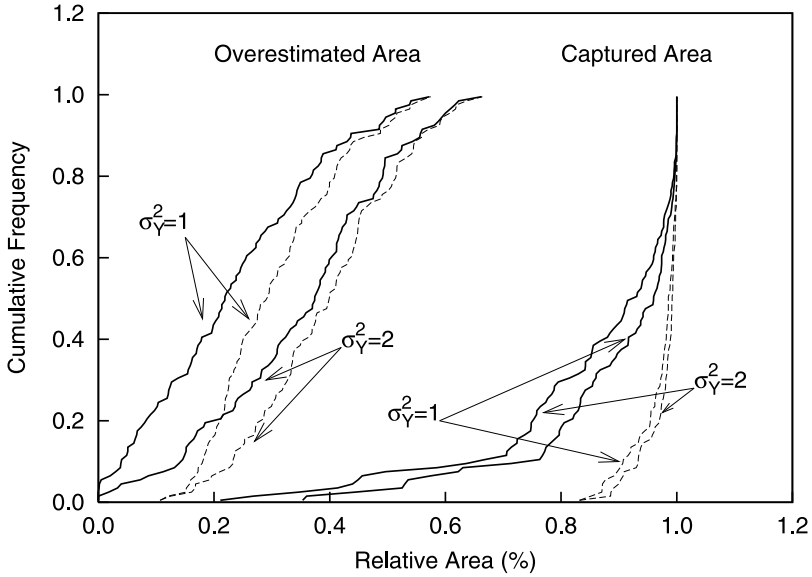

Figure 2. Cumulative distribution function of the performance indices obtained for $\sigma_Y^2 = 1$ and $\sigma_Y^2 = 2$. Thin lines correspond to the capture zone determined by (61), whereas dashed lines correspond to the capture zone estimated with the simplified overconservative approach given by (62).

were then used to calculate the related capture zones and performance indices obtained through (63) and (64). Here the assessment of τ_I' and τ_{cz}' assumes that the spatial distribution of the transmissivity field is completely known (without conditioning).

[60] Results presented in Figure 2 show the cumulative distribution function of the performance indices. Here thin lines correspond to the estimation of τ' by τ_I' (61), whereas dashed lines corresponds to τ_{cz}' given by (62). Remarkably, the fact that the approximation of transport connectivity τ_I' and τ_{cz}' provides relatively good predictions of the captured area for both degrees of heterogeneity, $\sigma_Y^2 = 1$ and $\sigma_Y^2 = 2$, indicates that the simplified transport connectivity indicators are rather robust for the flow condition examined. Notice, for instance, that 80% of the simulations corresponding to τ_I' were able to predict more than 80% of the real captured area. Furthermore, as expected by its definition which disregards low connectivity values, the use of the approximation $\tau' \approx \tau_{cz}'$ substantially improves the forecast of the captured area, which is now almost fully captured in most of the realizations. In this process, it is important to note that the overestimated area given by τ_{cz}' is still rather similar to that of τ_I'.

[61] In principle, the analytical solutions developed in section 2 are only valid for small degrees of heterogeneity (i.e., $\sigma_Y^2 < 1$). In this context, we note that because these results were obtained for $\sigma_Y^2 = 1$ and $\sigma_Y^2 = 2$, it adds confidence to the applicability of our analytical solutions to field problems with larger σ_Y^2.

7.2. Stochastic Simulation of Transport Connectivity

[62] We now illustrate the generation of conditional point-to-point transport connectivity maps honoring data obtained from different sources. For this purpose, an arbitrary individual aquifer realization of the random field with $\sigma_Y^2 = 1$ was chosen to represent the true aquifer. On the basis of this realization, we attempt to infer the transport connectivity patterns when only a limited amount of information is known. Here the only information required to generate conditional transport connectivity maps is the covariance function of Y, the location and features of the pumping well, and the data. Connectivity values are simulated over a regular grid formed by square cells of unit size. The support volume of the local transmissivity measurements is represented by the size of the grid cells. Because the example focuses on radial flow conditions, the transport weighting functions W_i are known analytically from (51) and (52). Yet we note that the methodology is general and that, for more complex flow systems, the weighting functions can also be directly computed by means of (20) and (34).

[63] The example is designed to evaluate the worth of including tracer data in a capture zone that has been designed using only available local transmissivity data. To perform this evaluation, the following scenarios were considered: (A) 16 equally spaced measurements of Y are known; (B) 16 equally spaced measurements of Y and 16 measurements of τ' are known at the same exact location (i.e., Y_i and τ'_i share the same location); (C) 16 equally spaced measurements of Y and 16 measurements of τ' are known at different locations; and (D) 100 equally spaced measurements of Y are known.

[64] Figure 3 compares the true transport connectivity map, numerically generated using particle tracking, with those obtained from the stochastic estimation and one Gaussian simulation of τ_I' in scenario D. Remarkably, both the stochastic estimation and the simulation of connectivity can be used to identify the three highly connected zones. As expected, the estimation produces globally smooth surfaces, while a better representation of the natural variability of connectivity is provided by the stochastic simulation.

[65] The method was then tested using the previously defined performance indices. The simulated area A_{sim} was

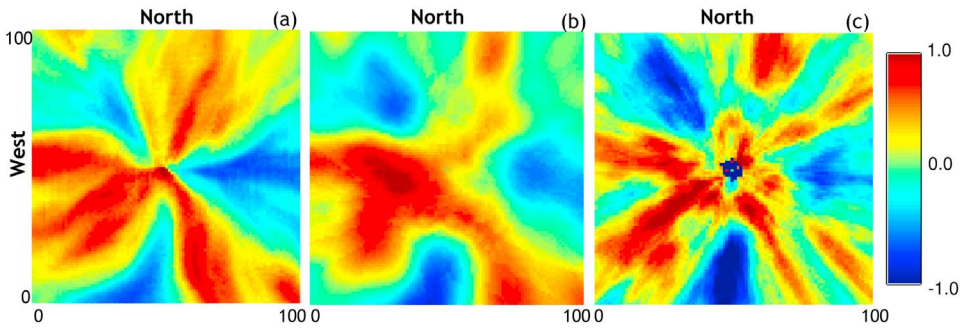

Figure 3. Visual comparison of (a) the real connectivity map obtained from particle tracking simulations with (b) the connectivity map obtained through stochastic estimation of τ'_I (scenario D) and (c) one realization of the stochastic simulation of connectivity τ'_I (scenario D).

obtained by the stochastic generation of τ'_I over 300 realizations, while the true captured area A_{real} was determined from backward particle tracking simulations. Figure 4 shows the cumulative distribution function of e_{miss} and e_{over} for all scenarios. As expected, when more data are available, the capture zone is better represented; the area not identified by the forecast e_{over} and its uncertainty (dispersion within the e_{fore} data) is smaller. This is clearly the case for the overestimated area (Figure 4a) but not for the predicted one (Figure 4b). In this case, the addition of a few τ'_i data values (scenarios B and C, Figure 4b) is shown to provide more information on connectivity than adding many Y_i values (scenario D).

[66] By looking at scenarios B and C in Figure 4, we see that when tracer data belong to a different location than that of the measurements of transmissivity a more faithful delineation of the capture zone is simulated. Notice that Figure 4 exhibits an important reduction of e_{fore} for the same degree of overestimation, e_{over}. This behavior is a direct consequence of the shape of the weighting function W_0, which in this case gives larger weights to the transmissivity data close to the injection location. Thus, introducing tracer test data at points where transmissivity is already known is found to be somewhat redundant.

[67] Figure 5 shows the isoprobability contour lines of the capture zone for all the scenarios considered. The isoprobability maps were calculated by estimating, at every centroid cell of the domain, the frequency at which the travel time is smaller than 50 days. In scenario A, where only 16 transmissivity measurements are known, the 0.5 isoline of the capture zone is almost circular, being close to the homogeneous solution (Figure 5a). When tracer data and transmissivity measurements share the same location (scenario B), a little improvement in the delineation of the capture zone is obtained. The reason is that, under convergent flow conditions, transport connectivity is mostly affected by the transmissivity data near the injection location and, therefore, transmissivity measurements and tracer data provide almost the same basic information. On the contrary, when tracer data and transmissivity measurements sample different areas (scenario C), the 0.5 isoline of the simulated capture zone approximates the real one, and the prediction is also less uncertain (Figure 5c).

[68] Whenever the number of transmissivity measurements available increases by an order of magnitude (scenario D), the accuracy of the method also increases dramatically (Figure 5d) and a good agreement is observed between the 0.5 isoline and the real perimeter of the capture zone. Nevertheless, those parts of the protection perimeter that correspond to highly connected zones still show high uncertainty, probably because of the lack of tracer test data in scenario D.

8. Conclusions

[69] We presented a general method for the stochastic simulation of point-to-point transport connectivity honoring data from three different sources: (1) travel time estimates obtained from field tracer tests, (2) estimates of flow connectivity indicators obtained from the relatively fast or slow response that is observed at a given location provided a flow impulse at another location, and (3) measurements of transmissivity at a local scale. The method can be used with generic flow configurations and it can easily integrate data obtained from different hydraulic tests, each one sampling different areas within the aquifer. In doing this, the following main findings are highlighted.

[70] 1. Point-to-point flow and transport connectivity are conceptually different but related. We generalized the mathematical approximate relationship between point-to-point transport connectivity and point-to-point flow connectivity given by *Trinchero et al.* [2008] for converging flow conditions to support other flow configurations and problem setups.

[71] 2. Point-to-point flow connectivity is seen as a weighted integral of transmissivity over the entire domain. Generally speaking, the weighting function is given by the sensitivity of heads with respect to the natural log of transmissivity per unit of aquifer volume. On the contrary, the point-to-point transport connectivity is a weighted integral along the particle path of the solute mass that involves two variables: transmissivity and flow connectivity. Each variable has its own weighting function. The weighting function of transmissivity is inversely proportional to the travel time and sampled velocity along the particle path.

[72] 3. For uniform flow conditions the weighting functions of the transport connectivity indicator assign equal

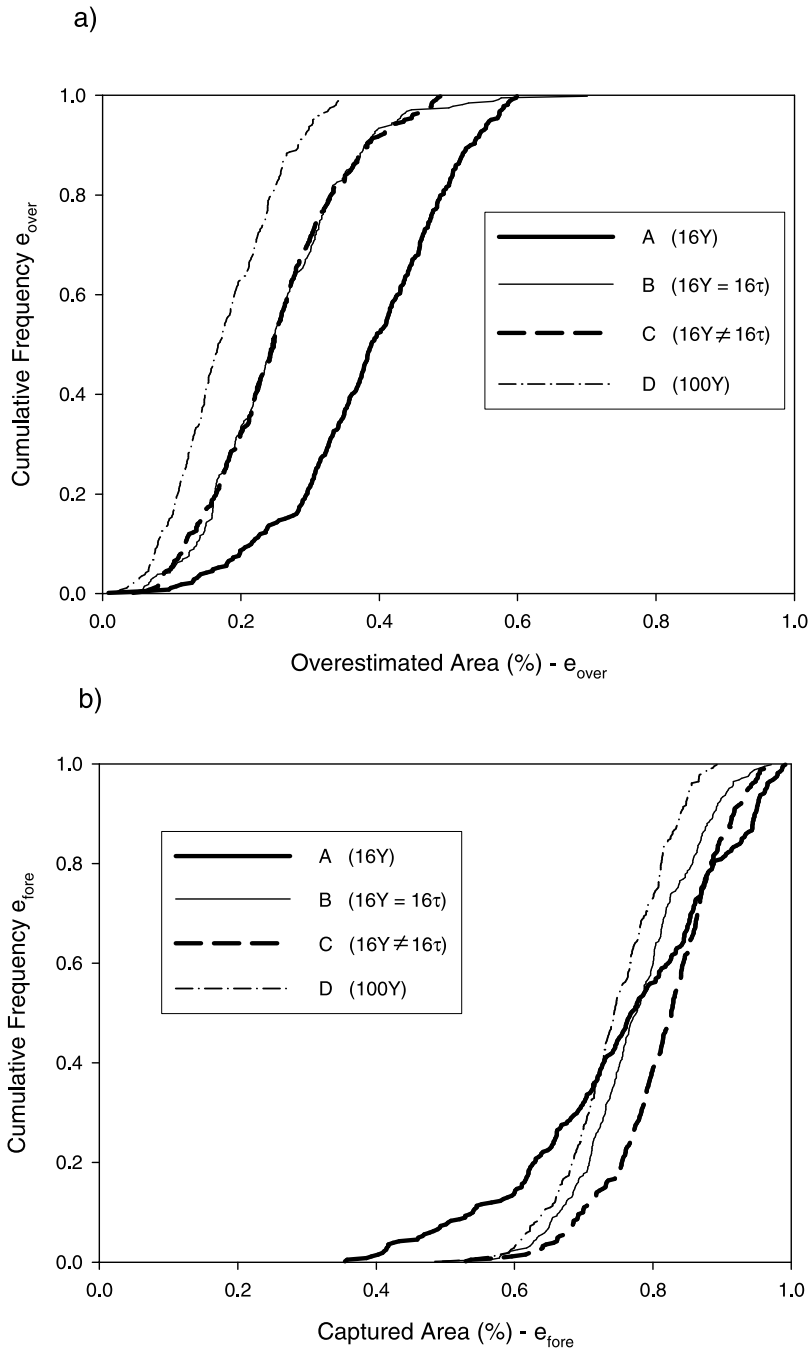

Figure 4. Cumulative distribution function of (a) e_{over} and (b) e_{fore} obtained over 300 conditional sequential Gaussian simulations of point-to-point transport connectivity for the four different scenarios considered.

weight to all the transmissivity values along the particle path. On the contrary, in radial flow conditions the transmissivity values farthest from the pumping well are the most influential.

[73] 4. Simplified expressions of point-to-point transport connectivity are provided to efficiently delineate time-related capture zones in heterogeneous aquifers. The approach allows one to easily incorporate flow connectivity measures (hydraulic response) obtained from pumping tests as well as transport connectivity measures (travel times) obtained from forced-gradient tracer tests into the delineation of capture zones. Monte Carlo simulations in moderately heterogeneous aquifers with $\sigma_Y^2 = 1$ and $\sigma_Y^2 = 2$ showed that the simplified expressions yield adequate estimates of the corresponding time-related capture zones, where τ'_{cz} is the most conservative one.

[74] 5. A stochastic framework conducive to generating conditional maps of point-to-point transport connectivity honoring data obtained from different sources is provided. Thanks to the general expression of point-to-point flow and

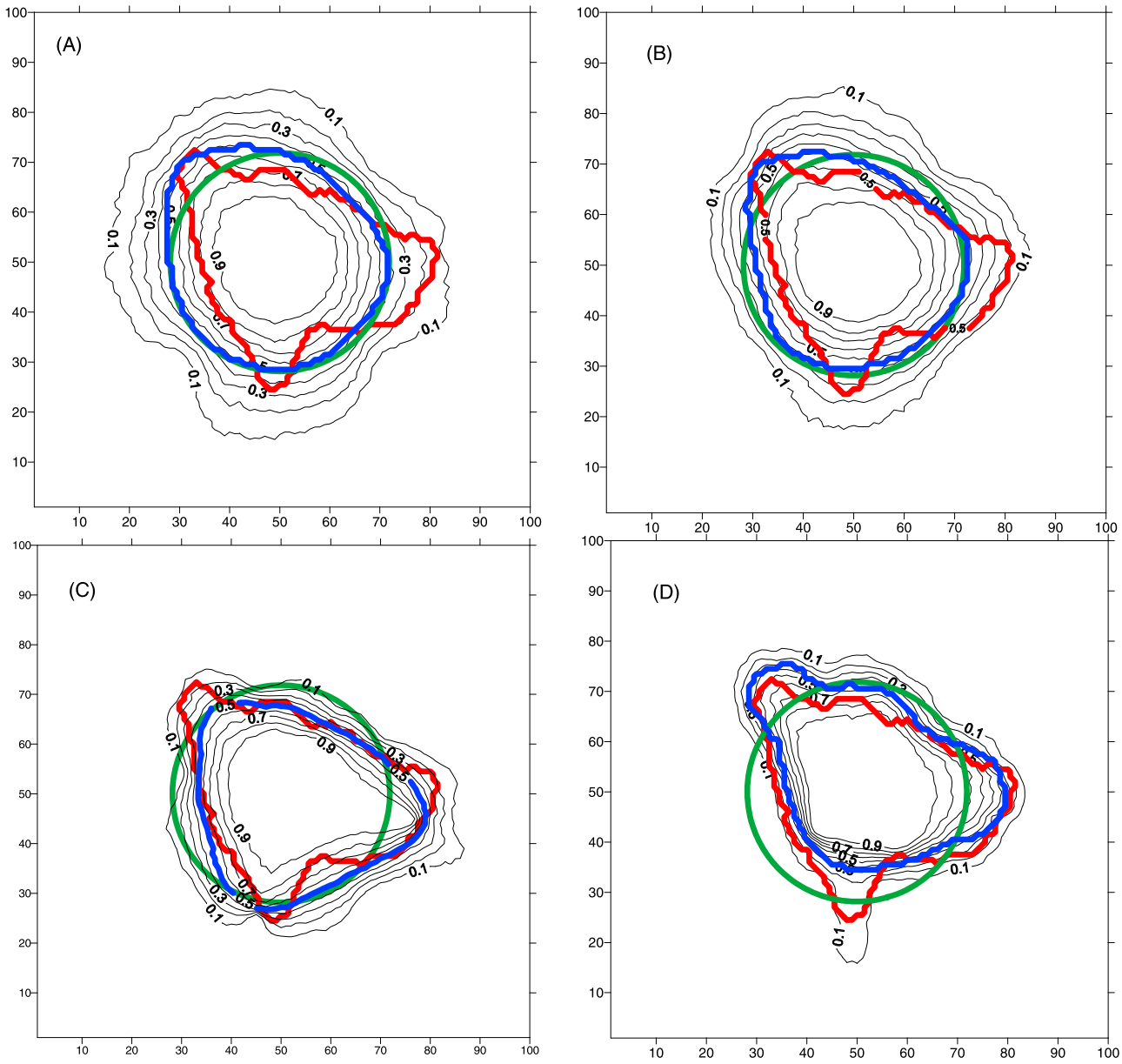

Figure 5. Isoprobability contour lines of the capture zone induced by a pumping well (thin lines) obtained from conditional stochastic simulations of connectivity τ'_i: red line, true capture zone numerically simulated through particle tracking; green line, capture zone of an equivalent homogeneous medium; blue line, 0.5 isoline.

transport connectivity, the method avoids the inference of cross-covariance functions between variables measured over different scales (which cannot be otherwise estimated because there are typically not enough data) by expressing them in terms of the covariance function of transmissivity.

[75] 6. An example of the method is provided to evaluate the worth of including tracer data in an otherwise transmissivity-based capture zone of an abstraction well. Several scenarios based on different available measures of transmissivity and travel times were envisioned. A visual inspection of the stochastically generated connectivity maps shows the ability of the method to identify the highly connected zones. Moreover, Monte Carlo simulations reveal that the worth of tracer data is a maximum whenever the travel time data stem from a tracer test performed at a different location than that of transmissivity measurements. The reason is that the weighting function of transmissivities associated with converging flow conditions gives larger weights to the injection location, making redundant the tracer test information whenever the injection is performed at a point where local transmissivity is known.

Appendix A: Covariance Functions

[76] Because too few pumping and tracer tests are available at a given site, we cannot typically estimate the covariance functions between $\{Y', \tau'\}$, $\{Y', \omega'\}$, $\{\tau', \tau'\}$, $\{\tau',$

$\omega'\}$, and $\{\omega', \omega'\}$. This problem can be overcome by employing the approximate analytical solution of τ' to express all the required covariance functions as a weighting function of C^{yy}. This is written as

$$C^{y\omega}(\mathbf{x}_i, \mathbf{x}_j) = -\int_{\Re^2} U(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}) C^{yy}(\mathbf{x}_i, \mathbf{x})\, d\mathbf{x}, \tag{A1}$$

$$\begin{aligned} C^{y\tau}(\mathbf{x}_i, \mathbf{x}_j) = &-\int_{\psi_{jp}} W_0(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0}) C^{yy}(\mathbf{x}_i, \mathbf{x}_{s_0}) ds_0 \\ &+ \int_{\psi_{jp}} W_1(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0}) C^{y\omega}(\mathbf{x}_i, \mathbf{x}_{s_0}) ds_0 \\ &- W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_j) C^{y\omega}(\mathbf{x}_i, \mathbf{x}_j) + W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{p_j}) \\ &\cdot C^{y\omega}(\mathbf{x}_i, \mathbf{x}_{p_j}), \end{aligned} \tag{A2}$$

$$\begin{aligned} C^{\omega\omega}(\mathbf{x}_i, \mathbf{x}_j) = &\int_{\Re^2} \int_{\Re^2} U(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}') U \\ &\times (\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}'') C^{yy}(\mathbf{x}', \mathbf{x}'')\, d\mathbf{x}'\, d\mathbf{x}'', \end{aligned} \tag{A3}$$

$$\begin{aligned} &C^{\tau\tau}(\mathbf{x}_i, \mathbf{x}_j) \\ &= \int_{\psi_{ip}} \int_{\psi_{jp}} W_0(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) W_0(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0''}) C^{yy}(\mathbf{x}_{s_0'}, \mathbf{x}_{s_0''}) ds_0' ds_0'' \\ &- \int_{\psi_{ip}} \int_{\psi_{jp}} W_0(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) W_1(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0''}) C^{y\omega}(\mathbf{x}_{s_0'}, \mathbf{x}_{s_0''}) ds_0' ds_0'', \\ &+ \int_{\psi_{ip}} W_0(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_j) C^{y\omega}(\mathbf{x}_{s_0'}, \mathbf{x}_j) ds_0' \\ &- \int_{\psi_{ip}} W_0(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{p_j}) C^{y\omega}(\mathbf{x}_{s_0'}, \mathbf{x}_{p_j}) ds_0' \\ &- \int_{\psi_{ip}} \int_{\psi_{jp}} W_1(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) W_0(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0''}) C^{\omega y}(\mathbf{x}_{s_0'}, \mathbf{x}_{s_0''}) ds_0' ds_0'' \\ &+ \int_{\psi_{ip}} \int_{\psi_{jp}} W_1(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) W_1(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0''}) C^{\omega\omega}(\mathbf{x}_{s_0'}, \mathbf{x}_{s_0''}) ds_0' ds_0'' \\ &- \int_{\psi_{ip}} W_1(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_j) C^{\omega\omega}(\mathbf{x}_{s_0'}, \mathbf{x}_j) ds_0' \\ &+ \int_{\psi_{ip}} W_1(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{p_j}) C^{\omega\omega}(\mathbf{x}_{s_0'}, \mathbf{x}_{p_j}) ds_0' \\ &+ \int_{\psi_{jp}} W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_i) W_0(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0''}) C^{\omega y}(\mathbf{x}_i, \mathbf{x}_{s_0''}) ds_0'' \\ &- \int_{\psi_{jp}} W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_i) W_1(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0''}) C^{\omega\omega}(\mathbf{x}_i, \mathbf{x}_{s_0''}) ds_0'' \\ &+ W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_i) W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_j) C^{\omega\omega}(\mathbf{x}_i, \mathbf{x}_j) \\ &- W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_i) W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{p_j}) C^{\omega\omega}(\mathbf{x}_i, \mathbf{x}_{p_j}) \\ &- \int_{\psi_{jp}} W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{p_i}) W_0(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0''}) C^{\omega y}(\mathbf{x}_{p_i}, \mathbf{x}_{s_0''}) ds_0'' \\ &+ \int_{\psi_{jp}} W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{p_i}) W_1(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{s_0''}) C^{\omega\omega}(\mathbf{x}_{p_i}, \mathbf{x}_{s_0''}) ds_0'' \\ &- W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{p_i}) W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_j) C^{\omega\omega}(\mathbf{x}_{p_i}, \mathbf{x}_j) \\ &+ W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{p_i}) W_2(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}_{p_j}) C^{\omega\omega}(\mathbf{x}_{p_i}, \mathbf{x}_{p_j}), \end{aligned} \tag{A4}$$

$$\begin{aligned} C^{\tau\omega}(\mathbf{x}_i, \mathbf{x}_j) = &\int_{\psi_{ip}} \int_{\Re^2} W_0(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) U(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}) C^{yy}(\mathbf{x}_{s_0'}, \mathbf{x})\, d\mathbf{x}\, ds_0' \\ &- \int_{\psi_{ip}} \int_{\Re^2} W_1(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{s_0'}) U(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}) C^{\omega y}(\mathbf{x}_{s_0'}, \mathbf{x})\, d\mathbf{x}\, ds_0' \\ &+ \int_{\Re^2} W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_i) U(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}) C^{\omega y}(\mathbf{x}_i, \mathbf{x})\, d\mathbf{x} \\ &- \int_{\Re^2} W_2(\mathbf{x}_i, \mathbf{x}_{p_i}, \mathbf{x}_{p_i}) U(\mathbf{x}_j, \mathbf{x}_{p_j}, \mathbf{x}) C^{\omega y}(\mathbf{x}_{p_i}, \mathbf{x})\, d\mathbf{x}. \end{aligned} \tag{A5}$$

[77] **Acknowledgments.** The authors acknowledge the financial support provided by ENRESA, by the Agència de Gestió d'Ajuts Universitaris i de Recerca of the Catalan Government, and by the Spanish CICYT (projects PARATODO and HEROS). Paolo Trinchero acknowledges the financial support provided by the Torres Quevedo Programme of the Spanish Ministry of Science and Innovation (PTQ-09-01-00695).

References

Bear, J., and M. Jacobs (1965), On the movement of water bodies injected into aquifers, *J. Hydrol.*, *3*, 37–57.

Bolster, D., M. Barahona, M. Dentz, D. Fernandez-Garcia, X. Sanchez-Vila, P. Trinchero, C. Valhondo, and D. M. Tartakovsky (2009), Probabilistic risk analysis of groundwater remediation strategies, *Water Resour. Res.*, *45*, W06413, doi:10.1029/2008WR007551.

de Dreuzy, J.-R., P. Davy, and O. Bour (2001), Hydraulic properties of two-dimensional random fracture networks following a power law length distribution: 1. Effective connectivity, *Water Resour. Res.*, *37*, 2065–2078.

Fernàndez-Garcia, D., X. Sanchez-Vila, and T. H. Illangasekare (2002), Convergent-flow tracer tests in heterogeneous media: Combined experimental-numerical analysis for determination of equivalent transport parameters, *J. Contam. Hydrol.*, *57*(1–2), 129–145.

Fernàndez-Garcia, D., T. H. Illangasekare, and H. Rajaram (2005), Differences in the scale dependence of dispersivity and retardation factors estimated from forced-gradient and uniform flow tracer tests in three-dimensional physically and chemically heterogeneous porous media, *Water Resour. Res.*, *41*, W03012, doi:10.1029/2004WR003125.

Fernàndez-Garcia, D., G. Llerar-Meza, and J. Gómez-Hernández (2009), Upscaling transport with mass transfer models: Mean behavior and propagation of uncertainty, *Water Resour. Res.*, *45*, W10411, doi:10.1029/2009WR007764.

Fogg, G. (1986), Groundwater flow and sand body interconnectedness in a thick, multiple-aquifer system, *Water Resour. Res.*, *22*(5), 679–694.

Frippiat, C. C., T. H. Illangasekare, and G. A. Zyvoloski (2009), Anisotropic effective medium solutions of head and velocity variance to quantify flow connectivity, *Adv. Water Resour.*, *32*, 239–249.

Gelhar, L. W., and C. L. Axness (1983), Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, *19*(1), 161–180.

Gómez-Hernández, J., and X. Wen (1998), To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology, *Adv. Water Resour.*, *21*(1), 47–61.

Gómez-Hernández, J. J., and A. G. Journel (1993), Joint simulation of multi-Gaussian random variables, in *Geostatistics Tróia '92*, vol. 1, edited by A. Soares, pp. 85–94, Kluwer, Dordrecht, Netherlands.

Harbaugh, A., E. Banta, M. Hill, and M. McDonald (2000), MODFLOW-2000 the U.S. Geological Survey modular ground-water model-user guide to modularization concepts and the ground-water flow process, *Open File Rep. U. S. Geol. Surv., 00-92*, 121 pp.

Kerrou, J., P. Renard, H. J. H. Franssen, and I. Lunati (2008), Issues in characterizing heterogeneity and connectivity in non-multigaussian media, *Adv. Water Resour.*, *31*, 147–159.

Knight, J. H., and G. J. Kluitenberg (2005), Some analytical solutions for sensitivity of well tests to variations in storativity and transmissivity, *Adv. Water Resour.*, *28*, 1057–1075.

Knudby, C., and J. Carrera (2005), On the relationship between indicators of geostatistical, flow and transport connectivity, *Adv. Water Resour.*, *28*(4), 405–421.

Knudby, C., and J. Carrera (2006), On the use of apparent hydraulic diffusivity as an indicator of connectivity, *J. Hydrol.*, *329*(3–4), 377–389.

LaBolle, E. M., and G. E. Fogg (2001), Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, *Transp. Porous Media*, *42*(1–2), 155–179.

Meier, P. M., J. Carrera, and X. Sanchez-Vila (1998), An evaluation of Jacob's method for the interpretation of pumping tests in heterogeneous formations, *Water Resour. Res.*, *34*(5), 1011–1025.

Milne, R. D. (1980), *Applied Functional Analysis: An Introductory Treatment*, Pitman, London.

Poeter, E., and P. Townsend (1994), Assessment of critical flow path for improved remediation management, *Ground Water*, *32*(3), 439–447.

Rajaram, H. (1997), Time and scale dependent effective retardation factors in heterogeneous aquifers, *Adv. Water Resour.*, *20*(4), 217–230.

Riva, M., L. Guadagnini, A. Guadagnini, T. Ptak, and E. Martac (2006a), Probabilistic study of well capture zones distribution at the Lauswiesen field site, *J. Contam. Hydrol.*, *88*(1–2), 92–118.

Riva, M., X. Sánchez-Vila, A. Guadagnini, M. De Simoni, and M. Willmann (2006b), Travel time and trajectory moments of conservative solutes in two-dimensional convergent flows, *J. Contam. Hydrol.*, *82*(1–2), 23–43.

Salamon, P., D. Fernàndez-Garcia, and J. J. Gómez-Hernández (2006), A review and numerical assessment of the random walk particle tracking method, *J. Contam. Hydrol.*, *87*(3–4), 277–305.

Sanchez-Vila, X., and J. Carrera (1997), Directional effects on convergent flow tracer tests, *Math. Geol.*, *29*(4), 551–569.

Sanchez-Vila, X., J. Carrera, and J. Girardi (1996), Scale effects in transmissivity, *J. Hydrol.*, *183*(1), 1–22.

Sanchez-Vila, X., P. M. Meier, and J. Carrera (1999), Pumping tests in heterogeneous aquifers: An analytical study of what can be obtained from their interpretation using Jacob's method, *Water Resour. Res.*, *35*(4), 943–952.

Sanchez-Vila, X., A. Guadagnini, and J. Carrera (2006), Representative hydraulic conductivities in saturated groundwater flow, *Rev. Geophys.*, *44*, RG3002, doi:10.1029/2005RG000169.

Strebelle, S. (2002), Conditional simulation of complex geological structures using multiple-point statistics, *Math. Geol.*, *34*(1), 1–21.

Trinchero, P., X. Sanchez-Vila, and D. Fernàndez-Garcia (2008), Point-to-point connectivity, an abstract concept or a key issue for risk assessment studies?, *Adv. Water Resour.*, *31*, 1742–1753.

Western, A. W., G. Bloschl, and R. B. Grayson (2001), Toward capturing hydrologically significant connectivity in spatial patterns, *Water Resour. Res.*, *37*(1), 83–97.

Willmann, M., J. Carrera, and X. Sanchez-Vila (2008), Transport upscaling in heterogeneous aquifers: What physical parameters control memory functions?, *Water Resour. Res.*, *44*, W12437, doi:10.1029/2007WR006531.

Zheng, C. M., and S. M. Gorelick (2003), Analysis of solute transport in flow fields influenced by preferential flowpaths at the decimeter scale, *Ground Water*, *41*(2), 142–155.

Zinn, B., and C. F. Harvey (2003), When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields, *Water Resour. Res.*, *39*(3), 1051, doi:10.1029/2001WR001146.

D. Fernàndez-Garcia and X. Sanchez-Vila, Department of Geotechnical Engineering and Geosciences, Technical University of Catalonia, c/Jordi Girona 1-3, E-08034 Barcelona, Spain. (daniel.fernandez.g@upc.edu)

P. Trinchero, Amphos XXI Consulting, S. L. Passeig Garcia Faria, 49-51, E-08019 Barcelona, Spain.