An automatic locally-adaptive method to estimate heavily-tailed breakthrough curves from particle distributions

Daniele Pedretti, Daniel Fernàndez-Garcia

Hydrogeology Group, Department of Geotechnical Engineering and Geosciences, Universitat Politecnica de Catalunya (UPC BarcelonaTech), Ed. D2/004/01, C/ Jordi Girona 1-3, E-08034 Barcelona, Spain

Abstract

Particle tracking methods to simulate solute transport deal with the issue of having to reconstruct smooth concentrations from a limited number of particles. This is an error-prone process that typically leads to large fluctuations in the determined late-time behavior of breakthrough curves (BTCs). Kernel density estimators (KDE) can be used to automatically reconstruct smooth BTCs from a small number of particles. The kernel approach incorporates the uncertainty associated with subsampling a large population by equipping each particle with a probability density function. Two broad classes of KDE methods can be distinguished depending on the parametrization of this function: global and adaptive methods. This paper shows that each method is likely to estimate a specific portion of the BTCs. Although global methods offer a valid approach to estimate early-time behavior and peak of BTCs, they exhibit important fluctuations at the tails where fewer particles exist. In contrast, locally adaptive methods improve tail estimation while oversmoothing both early-time and peak concentrations. Therefore a new method is proposed combining the strength of both KDE approaches. The proposed approach is universal and only needs one parameter (α) which slightly depends on the shape of the BTCs. Results show that, for the tested cases, heavily-tailed BTCs are properly reconstructed with $\alpha \approx 0.5$.

Keywords: Kernel density estimator, Particle tracking, Heterogeneity,

Email address: daniele.pedretti@upc.edu (Daniele Pedretti)

Anomalous solute transport, Random walk

1. Introduction

Among the several numerical methods that exist to solve the solute transport partial differential equation, Particle Tracking Methods (PTMs) (e.g., [1, 2, 3, 4]) provide a computationally efficient and versatile approach with many interesting applications. Processes such as non-Fickian dispersion [5, 6], rate-limited mass transfer [7, 8, 9, 10, 11, 12], incomplete mixing [13, 14, 15] and complex chemical reactions [16, 17, 18, 19, 20] can easily be addressed in a cost-effective manner without oversimplifying the description of the spatial variability of the physical and chemical properties. Thus, PTMs are especially indicated to simulate advection-dominated problems in natural heterogeneous formations, where large flow velocity contrasts typically translate into numerical problems for Eulerian methods, such as finite differences or finite elements (e.g., [21, 22, 23]).

A major disadvantage of PTMs relies on the need to convert particle distributions into smooth concentrations [18, 21]. Consequently, a challenge in this area is the inference of concentration breakthrough curves (BTCs) from particle attributes (travel times) determined at control locations. Since true concentrations can only be obtained in the limit, when an infinite number of particles are employed, there is always an estimation error in the final solution.

For computational reasons, a generic solute plume can be discretized only into a limited number of particles, usually ranging from 10^4 to 10^7 (e.g. [24, 25]). As a result, BTCs, obtained basically from histograms (counting particles in bins), normally produce discontinuous solutions that strongly depend on the number of particles used and the size of the bin selected. Large bins will yield biased estimates of the BTCs, while small bins will produce large fluctuations [26]. These fluctuations can be reduced by increasing the number of n particles. Unfortunately, the magnitude of these fluctuations is proportional to $n^{-1/2}$ and thereby an unfeasible large number of particles are necessary to minimize the estimation error.

The application of PTMs to many groundwater problems has been widely reported in the literature, especially with regards to risk assessment of polluted sites and solute transport phenomena [27, 28, 29, 30, 31, 32, 33]. A general conclusion from these works is that a proper quantification of the

actual threat that potential contaminants pose to human health and ecosystems requires an adequate description of the BTCs. Different parts of the BTC are important to different problems [34]. The early arrival of concentrations determines the potential hazard of a toxic compound, the central part controls human exposure to site-related contaminants during risk assessment, and the late arrival of concentrations is the key factor controlling the end-point of remediation [27, 35, 36]. In this context, particle tracking simulations of real aquifer systems often produce highly asymmetric BTCs with sustained low concentrations at large times [11, 24, 37, 38]. This late-time behavior is sometimes known as tailing. Here, we will refer to the complete distribution of concentrations associated to this behavior as 'heavily-tailed' BTCs. Heavily-tailed BTCs can be observed in porous media due to several mechanisms. Among them, heterogeneity [25, 38, 39], flow configuration [40], rate-limited mass transfer [35, 41, 42], kinetic reactions [43], and non-linear sorption [44] have been largely addressed. Accurate reconstruction of heavily-tailed BTCs predicted by particle tracking simulations in any of these situations will typically require an enormous number of particles (millions) to describe the entire concentration distribution [21].

Kernel density estimators (KDE) offer a valid option to properly reconstruct BTCs with many interesting advantages [18, 45]. It is a non-parametric method that can be used to infer the true structure of the BTC, it allows the automatization of the reconstruction process and it provides a better quality of results for the same number of injected particles [18]. The reader can find complete and exhaustive reviews of these methods in Härdle [46] and Silverman [26]. Although several KDE methods exist, two main categories can be distinguished depending on the parametrization chosen to characterize the key parameter controlling the smoothing degree, namely the bandwidth: Global KDE methods consider a constant bandwidth size, whereas, on the contrary, adaptive KDE methods allow the bandwidth to change with the particle location. According to Silverman [26], the selection of one of the two methods must depend on the specific goal of the study, since each of them is more likely to estimate a specific portion of the curves while penalizing the other parts. Adaptive methods are preferred to estimate the tails of heavily-tailed probability distributions, while global estimators can provide a more robust estimation of symmetric probability distributions.

This paper develops a general method that provides accurate estimates of all parts of the BTC. The approach combines the best of both KDE methods by means of the cumulative distribution function (CDFs) of the particles'

travel times, which is a natural output from each model run. The performance of all methods is analyzed by using benchmark solutions of heavily-tailed breakthrough curves.

The paper is structured as follows: Firstly, Section 2 gives some general concepts on particle tracking and KDE methods. Then, Sections 3 and 4 respectively introduce the global and adaptive kernel methods. Finally, after a careful examination of the limitation of both methods, Section 5 introduces a new approach, which is contrasted against the previous ones.

2. Particle Tracking Methods and Kernel Density Estimation

Particle tracking methods simulate transport by injecting a large number of particles into the system. Each particle carries a portion of the solute mass that moves with groundwater and reacts according to some fundamental mechanisms. A large variety of mechanisms can be chosen to efficiently simulate different transport phenomena. These mechanisms range from pure advection and random walk motions to complex reaction and mass transfer processes (e.g., [7, 23, 47, 48]).

In all cases, the outcome of PTMs is typically expressed as a distribution of particle travel times at control locations and/or positions observed at different times. Since this information is discrete in nature a reconstruction process is mandatory to finally convert particles into concentrations. This reconstruction process is normally seen as the main disadvantage of PTMs [18, 21]. This is based on the relationship between particle distributions and concentrations, which determines that concentrations are proportional to particle probability densities (PDFs) (see, e.g., [49, pag. 167]). Let us consider, for instance, the reconstruction of a BTC from a record of particle travel times obtained at a given well location. In this case, flux-concentrations, c_f, are related to the PDF travel times, $p(t)$, by

$$p(t) = \frac{Q}{m_0} c_f \qquad m_0 = \int_0^\infty Q c_f \mathrm{d}t, \tag{1}$$

where Q is the total flow at the extraction well, and m_0 is the mass underneath the BTC. The PDF travel time $p(t)$ is defined as the derivative of the travel time cumulative distribution function $P(t)$, where $P(t)$ is the probability that the travel time is smaller than t. For small time increments $\mathrm{d}t$, the term $p(t)\mathrm{d}t$ is the probability that, at time t, the particle travel time falls into the time interval $[t, t+\mathrm{d}t]$. Unfortunately, the application of such simple

relationship (1) is filled with difficulties. An infinite number of particles is never used to properly fulfill the validity of this equation.

Kernel Density Estimation (KDE) methods were introduced in the late '50 [50] in order to improve the density estimation of the probability density function $p(t)$. The need of KDE methods appears crucial when estimating smoothed PDFs with frequency histograms, which are determined by counting the number of particles within a given time interval $B_j = [t_0 + (j-1)h, t_0 + jh]$, where h is the time interval (bin size) and t_0 is the origin of the histogram. The frequency at the j-th bin is determined by

$$p_j \approx \widehat{p}_j \equiv \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} I\{t_i \in B_j\}, \tag{2}$$

where n is the number of particles (arriving at the well location), $\widehat{p}_j$ is the estimated density, t_i is the particle travel time data $(t_1, ..., t_n)$, and $I\{\cdot\}$ is an indicator function defined as $I = 1$ if t_i is inside the bin B_j and $I = 0$ otherwise. By definition, the use of a histogram provides a discontinuous box function that prevents the estimation of smooth probability density functions. Moreover, the histogram depends on its origin t_0 and the selection of the bin size h, which are arbitrary. KDE methods overcome these problems by replacing the fixed box model in the histogram by a moving weighting function K centered at the particle location,

$$p(t) \approx \widehat{p}(t) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} K\left(\frac{t - t_i}{h}\right), \tag{3}$$

where h is the bandwidth, and K is a weighting function that should be chosen from symmetric functions of unitary area, i.e.,

$$\int K(\tau) \mathrm{d}\tau = 1. \tag{4}$$

The density estimation can be seen as an attempt to calculate an histogram where every point is the center of the sampling lag interval h. The area of influence of the i-th datum is restricted to a fixed area, determined by arbitrarily choosing h. Typical kernel functions are reported in Table 1. The choice of a specific type of kernel does not substantially influence the final results. However, Gaussian kernels are usually preferred for mathematical advantages. We refer to [26] and [46] for details in this subject.

In the following sections we present three methods to evaluate h. The first two methods will be referred herein as the Global Bandwidth Method (GB) and the Adaptive Bandwidth Method (AB), and they are commonly used in statistics to reconstruct density distributions. The third method, which we called the Universal Adaptive Bandwidth Method (UAB), is a new approach that we developed to merge the capabilities of GB and AB methods to reconstruct heavily-tailed breakthrough curves.

3. The global bandwidth method

3.1. Mathematical statement

Parameter h controls the degree of smoothing and relates to the area of influence around a datum. If h is too small, there cannot be sufficient data to find the proper density function $p(t)$ and the outcome will fluctuate. If h is too large, distant particles can also contribute to the estimation, smoothing out the local structure of the data. Thus, an optimum value of h must exist. Selecting an optimum bandwidth h must be established based on an objective criterion about the goodness of the estimator given by (3). A common criterion is the Mean Integrated Squared Error (MISE), formally written as

$$MISE(h) = E \int \left(p(t) - \widehat{p}(t)\right)^2 \mathrm{d}t. \tag{5}$$

Knowing that the mean square error, $MSE = E[(p - \widehat{p})^2]$, can be expressed as the sum of the estimator bias, $\mathrm{bias}(\widehat{p}) = E(\widehat{p}) - p$, and its variance, $\mathrm{var}(\widehat{p}) = E([\widehat{p} - E(\widehat{p})]^2)$, $MISE$ can be written as

$$MISE(h) = \int \mathrm{bias}^2(\widehat{p}(t))\mathrm{d}t + \int \mathrm{var}(\widehat{p}(t))\mathrm{d}t. \tag{6}$$

Approximate solutions of $MISE$ can only be derived asymptotically, i.e., in the limit when $h \to 0$ and $nh \to \infty$. Assuming p to be twice differentiable, the bias and the variance can be approximated by [46]

$$\mathrm{bias}(\widehat{p}(t)) = \frac{h^2}{2} p''(t)\mu_2(K) + \mathcal{O}(h^2), \tag{7}$$

$$\mathrm{var}(\widehat{p}(t)) = \frac{1}{nh} \|K\|_2^2 p(t) + \mathcal{O}\left((nh)^{-1}\right), \tag{8}$$

where $\mu_2(K)$ is the second moment of K, $p''(t)$ is the second derivative of the estimated function, and the symbol $\|\cdot\|_2^2$ is the L_2-norm operator, i.e., $\|p(t)\|^2 = \int p(t)^2 \mathrm{d}t$. Note that while the bias error is quadratic in h, the variance error is inversely proportional to the product of the number of particles and the kernel bandwidth. Thus, large h values will reduce the variance (noise), while increasing the bias (oversmoothing).

An optimal value of h can be obtained by minimizing (6). Should the derivative of $MISE$ with respect to h be zero, in the limit when $h \to 0$ and $(nh) \to \infty$, and recalling that $\int p(t)\mathrm{d}t = 1$, then [51]

$$h_0 = \left(\frac{\|K\|_2^2}{n\|p''(t)\|_2^2 \mu_2(K)} \right)^{1/5}. \tag{9}$$

Equation (9) shows that h_0 is inversely proportional to the number of particles $n^{1/5}$. Thus, the larger the number of particles the smaller the bandwidth size should be. Moreover, h_0 depends on the second derivative of the density function $p''(t)$, i.e., on the shape of the density function in itself, which is unknown. To deal with the latter issue, several methods have been proposed in the literature [52, 53]. Among them, a common one is the so-called *plug-in* method, which has been applied for instance by Fernandez-Garcia and Sanchez-Vila [18] to reconstruct reactions rates from particle distributions. This method, which is described in detail in Appendix A, is used in the applications shown in following sections.

The global bandwidth (GB) method is obtained by simply substituting the optimal bandwidth h_0 with the kernel density estimator (3), that is

$$p(t) \approx \widehat{p}_0(t) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_0} K\left(\frac{t - t_i}{h_0} \right), \tag{10}$$

3.2. Method performance

The performance of the GB method is shown in Figure 1, which compares GB reconstructed BTCs obtained from particle tracking simulations with corresponding reference solutions. Transport simulations consider the one-dimensional homogeneous multirate mass transfer (MRMT) model [35], which simulates advection-dispersion solute transport with rate-limited mass transfer. In this model, heavily-tailed BTCs result from the diffusion of solute mass previously stored in low permeability areas with almost immobile water (back-diffusion). The details of the model are given in Appendix B.

To our purposes, the phenomena embedded in the model is not important. The model simply provides a way to generate both analytical solutions and particle tracking solutions of heavily-tailed BTCs with similar shapes as those observed in the field. Thus, in principle, the methods presented here can be equally used to reconstruct heavily-tailed breakthrough curves associated to different phenomena.

Analytical solutions of the multirate mass transfer model were generated with the STAMMT-L code [54] using a power-law distribution of mass transfer coefficients. Table 2 summarizes the parameters used to generate the reference solutions. These solutions were contrasted against reconstructed BTCs obtained from randomly distributed particles, of different population sizes, generated using a well-known particle tracking code, RW3D [25]. For completeness, the numerical algorithm of the random walk particle tracking method used here is also presented in Appendix C. We note that the chosen reference solutions are limited to power-law behavior and thereby other distribution shapes (e.g., multi-modal BTCs) are not specifically studied here.

Four dimensionless parameters control the general shape of the BTCs: the Peclet number (Pe), defined as the ratio of the rate of advection to the rate of dispersion; the Damköhler number (Da), defined as the ratio of the rate of mass-transfer to the rate of advection; the field capacity (β), defined as the volume ratio of immobile to mobile water; and the exponent of the power-law distribution of mass transfer rates (k). For a given Pe, when Da is small, BTCs exhibit very peaked shapes with long tails of the form $p \sim t^k$. In this case, the field capacity β determines the area of the BTC underneath the tail. As Da increases, mass transfer occurs more quickly and the process approaches equilibrium, which leads to BTCs with Gaussian-like behavior. In this case, the solute travel distance is retarded with respect to groundwater by a factor of $R = 1 + \beta$. Here, since the focus is on heavily-tailed BTCs, numerical simulations considered a constant Pe with varying Da, k and β values.

Two different numbers of particles were used to generate the distributions shown in Figure 1. The solid thick blue lines refer to the reconstructed BTCs obtained for $n = 10^4$, while the solid thin red lines illustrates those obtained for $n = 10^5$. We note that the GB method closely matches both the early arrival and the peak of the reference BTCs. This occurs for both particles datasets, and for any degree of symmetry of the reference distributions (i.e., any k-Da pairs). However, it performs poorly at late times (after the peak). Starting from an estimated concentration close to $1/n$, the numerical solution

oscillates as the estimation is performed based on a limited number of particles. Eventually, it tends to a fictitious asymptotic plateau at late times; in the case that the estimation is performed using only one particle, the probability of the plateau is $\widehat{p}_0(t) = (nh_0\sqrt{2\pi})^{-1}$, using a Gaussian kernel (Table 1). This effect is more pronounced for heavily-tailed distributions obtained with small Da and number of injected particles n. Nevertheless, fluctuations are less noticeable in symmetric BTCs with small tailing obtained for relatively large Da. This result suggests that the performance of the GB method is best for symmetric distribution of particle travel times. Unfortunately, this is usually not the case in subsurface hydrology applications.

4. Adaptive bandwidth method

4.1. Mathematical statement

Adaptive bandwidth (AB) methods are considered to be an efficient manner to deal with long-tailed distributions [26]. Long-tailed distributions suggest that the mass extends over a wider range of time values. Intuitively, as the number of particles used to estimate their density diminishes, the smoothing parameter h should increase to allow a smoother estimation in regions where data are lacking. This explains why errors appeared to be more pronounced at the BTCs tails in Figure 1. To partially overcome this problem, a common procedure consists in deterministically varying h as a function of the local data density. This can be formally written as

$$p(t) \approx \widehat{p}_1(t) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_1(t_i)} K\left(\frac{t-t_i}{h_1(t_i)}\right). \tag{11}$$

A common approach to find $h_1(t_i)$ is suggested by Silverman [26] and it is tested here against the reference distributions previously shown. This approach considers that h_i is a function of a normalized local data density measure λ_i, which in turn is a function of $\widehat{p}_0(t)$ (defined in (10)) . That is

$$h_1(t_i) = h_0\lambda_i, \qquad \lambda_i = \left(\frac{\widehat{p}_0(t_i)}{G}\right)^{-\alpha}, \qquad G = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\ln(\widehat{p}_0(t_i))\right), \tag{12}$$

where α is a sensitivity parameter fulfilling $0 \leq \alpha \leq 1$. Note that h_1 is no longer a constant value but encouraged by the first estimate to change from

particle to particle. Note that the new density estimate does not follow a second optimization process but relies on the first estimated value h_0. Choosing a particular value of α determines how important the shape of the first guess is with respect to the the second estimation. Although there is not a well-defined criterion to obtain a prior estimate of α, it is worth mentioning that Abramson [55] found a theoretical response given by $\alpha = 0.5$.

4.2. Method Performance

The comparison between the reference solutions and the particle tracking BTCs obtained by the AB method is shown in Figure 2 for $\alpha = 0.5$. The AB method performs reasonably well for heavily-tailed BTCs (low Da). However, an underestimation of peak values and departure from the analytical solution at late time is clearly visible for increasing Da (higher symmetry). The reason is related to the adaptive h. Let us focus for instance on the pair ($k = 3, Da = 2.71$). In this case the kernel estimation is erroneously overweighting the late-time concentrations, significantly sacrificing the early portions of the BTCs. A surprising effect is that the AB method seems to work better for a small number of particles n in the case of low Da (more asymmetric curves).

Quantitatively, the adequacy of the fit can be determined by the relative local ϵ error between the estimated and the reference (analytical) density functions, which is generally written as

$$\epsilon(t) = \left| \frac{\widehat{p}(t) - p(t)}{p(t)} \right|. \tag{13}$$

The results are shown in Figure 3, where the black lines with circles refer to the GB method and the dashed lines depict the errors of the AB method for different α values. Note that the scales are logarithmic to emphasize that ϵ can span over several orders of magnitude. The GB method shows minimum error around the concentration peaks (where particle densities are higher) but performs poorly at late times. This is especially true for heavily-tailed BTCs with low Da. The opposite is also true for the AB method. In this case, the reconstruction of tailing in heavily-tailed BTCs with $k = 1$ displays relatively small errors.

In short, the AB method is capable of improving the quality of highly skewed BTCs with low Da (with respect to the GB method), but their corresponding local errors are in general dramatically large for other cases. Especially, it causes a general oversmoothing of the concentration peaks posing serious concerns about the reliability of the reconstructed BTCs, as it can

give misleading solute transport parameters with a catastrophic impact on decision-making processes (e.g., in risk assessment or aquifer remediation activities). A solution to this problem is presented in the following section.

5. A universal adaptive method

5.1. Mathematical statement

The analysis of the two KDE methods described so far suggests that the choice of the reconstruction method depends on the type and goal of the problem. While the GB method can properly estimate the early arrival and peak of concentrations, the AB method can better reproduce the tail of heavily-skewed BTCs.

Unfortunately, in most hydrogeological applications the actual structure of the BTC is not known *a priori* and the selection of the kernel estimator is uncertain. During tracer tests performed in heterogeneous aquifers, for instance, the shape of BTCs measured after injecting solutes from different positions is highly variable (e.g., [56]) and can be drastically different from ideal parametric distributions. Dimensionless transport parameters controlling the shapes of the BTCs (such as Pe, Da and β factors) are obtained *a posteriori* by model fitting (e.g., [38]) and cannot be used to select the method *a priori*.

Therefore, the aim is to look for a common and general procedure independent from the final shape of the BTCs. In this case, we proposed a new KDE approach that combines the strength of both methods (GB and AB). In sum, as extracted from previous analysis, the bandwidth h_i should be similar to h_0 in (9) at the rising limb of the BTCs and around the concentration peak, but it also should increase at late times to smooth out the solution where particle information is scarce, similarly to h_1 in (12). This means finding a new equation that could incorporate this behavior of h_i as a function of time. A suitable and appealing function displaying these characteristics is the experimental cumulative density function, $P(t)$. A quick and reliable manner to obtain $P(t)$ is by sorting in ascending order the vector of arrival times ($t_1 < t_2 < ... < t_n$) so that

$$P(t_i) \approx \frac{i - C}{n}, \qquad i = 1, ..., n, \tag{14}$$

where C gives the functional form of P for a datum (e.g., for a triangle shape, $C = 0.5$; for a box shape, $C = 0$). $P(t_i)$ always takes values comprised

between $[0, 1]$, being lower around the early time and increasing in the regions where less mass is used to estimate the BTC for larger times. The new estimator for the normalized BTC can be this written as

$$p(t) \approx \widehat{p}_2(t) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_2(t_i)} K\left(\frac{t - t_i}{h_2(t_i)}\right). \tag{15}$$

where the new adaptive bandwidth method is

$$h_2(t_i) = (1 - P(t_i))h_0 + P(t_i)h_1(t_i). \tag{16}$$

We call this method universal adaptive bandwidth (UAB). The executable file of the code along with an example of its application is available on the GHS-UPC software web page at http://www.h2ogeo.upc.es/software/quern/index.htm.

5.2. Method Performance

Recalling that h_1 is sensitive to the α shape parameter (see (12)), we test this method using three different and increasingly values of α, ranging from $\alpha = 0.25$ to $\alpha = 1.0$ ($\alpha = 0$ corresponds to the GB method). The results are plotted in Figures 4, 5 and 6, respectively. We note from these figures that the UAB method substantially improves the estimation obtained for each combination of parameters, and works satisfactorily well for both heavily-tailed and less skewed BTCs.

More specifically, results from this numerical sensitivity analysis suggest that not only the oversmoothing effects around the concentration peak disappear (especially by using $\alpha = 0.5$ and $\alpha = 1$) but tailing is also properly captured, matching well with the reference analytical solutions. This is an effect of the shape of the new universal adaptive bandwidth, $h_2(t_i)$. Figure 7 compares the behavior of the latter function with the one of the AB method, $h_1(t_i)$. Both functions are normalized by h_0 and are functions of time, when $k = 2$, $\beta = 1$ and $Da = 0.009$. Note that the AB method produces extremely large normalized bandwidth values at early stages, giving raise to the oversmoothing effect on the reconstructed BTCs. On the contrary, the UAB method keeps the values close to h_0 at early times, allowing a good reconstruction of the early behavior and peak of the BTCs. At late times, both solutions approach the same values, so that the accuracy of the adaptive method in that region is maintained. Table 3 summarizes the different bandwidth solutions adopted in this analysis.

The critical question now is which α parameter to choose for the UAB method. The goal is to find a general parameter capable of reconstructing the entire BTC at once. From a quantitative perspective, the ϵ criterion helps discerning which α provides the lowest local error and a more accurate local estimation for each reference case. Figure 8 shows that at shorter times, all errors collapse to those of the GB method, while at larger times the UAB method is always the best solution. In general, results show that $\alpha = 0.5$ is the most reliable value for a universal and robust estimation:

- When $\alpha = 0.25$ (Figure 4), the UAB method provides very few benefits at late times, where the estimated BTCs still oscillate. In this case, the bandwidth h_2 in (16) is still controlled by the global bandwidth estimate h_0 rather than from its adaptive counterpart h_1 and thereby the reconstruction process gives results similar to those of the GB method.

- When $\alpha = 1.0$ (Figure 6), fluctuations almost disappear and local errors are in most cases small, but the solution departs from the reference distribution at large times for the most symmetric distributions (high Da). In this case, the bandwidth h_2 in (16) is mostly controlled by the adaptive bandwidth h_1 rather than from its global counterpart h_0 and thereby the reconstruction process gives results similar to those of the AB method.

- When $\alpha = 0.5$ (Figure 5), an excellent agreement is found for all cases. Even though residual fluctuations are still visible in BTCs with small Da and k, the solution is never biased (never departs from the reference distribution) and a minimum expression of fluctuations is obtained compared to the solution provided by other α values. Note that, despite we have not proved this theoretically for the UAB method, a selection of $\alpha = 0.5$ is consistent with the theoretical analysis of the AB method provided by Abramson [55].

Another important aspect of the UAB method is in terms of computation cost. Particle tracking simulations directly depends on the number of particles injected. So, the less particles injected the more efficient is the simulation. In this context, it is important to highlight that the proposed UAB method is capable of providing the same quality of results with less particles. A quick comparison of the reconstructed BTCs shown in Figures

1-6 illustrates that the use of the GB and AB methods with $n = 10^5$ perform less accurately than the UAB method with $\alpha = 0.5$ and $n = 10^4$. To quantitatively examine this in more detail, the dependence of the root mean square error $RMSE$ on the number of particles injected during simulations was analyzed,

$$RMSE(n) = \sqrt{\frac{1}{n}\left(\sum_{i=1}^{n}\left(\epsilon(t_i)\right)^2\right)}. \tag{17}$$

Results are shown in Figure 9. The UAB method with α values ranging from 0.5 to 1.0 always gives the smallest overall error, independently from the final shape of the BTC. Most interestingly and remarkably important is that, for the same quality of results, i.e., equal $RMSE$, the number of particles needed in the UAB method can be several orders of magnitude smaller than the GB method. On a computer equipped with Intel(R) Xeon(R) CPU with 2.80 GHz and 6 GB of RAM memory, the time needed to run one estimation with the UAB method is about 13 s using $n = 10^4$ and about 780 s using $n = 10^5$. Thus, performing a large number of Monte Carlo simulations with complex chemical reactions could dramatically reduce the computational cost.

6. Conclusions

Particle tracking methods constitute an attractive method to simulate a large variety of solute transport phenomena. Yet, severe fluctuations arise when converting a limited number of particles into concentrations. In this paper, the advantages and limitations of reconstructing BTCs were examined using two broad classes of kernel density estimators: global and adaptive methods. Results have shown that each method is more oriented to estimate a specific portion of the BTC. While global methods have offered a valid approach to estimate early travel times and peak concentrations, they still showed important fluctuations at the tails where fewer particles exist. In contrast, adaptive methods appeared to be capable of improving the tail of the BTCs while oversmoothing both early travel times and peak concentrations.

Based on this, a new method was proposed using the cumulative distribution of particles to locally adapt the smoothing parameter to the particle density at any given location. Thus, the smoothing parameter is progressively increased at the tails to allow a smoother estimation in regions where

particles are lacking. By doing this over a wide range of heavily-tailed BTCs, the proposed method has showed a more complete reconstruction of all parts of the BTCs at once, including the tail. Furthermore, the proposed method only needs one parameter α that slightly depends on the shape of the BTCs. A universal value of $\alpha \approx 0.5$ is shown to work well in all the tested cases.

The computational efficiency of the method was evaluated and contrasted against other KDE methods. Results showed that the proposed method can largely reduced CPU time during particle tracking simulations. Thus, it is shown that to get the same quality of results the number of particles needed by the new method may be several orders of magnitude smaller than with the rest of methods. Among the advantages of the method is the fact that that Monte Carlo transport simulations of complex chemical systems may now be feasible at a minimum computational cost.

Acknowledgments

This work has been supported by the Spanish Ministry of Science and Innovation through the projects Consolider-Ingenio 2010 CSD2009-00065 and FEAR CGL2012-38120. DP acknowledges the funding provided by the Spanish Ministry of Education through the FPU-MED 2009 Scholarship program.

Figure 1: Comparison between reference BTCs obtained from analytical solutions and those estimated using the global bandwidth (GB) method. Simulations are run using different numbers of particle datasets. Parameters k indicate the slope of the BTCs, β is the capacity coefficient and Da is the Damköhler number.

Figure 2: Comparison between reference BTCs obtained from analytical solutions and those estimated using the adaptive bandwidth (AB) method.

Figure 3: Absolute relative errors (ϵ) between the benchmark and estimated BTCs using global bandwidth (GB) method and adaptive bandwidth (AB) methods. The dotted line indicates the time of the concentrations peak ($\mathrm{T_P}$) of the corresponding BTCs.

Figure 4: Comparison between reference BTCs obtained from analytical solutions and those estimated using the universal adaptive method (UAB) with $\alpha = 0.25$.

Figure 5: Comparison between reference BTCs obtained from analytical solutions and those estimated with using the universal adaptive method (UAB) with $\alpha = 0.5$.

Figure 6: Comparison between reference BTCs obtained from analytical solutions and those estimated using the universal adaptive method (UAB) with $\alpha = 1$.

Figure 7: Shape of the bandwidths for AB and UAB methods to reproduce case $k = 2$, $\beta = 1$ using $n = 10^4$. Log-scale emphasizes the fact that the weights for the universal adaptive method (UAB) are closer to $h_j(t_i)/h_0 = 1$ (j=1,2 depending on the method) at early time and around the peak, while they tend to the adaptive method (AB) values as time elapses, for each α value. T_P indicates the temporal position of the density peak

Figure 8: Absolute relative errors (ϵ) between the benchmark and estimated BTCs using global bandwidth (GB) method and the universal adaptive method (UAB).

Figure 9: Root mean square error ($RMSE$) of the different methods as a function of the number of particles used to estimate the BTCs.

Table 1: Examples of kernel functions (from Silverman [26])

Type of kernel	$K(\tau)$
Biweight	$\begin{cases} \frac{3}{4\sqrt{5}}(1-\frac{\tau^2}{5}) & \text{for} \lvert\tau\rvert < 1 \\ 0 & \text{otherwise} \end{cases}$
Epanechnikov	$\begin{cases} \frac{15}{16}(1-\tau^2)^2 & \text{for} \lvert\tau\rvert < \sqrt{5} \\ 0 & \text{otherwise} \end{cases}$
Gaussian	$(2\pi)^{-1/2}\exp(\frac{-\tau^2}{2})$
Rectangular	$\begin{cases} \frac{1}{2} & \text{for} \lvert\tau\rvert < 1 \\ 0 & \text{otherwise} \end{cases}$
Triangular	$\begin{cases} 1-\lvert\tau\rvert & \text{for} \lvert\tau\rvert < 1 \\ 0 & \text{otherwise} \end{cases}$

Table 2: Parameters for analytical and numerical solutions: v= pore-scale velocity; a=longitudinal dispersivity; L=domain length; k=distribution power; β =capacity coefficient; w= mass transfer rates; Pe=Peclet Number; Da=Damköhler numbers.

Variable	Value [dimensions]
v	1 [LT^{-1}]
a	0.05 [L]
L	5 [L]
Pe	100 [-]
k	1; 2; 3 [-]
β	1; 5 [-]
w_{min}	10^{-4} [T^{-1}]
w_{max}	1 [T^{-1}]
Da (k=1,$\beta = 1$)	0.002 [-]
Da (k=1,$\beta = 5$)	0.005 [-]
Da (k=2,$\beta = 1$)	0.009 [-]
Da (k=2,$\beta = 5$)	0.023 [-]
Da (k=3,$\beta = 1$)	1.08 [-]
Da (k=3,$\beta = 5$)	2.71 [-]

Table 3: Bandwidth functions used in the three kernel density models. See the text for the definition of each parameter.

Method	Bandwidth function
Fixed bandwidth (GB)	h_0
Globally adaptive bandwidth (AB)	$h_0(\widehat{p}_0(t_i)/G)^{-\alpha}$
Universally adaptive bandwidth (UAB)	$(1 - P(t_i))h_0 + P(t_i)(\widehat{p}_0(t_i)/G)^{-\alpha}$

Appendix A. The plug-in method

The optimal bandwidth involves the unknown function p'' which needs to be further estimated. Here, we used the plug-in method presented by Engel [52], who employed another kernel estimate of p'' based on a different bandwidth b to estimate $\|p''\|_2^2$ by

$$\|p''\|_2^2 \approx R(b) = \frac{1}{n^2 b^5} \sum_{i,j} L_2 \left(\frac{X_{p,i} - X_{p,j}}{b} \right). \tag{A.1}$$

The functional L_2 is the convolution of K'' by itself, $L_2 = K'' \circ K''$. The bandwidth b is chosen to be a linear function of h_0 with the form $b(h_0) = Cn^q h_0$, where C and q are constants. Briefly, the algorithm presented by [52] is as follows: (1) Select an initial bandwidth, $h_0^{(0)} = Bn^{-r}$; (2) Iterate $i = 1, ..., I$ with $b^{(i-1)} = Cn^q h_0^{(i-1)}$ and $h_0^{(i-1)} = (\|K\|_2^2 / R(b^{(i-1)}) \mu_2^2(K) n)^{1/5}$; and finally, (3) set the bandwidth b to $b^{(I)}$. Tuning parameters $\{C, q, r, B, I\}$ required for the implementation of this algorithm are given in [52].

Appendix B. The multirate mass transfer model

The multi-rate mass transfer model [35] describes the porous media as a multiporosity system formed by a mobile domain, in which solute moves by advection and dispersion, and many immobile regions, in which water is mostly stagnant and solute transport is controlled by diffusion. This conceptualization allows one to represent a system that encompasses different mass transfer characteristics (different geometries, shapes, mineralogy,...). Provided that, initially, the total mass is only present in the mobile domain by the condition $p(t) = \delta(t)$, the governing equation of the MRMT can be formally written as

$$\frac{\partial p}{\partial t} + \beta \frac{\partial}{\partial t} \int_0^t g(t-\tau) p(\tau) \mathrm{d}\tau = -v \frac{\partial p}{\partial x} + av \frac{\partial^2 p}{\partial x^2} \tag{B.1}$$

where $g(t)$ is known as the memory function [41, 57], v is the velocity, β is the capacity coefficient (ratio between total immobile porosity and mobile porosity), and a is the local dispersivity. The second term of the equation is a source/sink term that represents the mass exchange from mobile to immobile regions. The memory function can be formally defined as the probability of residence times in the immobile region and relates to the probability density function of mass transfer rates $f(w)$ by

$$g(t) = \int_0^\infty w f(w) e^{-wt} \mathrm{d}w \tag{B.2}$$

Among the various probability density functions $f(w)$ described in the literature, we considered the truncated power-law function of the form [57],

$$f(w) = \mathcal{A}(k) w^{k-3}, \quad w_{min} < w < w_{max} \tag{B.3}$$

where k is the parameter of the memory function, and $\mathcal{A}(k)$ is defined as

$$\mathcal{A}(k) = \begin{cases} (k-2)/(w_{max}^{k-2} - w_{min}^{k-2}) & \text{if } k \neq 2 \\ \\ 1/\ln(w_{max}/w_{min}) & \text{if } k = 2 \end{cases} \tag{B.4}$$

The shape of the resulting BTCs is controlled by k, β and two dimensionless numbers, the Peclet number (Pe), which relates advective and dispersive characteristic times, and the Damköhler number (Da), which relates advective and mass-transfer characteristic times. Both numbers can be written

$$Pe = \frac{L}{a}, \qquad Da = \frac{(1+\beta) L w_h}{v}, \tag{B.5}$$

where L is the domain length, w_h is the harmonic mean of mass-transfer rates, and v is the groundwater velocity. If a power-law memory function is selected, w_h depends on w_{min}, w_{max} and k (this value was computed here numerically [54]). Note that for small Da values, the selection of a power-law memory function leads to BTCs with late-time behavior of the form $p \sim t^k$. The analytical solutions of the MRMT model can be found by numerical inversion of the Laplace transform of equation (B.1). A suitable code for this purpose is STAMMT-L [54], which is the one used here. Table 2 summarizes the parameters used to generate the reference solutions.

Appendix C. The random walk algorithm

This appendix describes the random walk algorithm used to simulate multirate mass transfer. The algorithm follows the method developed and implemented in the RW3D code by Salamon et al. [25] . The displacement of an ith particle X_i is given by a drift term that simulates advection and a superposed Brownian motion responsible for dispersion,

$$X_i(t+\Delta t) = X_i(t) + v\Delta t + \xi(t)\sqrt{2av\Delta t}, \tag{C.1}$$

where v is the groundwater velocity, a is the dispersivity, Δt is the time step, and $\xi(t)$ is an independent (uncorrelated in space and time), normally distributed random variable with zero mean and unit variance. RW3D simulates multirate mass transfer by also tracking in time the state of a particle, which defines the domain (mobile/immobile) at which the particle is present at a given time within the multiporosity system [25]. To apply this with a continuous distribution of mass transfer rates, the probability density function $f(w)$ must be first discretized into a set of N zones, so that $f(w) = \sum_{i=1}^{N} f_i \delta(w - w_i)$ and $\sum_i^N f_i = 1$. The individual probabilities f_i can be estimated as

$$f_i = \int_{w_i - \Delta w_i/2}^{w_i + \Delta w_i/2} f(w)\mathrm{d}w, \tag{C.2}$$

where w_i, $i = 1, ..., N$ is the discrete space of mass transfer rates, and Δw_i is the interval size of the discretization. A specific formulation is found for the power-law density function

$$f_i = \frac{(w_i + \Delta w_i/2)^{k-2} - (w_i - \Delta w_i/2)^{k-2}}{w_{max}^{k-2} - w_{min}^{k-2}}, \quad k \neq 0, \tag{C.3}$$

$$f_i = \frac{1}{\ln(w_{max}/w_{min})} \ln\left(\frac{2w_i + \Delta w_i}{2w_i - \Delta w_i}\right), \quad k = 2. \tag{C.4}$$

The change from one state to another is easily determined using transition probabilities. The transition probability, $P_{ij}(\Delta t)$, that a particle presently in state j will be in state i at a time $t + \Delta t$ is given by

$$\mathbf{P}(\Delta t) = \begin{pmatrix} -\beta \sum_{j=1}^{N} w_j f_j & w_1 & \cdots & w_N \\ \beta w_1 f_1 & -w_1 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ \beta w_N f_N & 0 & 0 & -w_N \end{pmatrix}. \tag{C.5}$$

If a particle is in the mobile domain, then it is susceptible to advection and dispersion, otherwise the particle is not allowed to move. Having calculated the transition probabilities, numerical implementation into particle

tracking is done without difficulties. For each time step a uniform $[0, 1]$ random number r is drawn for each particle and the new particle state is taken as that integer s for which

$$\sum_{i=1}^{s-1} P_{ij}(\Delta t) < r \leq \sum_{i=1}^{s} P_{ij}(\Delta t). \tag{C.6}$$

References

[1] S. W. Ahlstrom, H. P. Foote, R. C. Arnett, C. R. Cole, R. J. Serne, Multi-component mass transport model: theory and numerical implementation (discrete parcel random walk version)., Tech. rep., Rep. BNWL-2127. Battelle Pacific Northwest Lab., Richland, Washington (1977).

[2] W. Kinzelbach, The random walk method in pollutant transport simulation. advances in analytical and numerical groundwater flow and quality modelling., Tech. rep., In: Custodio, E., et al. (Eds.), NATO ASI Series C, vol. 224, pp. 227-246. (1987).

[3] D. W. Pollock, Semianalytical computation of path lines for finite-difference models, Ground Water 26 (1988) 743–750.

[4] T. A. Prickett, T. G. Naymik, C. G. Longquist, A random walk solute transport model for selected groundwater quality evaluations., Tech. rep., Illinois State Water Survey. Bulletin, vol. 65. 103 pp. (1981).

[5] B. Berkowitz, A. Cortis, M. Dentz, H. Scher, Modeling non-fickian transport in geological formations as a continuous time random walk, Reviews of Geophysics 44 (2006) RG2003. doi:10.1029/2005RG000178.

[6] Y. Zhang, E. M. LaBolle, K. Pohlmann, Monte carlo simulation of superdiffusion and subdiffusion in macroscopically heterogeneous media, Water Resources Research 45 (2009) W10417. doi:10.1029/2008WR007448.

[7] D. A. Benson, M. M. Meerschaert, A simple and efficient random walk solution of multi-rate mobile/immobile mass transport equations, Advances in Water Resources 32 (2009) 532–539.

[8] F. Delay, J. Bodin, Time domain random walk method to simulate transport by advection-dispersion and matrix diffusion in fracture networks, Geophys. Res. Lett. 28 (21) (2001) 4051. doi:10.1029/2001GL013698.

[9] M. Dentz, B. Berkowitz, Transport behavior of a passive solute in continuous time random walks and multirate mass transfer, Water Resour. Res. 39 (5) (2003) 1111–1131. doi:10.1029/2001WR001163.

[10] S. Painter, V. Cvetkovic, J. Mancillas, O. Pensado, Time domain particle tracking methods for simulating transport with retention and first-order transformation, Water Resources Research 44 (2008) W01406. doi:10.1029/2007WR005944.

[11] P. Salamon, D. Fernàndez-Garcia, J. Gomez-Hernandez, Modeling mass transfer processes using random walk particle tracking, Water Resour. Res. 42 (2006) W11417. doi:10.1029/2006WR004927.

[12] Y. W. Tsang, C. F. Tsang, A particle-tracking method for advective transport in fractures with diffusion into finite matrix blocks, Water Resources Research 37 (3) (2001) 831–835.

[13] D. Bolster, P. de Anna, D. A. Benson, A. M. Tartakovsky, Incomplete mixing and reactions with fractional dispersion, Advances in Water Resources 37 (2012) 86–93. doi:10.1016/j.advwatres.2011.11.005.

[14] A. M. Tartakovsky, D. M. Tartakovsky, P. Meakin, Stochastic langevin model for flow and transport in porous media, Physical Review Letters 101 (2008) 044502.

[15] Y. Zhang, D. A. Benson, M. M. Meerschaert, E. M. LaBolle, H. P. Scheffler, Random walk approximation of fractional-order multiscaling anomalous diffusion, Physical Review E 74 (2006) 026706.

[16] D. A. Benson, M. M. Meerschaert, Simulation of chemical reaction via particle tracking: Diffusion-limited versus thermodynamic rate-limited regimes, Water Resources Research 44 (7) (2008) W12201. doi:10.1029/2008WR007111.

[17] Y. Edery, H. Scher, B. Berkowitz, Modeling bimolecular reactions and transport in porous media, Geophys. Res. Lett. 36 (2009) L02407.

[18] D. Fernàndez-Garcia, X. Sanchez Vila, Optimal reconstruction of concentrations, gradients and reaction rates from particle distributions, Journal of Contaminant Hydrology 120-121 (2011) 99–114.

[19] J. Palanichamy, T. Becker, M. Spiller, J. Kongeter, S. Mohan, Multi-component reaction modelling using a stochastic algorithm, Computing and Visualization in Science 12 (2) (2009) 51–61. doi:10.1007/s00791-007-0080-y.

[20] A. F. B. Tompson, A. L. Schafer, R. W. Smith, Impacts of physical and chemical heterogeneity on contaminant transport in sandy porous medium, Water Resources Research 32 (4) (1996) 801–818.

[21] F. Boso, A. Bellin, M. Dumbser, Numerical simulations of solute transport in highly heterogeneous formations: A comparison of alternative numerical schemes, Advances in Water Resources 52 (2013) 178–189.

[22] G. Liu, C. Zheng, S. Gorelick, Limits of applicability of the advection-dispersion model in aquifers containing connected high-conductivity channels., Water Resources Research 40 (2004) W08308.

[23] P. Salamon, D. Fernàndez-Garcia, J. Gomez-Hernandez, A review and numerical assessment of the randomwalk particle tracking method, Journal of Contaminant Hydrology 87 (1-3) (2006) 277–305.

[24] M. Riva, A. Guadagnini, D. Fernandez-Garcia, X. Sanchez-Vila, T. Ptak, Relative importance of geostatistical and transport models in describing heavily tailed breakthrough curves at the Lauswiesen site, Journal of Contaminant Hydrology 101 (2008) 1–13. doi:10.1016/j.jconhyd.2008.07.004.

[25] P. Salamon, D. Fernàndez-Garcia, J. Gomez-Hernandez, Modeling tracer transport at the made site: The importance of heterogeneity, Water Resources Research 43 (2007) W08404. doi:10.1029/2006WR005522.

[26] B. W. Silverman, Density estimation for statistics and data analysis, Chapman & Hall CRC, 1986.

[27] R. Andricevic, V. Cvetkovic, Evaluation of risk from contaminants migrating by groundwater, Water Resources Research 32 (3) (1996) 611–621. doi:10.1029/95WR03530.

[28] R. Andricevic, Exposure concentration statistics in the subsurface transport, Advances in Water Resources 31 (2008) 714–725. doi:10.1016/j.advwatres.2008.01.007.

[29] A. Bellin, Y. Rubin, On the use of peak concentration arrival times for the inference of hydrogeological parameters, Water Resources Research 40 (7) (2004) 1–13. doi:10.1029/2003WR002179.

[30] S. Berglund, V. Cvetkovic, Pump and treat remediation of heterogeneous aquifers: Effects of rate-limited mass transfer, Ground Water 33 (4) (1995) 675–685. doi:10.1111/j.1745-6584.1995.tb00324.x.

[31] G. Dagan, V. Cvetkovic, A. M. Shapiro, A solute flux approach to transport in heterogeneous formations: 1. the general framework, Water Resources Research 28 (5) (1992) 1369–1376. doi:10.1029/91WR03086.

[32] A. M. Shapiro, V. Cvetkovic, Stochastic analysis of solute arrival time in heterogeneous porous media, Water Resources Research 24 (10) (1988) 1711–1718. doi:10.1029/WR024i010p01711.

[33] D. Zimmerman et al., A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow, Water Resources Research 34 (6) (1998) 1373–1413. doi:10.1029/98WR00003.

[34] D. Fernàndez-Garcia, J. J. Gomez-Hernandez, Impact of upscaling on solute transport: Traveltimes, scale dependence of dispersivity, and propagation of uncertainty, Water Resources Research 43 (2). doi:10.1029/2005WR004727.
URL http://dx.doi.org/10.1029/2005WR004727

[35] R. Haggerty, S. Gorelick, Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity, Water Resour. Res. 31 (10) (1995) 2383–2400.

[36] F. P. J. de Barros, Y. Rubin, A risk-driven approach for subsurface site characterization, Water Resour. Res. 58 (2008) W01414. doi:10.1029/2007WR006081.

[37] D. Pedretti, D. Fernàndez-Garcia, D. Bolster, X. Sanchez-Vila, On the formation of heavy-tailed breakthrough curves during convergent flow tracer tests in three-dimensional heterogeneous sandy aquifers, Water Resources Research.

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

[39] D. Fernàndez-Garcia, G. Llerar-Meza, J. J. Gomez-Hernandez, Upscaling transport with mass transfer models: Mean behavior and propagation of uncertainty, wrr 45 (2009) W10411. doi:10.1029/2009WR007764.

[40] A. Moench, Convergent radial dispersion: A laplace transform solution for aquifer tracer testing, Water Resour. Res. 25 (3) (1989) 439–447.

[41] J. Carrera, X. Sanchez-Vila, I. Benet, A. Medina, G. Galarza, J. Guimerà, On matrix diffusion: formulations, solution methods and qualitative effects., Hydrogeol. J. 6 (1998) 178–190.

[42] P. Gouze, T. Le Borgne, R. Leprovost, G. Lods, T. Poidras, P. Pezard, Non-Fickian dispersion in porous media: 1. Multiscale measurements using single-well injection withdrawal tracer tests, Water Resources Research 44 (2008) W06426. doi:10.1029/2007WR006278.

[43] M. Dentz, A. Castro, Effective transport dynamics in porous media with heterogeneous retardation properties, Geophys. Res. Lett. 36 (2009) L03403. doi:10.1029/2008GL036846.

[44] M. T. Van Genuchten, P. J. Wierenga, Mass transfer studies in sorbing porous media i. analytical solutions, Soil Sci. Soc. Am. J. 40 (4) (1976) 473–480.

[45] J. J. Starn, Bagtzoglou, G. A. Robbins, Methods for simulating solute breakthrough curves in pumping groundwater wells, Computer and Geosciences 48 (2012) 244–255. doi:10.1016/j.cageo.2012.01.011.

[46] W. Hardle, Smoothing Techniques with Implementation in S, Springer-Verlag, New York, 1990.

[47] M. Moroni, N. Kleinfelter, J. H. Cushman, Analysis of dispersion in porous media via matched-index particle tracking velocimetry experiments, Advances in Water Resources 30 (2007) 1–15.

[48] A. F. B. Tompson, L. W. Gelhar, Numerical-simulation of solute transport in 3-dimensional, randomly heterogeneous porous-media, Water Resources Research 26 (1990) 2541–2562.

[49] Y. Rubin, Applied stochastic hydrogeology, Oxford University Press, USA, 2003.

[50] M. Rosenblatt, Remarks on some nonparametric estimates of a density function, The Annals of Mathematical Statistics 27 (3) (1956) pp. 832–837.

[51] E. Parzen, On estimation of a probability density function and mode, Ann. Math. Statist. 33 (1962) 1065–1076.

[52] J. Engel, E. Herrmann, T. Gasser, An iterative bandwidth selector for kernel estimation of densities and their derivatives, Nonparametric Statistics 4 (1994) 21–34.

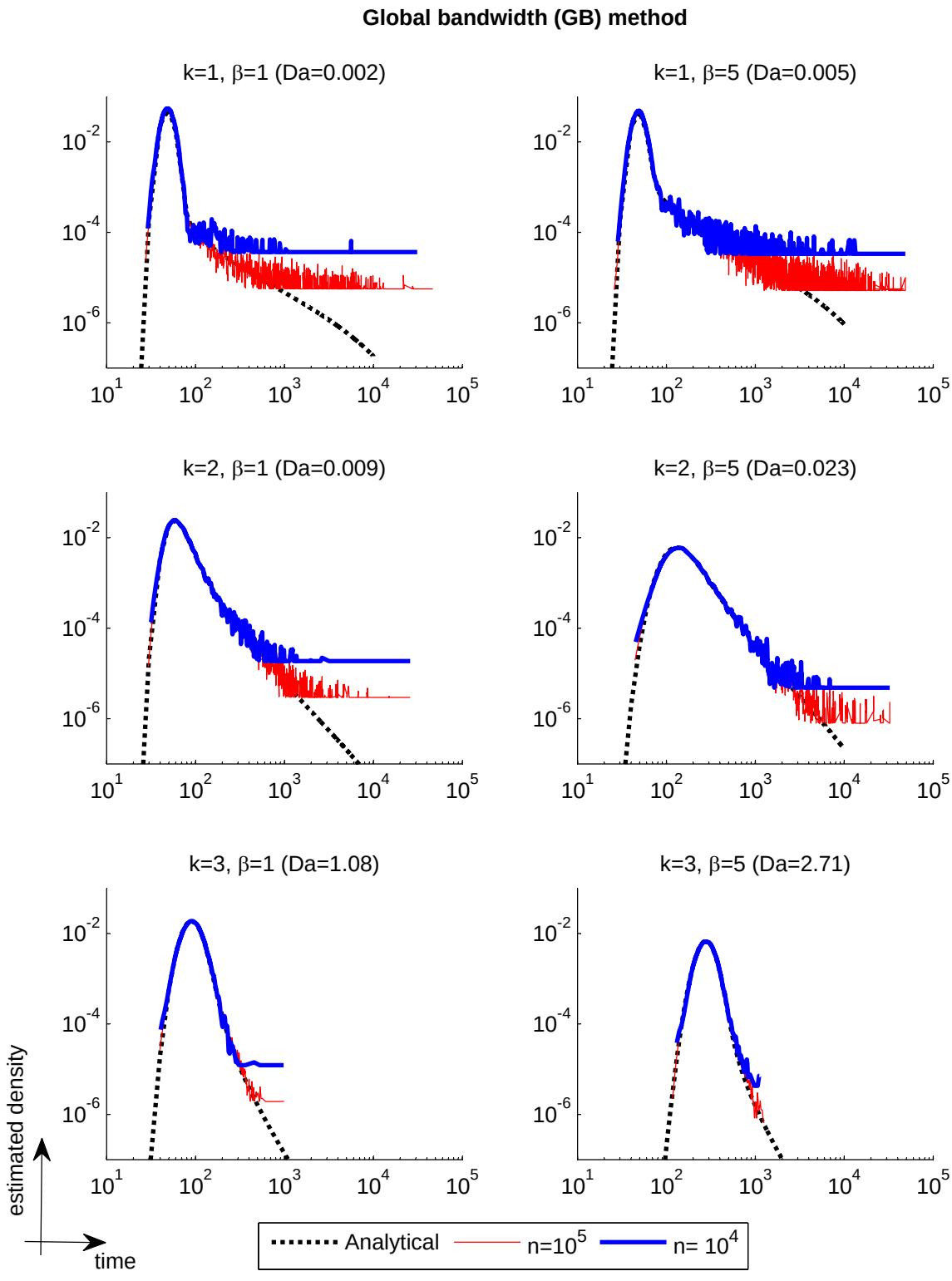
[53] U. Park, J. S. Marron, Comparison of data-driven bandwidth selectors, Journal of the American Statistical Association 85 (409) (1990) 66–72. doi:10.1080/01621459.1990.10475307.

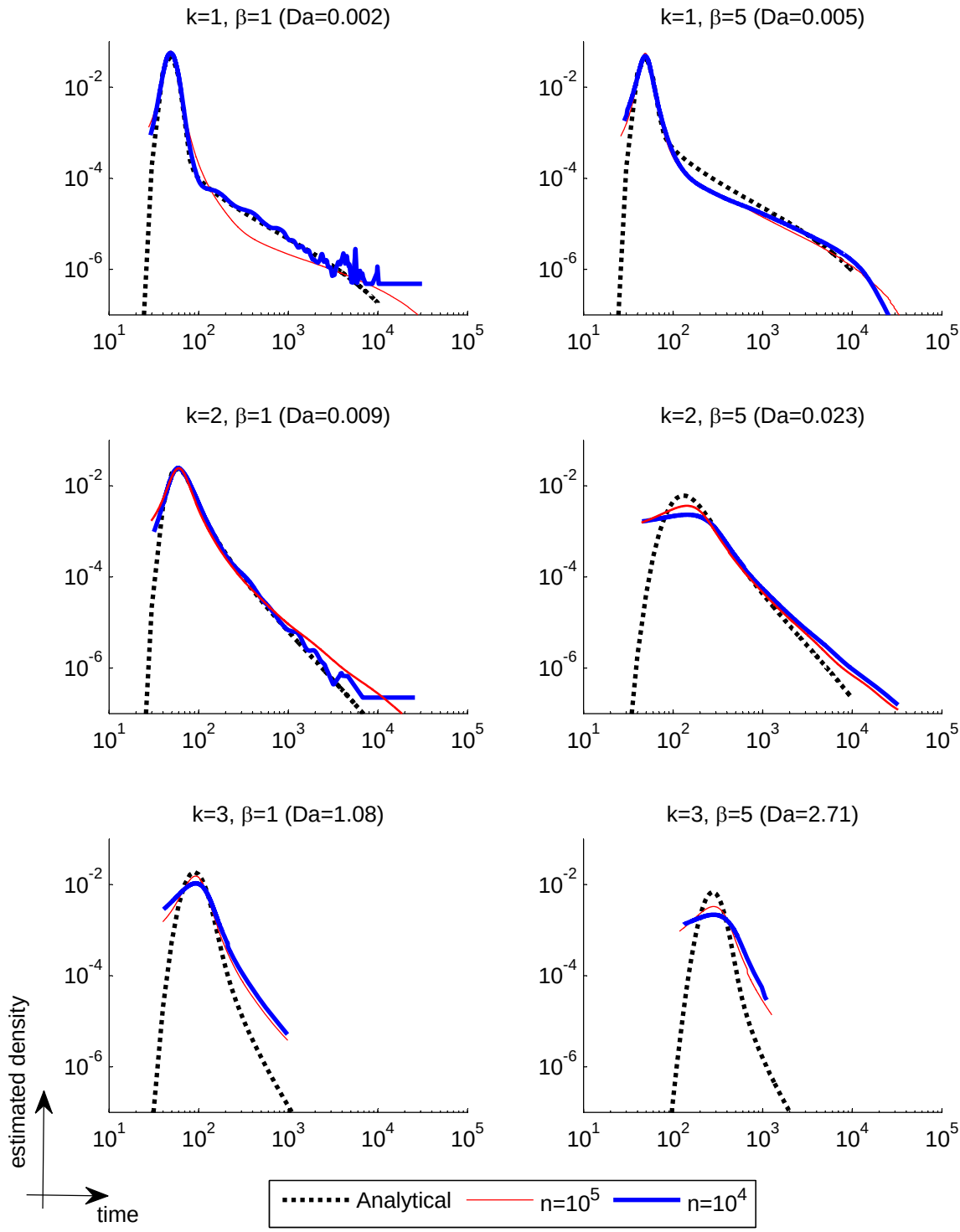
[54] R. Haggerty, P. C. Reeves, Stammt-l 1.0, formulation and user's guide, technical report erms 520308, Tech. rep., Sandia National Laboratories, Albuquerque, NM, USA (2002).

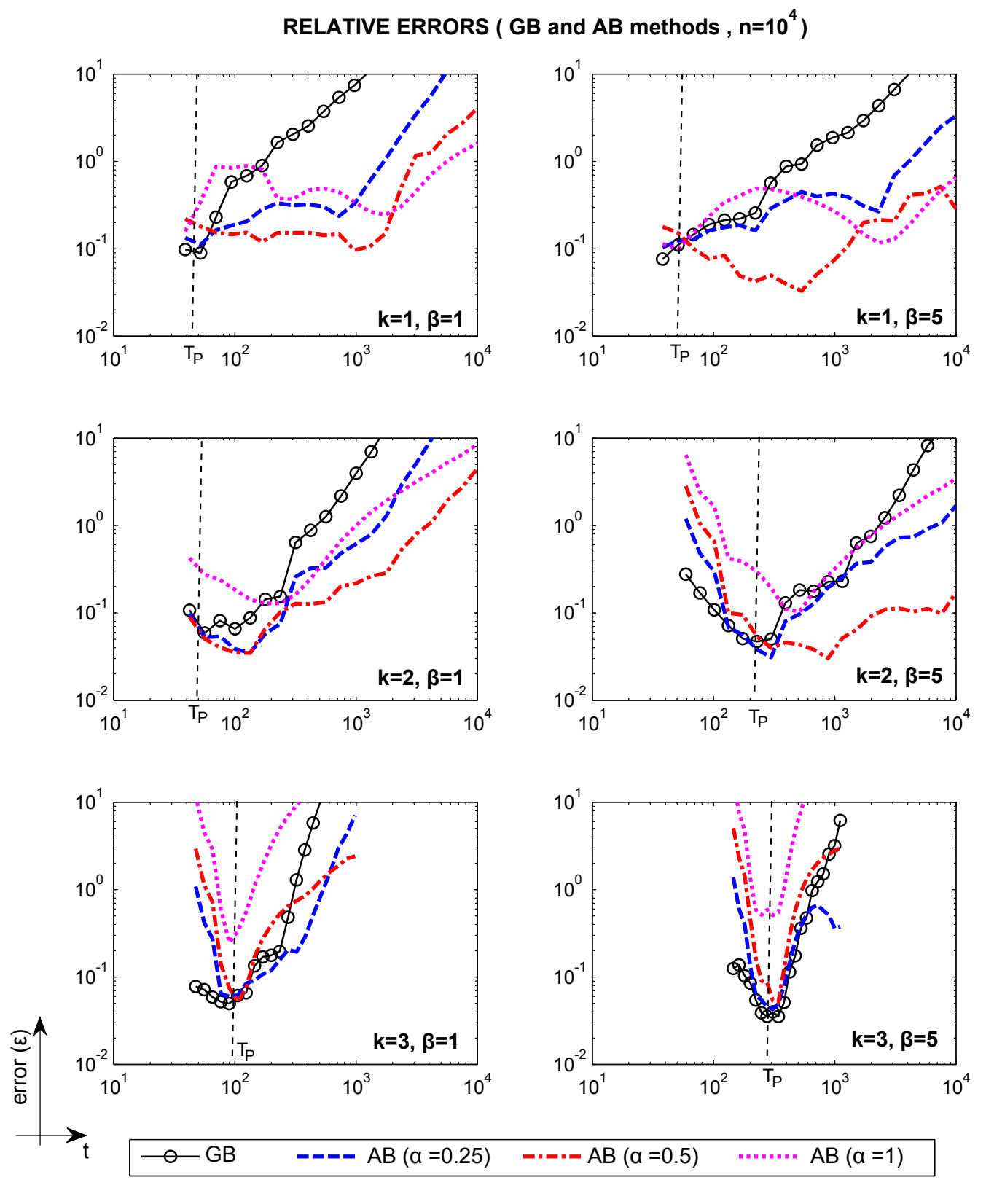
[55] I. S. Abramson, On bandwidth variation in kernel estimates-a square root law, The Annals of Statistics 10 (4) (1982) 1217–1223.

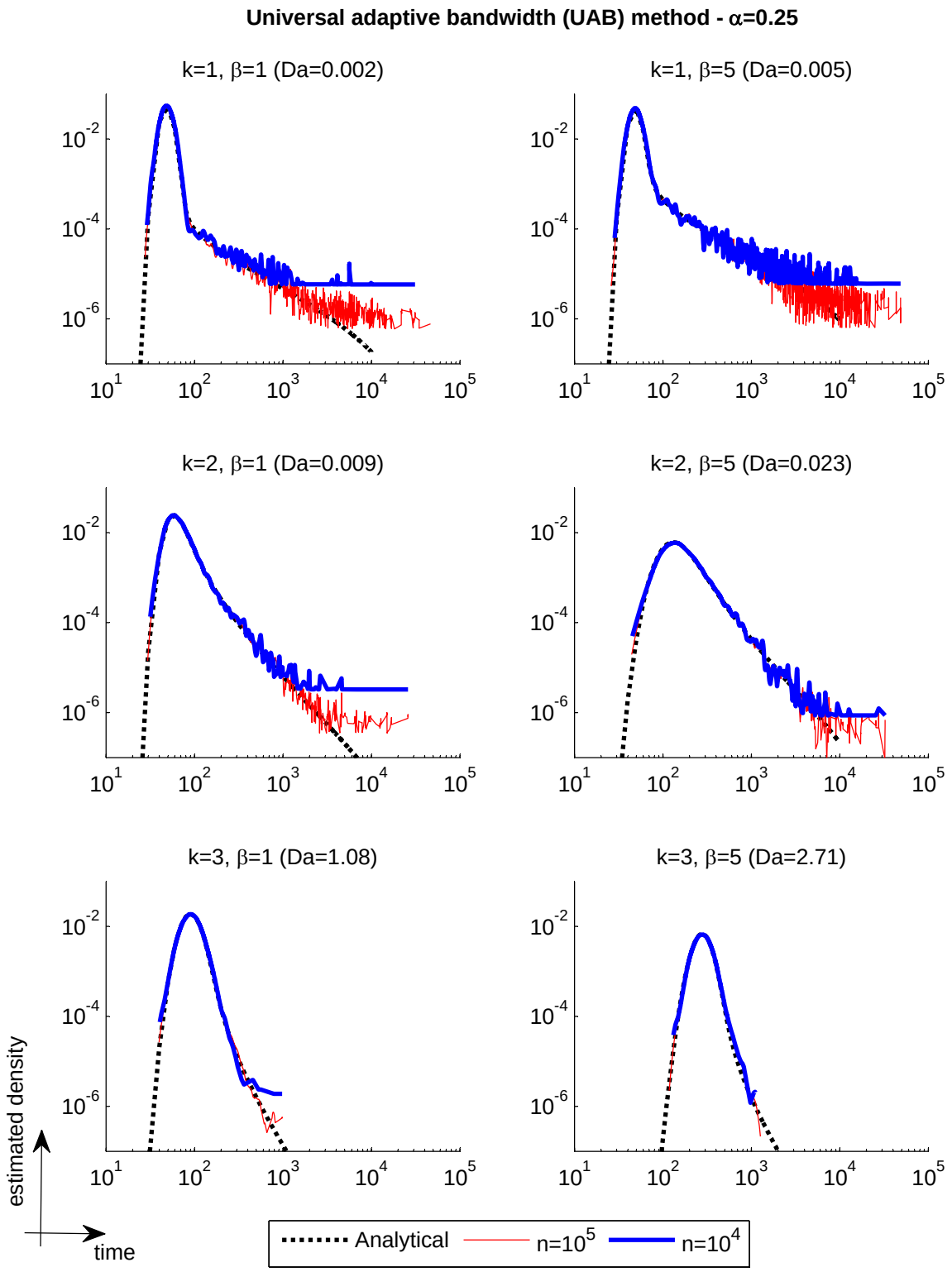
[56] D. Fernàndez-Garcia, T. H. Illangasekare, H. Rajaram, Conservative and sorptive forced-gradient and uniform flow tracer tests in a three-dimensional laboratory test aquifer, Water Resources Research 40 (2004) W10103. doi:10.1029/2004WR003112.

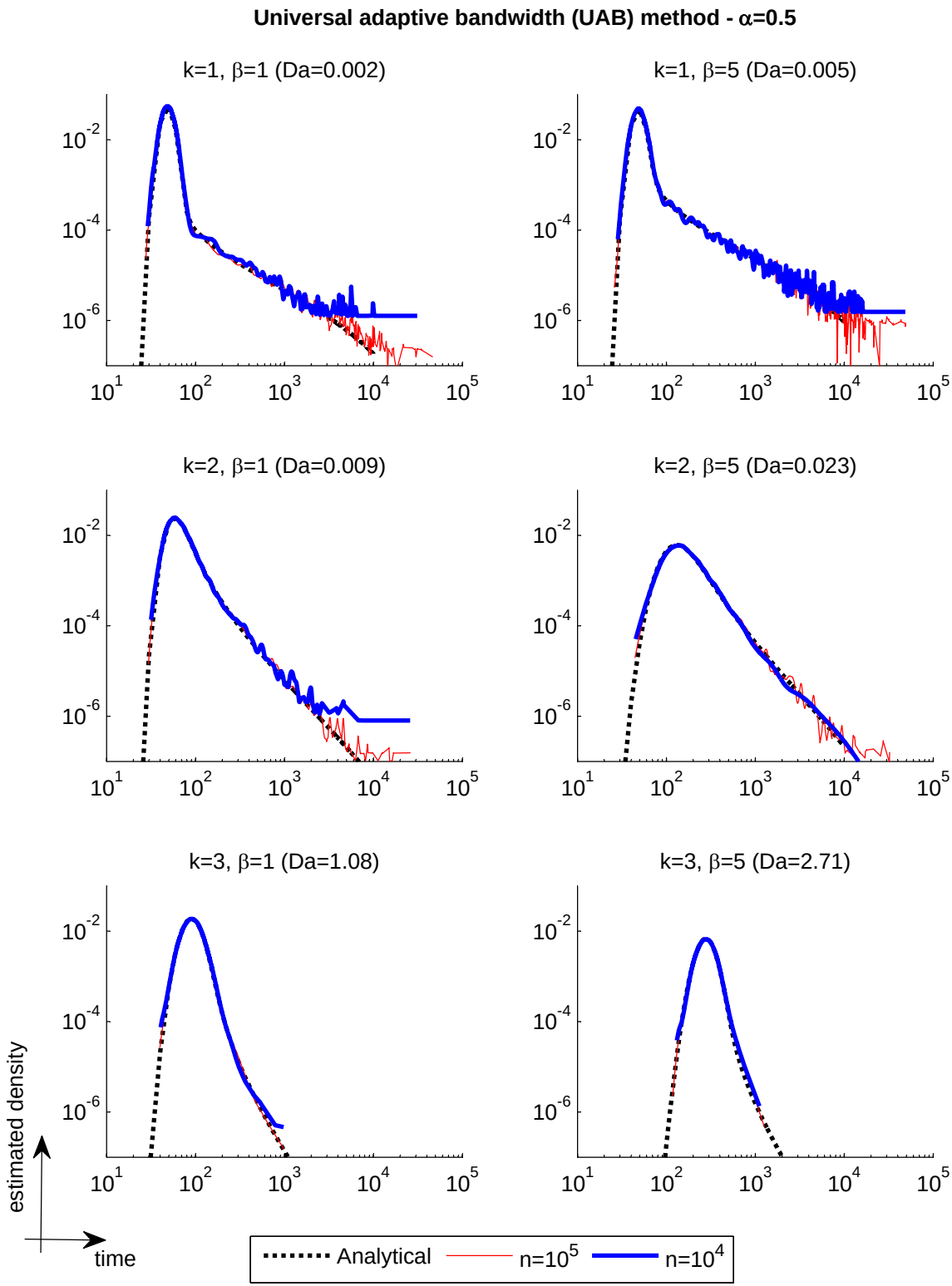
[57] R. Haggerty, S. A. McKenna, L. C. Meigs, On the late-time behavior of tracer test breakthrough curves, Water Resour. Res. 36 (12) (2000) 3467–3479.

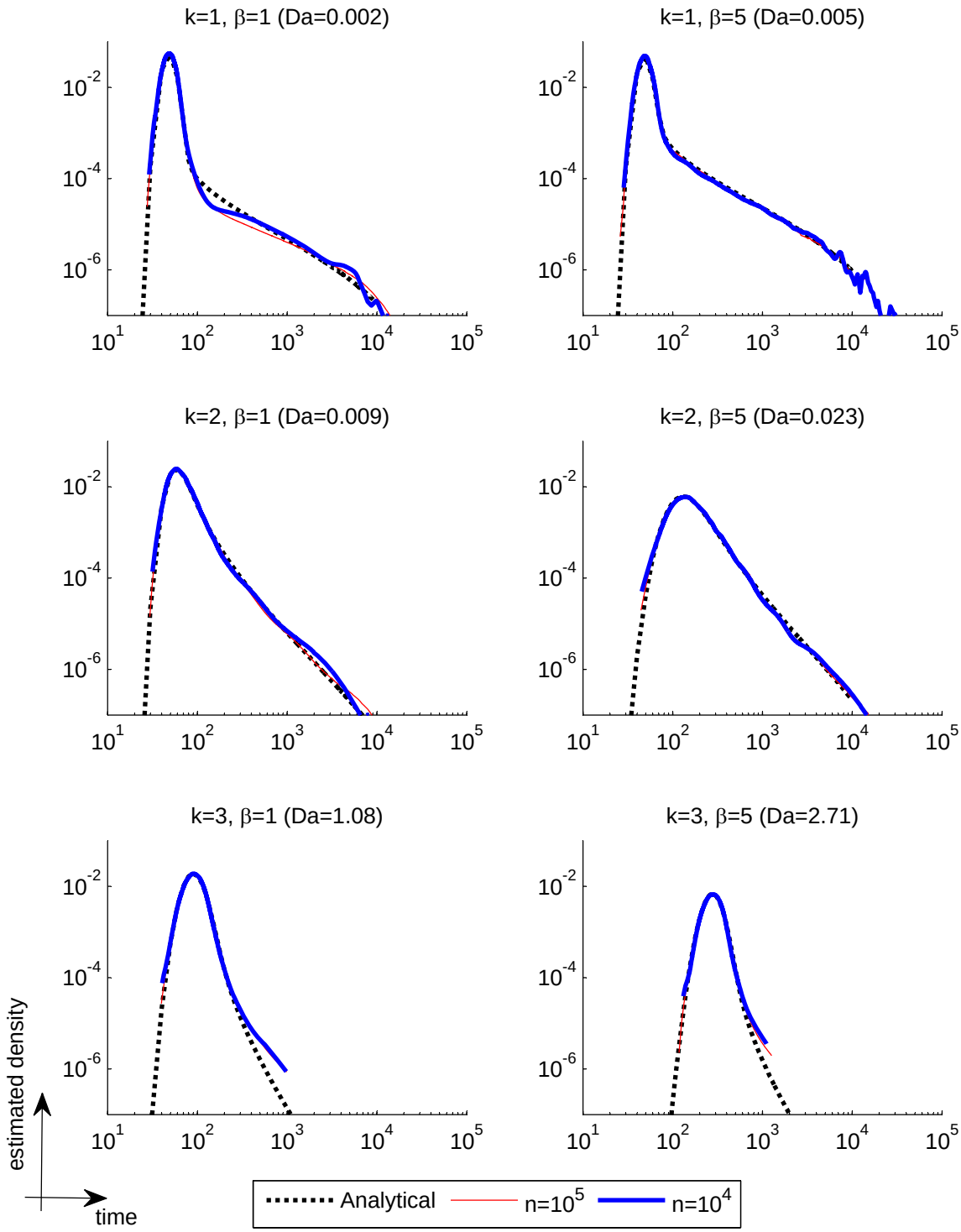

Global bandwidth (GB) method
k=1, β=1 (Da=0.002)
k=1, β=5 (Da=0.005)
k=2, β=1 (Da=0.009)
k=2, β=5 (Da=0.023)
k=3, β=1 (Da=1.08)
k=3, β=5 (Da=2.71)
estimated density
time
Analytical
n=10^5
n= 10^4

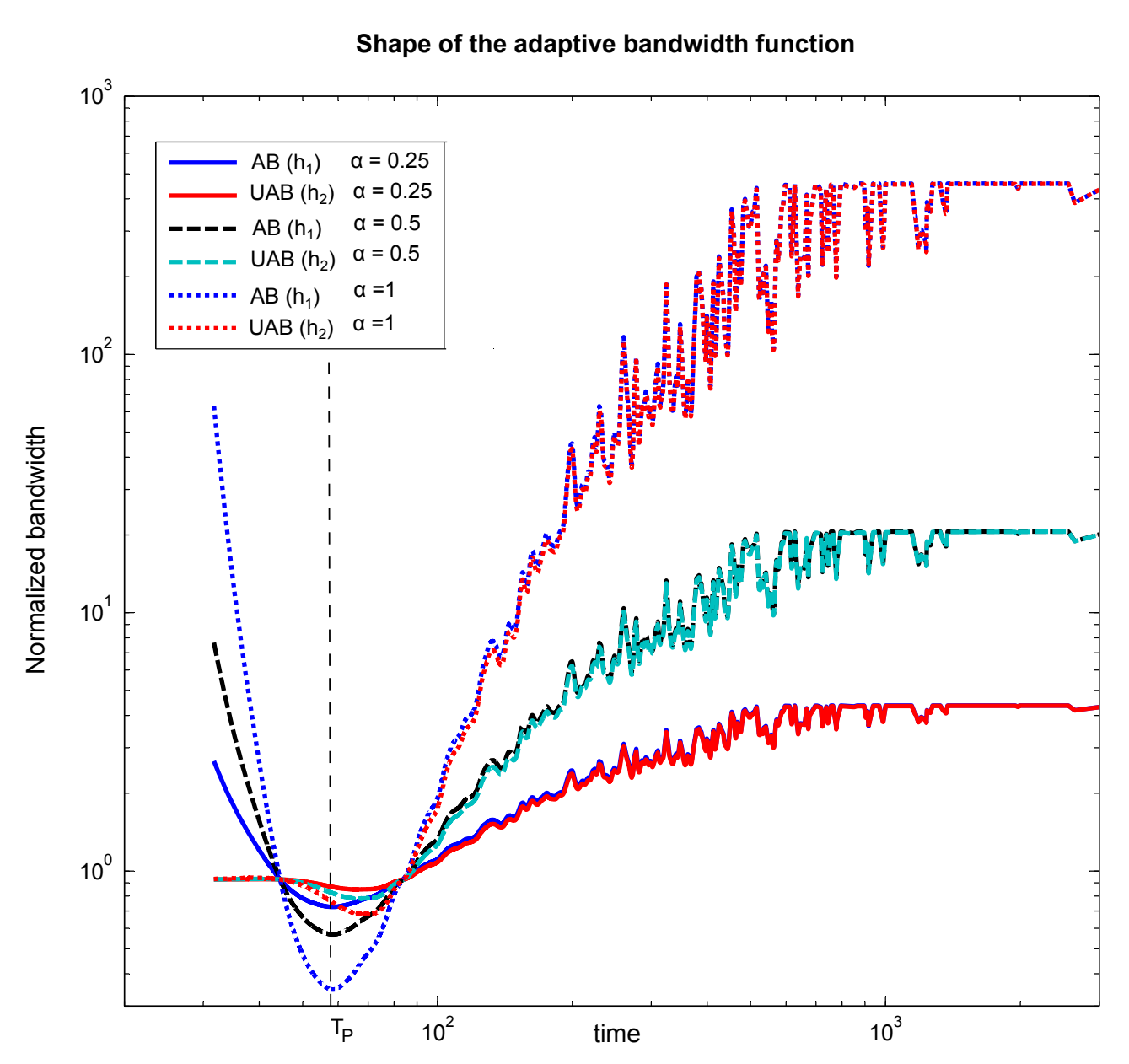

Adaptive bandwidth (AB) method - α=0.5
k=1, β=1 (Da=0.002)
k=1, β=5 (Da=0.005)
k=2, β=1 (Da=0.009)
k=2, β=5 (Da=0.023)
k=3, β=1 (Da=1.08)
k=3, β=5 (Da=2.71)
estimated density
time
Analytical
n=10^5
n=10^4

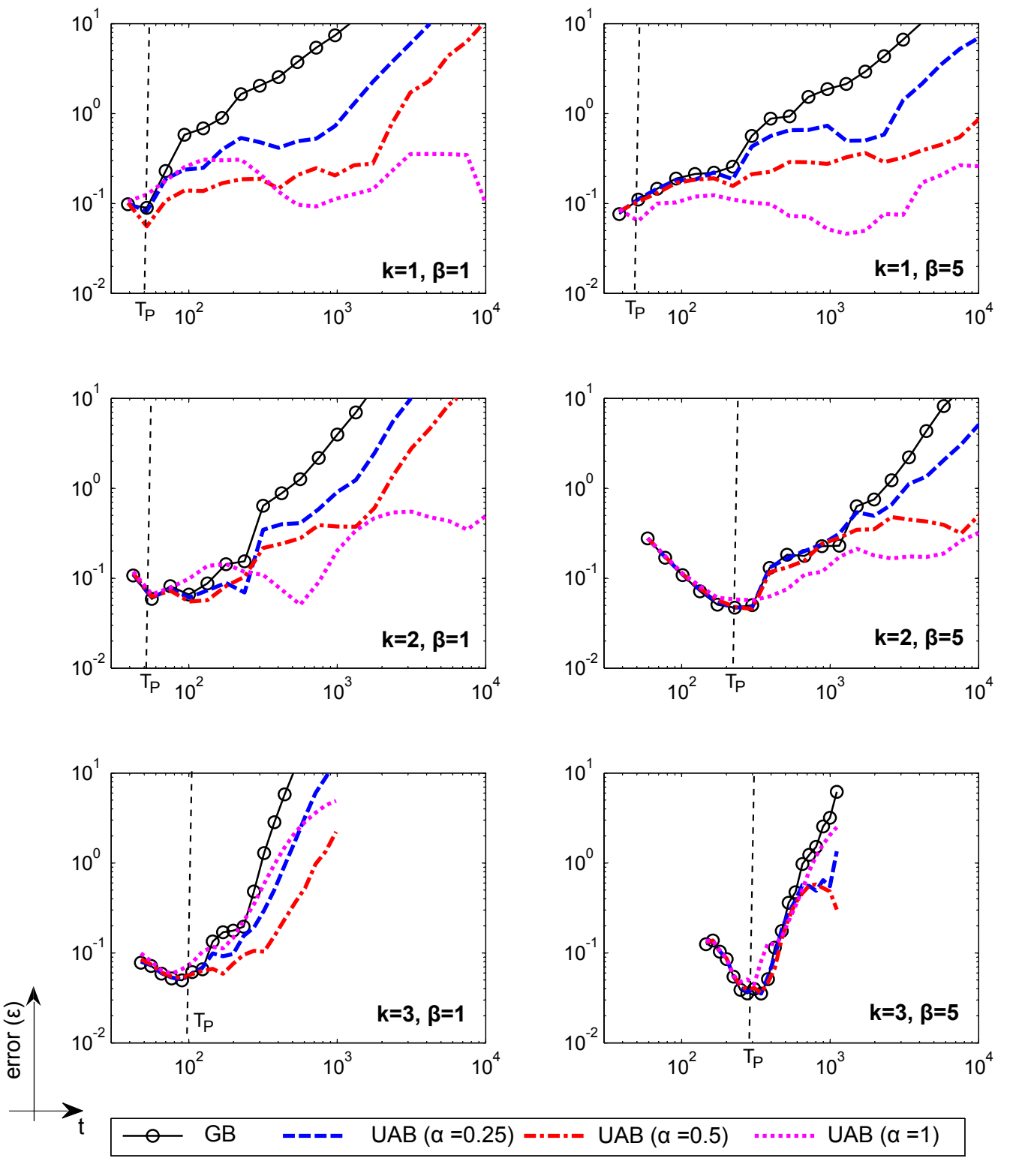

RELATIVE ERRORS (GB and AB methods , n=10^4)
k=1, β=1
k=1, β=5
k=2, β=1
k=2, β=5
k=3, β=1
k=3, β=5
T_P
error (ε)
t
GB
AB (α =0.25)
AB (α =0.5)
AB (α =1)

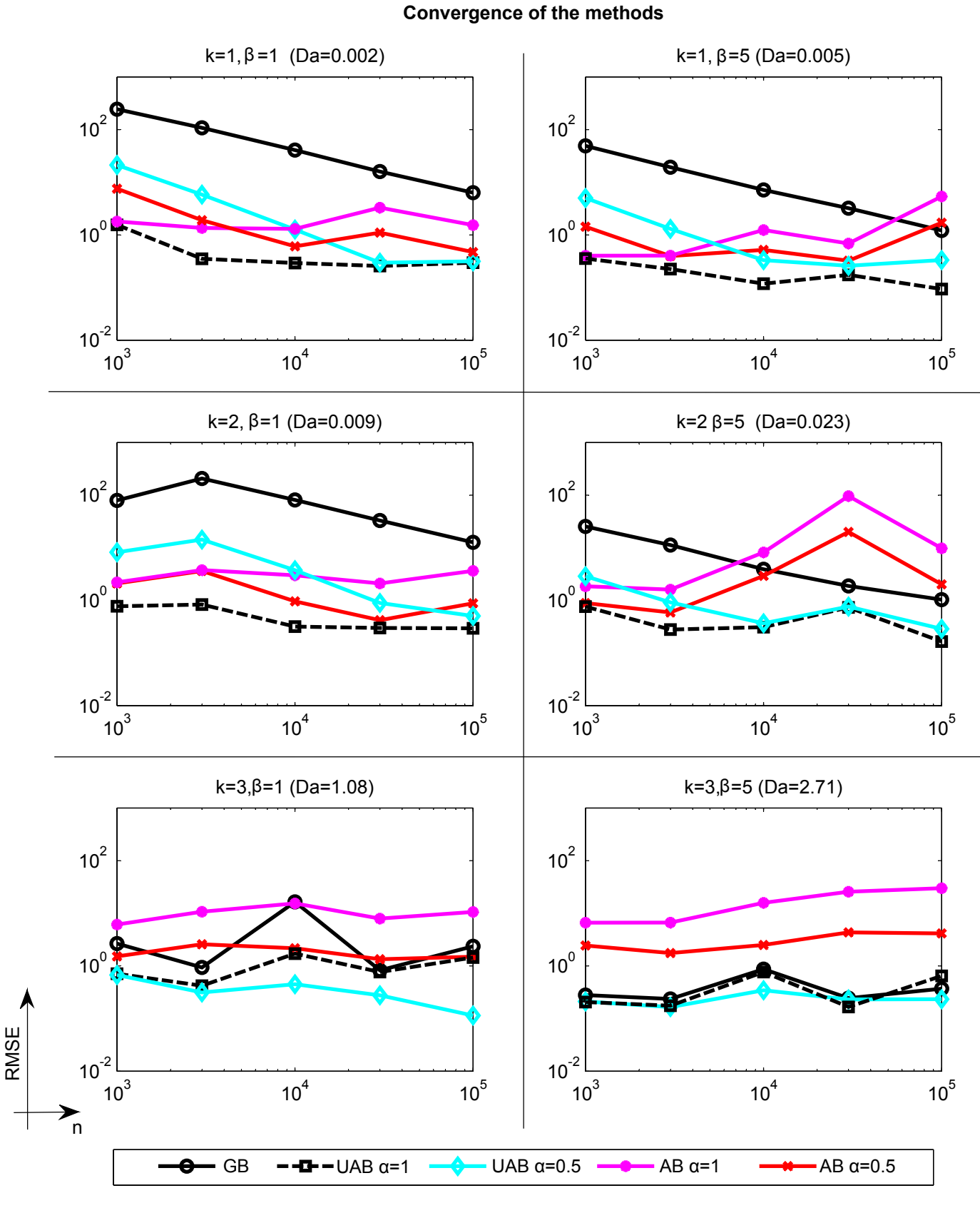

Universal adaptive bandwidth (UAB) method - α=0.25
k=1, β=1 (Da=0.002)
k=1, β=5 (Da=0.005)
k=2, β=1 (Da=0.009)
k=2, β=5 (Da=0.023)
k=3, β=1 (Da=1.08)
k=3, β=5 (Da=2.71)
estimated density
time
Analytical
n=10^5
n=10^4


Universal adaptive bandwidth (UAB) method - α=0.5
k=1, β=1 (Da=0.002)
k=1, β=5 (Da=0.005)
k=2, β=1 (Da=0.009)
k=2, β=5 (Da=0.023)
k=3, β=1 (Da=1.08)
k=3, β=5 (Da=2.71)
10^-2
10^-4
10^-6
10^1
10^2
10^3
10^4
10^5
estimated density
time
Analytical
n=10^5
n=10^4


Universal adaptive bandwidth (UAB) method - α=1
k=1, β=1 (Da=0.002)
k=1, β=5 (Da=0.005)
k=2, β=1 (Da=0.009)
k=2, β=5 (Da=0.023)
k=3, β=1 (Da=1.08)
k=3, β=5 (Da=2.71)
10^-2
10^-4
10^-6
10^1
10^2
10^3
10^4
10^5
estimated density
time
Analytical
n=10^5
n=10^4


Shape of the adaptive bandwidth function
AB (h1) α = 0.25
UAB (h2) α = 0.25
AB (h1) α = 0.5
UAB (h2) α = 0.5
AB (h1) α =1
UAB (h2) α =1
Normalized bandwidth
time
TP
10^3
10^2
10^1
10^0


RELATIVE ERRORS (GB and UAB methods , n=10^4)
k=1, β=1
k=1, β=5
k=2, β=1
k=2, β=5
k=3, β=1
k=3, β=5
T_P
error (ε)
t
GB
UAB (α =0.25)
UAB (α =0.5)
UAB (α =1)


Convergence of the methods
k=1, β=1 (Da=0.002)
k=1, β=5 (Da=0.005)
k=2, β=1 (Da=0.009)
k=2 β=5 (Da=0.023)
k=3,β=1 (Da=1.08)
k=3,β=5 (Da=2.71)
RMSE
n
GB
UAB α=1
UAB α=0.5
AB α=1
AB α=0.5