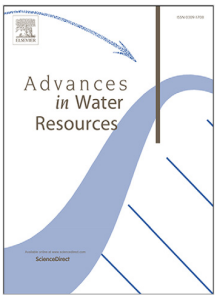
Upscaled models for time-varying solute transport: Transient spatial-Markov dynamics

Nicholas B Engdahl, Tomás Aquino

Journal Pre-proof



Upscaled models for time-varying solute transport: Transient spatial-Markov dynamics

Nicholas B. Engdahl, Tomás Aquino

PII: S0309-1708(22)00137-3
DOI: https://doi.org/10.1016/j.advwatres.2022.104271
Reference: ADWR 104271

To appear in: *Advances in Water Resources*

Received date : 16 March 2022
Revised date : 14 June 2022
Accepted date : 5 July 2022

Please cite this article as: N.B. Engdahl and T. Aquino, Upscaled models for time-varying solute transport: Transient spatial-Markov dynamics. *Advances in Water Resources* (2022), doi: https://doi.org/10.1016/j.advwatres.2022.104271.

Upscaled models for time-varying solute transport: Transient spatial-Markov dynamics

Nicholas B. Engdahl[1,*] and Tomás Aquino[2]

[1]Civil and Environmental Engineering, Washington State University, Pullman, WA, 99164-5818, United States

[2]Univ. Rennes, CNRS, Géosciences Rennes, UMR 6118, 35000 Rennes, France

[*]Corresponding author: nick.engdahl@wsu.edu

June 14, 2022

Abstract

Correlated velocity models (CVMs) have proven themselves to be effective tools for describing a wide range of solute transport behaviors in heterogeneous porous media. In particular, spatial Markov models (SMMs) are a class of CVMs where subsequent Lagrangian velocities along transport trajectories depend only on the current velocity, and not on past history. Such models provide a powerful tool for modeling transport in terms of a limited number of flow properties, such as the Eulerian point distribution of (flow) velocities, tortuosity, and the spatial scale of persistence of velocities. However, to date, all SMM modeling frameworks and applications have assumed that the underlying flow is steady-state. In this work, we extend SMMs to the case of time-varying flows. We propose, compare, and validate alternative numerical implementations, and we determine conditions for validity and efficiency based on standard physical quantities used to describe flow and transport at the Darcy scale. The models require additional information relative to a steady-state velocity SMM and we discuss the conditions under which this extra burden is warranted. We also provide clear, deterministic tests for the validity of the transient SMM, termed the "slow variation" and "fast propagation" criteria, which offer clear guidance on when transient, upscaled models are reasonable to employ. Our work forms the basis of a new framework allowing for the application of efficient upscaled models of transport to realistic transient flow conditions.

1 Introduction

The general aim of upscaled models of solute transport in porous media is to capture the impacts of inhomogeneities without explicitly representing the mechanisms that drive transport and/or their spatiotemporal variability (Dentz et al., 2020; Sund et al., 2019). The philosophy behind upscaled methods revolves around the notion that the computational and data-support burdens imposed by distributed models incur significant computational costs and lend sufficient uncertainty to predictions such that distributed models are not necessarily practical in every circumstance. A reduced-complexity strategy can be advantageous in many such cases. One of the promising upscaled transport frameworks is that of the continuous time random walk (CTRW), where the transition times between steps are modeled as a random variable (Berkowitz et al., 2006; Scher and Lax, 1973; Scher and Montroll, 1975). The model for the spatial increments and associated transition times distinguishes different flavors of CTRWs. A contemporary group of methods that have demonstrated broad applicability are correlated velocity models (CVMs), which in particular employ a fixed-length spatial step discretization.

In the conceptual model behind a CVM, travel times between adjacent steps in a CTRW are not independent and identically distributed events, due to correlations in the velocity field. Consider a Lagrangian particle moving through a natural system whose velocity is sampled at fixed spatial increments along its streamline. Natural media are often characterized by well-defined characteristic lengths, such as the mean lengths of hydrofacies (Carle and Fogg, 1996; Lee et al., 2007; Weissmann et al., 1999), and this means that a Lagrangian particle moving quickly along a preferential flow path is more likely to continue moving quickly than it is to abruptly slow down, though both options are possibilities. As the distance between sample locations increases, the Lagrangian (i.e., particle trajectory) velocity correlations decay proportionate to the spatial scales of the geological formations (Sherman et al., 2020), and the transitions eventually become uncorrelated. Models of the transition time to complete the "next" step in the random walk can leverage these correlations by conditioning the transition time based on the most recent step. This is precisely what is done in a Spatial Markov Model (SMM) (Dentz et al., 2016; Le Borgne et al., 2008a,b), where transitions are conditioned on the "previous" step only. An SMM can be parameterized in terms of a small number of properties with clear physical meaning, such as Eulerian velocity statistics, tortuosity, and spatial correlation lengths of Lagrangian velocities, which are related to the characteristic spatial scales of the hydrogeology (Aquino and Le Borgne, 2021; Dentz et al., 2016; Le Borgne et al., 2008). Although SMMs are Markovian in terms of the number of steps taken by a particle, and therefore in space due to the fixed

spatial increments, the resulting temporal dynamics may ultimately be non-Markovian as a result of broadly-distributed waiting times (De Anna et al., 2013; Holzner et al., 2015; Kang et al., 2014; Meyer and Saggini, 2016; Meyer and Tchelepi, 2010). The flexibility of the SMM allows complex transport phenomena to be modeled within its framework, resulting in significant conceptual and computational simplifications when compared to other CTRWs that otherwise require explicit modeling of nonlocal transport mechanisms, i.e., arbitrarily far particle jumps or dependency on long-term trajectory history (Berkowitz et al., 2006; Klages et al., 2008; Meerschaert and Sikorskii, 2012; Metzler and Klafter, 2004). Applications of SMMs to date have been diverse with compelling results obtained across a diverse spectrum of situations (Bolster et al., 2014; Comolli et al., 2019; Dentz et al., 2020; Hakoun et al., 2019; Kang et al., 2011; Puyguiraud et al., 2019a,b, 2021; Sherman et al., 2017, 2019; Sund et al., 2015a,b, 2017; Wright et al., 2019). However, one of the limitations of all SMM applications to date is that the transitions have been exclusively assumed to be stationary in both space and time, even in the case of multi-continuum formulations (Engdahl and Bolster, 2020; Kim and Kang, 2020).

The assumption of spatial stationarity often makes sense in the context of the linkages between SMM transitions and hydrogeologic correlations, and many studies have shown that stationary upscaled models are effective in certain heterogeneous media (Hakoun et al., 2019; Puyguiraud et al., 2019a). Allowing for spatial non-stationarities is not a particularly difficult issue to address, at least conceptually, because one could simply apply a different correlation model at different positions along the path of a Lagrangian particle (Aquino and Le Borgne, 2021). These correlation changes could be defined to coincide with known changes in the hydrogeology, so the only implementation barrier is developing different models of correlations for the different regions and deciding on the cutoffs for each. To do so may be time-consuming and require additional data, but it is not technically challenging, nor is it beyond the capabilities of current SMM frameworks.

The issue of temporal non-stationarities (transience) is significantly more involved because CVM formulations are based on connections between geological structure and spatial correlations. All work on CVMs has employed steady-state velocity fields, and it is unclear if such correlations between structure and velocity remain when the flow field varies in time. In reality, flow paths can change significantly due to transience, especially when flow is driven by spatially-distributed recharge or in unconfined settings (Engdahl, 2017). Transience can also impart non-uniqueness when an aggregated transport metric like a breakthrough curve is used. For example, two particles entering the same point of a distributed velocity field at two different times could take two different paths (drastically so in the case of variably-saturated flows, Engdahl and Bolster, 2020). Similarly, different particles

entering at different locations may ultimately have similar travel times to reach a fixed monitoring point because of transient changes in the flow field. These cases, and many more, would immediately invalidate assumptions of even weak stationarity (i.e., stationarity of increments), which would seem to deal a crippling blow to the conceptual underpinnings of all the current CVMs. One option to deal with these issues would be to relegate CVMs to cases of strict stationarity where transient effects are sufficiently averaged out. However, our perspective is that doing so would be unnecessarily limiting, because a more careful inspection of SMMs suggests that they can be adapted to accommodate at least some transient velocity fields if some care is taken. At a minimum, an upscaled representation of these transient processes should (i) be conditional to the "clock time" at which a particle entered the flow field, and (ii) somehow account for the temporal changes in upscaled velocity distributions, correlations, or both. As with any upscaled model, some simplifying assumptions are necessary, but in this case we will show that conditions for validity and numerical efficiency can be posed in terms of the typical physical parameters used to describe flow and transport in porous media at the Darcy scale.

The central questions addressed in this article are how to generalize (correlated) CTRWs to the case of transient velocities, and what conditions are necessary for these generalizations to be valid and practical. The motivation is to preserve the theoretical and computational benefits of SMMs when the underlying flow field is time-dependent. Several options of varying complexity are evaluated to accomplish this goal, and we consider their benefits and pitfalls in the context of analytic and numerically-defined transient velocity fields. We start by reviewing the basic concepts of the SMM and assessing its limitations regarding transient flow fields. Three approaches to accommodate transience are then developed, and we show that two of these are sufficiently robust for general applications. Specific criteria are developed for the validity of the transient SMM. The approach requires no further specific assumptions about the underlying flow field, but we focus here on flow through porous media at the Darcy (aquifer) scale. We validate our results against numerical simulations using both analytical and realistic flow fields where transience is induced by time-varying (periodic) boundary conditions. In the interest of compactness, the concepts and examples are demonstrated using a Bernoulli relaxation model for the Markov velocity process (Dentz et al., 2016), so we close with a discussion of how the approach can be generalized to other forms of transient CVMs. Collectively, the results advance the capabilities of CVMs to include transience and offer clear guidance regarding when these models would be appropriate and accurate.

2 Spatial-Markov model

SMMs are one of many CVMs that conceptualize (advective) transport in terms of Lagrangian particle trajectories, whereby a solute mass is discretized onto the particles (Sherman et al., 2020). Trajectories are usually modeled as a succession of steps of fixed length Δs along the streamlines of a flow and each step has a constant velocity, but the velocities may change as the particle completes successive steps. The basic concept is that the transition time (i.e., the step length divided by the velocity) distribution accounts for the heterogeneity in a flow field without explicitly modeling it, such that transport follows a stochastic-convective ensemble along streamlines. The step length corresponds to a choice of discretization of Lagrangian particle trajectories and the description converges to a continuum process in the limit of small $\Delta s \to 0$ (i.e., becomes independent of the discretization when it is sufficiently fine, as is expected of a properly-discretized model). Particle positions after k steps along a streamline (particle path) are denoted X_k with the corresponding times T_k to complete the kth step obey the stochastic recursion relations (Dentz et al., 2016)

$$X_{k+1} = X_k + \frac{\Delta s}{\chi}, \qquad T_{k+1} = T_k + \frac{\Delta s}{V_k}, \tag{1}$$

where V_k is the velocity magnitude during the kth step, which is constant throughout the step. Typically, the tortuosity χ is approximated by the average tortuosity, which is computed as the average of the Eulerian velocity magnitude divided by the average of its projection along the mean flow direction (Koponen et al., 1996),

$$\chi = \frac{\overline{v}}{\langle \boldsymbol{v} \cdot \hat{\boldsymbol{x}} \rangle}. \tag{2}$$

Here, $\boldsymbol{v}$ is the Eulerian velocity vector, $\hat{\boldsymbol{x}}$ is the unit vector along the mean flow direction, and $\langle \cdot \rangle$ denotes the average over space. The numerator represents the average of the Eulerian velocity magnitude, $\overline{v} = \langle |\boldsymbol{v}| \rangle$, so that $\chi \geqslant 1$. The initial time and position for each particle are often taken as $T_0 = 0$ and $X_0 = 0$, respectively (though nonzero positions and times are permissible), and the initial velocities V_0 are distributed according to the initial condition at this time.

The key ingredient of a spatial-Markov model is that the velocities V_k, seen as a function of k, form a Markov chain. The Markov property means that the probability of the next step having velocity V_{k+1} is conditional only on the most recent step's velocity V_k, and not on past history through earlier velocities. Under strict stationarity of the underlying flow field, the corresponding transition probabilities, given the current velocity, are constant in both space and time. Discretizing velocities

into classes, such that class i comprises velocities between b_i and b_{i+1} and has width $\Delta v_i = b_{i+1} - b_i$, the midpoint velocity $v_i = (b_{i+1} + b_i)/2$ is associated with class i. The velocity process is then characterized by the probabilities r_{ij} of transitioning to class i given that the current velocity is in class j.

In order for the velocities to correspond to a spatial-Markov process, the probability of transitioning to a different class must be proportional to the step length Δs, so that, for a given velocity, the spatial rate of transition (transition probability per unit distance) is constant and the transition probability decays exponentially with the step length (Van Kampen, 1992). The overall persistence of velocities is characterized by the correlation length ℓ_c of velocity magnitudes along streamlines, which at the Darcy scale is typically of the same order as the scale of spatial variability of permeability (Hakoun et al., 2019). Thus, taking into account that $\sum_i r_{ij} = 1$ for all classes j to conserve probability (a transition from any given velocity class j must end at *some* velocity class i), we write, for a small spatial step Δs compared to the correlation length ℓ_c (Aquino and Le Borgne, 2021),

$$r_{ij} = \frac{\Delta s}{\ell_c}\beta_{ij}(1-\delta_{ij}) + \left[1 - \frac{\Delta s}{\ell_c}(1-\beta_{ii})\right]\delta_{ij}, \tag{3}$$

where the dimensionless β_{ij} encode the velocity-dependence of the transition probabilities and δ_{ij} is a Kronecker delta. Thus, the term proportional to $(1-\delta_{ij})$ denotes the probability of changing to a different velocity class, whereas the term proportional to δ_{ij} denotes the probability of remaining in the same velocity class. As shown in Aquino and Le Borgne (2021), the corresponding dynamics result in a well-defined spatial-Markov process in the continuum limit of fine step discretization $\Delta s \to 0$, so long as the velocity class discretization associated with a given Δs is chosen such that the time increments $\Delta s / v_i \to 0$ for all classes i as $\Delta s \to 0$.

The full transition matrix of an SMM is an $N \times N$ matrix, where N is the number of velocity bins. This can be difficult to parameterize in practice, so we shall instead adopt an analytical model based on a discretized Bernoulli relaxation process for the velocities (Aquino and Le Borgne, 2021; Dentz et al., 2016; Sherman et al., 2020). We expect this approach to provide good results for quantities such as breakthrough curves at distances larger than a few correlation lengths (Hakoun et al., 2019; Puyguiraud et al., 2019a). Under this process, particle velocities persist on the scale of the correlation length ℓ_c. When a particle changes to a different velocity class in a given step, the probability of the new velocity being in class i is independent of the current velocity class j, and it is given by a prescribed equilibrium probability p_i^∞. In this sense, the Bernoulli process may be seen as the simplest Markov process that relaxes to a prescribed equilibrium distribution over a given characteristic scale.

This also provides a direct link to SMM paramaterizations based on Gaussian Copulas (Massoudieh and Dentz, 2020). Assuming that the probability of transition per unit length is constant and equal to $1/\ell_c$ implies that the probability of persistence is exponential (Feller, 2008; Van Kampen, 1992), and the transition probabilities are given by (Dentz et al., 2016)

$$r_{ij} = e^{-\Delta s/\ell_c}\delta_{ij} + \left(1 - e^{-\Delta s/\ell_c}\right) p_i^{\infty}. \tag{4}$$

Expanding in Taylor series for small $\Delta s/\ell_c \ll 1$ and comparing to Eq. (3), we obtain

$$\beta_{ij} = p_i^{\infty}, \tag{5}$$

independent of the current velocity class j as expected.

The probability p_i^{∞} must be defined in terms of flow properties in order for the Bernoulli process to relax to the correct velocity distribution for a given transport problem. To this end, we introduce the Eulerian velocity probability density function (PDF) p_E, defined such that $p_E(v)\,dv$ is the probability of finding a velocity in the infinitesimal vicinity dv of v at a uniformly-randomly chosen spatial location. In other words, the Eulerian velocity PDF represents the point velocity statistics of the underlying flow field, in terms of the spatial probability of occurrence. Note that the Eulerian mean velocity, which was introduced above as a spatial average, can also be computed from the Eulerian velocity PDF as $\overline{v} = \int_0^{\infty} dv\, v p_E(v)$. The equilibrium distribution of the Bernoulli process represents the distribution of velocities measured at a given downstream distance far from injection. Under the assumptions of flow incompressibility and ergodicity (*i.e.* velocity statistics sampled in time along a sufficiently long trajectory are the same as across the spatial domain), the corresponding equilibrium velocity PDF, called the s-Lagrangain velocity PDF in some works, is the flux-weighted Eulerian PDF (Dentz et al., 2016; Puyguiraud et al., 2019a),

$$p_F(v) = \frac{v p_E(v)}{\overline{v}}. \tag{6}$$

In the discretized description, p_i^{∞} is the probability associated with the discretized velocity class i,

$$p_i^{\infty} = \int_{b_i}^{b_{i+1}} dv\, p_F(v) \approx \Delta v_i p_F(v_i), \tag{7}$$

where the approximation holds for small velocity classes, $\Delta v_i/\overline{v} \ll 1$. The Bernoulli process is thus fully parameterized given knowledge of the Lagrangian (i.e., along streamlines) correlation length ℓ_c

and the Eulerian velocity PDF $p_E(v)$.

3 Non-stationary spatial-Markov model

Consider now how to generalize the previous description to situations where the underlying flow field depends on time. Specifically, we seek a spatial-Markov model that is (statistically) non-stationary in time, in order to reflect transience (i.e., time dependence) of the underlying flow field. In a real, distributed transport system, the local velocity of a Lagrangian particle depends on position and time, which change along particle trajectories; the particle transport paths may be changing as time passes and thus may not coincide with paths along instantaneous flow streamlines. A robust upscaled representation of general transport dynamics is hopeless, because this scenario implies that in general the position and transition time changes cannot be decoupled. This means that an SMM is not applicable unless some simplifying assumptions are made. Otherwise, the required three-dimensional random walk may have complexity comparable to a distributed model, defeating the purpose of upscaled modeling.

Conceptually, particle velocities in the upscaled model could be considered to change according to two mechanisms that represent the changes in a physical transport system: (i) As in the classical SMM, a particle moves according to the local velocity and then samples a new velocity at a different, nearby point in space; and (ii) The local velocity at a particle's position changes due to the time-dependent nature of the flow. In general, these two processes cannot be fully decoupled since they could be happening simultaneously, but under certain conditions an upscaled description remains possible. A critical evaluation reveals two criteria under which an SMM should remain valid and practical: (a) Slow (temporal) variation of velocities, and (b) Fast (spatial) propagation of velocity changes. Slow variation means that the temporal change in the flow distribution throughout the medium is sufficiently slow that many spatial transitions typically occur before appreciable changes in the local velocities. Fast propagation means that when substantial changes in the velocity field do occur, they act quickly throughout the spatial domain compared to transport processes, so that all changes in the velocity PDF can be safely approximated as synchronous, or instantaneous, throughout the domain. The latter has been a common assumption in many studies of transient transport behaviors (see Engdahl et al., 2016), suggesting it could also be adopted for SMM applications.

Even under these assumptions, the Eulerian velocity PDF representing spatial flow statistics still needs to be updated over time to reflect the transient changes. The remainder of this section is concerned with how, and how often, to do so, and the assumptions associated with these decisions.

In particular, the underlying Eulerian velocity PDF must be considered as transient in all of the specific cases analyzed below. The most practical approaches to achieving this consist in adopting parameterized PDFs where some or all of the parameters can be made functions of time. This important issue will be revisited in Section 5. For now, we merely posit that the transient Eulerian velocity PDF $p_E(v;t)$, describing point velocity statistics at each time t, is known, and we discuss three different candidates for implementing a discretized transient SMM.

3.1 Naïve explicit

The simplest version of a transient SMM is one where the velocity PDF is updated only at steps where velocity transitions occur. This "Naïve explicit" (NEX) scheme is still described by the recursion relations (1). The key difference is that the transition probabilities $r_{ij}(t)$ now depend on the current "clock time" of the random walker through the coefficients $\beta_{ij}(t)$, see Eq. (3). At each transition, the Eulerian velocity PDF is updated to $p_E(v;T_k)$, and the corresponding transition probabilities $r_{ij}(T_k)$ are calculated before determining the new velocity. Note that in the specific case of a Bernoulli random walk, particle velocities only change with a probability given by $\exp(-\Delta s/\ell_c)$, independent of the current velocity, but otherwise remain the same as in the previous step (see Eq. (4)); thus, in this case, the velocity PDF is only updated to accommodate transient changes when a Bernoulli-model change in velocity would occur. Thus, for a Bernoulli random walk, transience in-between transitions is effectively ignored.

The simplicity of this approach is appealing, but it suffers from significant limitations because it makes no attempt to identify when it is actually necessary to account for transient changes. As we will see, this means that it does not converge to the same solution as the more involved discretization schemes proposed below in the continuum limit of fine discretization $\Delta s \to 0$. When the flow field changes very slowly (in the sense of the slow-variation criterion developed in detail in what follows), the NEX model may provide sufficient accuracy in practice, but if the timescales of transience impart fluctuations faster than the travel times, which should occur often at low velocities, significant errors will accumulate because important transient changes are ignored. The necessary conditions for this NEX model to provide a realistic approximation may not be practical in many real-world situations. We nonetheless include it here for its conceptual simplicity and to highlight the role of the more subtle procedures developed for the following, more involved discretization schemes.

3.2 Turning point explicit

The problem with the primitive NEX model is that it is entirely oblivious to the rate at which the flow field changes. If the flow field changes quickly, many velocity updates are necessary in, potentially, a short time compared to standard SMM velocity transitions, especially for particles moving at low velocities. Thus, our goal is to find an approach where the time and number of velocity PDF updates are dictated by the magnitude of the temporal changes in the velocity PDF. Before continuing, recall that particle velocities in a transient SMM may change due to two mechanisms: (i) As before, a particle moves according to the local velocity, and samples a new velocity at a different, nearby point in space; and (ii) The local velocity at a particle's position changes due to the time-dependent nature of the flow.

In the context of a transient field, mechanism (i) requires a rule to determine the transition probabilities $r_{ij}(t)$ for times t over each time range between velocity changes. In turn, mechanism (ii) requires a rule to determine velocity transitions due directly to the change in the underlying flow field. First, we determine the time range characterizing appreciable velocity changes. Knowledge of the time-dependent Eulerian PDF $p_E(v;t)$, as a function of velocity v for each time t, implies knowledge of the mean Eulerian velocity as a function of time,

$$\overline{v}(t) = \int_0^\infty dv\, p_E(v;t) v. \tag{8}$$

Over a given time interval, which we call a variation window Δt_v, the difference in the average particle displacement associated with the change in mean velocity can be quantified through

$$\Delta s_v = |\overline{v}(t + \Delta t_v) - \overline{v}(t)| \Delta t_v. \tag{9}$$

The quantity Δs_v may be interpreted as the approximate error in the average particle displacement that would arise from not taking the mean velocity variability into account. The error in the usual discretized spatial-Markov description, associated with mechanism (i), is on the order of the discretization step length Δs. Thus, in order to obtain an error of the same order associated with discretizing mechanism (ii), we choose Δt_v such that $\Delta s_v = a \Delta s$, where $a \leqslant 1$ is a free parameter controlling the maximum step size under transience, and as such the magnitude of allowable errors. Note that this will in general correspond to a time-dependent variation window $\Delta t_v(t)$.

For given values of a and Δs, Eq. (9) can be solved numerically for Δt_v. The procedure leads to a

series of turning points $T_{v,k'}$ where variation of the Eulerian flow field is to be taken into account; for this reason, we call this approach the "Turning Point Explicit" (TPE) method. Specifically, we have

$$T_{v,k'+1} = T_{v,k'} + \Delta t_{v,k'}, \qquad T_{v,0} = T_0 = 0, \tag{10}$$

where $\Delta t_{v,k'} = \Delta t_v(T_{v,k'})$ is the variation window associated with the last turning point. Note that many transition times T_k associated with mechanism-(i) transitions are expected to occur between two turning points when the slow-variation condition (a) is met, as discussed in more detail below. A straightforward numerical procedure to determine the variation windows and associated turning points is described in Appendix A.

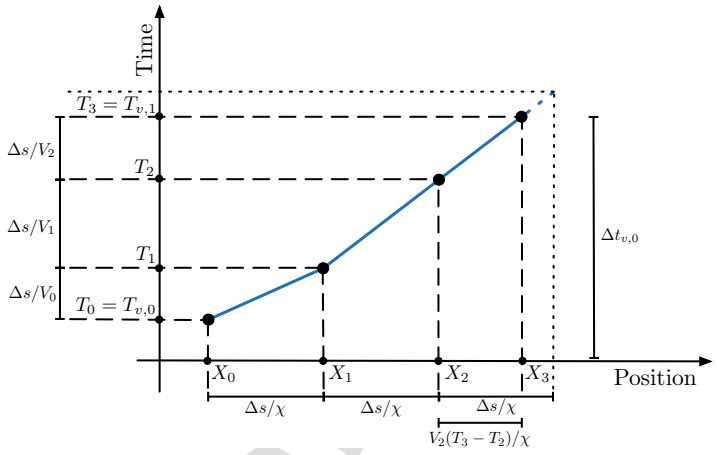


Figure 1: Illustration of the algorithm for mechanism (i), representing spatial-Markov transitions. As explained in detail in the text, starting at time $T_{v,0} = T_0$ and position X_0, the variation window $\Delta t_{v,0}$ is first computed based on mean flow velocity variability. The Markov transition times $\Delta s/V_i$ associated with steps of length $\Delta s/\chi$ are then computed, until the turning point time $T_{v,1} = T_{v,0} + \Delta t_{v,0}$ is reached. The portion of the last step corresponding to times exceeding $T_{v,1}$, represented by the dashed lines, is discarded, leading to the solid blue trajectory. Mechanism (ii) is then employed to find the new velocity at the new turning point, the Markov transition probabilities are updated according to the transient Eulerian PDF evaluated at $T_{v,1}$, and the algorithm is repeated.

We now formalize mechanism (i). Starting at the time $T_{v,k'}$ of the last flow-variation transition, determine the next variation window $\Delta t_{v,k'}$. Then, employ Eq. (3) for the transition probabilities $r_{ij}(t) = r_{ij}(T_{v,k'})$, together with the transition coefficients $\beta_{ij}(t) = \beta_{ij}(T_{v,k'})$, which depend on the choice of spatial-Markov process. The transition probabilities remain constant throughout the variation window. Next, update particle positions and times according to Eq. (1). However, when during some step k a particle's time would exceed the next turning point time $T_{v,k'+1}$ associated velocity variation

(Eq. (10)), the new time and position are determined according to

$$X_{k+1} = X_k + V_k \frac{T_{k+1} - T_k}{\chi}, \qquad T_{k+1} = T_{v,k'+1}, \tag{11}$$

in order to account for the partial completion of the step. Note that the remainder of the last transition distance and duration are discarded. Having determined that the next Markovian velocity transition has not occurred by time T_{k+1}, we know the velocity remains constant and equal to V_k during the partial step. The turning point corresponding to the next Markovian change in velocity can simply be recomputed in the next iteration without further assumptions due to the lack of memory of Markov processes (see, e.g., Van Kampen, 1992). The algorithm for mechanism (i) is illustrated in Fig. 1. This procedure is to be applied to all particles, followed by mechanism (ii), described below, and then repeated. Note that the NEX scheme proceeds similarly regarding the turning point times T_k and positions X_k, but does not require explicit variation windows. Rather, the position increments are always $\Delta s/\chi$, and the underlying transition probabilities are updated only when a velocity transition to a different class occurs.

Next, we turn to mechanism (ii), which corresponds to determining the new velocity at the flow-variation turning point times $T_{v,k'}$. Consider the transition probabilities of Eq. (3). Under a change in the flow field, these may change through the correlation length ℓ_c and/or through the velocity-dependent coefficients β_{ij}. In order to set up mechanism (ii) in a simple and physically-reasonable manner, we assume that the flow structure remains unchanged, maintaining the correlation length ℓ_c and the tortuosity χ constant, but the Eulerian velocity PDF may change in time, keeping its functional form but changing its mean through a rescaling. As familiar examples, this is the case at the pore scale when the underlying velocity field corresponds to Stokes flow, and at the Darcy scale when the hydraulic conductivity structure remains the same but the average head gradient driving the flow is rescaled. Once a transition due to velocity variation happens, at some transition time which we again name T_k (now with $k \geqslant 1$) for convenience, the local flow velocity at the particle's position is likely to have changed appreciably. To take this into account, mechanism (ii) consists of rescaling the particle's previous velocity according to the change in mean velocity,

$$V_{k+1} = \frac{\overline{v}(T_{v,k'+1})}{\overline{v}(T_{v,k'})} V_k, \tag{12}$$

or the corresponding class velocity in the discretized picture. This choice corresponds to assuming that the change in the velocity statistics can be approximated by an overall rescaling of the point

velocities, in line with the assumptions discussed above. This mechanism is applied to all particles, and the procedures described for mechanisms (i) and (ii) are then repeated. Note that, at the begin of procedure (i), the Markov transition probabilities are recomputed according to the velocity distribution at the new turning point time.

The correlation length and tortuosity are determined by the flow structure but can change in a given medium with an unchanging structure, for example due to the formation of preferential flow paths. While the mechanism (ii) rule can be applied to a case where the flow structure also varies, its physical significance is more difficult to justify. A more complex transition rule may be necessary in such cases, which we do not discuss further here.

3.3 Fully-implicit model

So far we have considered one method that only updates transition probabilities each time a velocity change takes place (and not at turning points where velocity remains the same), and one that automatically "detects" when updates are needed, which, in the process, may cause the step sizes to change (i.e., TPE). Another possibility is one where the spatial step size is chosen and fixed, but transient changes are always accommodated, no matter how big or small the transient fluctuation(s) may be. In practice, the concept of a variation window introduced for TPE subtly implies that, for a given finite step size Δs, the changes of the velocity PDF during a step are small enough that stochastic variations compensate for any inaccuracies imposed by the use of a constant velocity. In other words, the "true" velocity might be slightly higher/lower over any given step, but the average remains representative. An alternative interpretation of this nuanced point is that it assumes that small changes to the *probability* associated with a given velocity are insignificant inside an appropriately-sized variation window. Transposing this argument, one could instead assume that small changes to a velocity have an insignificant impact on its probability over the time of the transition, which leads us to the third strategy.

The key assumption for the following approach is that the cumulative probability associated with a particle velocity,

$$P(v) = \int_0^v dv\, p_E(v), \tag{13}$$

does not change during a spatial step, or that a particular particle's velocity *rank* on the cumulative density function (CDF) remains constant over any given step. This is similar to the assumption made under the TPE method, where changes in the underlying flow field were modeled as a constant rescaling

of the velocity PDF due to change in the mean velocity. For example, at $t = 0$, perhaps $v = 0.1$ has cumulative probability $P(v) = 0.8$ (20% of velocities above 0.1), but at $t = 1$ the overall flow increases such that $v = 0.15$ now corresponds to $P(v) = 0.8$ (20% of velocities above 0.15); in other words, a particle that begins moving with $P(v) = 0.8$ holds this rank throughout a step even as the velocity associated with this rank evolves.

Discretizing velocities in terms of rank, and denoting the velocity of a random walker conditional to a particular probability value (or rank) as $v_p(t)$, where p denotes the associated rank class, we can consider the trajectory of a particle along the SMM path as an equation of motion for each step. Since within a transition the particle velocity is allowed to change but the rank remains fixed, each step in the 1d random walk is described by the ordinary differential equation (ODE)

$$\frac{dX_p(t)}{dt} = \frac{v_p(t)}{\chi}, \tag{14}$$

where $X_p(t)$ is the downstream position, and $v_p(t)$ is a time-dependent function that describes the transient velocity as a function of clock time for a given probability rank class, p. For a step of known length Δs, this separable ODE has the general solution

$$\Delta s = \int_{T_k}^{T_{k+1}} dt\, v_p(t), \tag{15}$$

where $\Delta s = \chi[X_p(T_{k+1}) - X_p(T_k)]$ is the imposed displacement along particle paths, T_k is the clock time at the beginning of the step, and T_{k+1} is the unknown final time. Thus, particle positions in terms of step number k remain given by $X_{k+1} = X_k + \Delta s/\chi$, but transition times are determined according to an implicit equation.

Given a function for $v_p(t)$, the left-hand side of (15) is known and the right-hand side will be a function of T_{k+1} only, the unknown time when the step is finished, to be found via an implicit solution. The resulting equation will likely be nonlinear, but the solution of (15) for the final time, T_{k+1}, gives an exact solution when $v_p(t)$ may be approximated analytically, subject to the simplifying assumptions. We term this approach the "Fully-implicit model", since it requires the solution of an implicit (possibly nonlinear) equation for every particle in the random walk at every step. Note that this approach is an exact expression for the travel time when $v_p(t)$ is known analytically, with the single assumption that the probability associated with the velocities is constant for the duration of the step. Once a step k is completed, a new probability rank class may be determined analogously to before, according to the transition probabilities (3) associated with the step (*i.e.*, via a transition matrix or analytical

Markov process). The Eulerian velocity PDF is made a function of clock time, and the transitions probabilities are computed according to its form at the beginning of the step, $p_E(v;T_k)$, as for the previous methods.

Consider Fig. 1 for the turning points and variation windows of the TPE scheme. Like the NEX scheme, the fully-implicit method, as well as the approximations developed below, does not require the computation of variation windows, and the position increments are always $\Delta s/\chi$. Unlike NEX, however, transient changes are reflected in the transition probabilities at every step, and not only when velocity changes occur. Furthermore, as already discussed, velocity variability of a particle due to the transient changes within a transition can be captured, in which case the time increments are obtained implicitly via Eq. (15) rather than given directly by $\Delta s/V_k$. The stationary SMM case is easily recovered by defining, within step k, $v_p(t) = V_{p,k}$, where $V_{p,k}$ is constant within the step. Then, $\Delta s = V_{p,k}(T_{k+1} - T_k)$, and we conclude that the transition time is $T_{k+1} - T_k = \Delta s/V_{p,k}$, as expected. Next, consider a simple example of transience by assuming a linear increase in velocity over time: $v_p(t) = V_{p,k} + \alpha(t - T_k)$, where α is a constant growth rate. This gives the quadratic $\Delta s = \alpha(T_{k+1} - T_k)^2/2 + V_{p,k}(T_{k+1} - T_k)$, which has one real-valued, positive solution for the transit time, $T_{k+1} - T_k = (\sqrt{1 + 2\alpha\Delta s/V_{p,k}^2} - 1)V_{p,k}/\alpha$. Note that this reduces to the previous case when the growth rate or step size are sufficiently small such that $2\alpha\Delta s/V_{p,k}^2 \ll 1$. A similar technique can in principle be employed for any integrable function that defines $v_p(t)$, but we reiterate that solutions will likely need to be approximated using a nonlinear solver.

3.3.1 Fully-Transient explicit approximation

The fully-implicit scheme has the advantage of accounting for all the changes in $v_p(t)$ when the latter is known, but it should be evident that an implicit nonlinear solution for every particle at every step will be computationally demanding. The most obvious simplification is to use the velocity from the beginning of the time step in an explicit, first-order scheme that always updates the PDF for transience, so we abbreviate this Fully-Transient Explicit approximation as FTE.

A single evaluation of the velocity is used for every step, so that transient changes *during* the step are strictly ignored. The FTE then proceeds according to the recursion relations (1), with the velocities V_k selected according to (3), computed according to an Eulerian velocity PDF that is a function of clock time, updated as $p_E(v;T_k)$ at each particle step as before. The advantage of this approach is speed and simplicity but, like the NEX scheme, the cost is that it makes no attempt to account for transient changes during a spatial step. However, the velocity associated with each rank is updated at

every step, whereas the NEX scheme only accounts for changes in the transition probabilities when a transition leads to a *change in velocity class*, so TPE will naturally have advantages from an accuracy standpoint. Note that NEX and FTE can differ significantly under the Bernoulli relaxation model, because the probability of remaining in the same velocity class may be significant. Unlike the NEX scheme, for which velocity changes occur on the order of the correlation length ℓ_c, FTE accounts for velocity changes within a step length Δs. Thus, the accuracy in capturing the transient field increases when the spatial discretization is refined, leading to appropriate convergence of FTE. For a given Δs, the accuracy will be dependent on the nature of the transient signal, and significant errors should be expected anytime the velocity changes are large relative to the magnitude of the transition time and/or spatial steps.

3.3.2 Runge-Kutta 3 integration

Additional accuracy for an explicit approximation of the fully-implicit scheme can be obtained by adding more evaluations of the velocity distribution to create a Runge-Kutta predictor-corrector scheme. The velocity rank (cumulative probability) is required to be constant during each spatial step, since the velocity will be evaluated at multiple times for the "predictor" steps, but the associated velocity distributions being evaluated may be any arbitrary transient PDF that is defined as a function of time. This goes beyond the NEX and FTE schemes by accounting for transient changes *during* each step of the SMM, but avoids solving a nonlinear equation as is typically required for the fully-implicit scheme.

A good balance of accuracy and numerical cost is provided by the standard 3rd-order Runge-Kutta (RK3) scheme (Pozrikidis et al., 1998). For the trajectory of a temporally non-stationary random walker during step k and for a given rank p, the time at the end of the step, T_{k+1}, is computed according to

$$t^* = T_k + \frac{\Delta s}{2v_p(T_k)}, \tag{16a}$$

$$t^{**} = T_k + \Delta s \left[\frac{2}{v_p(t^*)} - \frac{1}{v_p(T_k)}\right], \tag{16b}$$

$$T_{k+1} = T_k + \frac{\Delta s}{6}\left[\frac{1}{v_p(T_k)} + \frac{4}{v(t^*)} + \frac{1}{v(t^{**})}\right], \tag{16c}$$

where t^* and t^{**} are the first and second predictor estimates of the time to complete the step. As before,

the new velocity at the end of each step k is computed according to the transition probabilities (3) associated with the step, obtained based on the time-dependent Eulerian PDF $p_E(v;T_k)$ updated at the beginning of the step.

We reiterate that the rank (probability) of the velocity is assumed constant throughout the spatial step but that the value associated with each predicted transition time can change; note that the assumption of constant rank is also necessary in other stochastic Runge-Kutta schemes (Engdahl and Aquino, 2018; Honeycutt, 1992). The three evaluations of the transient velocity lead to an accuracy of the scheme scaling as $\mathcal{O}(\Delta s^3)$; note that this is the accuracy of the *estimated transition time* $T_{k+1}-T_k$, and not the overall accuracy of a simulated breakthrough curve. We found that this RK3 scheme gives accuracy comparable to a direct, implicit solution of (15), while providing a computationally-efficient approach (but note that this may not hold when the velocity changes are not smooth and slowly-varying). As such, the fully-implicit solution is omitted below for the sake of brevity.

4 Applicability conditions

The four different explicit methods described above (NEX, TPE, FTE, and RK3) each have slightly different assumptions and conceptual models, but some general criteria must be met by the flow field for these methods to provide reasonable approximations of transport in a computationally-efficient manner. We posited in Section 3 that a transient SMM should be valid and efficient under (a) slow variation of the velocity, and/or (b) fast propagation of transient changes to the velocity field. Each of these merits some additional discussion in the context of Darcy-scale flow in aquifers.

4.1 Physical mechanics of the flow

Consider a section of a confined aquifer that has a well defined mean flow direction along which the head decreases. As long as the source of any transience is imposed outside the section under consideration (i.e. changes in recharge are applied some distance upstream of the section in question), we may quantify the impacts of those changes simply in terms of time-varying heads observed at each end (longitudinally) of the section and ignore the specific cause(s) of the transience. The impacts of the head changes at the boundaries on the velocity distribution within the aquifer depend on the heterogeneity of the various properties in the domain, but generally what matters is how much of the field is affected and how quickly.

Consider what happens in the case where a pressure pulse rapidly propagates through the aquifer.

Assuming the hydraulic conductivity structure of the medium remains unchanged, and no source/sink terms, changes in flow velocity across the medium are due to variations in head the reflect the underlying hydraulic conductivity field. The piezometric head, h, obeys

$$S_s \frac{\partial h}{\partial t} = \nabla \cdot \boldsymbol{K} \nabla h, \tag{17}$$

where S_s $[1/L]$ is the specific storage and $\boldsymbol{K}$ is the hydraulic conductivity tensor $[L/T]$, subject to appropriate boundary and initial conditions (Charbeneau, 2006). Assuming constant S_s (spatial variation produces an advective-type term), this is a diffusion equation for h, with the role of the diffusion coefficent played by the hydraulic diffusivity $[L^2/T]$

$$\boldsymbol{D}_H = \frac{\boldsymbol{K}}{S_s}. \tag{18}$$

Hereafter, we assume for simplicity a locally-isotropic $\boldsymbol{K}$ field, so that it is sufficient to consider the scalar (diagonal) values K and D_H. The conductivity K can vary spatially, so it is convenient to consider an average value for D_H that realistically homogenizes spatial heterogeneities, D_H^*, which could be computed, e.g., as a geometric (power) mean over $K(x, y, z)$ (Charbeneau, 2006). The timescale associated with the propagation of head perturbations across a distance ℓ, and associated flow variations, is then the diffusive timescale $\tau_H = \ell^2/(2D_H^*)$.

Over a given longitudinal length scale of interest, ℓ, the timescale associated with (advective) transport can be estimated as $\tau_A = \ell/\overline{v}$. The fast-propagation condition (b) can now be translated as the requirement that flow variations must propagate much faster than solute transport, $\tau_H \ll \tau_A$, corresponding to

$$\ell \ll \frac{2D_H^*}{\overline{v}}. \tag{19}$$

We have assumed that the main limiting factor is the propagation along the longitudinal direction, but a similar criterion could be developed that includes any propagation speed contributions from the lateral components.

We take ℓ equal to the length of the domain of interest. In that case, if condition (19) holds, the perturbation may be assumed to travel instantaneously across the domain, or that all velocities change instantly when a head change is applied at the boundaries. For large domains, this criterion could be relaxed by estimating ℓ according to the characteristic size of the solute plume through its longitudinal dispersion σ_x^2, such that $\ell \sim \sigma_x$. In this case, the perturbation can be assumed to cross the

entire plume instantaneously, but it may be necessary to delay the change in transition probabilities according to the time it takes the perturbation to reach the plume.

Finally, note that the Darcy equation itself does not dictate the velocity and it is assumed that the average local flow velocity is proportional to the local hydraulic conductivity K and the porosity. Since K typically exhibits much broader variability than specific storage, low hydraulic diffusivity D_H is commonly associated with low flow velocities, so that velocity in the lower-D_H regions is expected to be slower than the mean value, and conversely for the higher-D_H regions. Thus, we expect that employing an appropriate average, such as D_H^*, in Eq. (19) will usually lead to a reasonable estimate of the applicability of the fast propagation criteria, but factors like connectivity and extreme degrees of heterogeneity could impact this criterion.

4.2 Slow-variation criterion

As long as the velocity changes propagate across the domain sufficiently fast, the TPE and RK3 methods will provide good approximations of the transient PDF. However, it is advantageous from a computational standpoint if Eulerian velocities across the domain change sufficiently slowly in time that many transitions occur within a variation window Δt_v. In order to estimate Δt_v in terms of the variability in the mean velocity, consider the limit of small Δs, under which Δt_v is expected to be small. Then, Taylor expansion of Eq. (9) yields

$$\Delta s_v \approx \left|\frac{d\overline{v}}{dt}\right| \Delta t_v^2, \tag{20}$$

and, solving for Δt_v,

$$\Delta t_v \approx \sqrt{\frac{a\Delta s}{|d\overline{v}/dt|}}. \tag{21}$$

Note that the Taylor expansion leading to this result is inaccurate near local temporal extrema of the mean velocity, where $|d\overline{v}/dt| = 0$, which is why we employ the more robust numerical procedure described in Appendix A to compute Δt_v. However, this approximation provides a useful estimate of the role of flow variability. The number of mechanism (i) transitions within Δt_v is of order $\overline{v}\Delta t_v/\Delta s$, which we wish to be large. We thus obtain for the slow-variation condition (a):

$$\left|\frac{d\overline{v}}{dt}\right| \ll \frac{a^2\overline{v}^2}{\Delta s}. \tag{22}$$

In particular, for the spatial-Markov description to adequately resolve transport, we need $\Delta s \lesssim \ell_c$, and we must have $a \leqslant 1$. Thus, the minimal requirement for condition (a) to be met may be expressed as

$$\left|\frac{d\overline{v}}{dt}\right| \ll \frac{\overline{v}^2}{\ell_c}. \tag{23}$$

This is a time-dependent criterion, and the procedure may remain practical even if it does not hold for certain times. If this constraint holds, Eq. (22) may be used to choose

$$\frac{a^2\overline{v}^2}{|d\overline{v}/dt|} < \Delta s < \ell_c, \tag{24}$$

in order to ensure the method is both accurate and efficient. In practice, Δs can be chosen as the minimum of given multiples of the left and right terms in the inequality, e.g., $\Delta s = \min\{5\overline{v}^2/(|d\overline{v}/dt|), \ell_c/10\}$. Note also that Δs may be chosen adaptively, according to the temporal variation of the mean velocity, or constant according to a specific value such as the maximum or average of $\overline{v}^2/|d\overline{v}/dt|$ over the times of interest.

Combining the slow-variation condition, Eq. (23), and the fast-propagation condition, Eq. (19), we obtain

$$\sqrt{\ell_c\left|\frac{d\overline{v}}{dt}\right|} \ll \overline{v} \ll \frac{2D_H^*}{\ell}. \tag{25}$$

Given the spatial mean $\overline{v}(t)$ of the underlying flow field as a function of time, this result represents the conditions for practical applicability (accuracy and efficiency) of the transient spatial-Markov model, in terms of the velocity correlation length ℓ_c, the longitudinal scale of interest ℓ, and the (average) hydraulic diffusivity D_H^*.

5 Examples and cross-comparison

Existing analytical models for transport under transient velocities assume spatially-uniform flow fields (see Engdahl et al., 2016), and there are no closed-form analytical solutions for the transient, heterogeneous velocity fields that would lead to correlated transport. Accordingly, this section provides cross-comparison of the different transient SMM models under varying degrees of transience. We first compare the behavior of the four methods using a simplified analytical flow field, before moving on to numerical validation against direct simulations based on numerically-computed, spatially-heterogeneous velocity fields. We exclusively consider the Bernoulli process SMM hereafter. Recall that the Bernoulli SMM admits a minimal parameterization in terms of the Eulerian velocity PDF

and a velocity correlation length, providing a simple and parsimonious model. Nonetheless, any SMM transition mechanisms could be employed with minor modifications involving only the model's parameterization (see Sherman et al., 2020). Under the present choice, the coefficients $\beta_{ij}(t)$, which fully characterize the transition probabilities $r_{ij}(t)$ through Eq. (3), are obtained from the Eulerian PDF of point velocity magnitude statistics at a given time through Eq. (7). To further simplify the demonstrations, we also adopt a gamma PDF of Eulerian velocities with various prescribed time-dependent mean velocities $\overline{v}(t)$,

$$p_E(v;t) = \left[\frac{\theta v}{\overline{v}(t)}\right]^{\theta} \frac{e^{-\theta v/\overline{v}(t)}}{v\Gamma(\theta)}, \tag{26}$$

where $\Gamma(\cdot)$ is the gamma function. This type of PDF combines low-velocity power-law behavior (with scaling $v^{\theta-1}$, $\theta > 0$) with an exponential cutoff at high velocities. These features control long-term tailing of the resulting transit time distributions due to retention in low velocity zones as well as mean transit times, which in turn control key transport features such as mean plume displacement and longitudinal dispersion (Aquino and Le Borgne, 2021; Dentz et al., 2016). The gamma PDF has been employed to model Eulerian velocity PDFs in porous media both at the pore and the Darcy scales (Alim et al., 2017; Aquino and Le Borgne, 2021; Dentz et al., 2016; Holzner et al., 2015). The corresponding flux-weighted Eulerian (or s-Lagrangian) PDF, Eq. (6), is again gamma,

$$p_F(v;t) = \left[\frac{\theta v}{\overline{v}(t)}\right]^{\theta} \frac{e^{-\theta v/\overline{v}(t)}}{\overline{v}(t)\Gamma(\theta)}, \tag{27}$$

with the same exponential cutoff and a low-velocity dependency $\propto v^{\theta}$. Alternative parameterizations of the gamma PDF, along with fitting procedures, are discussed in Appendix B.

5.1 Analytically-defined velocities

The analytical cross-validation exercise assumes that *i*) a gamma distribution of velocities exists within the domain, and *ii*) the Eulerian mean velocity is described by a periodic function of the form

$$\overline{v}(t) = v_0 \left[1 + \eta \sin\left(\frac{2\pi(t+t_0)}{\tau}\right)\right], \tag{28}$$

where v_0 $[L/T]$ is a long-term mean velocity, η $[-]$ scales the magnitude of the velocity fluctuation (subject to $0 < \eta < 1$ so velocities remain positive), τ $[T]$ is the period of the transient cycle, and t_0 $[T]$ is a temporal shift. The corresponding transient gamma distribution for the SMM is then given by (27).

The example problem is defined by a domain length $L = 100$ $[L]$, tortuosity $\chi = 1$ for simplicity, velocity correlation length $\ell_c = 10$ $[L]$, gamma PDF exponent $\theta = 5$, long-term average velocity $v_0 = 0.04$ $[L/T]$, and temporal shift $t_0 = 0$. The four approaches (NEX, TPE, FTE, and RK3) are assessed under different τ and η combinations (Figure 2), and then at different discretizations to demonstrate convergence (Figure 3). Any number of parameter combinations could be used, but our goal is to demonstrate how transience impacts the model relative to a steady-state approximation. In each case, we provide comparison to a stationary SMM, which is obtained by setting $\overline{v}(t) = v_0$ and $\eta = 0$. We choose $\Delta s = 1$, so a random walker crosses a velocity correlation length in $\ell_c/\Delta s = 10$ steps and the full domain in $L/\Delta s = 100$ steps. We use 5000 random walkers; higher particle numbers did not have a significant impact on the results since we focus on mean behaviors, not on capturing tailing. The (cumulative) breakthrough curves (BTCs) at the downstream domain boundary for different parameter combinations of low/high magnitude (η) and small/large period (τ) of transience are shown in Fig. 2, with specific values shown in each panel. The time scales of transience were defined in terms of the average velocity (v_0) and domain length (L), corresponding to the typical time for a particle to cross the domain. In all of these plots, the FTE curve is under the RK3 curve at this scale, and both are usually close to the TPE curve. Only the NEX and SS (steady-state) curves are visibly distinct from the other transient models at all times.

An observation that can be made from Fig. 2 is that there are some cases where TPE differs from RK3. The reason for this is the parameter $\Delta s_v = a\Delta s$ in the TPE model; a value of a must be specified, which controls the magnitude of the "allowable" errors. Fig. 2 used $a = 0.5$, and this can be reduced to increase accuracy, at the cost of requiring more steps. Given sufficiently small a, and thus Δs_v, the TPE and RK3 results are essentially identical if the spatial discretization $\Delta s/\ell_c$ is also sufficiently small. This is shown via a convergence analysis in Fig. 3 with $a = 0.1$, where TPE, FTE, and RK3 all exhibit nearly identical mean travel times as $\Delta s/\ell_c$ is decreased (i.e., the number of steps needed to cross a correlation length is increased, so that all relevant structure in the flow field is resolved). Similar behaviors can be found for any fixed level of the BTC, but we only show convergence of the median arrival time for brevity.

5.2 Spatially-heterogeneous flow field

Our final example considers flow in a 2d, heterogeneous flow field subjected to transient boundary conditions. Here we simulate transport explicitly using fully-resolved Lagrangian random walk particle tracking (RWPT), and then compare the result to the proposed, upscaled, transient SMM schemes.

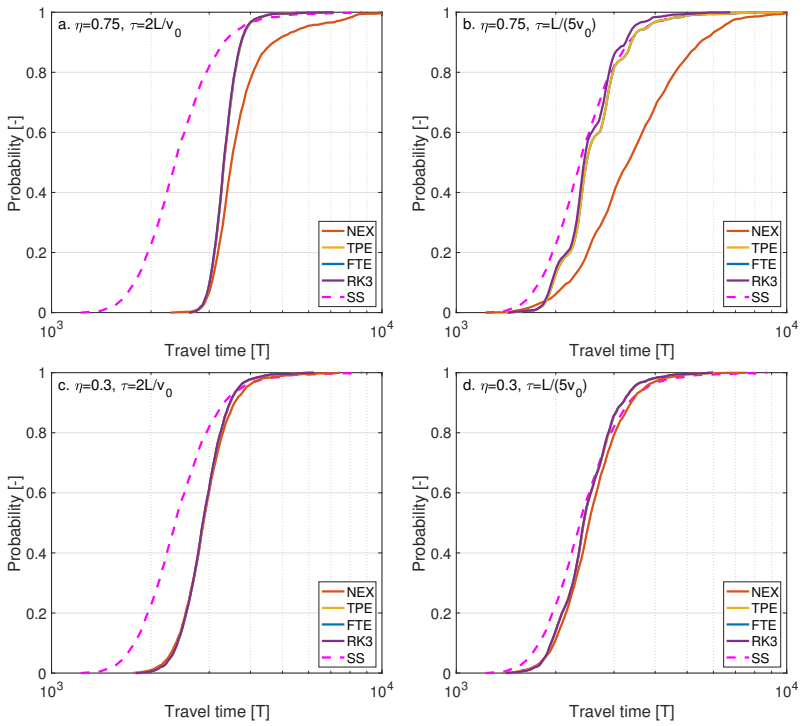


Figure 2: Comparison of the transient SMM models to a steady-state (SS) approximation for different magnitudes and periods of transient velocity changes. Small fluctuations with long periods may not require transient corrections, but it should be clear that as the frequency and magnitude of transient deviations increase the transient models depart significantly from the SS curve. Note also that three of the methods (TPE, FTE, RK3) generally agree with each other, whereas NEX is only reasonable under low-magnitude transience (small η).

The flow domain was defined to have a length $L = 100$ along the mean flow direction $[L]$ and an aspect ratio of $2:1$ (length to width). The hydraulic conductivity tensor was locally isotropic, and the scalar conductivity K in the domain was a log-normal multi-Gaussian random field with major and minor correlation length scales of $\lambda_1 = 10$ $[L]$ and $\lambda_2 = 6$ $[L]$, a geometric mean of $K_* = 0.2$ $[L/T]$, and unit variance of the log-K field. The specific storage and porosity were taken to be spatially constant and given by $S_s = 1.0 \times 10^{-5}$ $[1/L]$ and $\phi = 0.3$. A longitudinal spreading scenario was created by assigning zero-flux boundaries at the extents of the minor axes and Dirichlet boundaries at the ends of the major axes. The transient head changes were applied at the upstream boundary according to a periodic sine function that varied the gradient across the domain from 2% to 4% with a period of $\tau = 4000$ $[T]$. This fluctuation and parameter definitions satisfy the fast propagation and slow variation criteria, and are also representative of the kinds of fluctuations one can expect in real,

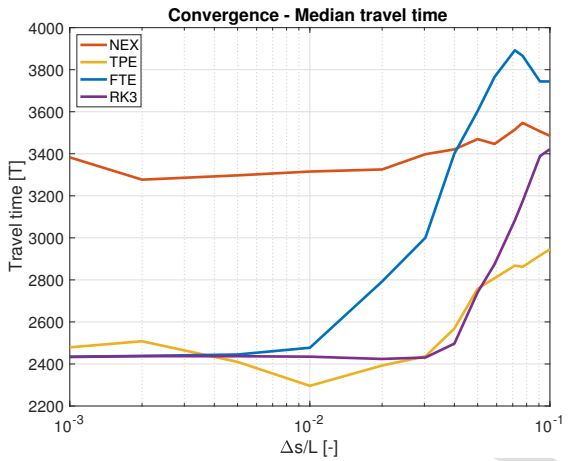


Figure 3: Convergence test of the proposed methods. As Δs is decreased, all but the NEX scheme approach the same mean behavior showing that they are solving the same system. TPE, FTE, and RK3 all converge to the same result for sufficiently refined discretizations, though they have different computational costs and assumptions.

undeveloped aquifers (Engdahl, 2017; McCallum and Shanafield, 2016).

The transient groundwater flow equation was solved using 2nd-order implicit finite differences and the domain was uniformly discretized into square cells of size $\Delta x = \Delta y = 1$ $[L]$. A snapshot of the velocity field, head contours, and streamlines is shown in Fig. 4. The time step of the transient model was $\Delta t = 1$ $[T]$ and flow and transport were solved sequentially at each time step using an operator-splitting scheme. The reference, fully-resolved, RWPT transport simulation used a flux-weighted initial condition of 10^5 particles released at $x = 5$ and tracked forward over time to $x = 95$ (corresponding to a length $L = 90$ for transport) to avoid any potential boundary impacts. Standard advective particle tracking methods were used, integrated in time with a 2nd-order Runge-Kutta scheme.

A gamma distribution was fit to the Eulerian velocity PDF at each time step of the transient flow simulation to simplify the parameterization of the SMM. The error of the fitted to the simulated distribution was computed to confirm that the simplified model was reasonable. Root mean squared relative errors of the fitted CDFs were small (≈ 0.011 over all times) and the worst linear correlation coefficient across all fits in time was $\rho = 0.991$; this shows the gamma PDF is a good approximation for this flow field, although it is not exact. Further, the Eulerian PDF was well described by Eq. (26) with fixed $\theta = 4.14$, and the effect of transient changes at the boundary on the transient mean velocity were modeled well by (28), with $v_0 = 2.06 \times 10^{-2}$ $[L/T]$, $\eta = 0.33$, $\tau = 4000$ $[T]$, and $t_0 = -\tau/2$. A comparison of the simulated and fitted transient velocity PDFs for 4 times is shown in Fig. 5. The time

scale of the transient changes, τ, was identical to that of the prescribed head changes, providing more evidence that the fast propagation assumption is valid in this case. In addition, the average gradient was 3%, so the $\eta = 0.33$ factor represents a fluctuation of $\pm 1\%$. This value matches the specified range of a 2% to 4% gradient and shows that the transient velocity model can be inferred from the transient boundaries.

The correlation length was estimated based on the multi-Gaussian field as $\ell_c = \lambda_1 = 10$ $[L]$. The upscaled Bernoulli SMM uses $\Delta s = 1$ $[L]$, corresponding to $\ell_c/\Delta s = 10$ and 90 steps to traverse the domain of transport, and tortuosity $\chi = 1.12$, which was computed directly from the flow field. An ensemble of 5000 random walkers were used for the SMM and the resulting BTCs for all four proposed transient SMM explicit schemes are shown in Fig. 6a, along with a steady-state SMM and the simulated BTC for comparison. The blue bars represent the (distributed, 2d) reference simulation. All data is binned according to the bars shown for the resolved simulations to make the comparison clearer; the value of each bar applies at its mid-point along the horizontal axis. The PDFs for the two best methods (TPE and RK3) are also shown in Fig. 6b, along with the steady-state simulation SS and the reference RWPT model; the main difference is that RK3 captures some of the secondary peak in the falling limb of the BTC. Note that, for the TPE method, $\Delta s_v = a\Delta s$ with $a = 1$ was used as for the analytical examples. As before, use of sufficiently small a and Δs would lead to similar results for TPE and RK3, at the cost of increased computational expense, but either of these schemes provides a good upscaled approximation of the simulated BTC.

The similarity of the different SMM approximations to the simulated BTC was assessed using root mean square error (RMSE) and the Hellinger distance (HD) metric (Hellinger, 1909), both applied to the PDF of travel times for the upscaled SMM simulations. The HD metric quantifies the overlap or similarity between the different PDFs, relative to the RWPT simulation. Values close to zero indicate strong similarity and values near one indicate high degrees of difference; smaller values mean better reproduction of the target distribution. Bianchi Janetti et al. (2020) used the HD metric to assess the performance of a trajectory-based SMM, demonstrating its utility in assessing SMMs. The RMSE and HD values are shown in Table 1 and demonstrate quantitatively that the transient versions all out-perform the steady-state SMM. The approximations of the fully-implicit model (see Section 3.3, FTE and RK3) offer a slight advantage but smaller values of $\Delta s_v = a\Delta s$ (through using smaller values of the free parameter a) would increase the accuracy of TPE to a similar extent. In any case, all the transient SMMs are considered overall good approximations. It is worth noting that the magnitude of the transient changes in this example are not as severe as some of those seen in Section 5.1, but there

Model	SS	NEX	TPE	FTE	RK3
HD	1.01×10^{-1}	9.63×10^{-2}	7.94×10^{-2}	5.50×10^{-2}	5.52×10^{-2}
RMSE	1.53×10^{-2}	1.20×10^{-2}	1.21×10^{-2}	6.46×10^{-3}	6.81×10^{-3}

Table 1: Hellinger distance (HD) metric and RMSE for the different SMM approximations of the simulated BTC. All transient SMM models show better performance than the steady-state model.

are clear departures from the steady-state model. This flow field is weakly heterogeneous due to its low log-K variance, so a higher degree of heterogeneity and greater contrast in the K field would likely lead to more significant departures. Whether or not these departures are significant enough to justify a fully-transient upscaled model leads directly into our discussion.

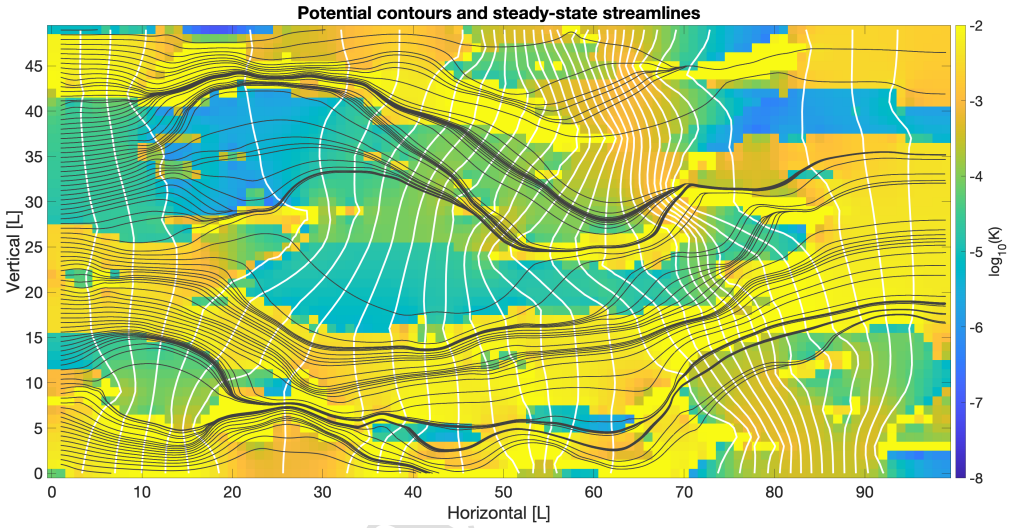


Figure 4: Heterogeneous, 2d flow field used for the transient model evaluation. White solid lines are contours of the potential field, black lines are the streamlines from the steady-state simulation, and the colors represent the base-10 logarithm of the velocity magnitude. The hydraulic conductivity field was generated using a hierarchical combination of transition probability geostatistics and stochastic multi-Gaussian fields.

6 Discussion and conclusions

The main purpose of this manuscript has been to determine if transient versions of spatial Markov models can be developed, and in this we have been successful. The heterogeneous velocity field example (Figure 4) with a time-dependent Dirichlet boundary condition verifies that transient SMM schemes can offer good upscaled approximations of key quantities such as breakthrough curves. As clearly seen in Fig. 6b, the proposed RK3 scheme most accurately captured the BTC, in particular regarding both

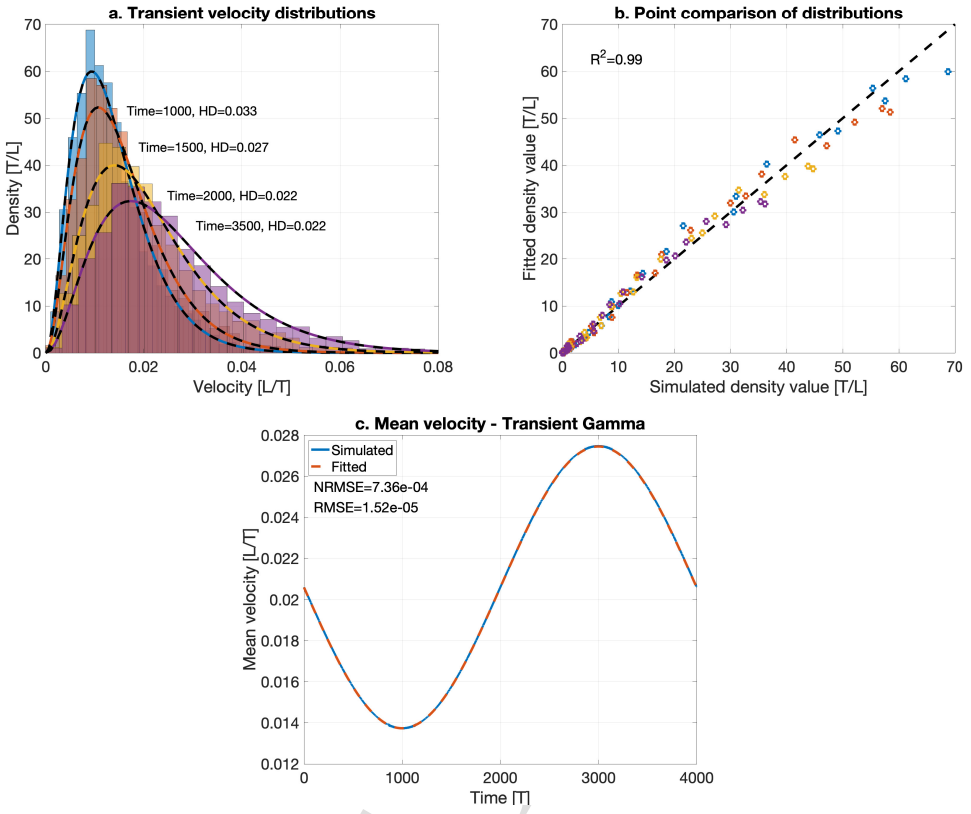


Figure 5: a. Simulated (bars) and fitted (lines) transient velocity PDFs in the 2d simulation domain for four times spanning the range of the transient cycle. HD denotes the Hellinger distance. b. Point-wise comparison of the fitted transient gamma distributions in panel (a) versus their simulated values. c. Comparison of the simulated and fitted mean velocity for the example flow field for one period of the transient cycle. Root mean squared (RMSE), normalized root mean squared errors (NRMSE), R^2, and HD confirm the accuracy and effectiveness of this functional approximation.

the maximum and transient-induced secondary peak. Nonetheless, both the TPE and RK3 methods offer accuracy for a reasonable increase in computational cost over stationary SMMs, and both TPE and FTE converge to the same answer as RK3 when the spatial step is sufficiently refined (see Fig. 3).

However, there is one major concern that cannot be overlooked, which is not unique to this study. A key question regarding practical application of any upscaled model is, can the model parameters be inferred reliably? In this case, the bare-minimum required elements for the transient SMMs are: i) the correlation length scale for the Bernoulli relaxation process, ii) a reference Lagrangian velocity distribution, and iii) a model for how that distribution changes over time. Each of these is considered independently in the following paragraphs.

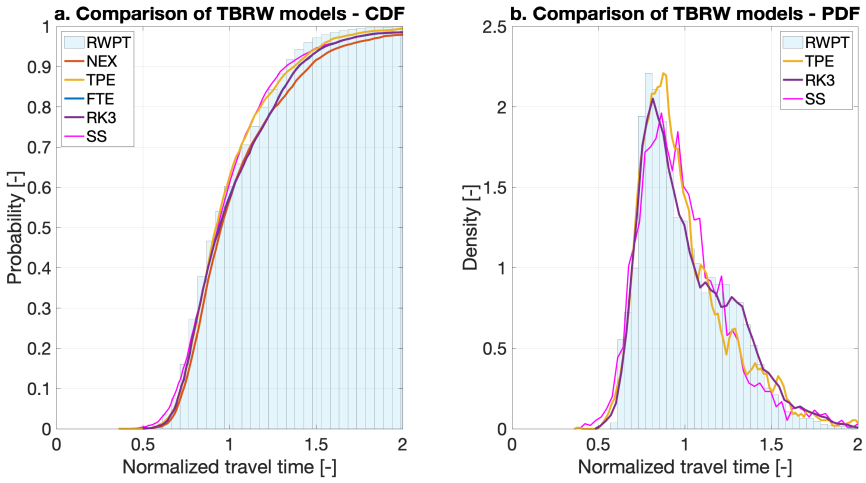


Figure 6: Comparison of the BTCs computed from the (fully-resolved) RWPT-based and (upscaled) transient SMM simulations. This example uses a small, but realistic, transient forcing that might be expected in natural aquifers. The advantage of the transient model is evident in the PDF plot, where the small secondary peak in the RWPT-BTC is captured by the RK3 scheme but completely missed by the steady-state (SS) approximation. As discussed in detail in the text, better results for the different transient methods can be obtained by further refining the discretization.

The first item is the model for the SMM transitions. We have assumed the spatial Markov correlations do not vary over time. There are strong connections between geological structure and the spatial correlations (Sherman et al., 2020), and geological structures generally change on time scales orders of magnitude larger than solute transport, so modeling the correlations does not represent a unique or undue burden to the transient random walk. Furthermore, if a full SMM transition matrix (e.g Engdahl and Bolster, 2020) was used instead of a Bernoulli relaxation process there are only a few more operational issues to consider. One is whether the initial and final bins change simultaneously as the Lagrangian velocity PDF evolves; we see no reason they would not evolve jointly, particularly since the model would become intractable if they did not. Another concern is whether the bounds on the individual velocity bins should evolve over time. The development of the fully-implicit scheme (Section 3.3) required that small changes in velocity cannot significantly impact the velocity rank, so the extension of this for broader validity is that slow changes to the velocity field cannot change the ranks of the distribution; this is merely another way of looking at the slow-variation criterion (23). In the absence of a strong transient boundary, our view is that a "shock" to the system would be necessary to invalidate the assumption of stable velocity ranks, such as abruptly turning on a large pumping well. If this were the case, the slow variation criterion would clearly be violated, negating

the benefits of application of the upscaled model in the first place. Aside from these, we see no obvious additional considerations necessary to adopt a full SMM transition matrix instead of a Bernoulli relaxation process.

The second item to consider is how to obtain the reference Lagrangian velocity distribution. This is arguably the most important yet difficult component to obtain. The best one could be expected to do is to use a data-driven, geostatistical description of the expected hydraulic conductivity field that is subjected to the anticipated boundary conditions for flow and transport. Evaluating this expected distribution might require methods like a stochastic Monte-Carlo ensemble, but each realization would be steady-state and so the ensemble should run quite fast. From these, the expected behaviors of the reference velocity distribution can be obtained, or any other threshold value (such as percentiles) to assess the uncertainty range, and the slow variation and fast propagation criteria (25) could easily be assessed at the same time. The resulting velocity distributions could then be used in a transient SMM in lieu of a large ensemble of transient Monte-Carlo simulations, which would surely offer large computational savings. We consider this a reasonable compromise, but it must be noted that uncertainties in the geostatistical description, including unresolved heterogeneities or non-stationarities, will propagate into the upscaled model as will uncertainties in the boundary conditions. It is also possible to estimate SMM model parameters from breakthrough curves alone (see Sherman et al., 2017), though doing so in aquifers would be hampered by incomplete sampling or recovery of a tracer. Estimating the velocity PDF remains challenging but methods exist by which it can be reasonably approximated, which is all one should expect when using an upscaled model.

The third item is the model for how the reference velocity distribution changes over time. The model for changes is at least "plausibly obtainable" because of the fast propagation criteria. The key point is that if (19) is satisfied then the relative changes at the boundaries of the flow field can be used to approximate the changes in the velocity PDF. Engdahl (2017) considered a system where combinations of transient Dirichlet boundaries were used at the ends of a confined, longitudinal domain where transport was simulated using the fully transient velocity fields. The results showed strong correlation between the transient forcing and the velocity fluctuations, meaning that relative changes in the mean can be inferred, hence our definition of (28). Long-term shifts in the mean may also be accommodated (e.g Massoudieh, 2013), which can quickly overwhelm higher frequency impacts on the mean. So, depending on the time scales of transport, it may be more important to capture long-term trends, which can be accurately inferred from observation well data, though models would be needed for forecasting. Some inaccuracies are inevitable, but as long as the estimated transient

signal is representative of a system's overall changes, reasonable results can be expected. Our example from Section 5.2 illustrates this idea: the fitted model for the velocity transience was based solely on the transience at the boundaries, and the model performed well.

Upscaled models should not strive to be perfect reproductions of transport behaviors, as this would invalidate their purpose of being large-scale approximations through over-fitting. The goal of the transient Spatial Markov models proposed herein is to balance the complexities of transient velocity fields with the simplicity of upscaled models using a framework that leverages recent advances in correlated velocity models. The main point of this discussion is that our definitions of the slow variation and fast propagation criteria (25) provide all the necessary evaluation criteria to assess the validity and usefulness of the proposed models for a given scenario. There is a need for site-specific data in order to evaluate those criteria, and the decision to use transient upscaled models likely comes down to the subjective question of sufficient data abundance: is there enough data to confidently build the desired model? To this we can offer no new insights because every case is unique. We can say that the data requirements for transient SMMs falls between those of steady-state SMMs and spatially explicit, distributed models. There are benefits to accuracy (Section 5) relative to the former, and clear advantages of speed relative to the latter, but ultimately the data dictate which models should be used for a given purpose.

A Numerical determination of the flow variation window

In this appendix, we describe a straightforward numerical approach to obtain the variation window Δt_v according to Eq. (9). Note that more sophisticated root-finding techniques could also be employed.

In order to sequentially determine the $\Delta t_{v,k'}$ associated with each of the turning point times $T_{v,k'}$, see Eq. (10), we consider a time resolution for step k' given by

$$\Delta t_{k'} = \frac{\Delta s_v}{\overline{v}(T_{v,k'})} = \frac{a\Delta s}{\overline{v}(T_{v,k'})}. \tag{29}$$

This resolution represents the time necessary to cross the spatial variation threshold $\Delta s_v = a\Delta s$ of Eq. (9) at the current mean velocity. We expect this choice to provide a good compromise between speed and accuracy, especially when the slow-variation condition (a) is met (see Section 3), but note that a finer or coarser resolution could be employed. The variation window $\Delta t_{v,k'} = n_{k'}\Delta t_{k'}$ is then determined in terms of the number $n_{k'}$ of time-resolution steps required to exceed the allowed variation

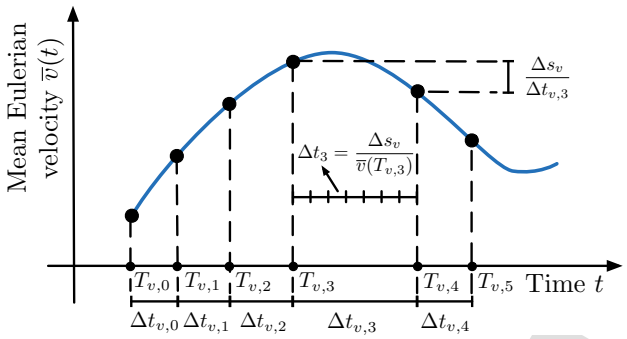


Figure 7: Illustration of the algorithm to determine the variation windows $\Delta t_{v,k'}$ associated with temporal variation of the Eulerian mean velocity. The variation windows $\Delta t_{v,k'} = T_{v,k'+1} - T_{v,k'}$ determine the turning point times $T_{v,k'}$, starting at $T_{v,0} = T_0$, at which flow velocity variations are taken into account. Each $\Delta t_{v,k'}$ is determined so that the mean velocity variation $\Delta v_{k'} = |\overline{v}(T_{v,k'} + \Delta t_{v,k'}) - \overline{v}(T_{v,k'})|$ is such that $\Delta v_{k'} \Delta t_{v,k'} = \Delta s_v$, where Δs_v is related to the spatial-Markov step size by a factor $a \leqslant 1$, $\Delta s_v = a\Delta s$. In order to determine these variation windows numerically, we consider a step-dependent maximum resolution $\Delta t_{k'} = \Delta s_v / \overline{v}(T_{v,k'})$, as illustrated for $\Delta t_{v,3}$. Then, $\Delta t_{v,k'}$ is approximated the smallest integer multiple of $\Delta t_{k'}$ such that $\Delta v_{k'} \Delta t_{v,k'}$ exceeds Δs_v.

Δs_v. Numerically, $n_{k'}$ can be computed as the smallest integer n such that

$$|\overline{v}(T_{v,k'} + n\Delta t_{k'}) - \overline{v}(T_{v,k'})| n\Delta t_{k'} > \Delta s_v. \tag{30}$$

In the simplest implementation, the value of the mean velocity over time is scanned sequentially, at a temporal resolution of $\Delta t_{k'}$, until the prescribed tolerance Δs_v is exceeded. This procedure is illustrated in Fig. 7.

B Parameterization and fitting of the gamma velocity PDF

The gamma PDF is typically parameterized in terms of a shape parameter α and a rate parameter ξ, defined such that

$$p_\Gamma(x; \alpha, \xi) = \frac{\xi^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\xi x}, \tag{31}$$

where, for a random variable with this distribution, $p_\Gamma(x; \alpha, \xi)\, dx$ is the probability of a value in the infiniteseimal vicinity dx of x. This PDF can be fit to velocity data directly by applying a standard minimum-square criterion to determine α and ξ.

In the present application, where the Eulerian velocities are taken to be gamma-distributed, it is convenient to choose a parameterization that emphasizes features that are key to solute transport.

The scale parameter α controls the tailing properties at low velocities, which control the large-time tailing of transit times and thus the late-time dispersion behavior (Aquino and Le Borgne, 2021; Dentz et al., 2016). Thus, we choose to keep $\theta = \alpha$ as a parameter. On the other hand, the mean α/ξ of the gamma distribution has a clear physical meaning in our context: it represents the spatial average of the velocity at a given time. Thus, we parameterize our Eulerian velocity PDF by setting $\xi = \alpha/\overline{v}(t)$, i.e.,

$$p_E(v;t) = p_\Gamma\left[v;\alpha,\frac{\alpha}{\overline{v}(t)}\right], \tag{32}$$

which corresponds to Eq. (26). To fit this form to velocity data at a given time, we fit α to the low-velocity behavior of the data PDF, and set $\overline{v}(t)$ to the spatial mean of the data.

Alternatively, we could enforce the correct average velocity $\overline{v}(t) = \alpha/\xi$ and velocity variance $\sigma_v^2 = \alpha/\xi^2$, which can be achieved by setting

$$p_E(v;t) = p_\Gamma\left[v;\frac{\overline{v}^2(t)}{\sigma_v^2(t)},\frac{\overline{v}(t)}{\sigma_v^2(t)}\right]. \tag{33}$$

To fit this form, we would simply set the mean and variance according to the data.

These three parameterizations are formally equivalent. If the true Eulerian velocity distribution were gamma, the three fitting procedures would also be equivalent. However, if the latter are applied to arbitrary data, they may produce different results, as they focus on constraining different quantities given the two degrees of freedom (independent parameters) that characterize a gamma distribution. The first aims to provide the "overall best" fit for the PDF itself, while the second enforces the correct mean velocity and large transit time (low velocity) tailing, and in turn the third captures mean velocity and velocity variance exactly.

C Implementing the flux-weighted CDF

Flux-weighting of the gamma PDF in Eq. (31) corresponds to multiplication by $v/\bar{v}$, from which we can obtain the associated cumulative distribution function by integration:

$$\begin{aligned} P_\Gamma(v;\alpha,\xi) &= \frac{1}{\bar{v}\Gamma(\alpha)}\int_0^v (\xi v')^\alpha e^{-\xi v'} dv', \\ &= \frac{\gamma(\alpha+1,\xi v)}{\bar{v}\Gamma(\alpha)}, \end{aligned} \tag{34}$$

where $\gamma(\cdot,\cdot)$ is the lower incomplete gamma function. Recalling that $\xi = \alpha/\bar{v}$, we obtain

$$P_L(v;t) = P_\Gamma[v, \alpha, \alpha/\bar{v}(t)] = \frac{\gamma[\alpha+1, \alpha v/\bar{v}(t)]}{\bar{v}(t)\Gamma(\alpha)}. \tag{35}$$

This form of the flux-weighted CDF allows standard, well-known functions to be used to approximate the SMM numerically.

Acknowledgements

TA was supported by a Marie Skłodowska Curie Individual Fellowship, funded by the European Union's Horizon 2020 research and innovation programme under the project *ChemicalWalks* 838426. NE is supported by the U.S. National Science Foundation under awards EAR-2049687 and CBET-2129531.

References

Alim, K., Parsa, S., Weitz, D. A., and Brenner, M. P. (2017). Local pore size correlations determine flow distributions in porous media. *Phys. Rev. Lett.*, 119(14):144501.

Aquino, T. and Le Borgne, T. (2021). The diffusing-velocity random walk: a spatial-markov formulation of heterogeneous advection and diffusion. *J. Fluid Mech.*, 910:A12.

Berkowitz, B., Cortis, A., Dentz, M., and Scher, H. (2006). Modeling non-Fickian transport in geological formations as a continuous time random walk. *Rev. Geophys.*, 44(2).

Bianchi Janetti, E., Sherman, T., Guédon, G. R., Bolster, D., and Porta, G. M. (2020). Upscaling of solute plumes in periodic porous media through a trajectory-based spatial markov model. *Water Resources Research*, 56(12):e2020WR028408.

Bolster, D., Méheust, Y., Borgne, T. L., Bouquain, J., and Davy, P. (2014). Modeling preasymptotic transport in flows with significant inertial and trapping effects - the importance of velocity correlations and a spatial Markov model. *Adv. Water Resour.*, 70:89 – 103.

Carle, S. F. and Fogg, G. E. (1996). Transition probability-based indicator geostatistics. *Mathematical Geology*, 28(4):453–476.

Charbeneau, R. J. (2006). *Groundwater hydraulics and pollutant transport*. Waveland Press.

Comolli, A., Hakoun, V., and Dentz, M. (2019). Mechanisms, upscaling, and prediction of anomalous dispersion in heterogeneous porous media. *Water Resour. Res.*, 55(10):8197–8222.

De Anna, P., Le Borgne, T., Dentz, M., Tartakovsky, A. M., Bolster, D., and Davy, P. (2013). Flow intermittency, dispersion, and correlated continuous time random walks in porous media. *Phys. Rev. Lett.*, 110(18):184502.

Dentz, M., Comolli, A., Hakoun, V., and Hidalgo, J. J. (2020). Transport upscaling in highly heterogeneous aquifers and the prediction of tracer dispersion at the made site. *Geophys. Res. Lett.*, 47(22):e2020GL088292.

Dentz, M., Kang, P. K., Comolli, A., Le Borgne, T., and Lester, D. R. (2016). Continuous time random walks for the evolution of Lagrangian velocities. *Phys. Rev. Fluids*, 1(7):074004.

Engdahl, N. (2017). Transient effects on confined groundwater age distributions: Considering the necessity of time-dependent simulations. *Water Resources Research*, 53(8):7332–7348.

Engdahl, N. B. and Aquino, T. (2018). Considering the utility of backward-in-time simulations of multi-component reactive transport in porous media. *Advances in water resources*, 119:17–27.

Engdahl, N. B. and Bolster, D. (2020). Markovian transport processes in a heterogeneous, variably saturated watershed: A multi-domain spatial Markov model. *Advances in Water Resources*, 138(February):103555.

Engdahl, N. B., McCallum, J. L., and Massoudieh, A. (2016). Transient age distributions in subsurface hydrologic systems. *Journal of Hydrology*, 543:88–100.

Feller, W. (2008). *An introduction to probability theory and its applications*, volume 2. John Wiley & Sons, Woodbine.

Hakoun, V., Comolli, A., and Dentz, M. (2019). Upscaling and prediction of Lagrangian velocity dynamics in heterogeneous porous media. *Water Resour. Res.*, 55(5):3976–3996.

Hellinger, E. (1909). Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen. *Journal für die reine und angewandte Mathematik*, 1909(136):210–271.

Holzner, M., Morales, V. L., Willmann, M., and Dentz, M. (2015). Intermittent Lagrangian velocities and accelerations in three-dimensional porous medium flow. *Phys. Rev. E*, 92(1):013015.

Honeycutt, R. L. (1992). Stochastic runge-kutta algorithms. i. white noise. *Physical Review A*, 45(2):600.

Kang, P. K., De Anna, P., Nunes, J. P., Bijeljic, B., Blunt, M. J., and Juanes, R. (2014). Pore-scale intermittent velocity structure underpinning anomalous transport through 3-D porous media. *Geophys. Res. Lett.*, 41(17):6184–6190.

Kang, P. K., Dentz, M., Le Borgne, T., and Juanes, R. (2011). Spatial Markov model of anomalous transport through random lattice networks. *Phys. Rev. Lett.*, 107(18):180602.

Kim, J. S. and Kang, P. K. (2020). Anomalous transport through free-flow-porous media interface: Pore-scale simulation and predictive modeling. *Adv. Water Resour.*, 135:103467.

Klages, R., Radons, G., and Sokolov, I. M. (2008). *Anomalous transport: foundations and applications*. John Wiley & Sons.

Koponen, A., Kataja, M., and Timonen, J. v. (1996). Tortuous flow in porous media. *Phys. Rev. E*, 54(1):406.

Le Borgne, T., Dentz, M., and Carrera, J. (2008a). Lagrangian statistical model for transport in highly heterogeneous velocity fields. *Physical Review Letters*, 101(9):1–4.

Le Borgne, T., Dentz, M., and Carrera, J. (2008b). Spatial Markov processes for modeling Lagrangian particle dynamics in heterogeneous porous media. *Physical Review E*, 78(2):1–9.

Le Borgne, T., Dentz, M., and Carrera, J. (2008). Spatial Markov processes for modeling Lagrangian particle dynamics in heterogeneous porous media. *Phys. Rev. E*, 78(2):026308.

Lee, S. Y., Carle, S. F., and Fogg, G. E. (2007). Geologic heterogeneity and a comparison of two geostatistical models: Sequential Gaussian and transition probability-based geostatistical simulation. *Advances in Water Resources*, 30(9):1914–1932.

Massoudieh, A. (2013). Inference of long-term groundwater flow transience using environmental tracers: A theoretical approach. *Water Resources Research*, 49(12):8039–8052.

Massoudieh, A. and Dentz, M. (2020). Upscaling non-linear reactive transport in correlated velocity fields. *Advances in Water Resources*, 143:103680.

McCallum, J. L. and Shanafield, M. (2016). Residence times of stream-groundwater exchanges due to transient stream stage fluctuations. *Water Resources Research*, 52(3):2059–2073.

Meerschaert, M. M. and Sikorskii, A. (2012). *Stochastic models for fractional calculus*, volume 43. Walter de Gruyter, Göttingen.

Metzler, R. and Klafter, J. (2004). The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *J. Phys. A: Math. Gen.*, 37(31):R161.

Meyer, D. W. and Saggini, F. (2016). Testing the Markov hypothesis in fluid flows. *Phys. Rev. E*, 93(5):053103.

Meyer, D. W. and Tchelepi, H. A. (2010). Particle-based transport model with Markovian velocity processes for tracer dispersion in highly heterogeneous porous media. *Water Resour. Res.*, 46(11).

Pozrikidis, C. et al. (1998). *Numerical computation in science and engineering*, volume 6. Oxford university press New York.

Puyguiraud, A., Gouze, P., and Dentz, M. (2019a). Stochastic dynamics of Lagrangian pore-scale velocities in three-dimensional porous media. *Water Resour. Res.*, 55(2):1196–1217.

Puyguiraud, A., Gouze, P., and Dentz, M. (2019b). Upscaling of anomalous pore-scale dispersion. *Transport Porous Med.*, 128(2):837–855.

Puyguiraud, A., Gouze, P., and Dentz, M. (2021). Pore-scale mixing and the evolution of hydrodynamic dispersion in porous media. *Phys. Rev. Lett.*, 126(16):164501.

Scher, H. and Lax, M. (1973). Stochastic transport in a disordered solid. I. Theory. *Phys. Rev. B*, 7(10):4491.

Scher, H. and Montroll, E. W. (1975). Anomalous transit-time dispersion in amorphous solids. *Phys. Rev. B*, 12(6):2455.

Sherman, T., Engdahl, N. B., Porta, G., and Bolster, D. (2020). A review of spatial Markov models for predicting pre-asymptotic and anomalous transport in porous and fractured media. *Journal of Contaminant Hydrology*, 236:103734.

Sherman, T., Fakhari, A., Miller, S., Singha, K., and Bolster, D. (2017). Parameterizing the spatial Markov model from breakthrough curve data alone. *Water Resour. Res.*, 53(12):10888–10898.

Sherman, T., Paster, A., Porta, G., and Bolster, D. (2019). A spatial Markov model for upscaling transport of adsorbing-desorbing solutes. *J. Contam. Hydrol.*, 222:31–40.

Sund, N., Aquino, T., and Bolster, D. (2019). Effective models for transport in complex heterogeneous hydrologic systems. In Maurice, P., editor, *Encyclopedia of Water: Science, Technology, and Society*. John Wiley & Sons.

Sund, N., Bolster, D., and Dawson, C. (2015a). Upscaling transport of a reacting solute through a peridocially converging–diverging channel at pre-asymptotic times. *J. Contam. Hydrol.*, 182:1–15.

Sund, N., Bolster, D., Mattis, S., and Dawson, C. (2015b). Pre-asymptotic transport upscaling in inertial and unsteady flows through porous media. *Transport Porous Med.*, 109(2):411–432.

Sund, N. L., Porta, G. M., and Bolster, D. (2017). Upscaling of dilution and mixing using a trajectory based spatial Markov random walk model in a periodic flow domain. *Adv. Water Resour.*, 103:76–85.

Van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry*, volume 1. Elsevier.

Weissmann, G. S., Carle, S. F., and Fogg, G. E. (1999). Three-dimensional hydrofacies modeling based on soil survey analysis and transition probability geostatistics. *Water Resources Research*, 35(6):1761–1770.

Wright, E. E., Sund, N. L., Richter, D. H., Porta, G. M., and Bolster, D. (2019). Upscaling mixing in highly heterogeneous porous media via a spatial Markov model. *Water*, 11(1):53.

Highlights

Transient versions of the spatial Markov model are developed

Each flavor has tradeoffs in simplicity and robustness suited to different cases

Robust validity criteria to asses model applicability are derived

Author statement

Engdahl - Conceptualization, Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization

Aquino - Conceptualization, Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: