CAUSAL FACTORS OF NON-FICKIAN DISPERSION EXPLORED THROUGH MEASURES OF AQUIFER CONNECTIVITY

Katherine A. Klise[1] (kaklise@sandia.gov), Sean A. McKenna[1], Vincent C. Tidwell[1], Jonathan W. Lane[2], Gary S. Weissmann[3], Tim F. Wawrzyniec[3], Elizabeth M. Nichols[3]

[1] Sandia National Laboratories, Geoscience Research & Applications Group, PO Box 5800, MS 0701, Albuquerque, NM 87185
[2] Rice University, Department of Statistics, PO Box 1892, Houston, TX 77251
[3] University of New Mexico, Department of Earth and Planetary Sciences, MSCO3 2040, 1 University of New Mexico, Albuquerque, NM 87131

Abstract

While connectivity is an important aspect of heterogeneous media, methods to measure and simulate connectivity are limited. For this study, we use natural aquifer analogs developed through lidar imagery to track the importance of connectivity on dispersion characteristics. A 221.8 cm by 50 cm section of a braided sand and gravel deposit of the Ceja Formation in Bernalillo County, New Mexico is selected for the study. The use of two-point (SISIM) and multipoint (Snesim and Filtersim) stochastic simulation methods are then compared based on their ability to replicate dispersion characteristics using the aquifer analog. Detailed particle tracking simulations are used to explore the streamline-based connectivity that is preserved using each method. Connectivity analysis suggests a strong relationship between the length distribution of sand and gravel facies along streamlines and dispersion characteristics.

Introduction

Geologic heterogeneity, consisting of both connected and randomly distributed structures, has a large impact on solute transport behavior. Solute plumes are believed to disperse in a non-Fickian manner due to variable velocity and small-scale heterogeneity that create preferential pathways (Tsang and Neretnieks, 1998; Trefry et al., 2003; Cortis et al., 2004). Non-Fickian dispersion is characterized by (1) earlier and later arrival of solute at a reference plane (breakthrough curve) and (2) faster spreading rates when compared to dispersion that results in a Gaussian plume (Zhang et al., 2007). Stochastic simulation methods are often used to generate geologic heterogeneity for the purpose of predicting solute transport behavior. Understanding how connectivity of natural heterogeneous media influences dispersion is an important step towards building realistic stochastic models over a variety of scales.

Small scale heterogeneity is often simulated through use of sequential simulation algorithms using random field generators based on two-point covariance or transition probability functions. Sequential simulation, such as sequential Gaussian simulation (SGSIM) and sequential indicator simulation (SISIM) (Deutsch and Journel, 1998), requires knowledge of the histogram and variogram of the desired field. SISIM is commonly used for cell-based facies modeling (Yao, 2002; Zappa et al., 2006). While aquifer properties based on random field generators can produce preferential flow pathways (Trefry et al., 2003), it is unknown how well the transport results emulate dispersion through natural heterogeneous media. Questions regarding the ability of a two-point variogram to capture the complex structures seen in geologic media have been raised (Okabe and Blunt, 2004; Knudby and Carrera, 2005; Lee et al., 2007). Indicator variograms are thought to preserve connectivity, though this has not been proven (see review in Western et al. 1998). Two-point covariance functions present difficulty in replicating strongly connected, curvilinear structures particularly in multimodal distributions (Feyen and Caers, 2004). More accurate reproduction of geologic heterogeneity, and resulting solute transport, may be possible using stochastic methods that rely on the reproduction of multiple point patterns observed in training images (Caers et al., 2000). These methods preserve local heterogeneity structures that contribute to the overall connectivity of the media. Several multipoint methods exist, including Snesim (Strebelle, 2002) and Filtersim (Zhang et al., 2006). To explore the relationship between aquifer connectivity and dispersion characteristics, this

research compares how two-point and multipoint geostatistical simulation methods preserve solute transport characteristics and related measures of connectivity using a natural aquifer analog.

Lidar Scans and Segmentation

A braided stream deposit of the Tertiary Ceja Formation in Bernalillo County, New Mexico was selected to build a natural aquifer analog (Figure 1). For this study, we collected a 221.8 cm long by 50.0 cm high lidar scan at approximately 2-mm point spacing, taken from a range of 22 m using the University of New Mexico Lidar Lab's Optech Ilris 3d lidar scanner. Each laser shot generates a point with x, y, and z coordinates and the intensity of the reflected pulse off the surface of the target outcrop. Intensity is a function of the strength of reflected infrared energy, and is most strongly influenced by target surface properties, angle of incidence, and the scanner's optical properties (Kaasalainen et al., 2008). There is no direct relationship between intensity and hydraulic conductivity, yet intensity can be used to identify lithologic units in outcrop where zones of lithologic character are expressed as variation in visible color or surface roughness.

Point clouds from seven scans of the same region were projected onto a vertical plane in a process similar to stacking of geophysical signals. The mean and a standard deviation of intensity values were calculated for each pixel or grid location using the multiple scans (Figure 2). The reflectivity and angle of incidence of the gravel are more variable than those of the sand, thus providing high contrast between these sediment types in the standard deviation map. Therefore, the image of the calculated standard deviation, which shows this variation, was used to segment the outcrop into a sand and gravel facies map (Figure 3). Finally, reasonable hydraulic conductivity values of 0.001 and 0.1 cm/s were assigned to the sand and gravel facies, respectively. The resulting segmented lidar (SL) hydraulic conductivity field was used in our simulations to represent natural media. For more detail on the lidar scan and segmentation process, see Klise et al. (2009) and Nichols et al. (2008).

Figure 1. Digital photo of cross-bedded sand and gravel of the Ceja Formation.

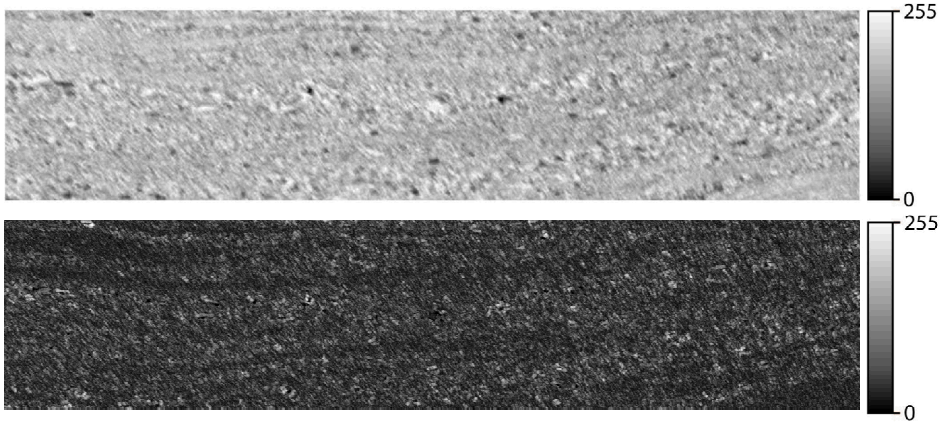

Figure 2. Mean (top) and standard deviation (bottom) of lidar intensity.

Figure 3. The segmented lidar (SL) field results in a sand and gravel facies map (black = sand, white = gravel).

Two-point and multipoint simulation

Correlated random fields, based on the two-point covariance function of the SL field, are generated using SISIM. The SISIM program is included in the publicly available Geostatistical Software LIBrary (GSLIB) (www.gslib.com, Deutsch and Journel, 1998). SISIM generates an indicator variable at each cell in the modeled region based on the spatial correlation of that variable. To create spatially correlated random fields using SISIM that characterize the SL field, the variogram function, including the principle dip direction, correlation length, and anisotropy, along with the percent of sand and gravel must be defined. A variogram map (Deutsch and Journel, 1998) is used to define the principle orientation direction that best captures the cross-bedded structure. While the cross-bedding is not uniformly oriented throughout the lidar scan, a predominant dip direction of 5 degrees counterclockwise from horizontal can be inferred based on the variogram map and through visual inspection of the SL field. Based on a principal and secondary direction of 5 and 95 degrees, experimental variograms are computed. The experimental variograms are best fit to a nested model using 2 exponential structures. The first structure has a range of 3.29 cm, a sill of 0.20 and anisotropy of 0.3. The second structure has a range of 50 cm, a sill of 0.052, and anisotropy of 0.1. Additional detail regarding unconditioned and conditioned SISIM simulation is outlined in Klise et al. (2009).

Two multi-point simulation methods are used to create simulated fields based on patterns extracted from the SL field. Both methods, Snesim and Filtersim, are publically available as part of the Standford Geostatistical Modeling Software (SGeMS) (http://sgems.sourceforge.net/, Remy et al., 2009).

The single normal equation simulation, or Snesim, is a multiple-point simulation algorithm developed by Strebelle (2002). The algorithm draws information from a given training image, or training images, to determine the conditional probability values to be used in the simulation by using a search template to draw this geological data. This data is then stored in a search tree. To simulate a node, the algorithm examines the existing data in the area around the node, defined by the search template size and shape, to determine the conditional probabilities associated with that node. Each node is visited sequentially in a random path and previously simulated nodes are used as conditional data for the subsequent nodes. Several parameters affect the sensitivity of the Snesim algorithm. Based on our own observations, and studies done by Liu (2006), we selected input parameters that visually best reproduce the SL field. The same multi-grid search used with SISIM is used for simulation. The target marginal distribution is set to match the proportions of the Lidar scan by having a 48.1% target probability of sand and a 51.9% target probability of gravel.

Simulation using filter scores, or Filtersim, is a multiple-point simulation algorithm developed by Zhang et al. (2006). Like Snesim, the algorithm also uses a search template to draw heterogeneity data from a training image. However, rather than using a search tree, the algorithm applies pre-defined filters to each searched object and then uses the calculated filter scores to place the objects into bins. Filters include an average, curvature, and gradient calculation in both the horizontal and vertical directions. During the simulation process, the algorithm follows a random path that covers the entire simulated image, calculating filter scores based on the existing data. Entire objects are then drawn from the appropriate bin and applied to the simulated image. The Filtersim algorithm assigns multiple nodes at the same time, as opposed to single node simulation used in both SISIM and Snesim. Filtersim also has several input parameters that affect its sensitivity; we choose parameters from Zhang et al. (2006) to

define the simulations. The same multi-grid search used with SISIM is used for simulation. Currently, there is no way to set the marginal distribution for facies type in Filtersim. We find that the resulting simulations regularly overestimate the percent gravel. A subset of the two-point and multipoint simulated fields are compared to the SL heterogeneity in Figure 4.

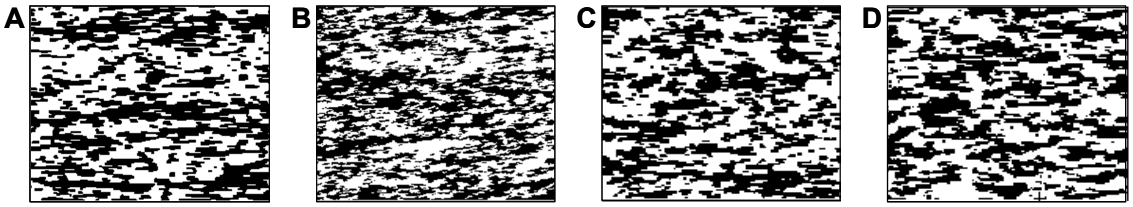

Figure 4. A 60 cm by 50 cm subset of A) the SL field, B) an example SISIM realization, C) an example Snesim realization, and D) an example Filtersim realization (black = sand, white = gravel).

Particle tracking simulations

Steady-state two-dimensional, solute transport simulations span the domain of the lidar scan, with cell size equal to the lidar point spacing (2 mm). This resolution defines over 1 million cells within the 221.8 by 50.0 cm region. The random walk particle tracking code RWHet (LaBolle et al., 1996, LaBolle, 2000) is used to simulate solute transport across the realistic and simulated hydraulic conductivity fields. The US Geological Survey model MODFLOW-2000 (McDonald and Harbaugh, 1988; Harbaugh et al., 2000) is used to calculate steady state head values for use in RWHet. Particle tracking simulations use advective and diffusive transport only. A dispersion parameter is not added to the model as to isolate the influence that heterogeneity alone has on dispersing particles.

A constant porosity of 30% is assumed for the simulation. Sand is assigned a hydraulic conductivity value of 0.001 cm/s while gravel is assigned a value of 0.1 cm/s. Inflow and outflow boundaries are defined using general head boundary conditions with a gradient of -0.0045 (dh/dl = 1 cm/221.8 cm) across the simulated domain. Boundaries parallel to flow are defined with no flow boundary conditions. The diffusivity coefficient is set to $1*10^{-7}$ cm^2/s to ensure that particles traveling through gravel remain advective dominated, while solute transport through the sand is influenced by diffusion. To capture detailed particle velocity, the transport time step is calculated based on a Courant number less than 1. This results in a time step of 120 seconds.

Particle breakthrough is monitored at the down-gradient boundary using 1,000 particles that are instantaneously released according to a flux weighted start location. Particles are released along a line 30 cm long, located 10 cm from the inflow boundary and no-flow boundaries. Hydraulic conductivity and velocity are then analyzed in a Lagrangian framework (along particle paths). Additional information on the setup for particle tracking simulations can be found in Klise et al. (2009).

Solute transport simulations in the SL field show strong solute focusing through gravel-dominated strata, resulting in a heavy-tailed, non-Fickian, breakthrough. The results of particle tracking simulations shows that transport through the SL field is characterized by large number of particles traveling through the media at early time followed by a long tail of subsequent particle moving through the media (Figure 5). Results using the simulated fields are averaged over 10 realizations, unless otherwise noted. While Filtersim simulations best approximate the solute transport in the SL field, none of the simulation methods tested here replicate both early and late time dispersion characteristics from the SL field.

To explore the differences in solute transport through the realistic and simulated media, we start by comparing bulk streamline velocity and continuous travel path lengths. The distribution of velocity along streamlines shows slight differences between the SL and simulated media (Figure 6). The simulated media results in slightly less fast velocity values around 10^{-3} cm/s and slightly increased number of intermediate velocity values near $10^{-3.85}$ cm/s. The distribution of continuous travel paths in sand and

gravel along streamlines was also compared (Figure 7). An increase number in small-scale features are inherent when using SISIM and add to the lack of connectivity between facies. Other than that difference, the SL and simulated media sample from very similar distribution of continuous sand and gravel paths.

While solute transport through the realistic and simulated fields sample from similar distributions of facies lengths, the bulk streamline-based velocity and resulting breakthrough curves are different. We contend that there must be a measurable difference in the way these continuous facies lengths are organized in space to cause the differences seen in the velocity and travel time (i.e. breakthrough curve) distributions. In the next section, we explore various measures of connectivity to investigate differences in solute transport characteristics.

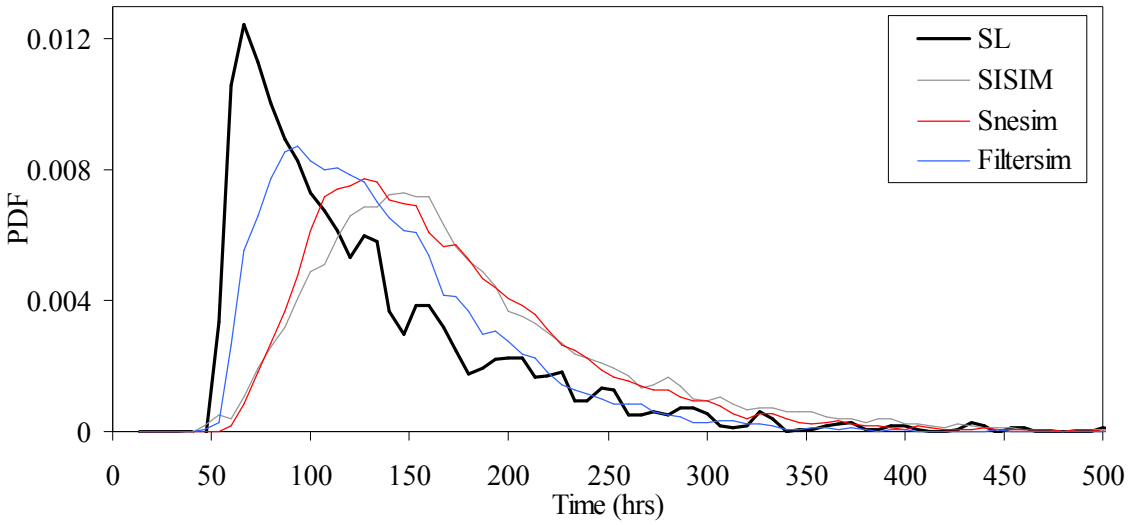

Figure 5. Breakthrough curves for solute transport through realistic and simulated media.

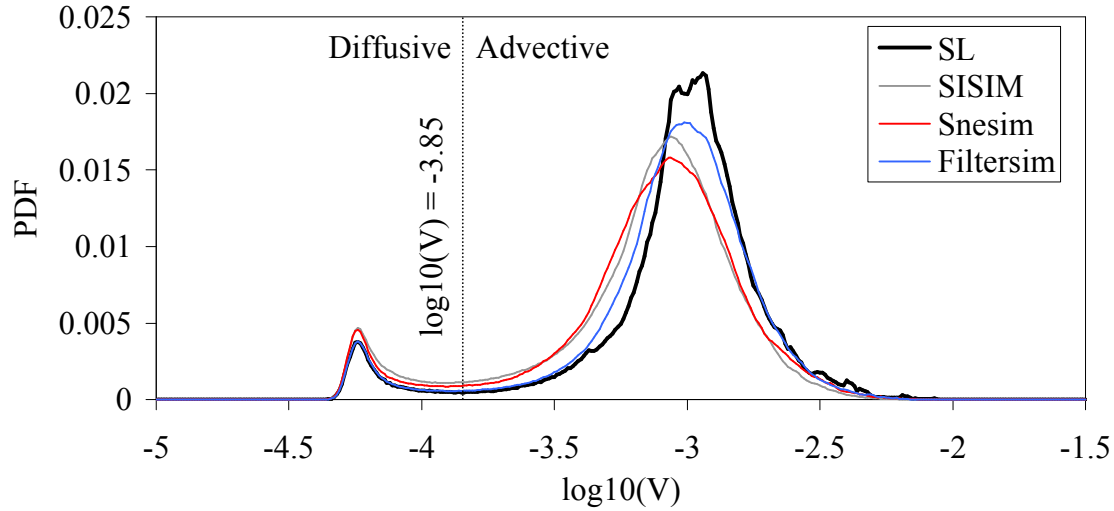

Figure 6. Streamline-based velocity distribution for all particles. A threshold value of log10(V) = -3.85 separates advective from diffusive velocities. This threshold is used in the connectivity analysis described below.

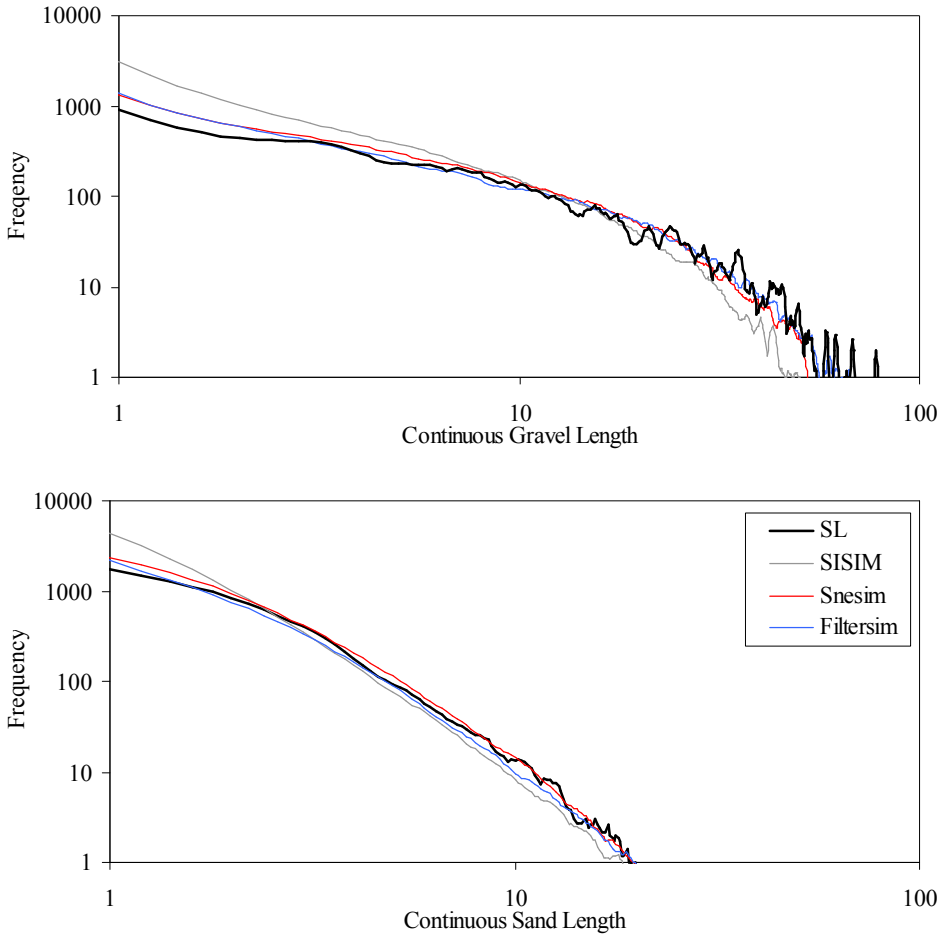

Figure 7. Continuous gravel and sand length distribution for all particles. Note that SISIM results in more small continuous gravel and sand paths, otherwise solute transport though the realistic and simulated media sample from similar distributions of continuous gravel and sand paths.

Connectivity analysis

Connectivity describes interconnected paths through a spatial pattern. This interconnected path can describe high conductivity paths that cause early solute arrival or low conductivity paths that slow solute and contribute to late time arrival. While it is clear that connectivity of geologic media heavily influences dispersion characteristics, there is no set definition on how to measure connectivity.

Knudby et al. (2006) show the importance of using connectivity measures to upscale hydraulic conductivity. They define connectivity using the average inclusion length and the mean distance between inclusions along streamlines. We use a similar measure here, defining connectivity by the percent gravel along streamlines. We also calculate velocity connectivity, defined as the percent advective velocity along streamlines. Additionally, we calculated the number of "runs" (e.g., Davis, 2002) along each streamline to quantify the tendency of facies type and velocity values to cluster together along particle paths. A "run" is defined as a consecutive sequence of equal value parameters. A low number of runs indicate that streamlines tend to remain in the same facies or velocity category for greater distances, while a high number of runs indicate that observations along the streamline rapidly switch between the two states. We track these two measures of connectivity to see how they relate to dispersion characteristics and to explore relationships between velocity and facies type.

For these measures of connectivity, the observations must be segmented into distinct categories, or states. Facies (or hydraulic conductivity) are already separated into two categories (sand and gravel), but streamline-based velocity requires a threshold value to define high and low values. We choose a threshold of $10^{-3.85}$ cm/s, based on the break in velocity values between advective dominated and diffusive dominated flow (Figure 6).

Probability distributions for each measure of connectivity are plotted in Figure 8. Connectivity measures using velocity closely approximate connectivity measures using facies type (compare Figure 8A to 8B, compare Figure 8C to 8D). This similarity reflects the correlation between facies type (gravel vs sand) and velocity value (advective vs. diffusive). While velocity profiles across individual streamlines are complex, the general relationship between facies type and velocity makes the facies profile a good proxy for the complex velocity distribution. Additionally, we note similarity between the connectivity distributions and the breakthrough curve distributions. In all cases, the SL field produces highly skewed distributions, followed by Filtersim, then Snesim, and finally SISIM. By comparing travel time with facies-based measures of connectivity, we see a strong correlation with the percent distance traveled in gravel and travel time. However, the number of runs is less correlated with travel time (Figure 9). This means that the number of times a particle alternates between sand and gravel is not as important to the overall travel time as the total distance traveled in gravel. Figure 9B shows how the percent travel distance in gravel along streamlines for each simulation method is a good indicator of the travel time distribution. Failure to replicate the highly skewed distribution of travel distance in gravel leads to poor prediction of dispersion characteristics using the 2-point and multipoint methods.

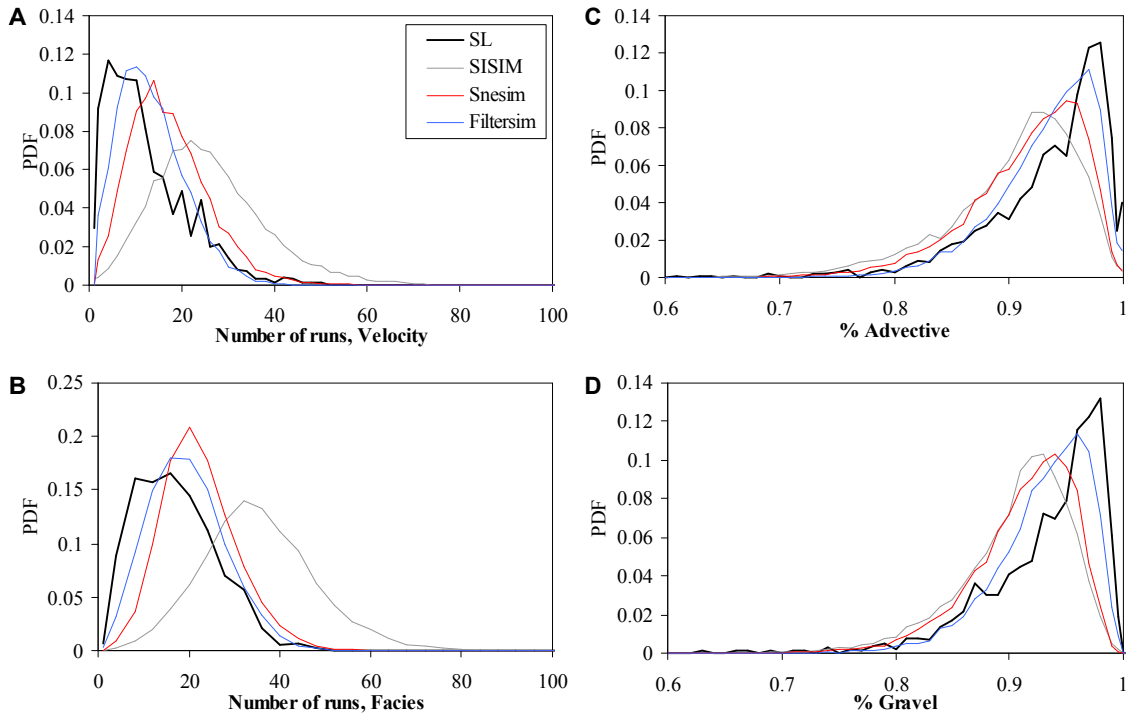

Figure 8. Velocity and facies connectivity in the realistic and simulated media. A) Number of velocity runs, B) Number of facies runs, C) percent advective, D) percent gravel.

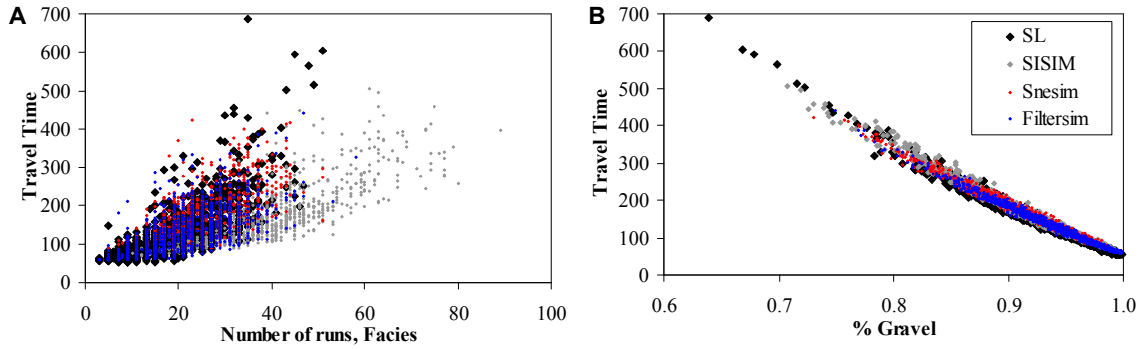

Figure 9. Connectivity measure compared to travel time. A) Number of facies runs, B) percent gravel. SISIM, Snesim, and Filtersim results are based on one realization.

Discussion

We compare two-point and multipoint geostatistical simulation methods to determine how each method preserves dispersion characteristics and streamline-based measures of connectivity. This study uses a natural aquifer analog as a training image and as "ground truth" for comparing simulation methods. Results indicate that simulated fields do not capture salient features of the complex geologic media necessary to replicate dispersion characteristics.

Using both a velocity-based and facies-based measure of connectivity, we find that the simulation methods tested here do not preserve connectivity measured in the SL field. In general, multipoint methods (Snesim and Filtersim) better replicate connectivity as compared to 2-point methods (SISIM). Velocity-based measures of connectivity are highly correlated with facies-based measure of connectivity. We conclude that preserving the percent travel distance in gravel is a key component to preserving dispersion characteristics in the simulated media.

We found that while the simulated and lidar-based heterogeneity sample from the same distribution of continuous sand and gravel travel path lengths along particle paths, the dispersion characteristics and velocity profiles differ. This discrepancy is a function of differences in the way continuous travel paths are organized along individual streamlines. Understanding the relationship between the streamline-based connectivity and the travel time distribution could enhance the ability to estimate dispersion characteristics from the underlying heterogeneity.

Breakthrough and connectivity results are dependent on the difference in hydraulic conductivity, the dispersion coefficient, and the gradient chosen for simulation. Further numerical analysis is needed to better understand how these parameters influence the relationship between connectivity and dispersion. We are currently investigating these questions to better understand the causal factors of non-Fickian dispersion.

Acknowledgements

This work was funded by a Department of Energy Basic Energy Science grant, a Department of Energy Office of Science (BER) grant, and by the Sandia-University Research Program. The authors greatly acknowledge Louis Scuderi, for work on lidar segmentation, and Eric LaBolle and Josh Christian for contributions to solute transport simulations, and three anonymous reviewers. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

Caers, J. K., S. Srinivasan, and A. G. Journel (2000), Geostatistical quantification of geological information for a fluvial-type North Sea reservoir, SPE Reservoir Eval. Eng., 3(5), 457–467.

Cortis, A., C. Gallo, H. Scher, and B. Berkowitz (2004), Numerical simulation of non-Fickian transport in geological formations with multiple-scale heterogeneities, Water Resour. Res., 40, W04209.

Deutsch, C. V., and A. G. Journel (1998), GSLIB Geostatistical Software Library and User's Guide, 2nd ed., Oxford Univ. Press, New York.

Feyen, L., and J. Caers (2004), Multiple-point geostatistics: A powerful tool to improve groundwater flow and transport predictions in multi-modal formations, in geoENV V: Geostatistics for Environmental Applications, edited by P. Renard, H. Demougeot-Renard, and R. Froidevaux, pp. 197– 208, Springer, Berlin.

Harbaugh,A., E. Banta, M.Hill, andM.McDonald (2000), MODFLOW-2000, the U. S. Geological Survey modular ground-water model—User guide to modularization concepts and the ground-water flowprocess, U.S. Geol. Surv. Open File Rep., 00-92.

Kaasalainen, S., A. Kukko, T. Lindroos, P. Litkey, H. Kaartinen, J. Hyyppa, and E. Ahokas (2008), Brightness measurements and calibration with airborne and terrestrial laser scanners, IEEE Trans. Geosci. Remote Sens., 46, 528– 534.

Klise, K. A., G. S. Weissmann, S. A. McKenna, E. M. Nichols, J. D. Frechette, T. F. Wawrzyniec, and V. C. Tidwell (2009), Exploring solute transport and streamline connectivity using lidar-based outcrop images and geostatistical representations of heterogeneity, Water Resour. Res., 45, W05413.

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

Knudby, C., J. Carrera, J. D. Bumgardner, and G. E. Fogg (2006), Binary upscaling—The role of connectivity and a new formula, Adv. Water Resour., 29(4), 590– 604.

LaBolle, E. M. (2000), RWHet: Random walk particle model for simulating transport in heterogeneous permeable media, version 2.0, user's manual and program documentation, 27 pp., Univ. of Calif., Davis.

LaBolle, E. M., G. E. Fogg, and A. F. B. Thompson (1996), Random-walk simulation of transport in heterogeneous porous media: Local massconservation problem and implementation methods, Water Resour. Res., 32(3), 583–593.

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

Liu, Y. (2006) Using the Snesim program for multiple-point statistical simulation, Computers & Geosciences 32, 1544–1563.

McDonald, M. G., and A. W. Harbaugh (1988), A modular threedimensional finite-difference groundwater flow model, U.S. Geol. Surv. Tech. Water Resour. Invest., Book 6, Chap. A1, 586 pp.

Nichols, E. M., G. S. Weissmann, T. F. Wawrzyniec, L. A. Scuderi, J. D. Frechette, and K. A. Klise (2008), Lidar imagery based 2D and 3D representations of outcrop heterogeneity for groundwater modeling, Eos Trans. AGU, 89(53), Fall Meet. Suppl., Abstract H41A-0833.

Okabe, H., and M. J. Blunt (2004), Prediction of permeability for porous media reconstructed using multiple-point statistics, Phys. Rev. E, 70, 066135.

Remy, N., A. Boucher, and J. Wu (2009), Applied geostatistics with SGeMS: A users guide, 1st edition, Cambridge University Press, 288 pages.

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

Trefry, M. G., F. P. Ruan, and D. McLaughlin (2003), Numerical simulations of preasymptotic transport in heterogeneous porous media: Departures from the Gaussian limit, Water Resour. Res., 39(3), 1063.

Tsang, C. F., and I. Neretnieks (1998), Flow channeling in heterogeneous fractured rocks, Rev. Geophys., 36(2), 275 – 298.

Western, A. W., G. Bloschl, and R. B. Grayson (1998), How well do indicator variograms capture the spatial connectivity of soil moisture?, Hydrol. Processes, 12, 1851 – 1868.

Yao, T. (2002), Integrating seismic data for lithofacies modeling: A comparison of sequential indicator simulation algorithms, Math. Geol., 34(4), 387– 403.

Zappa, G., R. Bersezio, F. Felletti, and M. Giudici (2006), Modeling heterogeneity of gravel-sand, braided stream, alluvial aquifers at the facies scale, J. Hydrol., 325(1 – 4), 134 – 153.

Zhang, T., P. Switzer, and A. Journel (2006), Filter-based classification of training image patterns for spatial simulation, Math. Geol., 38(1), 63– 80.

Zhang, Y., D. A. Benson, and B. Baeumer (2007), Predicting the tails of breakthrough curves in regional-scale alluvial systems, Ground Water, 45(4), 473– 484.