University of Alberta

Geostatistical Modeling of Unstructured Grids for Flow Simulation

by

Johnathan Gregory Manchuk

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mining Engineering

Department of Civil and Environmental Engineering

Fall 2010
Edmonton, Alberta

Examining Committee

Clayton V. Deutsch, Civil and Environmental Engineering

Jozef Szymanski, Civil and Environmental Engineering

Jeff Boisvert, Civil and Environmental Engineering

Carl Mendoza, Earth and Atmospheric Science

Guillaume Caumon, Nancy School of Geology, Nancy Université

To my wife, Erin, for your love, support, and patience

... and to Montgomery

Abstract

A challenge in petroleum geostatistics is the application of modeling algorithms such as Gaussian simulation to unstructured grids that are being used for flow simulation. Geostatistical modeling is typically applied on a fine scale regular grid and then upscaled to the unstructured grid. This work proposes a fine scale unstructured grid. The grid is designed so that its elements align with the flow simulation grid elements, eliminating the occurrence of intersections between the two grids. Triangular and tetrahedral grids are used for the fine scale grid; however, they introduce a variety of element scales. The approach developed in this work populates the fine scale grid based on the scale of conditioning data. The resulting error due to scale discrepancy is quantified and mitigated though the upscaling process. A methodology to assess the error in upscaled properties is developed and used to control the design of the fine scale grid. Populating the fine scale grid with reservoir properties requires modification of existing geostatistical algorithms. The set of spatial locations for modeling is irregular and three differences that result from this are addressed: random path generation; spatial search; and the covariance lookup table. Results are compiled into an algorithm for sequential indicator and sequential Gaussian simulation on irregular point sets. Checking variogram reproduction on large irregular point sets is a challenge. An algorithm that efficiently computes the experimental variogram for these cases is developed. A flow based upscaling method based on the multipoint flux approximation is developed to upscale permeability models from the fine scale unstructured grid to the flow simulation grid. Triangular grids are assumed for the fine scale. Flow simulation results using the upscaled transmissibilities are very similar to results obtained using traditional flow simulation on high resolution regular grids.

Acknowledgments

I would like to thank my supervisor, Dr. Clayton Deutsch, for his invaluable support throughout the development of this research. I would like to express my gratitude to Dr. Jeff Boisvert for taking the time to review and provide input for all aspects of this thesis. I would also like to thank Dr. Martin Mlacnik for conveying his knowledge of flow simulation and providing insight for developing the flow based upscaling technique in this thesis.

I would like to thank the member companies of the Centre for Computational Geostatistics for their financial support.

Finally, I would like to thank my parents for their support and guidance.

Table of Contents

Chapter 1 Introduction

Chapter 2 Methodology

Chapter 3 Sequential Simulation

Chapter 4 Upscaling

Chapter 5 Demonstration and Case Study

Chapter 6 Concluding Remarks

List of Tables

List of Figures

List of Symbols, Nomenclature, or Abbreviations

A	area
AI	acoustic impedance
ALR	additive log ratio
FMI	Formation micro image
ft	feet
FEM	finite element method
FVM	finite volume method
GR	gamma ray
HCV	hydrocarbon volume
m	meters
mD	millidarcy
MFEM	mixed finite element method
MPFA	multipoint flux approximation
MSE	mean squared error
NPHI	neutron porosity
OIP	oil in place
PEBI	perpendicular bisector
PLC	piecewise linear complex
PVI	pore volume injected
RD	deep resistivity
REV	representative elementary volume
RHOB	bulk density
RS	shallow resistivity
SAGD	steam assisted gravity drainage
SGS	sequential Gaussian simulation
SIS	sequential indicator simulation
SSE	sum squared error
TPFA	two-point flux approximation
V	volume

Vsh	volume of shale
∇	gradient operator
$\cap$	intersection operator
$\in$	is an element of
α	dip angle
$d\alpha$	dip angle tolerance
β	linearization exponent
$\gamma(\mathbf{h})$	variogram function
$\bar{\gamma}(\mathbf{h})$	average variogram function
$\bar{\gamma}(v,V)$	average variogram between elements v and V
$\delta_{z,k}$	indicator transform function
Δabc	triangle between points a, b, and c
ε^2	squared error function
θ	azimuth angle
$d\theta$	azimuth angle tolerance
$\lambda_i, \boldsymbol{\lambda}$	kriging weight assigned to point i; vector of kriging weights
μ	fluid viscosity
ρ	fluid density
$\rho(\mathbf{u})$	density function
σ^2	variance
σ_{sk}^2	simple kriging variance
σ_y^2	variance of a variable y
Σ	variance matrix
τ	linearization constant
φ	proportion of an indicator variable
ϕ	porosity
Ω	domain or area of interest
a	variogram range

$\overline{ab}$	edge between points a and b
A_0, A_1, A_2	linear regression coefficients
b	center point of an interaction region
B	power average exponent
b_h	horizontal bandwidth
b_v	vertical bandwidth
$\mathbf{c}, \mathbf{C}$	vector of covariances; covariance matrix
$\bar{\mathbf{c}}, \bar{\mathbf{C}}$	vector of average covariances; average covariance matrix
C_i	sill of nested variogram structure i
C_0	nugget effect
$C(\mathbf{h})$	covariance function
$C(\mathbf{u}_i, \mathbf{u}_j)$	covariance between two positions
$\bar{C}(v, V)$	average covariance between elements v and V
$Cov\{a, b\}$	covariance between two variables a and b
C_T	covariance lookup table
d	dimensionality
$D^2(v, V)$	dispersion variance of scale v within V
E	Voronoi grid
$E\{a\}$	expected value of variable a
f_σ	variance reduction
F_w	fractional flow of water
$G(Z)$	Gaussian transform of the variable Z
$\mathbf{h}$	distance or lag vector
$d\mathbf{h}$	lag separation tolerance
$\mathbf{I}$	index vector of a regular grid element
$d\mathbf{I}$	relative index vector between two regular grid elements
$\mathbf{k}$	permeability tensor
k_A	arithmetic averaged permeability
k_B	power averaged permeability
k_G	geometric averaged permeability
k_H	harmonic averaged permeability

$\mathbf{L}$	lower triangular matrix
$\boldsymbol{L}$	matrix of barycentric coordinates
m	midpoint of an edge
$m(\mathbf{u})$	mean at position $\mathbf{u}$
n, N	number of elements or items
$\mathbf{n}$	normal vector to a surface
n_x, n_y, n_z	number of elements in a regular grid in Cartesian coordinates
nc	number of categories
O	interaction region cluster
$p, \mathbf{p}$	pressure; vector of pressures
P	probability
p	vertices of a triangulation
p_c	centroid of a grid element
$q, \mathbf{q}$	flux; vector of fluxes
r	random number
$\mathbf{r}$	random vector
R_o	additive log ratio transformed oil saturation
R_w	additive log ratio transformed water saturation
S	surface
S_f	fluid saturation
S_o	oil saturation
S_w	water saturation
T	triangulation
T_{ij}	two-point transmissibility between elements i and j
t_{ij}	multi-point transmissibility
u, v, t	Cartesian depositional coordinates
$\mathbf{u}$	spatial position vector
$\mathbf{u}_0$	origin position of a regular grid
$d\mathbf{u}_0$	extents of a regular grid element
$\mathbf{v}$	normal vector
v, V	grid elements, $V > v$

$\int_V \; d\mathbf{u}$	integral over the element V
∂V	boundary of the element V
$\int_{\partial V} d\mathbf{u}$	integral over the boundary of element V
$\overleftrightarrow{V}$	extent of an element in a specified direction
$\lvert V \rvert$	volume of a grid element V
$Var\{a\}$	variance of variable a
w	weight used in numerical integration
x, y, z	Cartesian axes
Y	a Gaussian distributed random variable
$y(\mathbf{u})$	random variable of spatial position $\mathbf{u}$
$\hat{y}(\mathbf{u})$	estimate of the function $y(\mathbf{u})$
$y_s(\mathbf{u})$	random simulated value of the function $y(\mathbf{u})$
z_{cb}	z-coordinate of a correlation base surface.
z_{ct}	z-coordinate of a correlation top surface.
z_s	stratigraphic depth coordinate
Z_v	random variable with scale v

Chapter 1 Introduction

This chapter begins with a discussion on the problem of modeling reservoirs with unstructured grids using geostatistics. Section 1.1 describes the components involved and why they are important, including the need and benefit of generating numerical reservoir models, the motivation for using geostatistics to generate those models, and the motivation for using unstructured grids. The purpose of this thesis and a list of contributions are covered in Section 1.2. Background in Section 1.3 describes the geometry of unstructured grids and scaling laws of reservoir properties. Sections 1.4 and 1.5 give a review of the relevant literature and summarize the outline of this dissertation.

1.1 Problem Setting

1.1.1 Reservoir Modeling and Geostatistics

Several important commodities originate from reservoirs including hydrocarbons – detailed numerical models help assess our ability to exploit these resources. Reservoir models are constructed based on known information and they are used in a variety of ways. One use is predicting original volumes of hydrocarbons or other fluids in the reservoir, which is important for assessing the economic viability of producing them and for estimating initial capital requirements in the form of land, materials, and facilities. Another use of reservoir models is in selecting methods of production. Several techniques are available for hydrocarbon production and the viability of each can be assessed using flow simulation to determine which is optimal for a given reservoir. Models are also used to choose economically optimal well locations, configurations of multiple wells, and number of wells required (Deutsch, 2002). Generating reservoir models that consider all available information can be accomplished through geostatistics.

Understanding the geological heterogeneity of a reservoir is crucial for making decisions since the heterogeneity directly affects the flow response. Geostatistics treats properties such as porosity and permeability as spatially correlated random variables. Reservoirs are treated as probabilistic and our lack of understanding of the true heterogeneity is reflected through multiple equally probable realizations of what could be the truth. Each realization reveals a different flow response. Together, the set of realizations and flow responses quantifies uncertainty, that is, a measure of our incomplete knowledge.

Geostatistics is a powerful aid to classical deterministic approaches involving geophysics and geology. Methods such as inverse distance interpolation and contour mapping provide a single model. The advantage of geostatistics is that uncertainty is quantified and it can be used as an additional measure for decision making. It is possible to assign risk to decisions and optimize objectives such as profitability, recovery and sustainability. Geostatistics requires a large amount of domain specific knowledge and can be demanding for practitioners, requiring them to become familiar with many complex techniques and available data. This level of involvement leads to an improved understanding of the reservoir and better models when implemented correctly. Proper use of geostatistics requires all relevant data to be integrated into the final models; therefore, resulting models reproduce many sources of data such as core, log, and seismic.

1.1.2 Unstructured Grids

The focus of this dissertation is on populating unstructured grids with reservoir properties using geostatistical techniques. Reservoirs are discretized by a grid so that the flow equations can be solved and an analysis of a reservoirs response to different flow conditions is possible. The response depends on many factors including reservoir structure, geologic heterogeneity, well locations, and flow conditions at the wells. Our ability to predict the response also depends on the grid design. A good design accounts for the factors affecting flow and increases our confidence in response predictions.

Designing a grid to account for reservoir structure is a demanding task as the structures may form complex configurations of stratigraphy and faults. Tying everything together with a grid that also meets requirements of available flow simulators requires a flexible data structure. Unstructured grids are a natural response to this problem. They can be constrained to flow simulator requirements, reservoir structures, well locations, and controlled locally with reservoir heterogeneity. Unstructured grids also offer the ability to control grid resolution that can be made high where it is needed such as near flowing wells and low in areas where little flow is expected. With good grid design, increased resolution results in increased accuracy in important high flow areas. Flow simulation is computationally demanding and if available resources limit the size of a grid that can be used, having flexibility in resolution becomes an advantage.

Four examples of unstructured grids are provided in Figure 1-1. Specific details regarding the geometry of these grids are found in the background section of this chapter. Two grid types are shown: perpendicular bisector (PEBI) grids in A, C and D and a tetrahedral grid in B. The PEBI grid in A is shaded by mean flow velocity, which is computed using a high resolution permeability model for flow based grid generation (Prevost et al, 2005). Several goals were targeted in this grid design including: reproducing geologic features, which in this case is a single fault; achieving higher grid resolution in regions where higher flow velocities are anticipated; and maintaining orthogonality between elements. Orthogonality is an important feature for commercial flow simulators that use the two-point flux approximation. The primary goal for the grid in B is to represent a complex network of faults (Prevost et al, 2005). Since tetrahedral elements are used, a different flow simulation method is necessary, such as finite elements or the multipoint flux approximation. In C, a single fault, multiple horizons, and a deviated well are incorporated into the grid design (Verma, 1998). Such a grid is useful in analysing the effectiveness of a newly proposed well. Notice in this grid that only the fault and the well path are reproduced accurately. The horizons shaded with different colors are only approximately captured as the element interfaces do not align with the color transitions. Slightly different from the grids in A and C, the grid in D is a 2.5D PEBI grid: the PEBI portion of the grid is designed in two dimensions in plan view to reproduce fault lines, then the result is projected along depth to match horizons. This grid accurately reflects faults and horizons, but it introduces partial element connections along displaced faults, that is, elements do not perfectly align in depth as shown by Figure 1-2. Partial connections do not occur in the other grid examples.

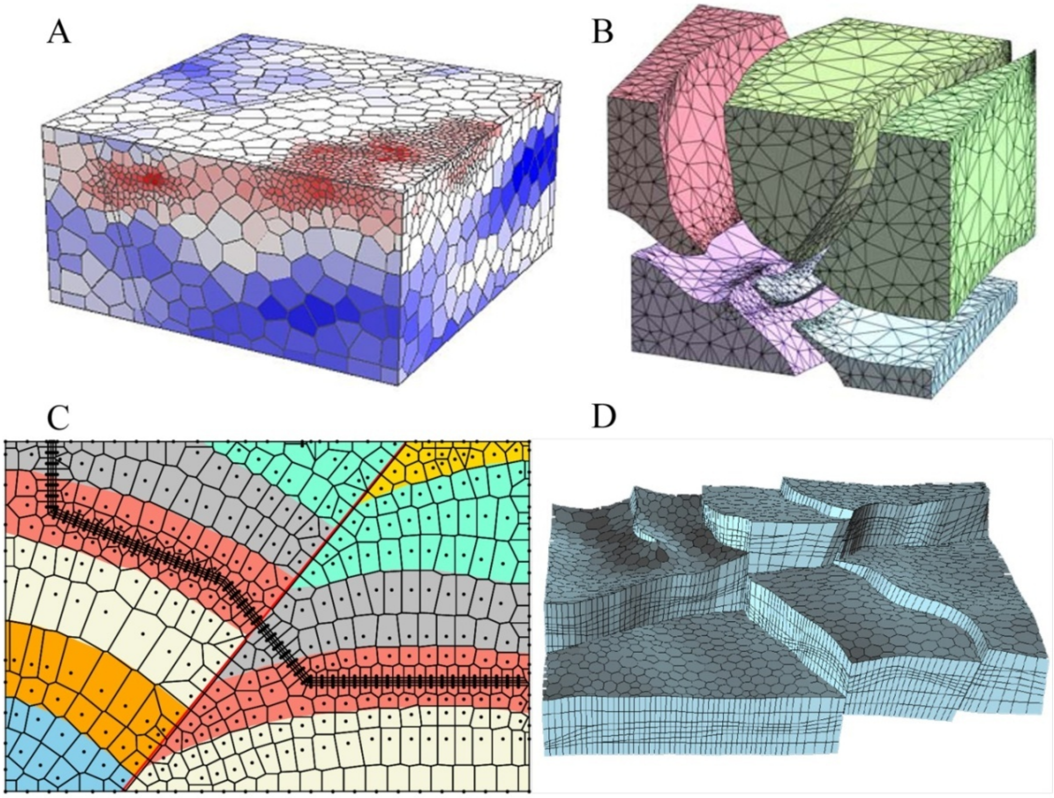


Figure 1-1: Example grids: A – PEBI grid with fault, blue to red is low to high velocity (Prevost et al, 2005); B – Tetrahedral grid with multiple intersecting faults (Prevost et al, 2005); C – PEBI grid with multiple stratigraphic layers, a fault, and a deviated well (Verma, 1998); D – 2.5D faulted PEBI grid (Levy et al, 2001). Not to scale.

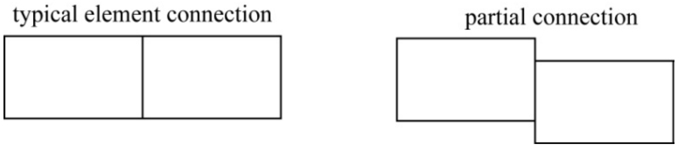


Figure 1-2: Element connections.

Consider designing grids for examples A to D using structured grids. A structured grid is shown in Figure 1-3 for comparison. Achieving high resolution in grid A where mean flow velocity is high using a structured grid is only possible if a high resolution is assigned to all parts of the grid. This increases the number of grid elements dramatically and increases CPU time for flow simulation. Moreover, to represent the fault either the orthogonality property would be lost or the accuracy of the fault would be reduced. The same issue is encountered for B in attempting to represent the faults with a structured grid, and in C trying to represent the deviated well path, horizons, and fault. Grid D is similar to existing structured grid design with the only difference being that in plan a PEBI grid is used. The design process is also similar in that grid structure is designed in plan view and projected in depth.

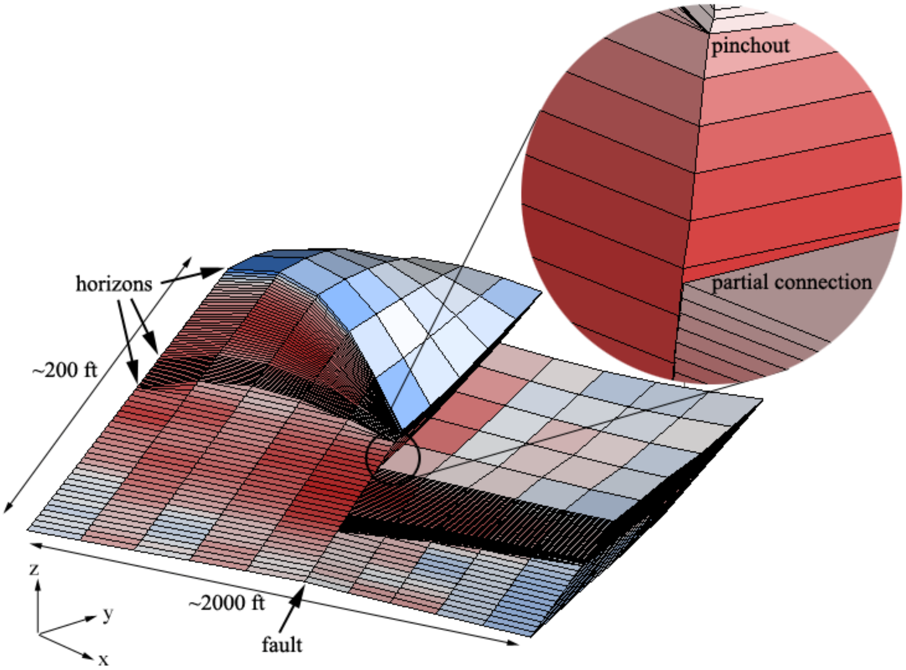


Figure 1-3: Example structured grid. Not to scale.

Geostatistical algorithms are designed for regular grids and cannot be directly applied to unstructured grids. Available geostatistical approaches to characterize unstructured grids involve an underlying structured grid. Take the examples in Figure 1-1, geological models are built using an underlying regular grid and equivalent properties are computed for the unstructured grid elements. The process of obtaining equivalent properties is often called upscaling since the volumes of regular grid elements are usually smaller than the volumes of unstructured grid elements. The regular grid forms a discretization of the unstructured grid. Properties for the unstructured grid elements are calculated without having to account for the geometry of each element directly. This dissertation claims that a regular grid is not the most appropriate discretization of an unstructured grid for upscaling purposes.

1.1.3 Discretization

Flow simulation grids are often limited by computer resources and time and therefore involve an order of magnitude fewer elements than geostatistical modeling grids. The majority of unstructured grid elements are larger than those used for geostatistical modeling especially in areas of low flow. Accounting for the scale and shape of unstructured elements as well as the geological heterogeneity within them is accomplished by discretization followed by upscaling. The grid used for discretization is constrained by the upscaling processes involved. In cases where flow equations are solved to upscale permeability, the discretization grid must agree with the method used to solve the equations. Regular grids are used due to limitations of geostatistical algorithms. Regular grids are a good choice when the flow simulation grid is also regular because a fine regular grid forms an ideal partition of a coarse regular grid when the grids overlap, see Figure 1-4. Such a partition is ideal because:

1. The coarse and fine grids are both amenable to existing flow simulation algorithms permitting the use of flow based upscaling, which involves flow simulation on the fine scale grid and flow simulation of the coarse model. Although this seems somewhat trivial, the ability to flow simulate a grid must be extended to all discretization regimes and this restricts the set of feasible discretizations.
2. Every coarse element is represented by the same number of fine elements resulting in a consistent level of accuracy used to describe the heterogeneity within coarse elements. For upscaling, no coarse element is over or under-represented by fine elements; therefore any upscaling error that exists is constant throughout the whole grid.
3. Fine grid elements do not overlap one another and are fully embedded in coarse elements. This is in accordance with dispersion variance concepts in geostatistics, which is based on a disjoint partition of equal volume elements. In some cases, not all fine scale elements represent the same volume even with regular grids and the equal volume requirement is an approximation. Grid alignment also simplifies computing average values for upscaling variables that scale arithmetically including facies, porosity, fluid saturations, and pressure.
4. The interfaces of fine elements are aligned with coarse element interfaces. Although closely related to point 3, the implications of this are important for upscaling permeability and transmissibility using flow based techniques. Flow based upscaling relies on solving the flow equations on the fine grid to obtain equivalent flow parameters on the coarse grid. In finite volume formulations solving the flow equations results in fluxes defined across element interfaces. The fluxes across coarse element interfaces are immediately computed as the sum of fluxes across all coincident fine element interfaces as shown in Figure 1-5.

One goal of this dissertation is to implement a discretization that achieves the features listed in points 1 to 4 for unstructured grids. The discretization will not be a regular Cartesian grid. A more flexible discretization structure is needed given the range of geometric configurations that are possible with unstructured grids. One discretization style that meets conditions of all four points is n-simplicial grids, which are triangulations for $n = 2$ and tetrahedralizations for $n = 3$. These are referred to as simplex grids to avoid any ambiguity with dimension. Figure 1-6 shows a pair of two dimensional PEBI grid elements discretized by both a regular grid and a simplex grid.

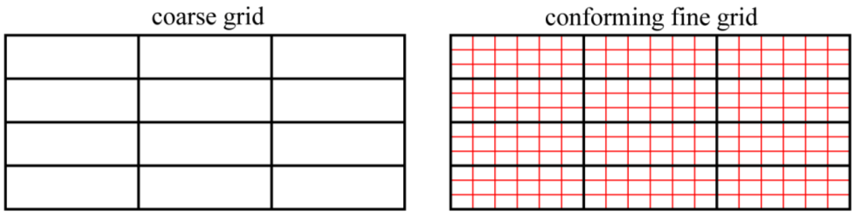


Figure 1-4: Coarse regular grid and conforming discretization.

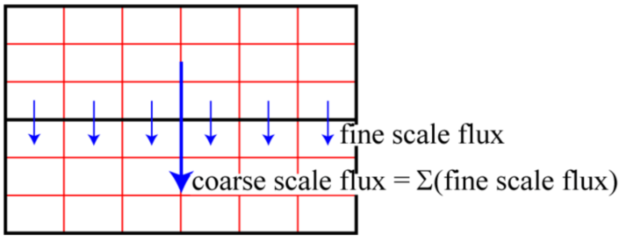


Figure 1-5: Fine and coarse scale flux in flow based transmissibility upscaling.

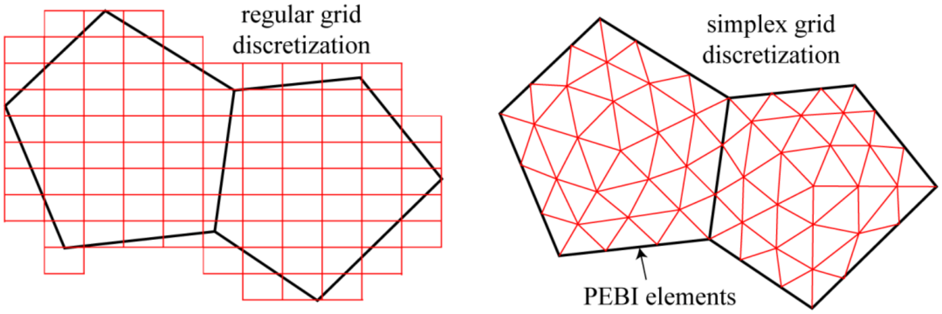


Figure 1-6: PEBI elements discretized by a regular grid and a simplex grid.

The geometric complexity of unstructured grids and simplex discretization make achieving an equal number of fine elements per coarse element challenging. A more appropriate measure than number is accuracy: how effectively the discretization represents the upscaling process. This is similar to measuring the accuracy of numerical integration techniques (Kythe and Schaferkotter, 2004). Simplex grids also form a disjoint partition; however, there is some variation in the scale of each simplex that has implications in regards to geostatistical scaling laws. Because the simplex discretization is aligned with the coarse grid, computing arithmetic averages and resolving interface flux in flow based upscaling is straightforward; however, computing the interface flux requires solving the flow equations on a simplex grid.

1.2 Statement of Purpose and Contributions

The problem addressed in this dissertation is the population of unstructured grids by geostatistical techniques for reservoir flow simulation. The function through which uncertainty is measured is flow response. Therefore geostatistical models are generated with the knowledge they will be used in upscaling processes and flow simulation.

The most important variable involved is permeability. It is a second order tensor property that describes the ease that reservoir rock transmits fluids. The permeability tensor, **k**, links fluid flux, **q**, to reservoir pressure, *p*, through Darcy's law (Equation 1.1), where μ is the fluid viscosity and ∇ is the gradient operator. Tensors can take on several forms as shown in Equation 1.2, all of them are symmetric and only the upper triangular components are shown. Developments made in this thesis assume that a tensor is isotropic or transversely isotropic for populating fine scale simplex grids. This is a common assumption for geostatistical modeling of permeability. As an alternative to permeability, transmissibility is also used in flow simulation and it relates the flux through a surface to the pressure drop across the same surface.

$$\mathbf{q} = -\frac{1}{\mu}\mathbf{k}\nabla p \qquad 1.1$$

$$\begin{array}{ccc} \text{isotropic} & \text{transverse} & \text{symmetric} \\ \begin{bmatrix} k & 0 & 0 \\ & k & 0 \\ & & k \end{bmatrix} & \begin{bmatrix} k_h & 0 & 0 \\ & k_h & 0 \\ & & k_v \end{bmatrix} & \begin{bmatrix} k_{xx} & k_{xy} & k_{xz} \\ & k_{yy} & k_{yz} \\ & & k_{zz} \end{bmatrix} \end{array} \qquad 1.2$$

The purpose of this thesis is to develop mathematical and computational methods for populating unstructured grids with reservoir properties to be used in flow simulation studies. Underlying this are discretization, geostatistical modeling, and upscaling.

The main contributions of this dissertation are a workflow and associated tools for modeling reservoirs with geostatistics on unstructured grids. It is different than existing work in that an unstructured grid is used on the fine scale as opposed to a regular grid. Grid design is not the main focus of this dissertation; however, guidelines and methods are provided. The data scale becomes a concern since all grid elements may have different geometry and volume. General guidelines on upscaling from core and log scale to another consistent scale for property modeling are provided as well.

One step in the workflow is discretization. It involves generating a fine scale grid that conforms to a coarse scale grid, both being unstructured. For variables that average arithmetically, a method to choose how many fine scale elements to use in each coarse scale element is developed. The approach results in an analytical function relating element volume to number of discretization elements that can be used with existing grid generation tools.

Reservoir property modeling follows discretization. A contribution in this area is the development of a sequential simulation tool that does both sequential indicator simulation (SIS) for facies and sequential Gaussian simulation (SGS) for continuous properties on unstructured grids. It is shown that the optimal point to use for each grid element is their centroid. This minimizes the error between the estimation variance at the scale of the conditioning data being modeled by the sequential simulation algorithm and the scale of the unstructured grid element it represents. It is further demonstrated that when the grid being modeled is upscaled to a coarser grid, the error is reduced substantially. For variables that average arithmetically, it is shown that the correct estimate, estimation variance, and covariance are recovered at the coarse grid scale. Another contribution is a tool to compute the experimental variogram from potentially very large non-uniform point sets. It was adapted from (Deutsch and Journel, 1998) and is intended to validate the variogram for models generated on unstructured grids.

A new method for upscaling permeability from the fine scale grid to transmissibility on the coarse grid is developed. It is based on the multipoint flux approximation and is developed for two dimensions. Extending the theory to three dimensions is straightforward; however, implementation may be more complex. The approach is validated for homogeneous, layered, and heterogeneous media; however, poor grid design or highly anisotropic permeability may still yield non-physical flow results.

Additional contributions include a description and some usage information for existing tools for unstructured grid discretization and visualization. Two programs used throughout this dissertation are: *Triangle* by Shewchuk (1996) for generating two dimensional triangular grids and *TetGen* by Si (2006) for generating three dimensional tetrahedral grids. For visualization, an open-source application called Paraview is used (Squillacote, 2008).

1.3 Background

1.3.1 Grid Geometry

A variety of different grid specifications exist and they have common geometric characteristics. Grids consist of a set of elements that consist of a set of faces, edges, and vertices as depicted in Figure 1-7. To meet the constraints of reservoir simulators and flow based upscaling techniques that use them, elements are also linear and convex, see Figure 1-8. Linear implies the edges and faces are lines and planes rather than curves and curvilinear surfaces. Non convex and curvilinear elements are not discussed. Simulation techniques involving the two-point or multipoint flux approximation require additional geometric information including centroids, centroid connections, and face midpoints. These details are reserved for Chapter 4 that discusses flow based upscaling.

For geostatistical modeling purposes, the geometric information of interest is the vertices for simulators requiring vertex parameterization or centroids for simulators requiring element distributed parameterization. Mixed finite elements may be specified either way, whereas two-point and multipoint simulation require properties at the centroid or some other interior point representing the whole element. Other geometric components are used after vertices or centroids have been populated with reservoir properties. For example, assessing global statistics requires element volumes and permeability upscaling requires volumetric and connectivity information.

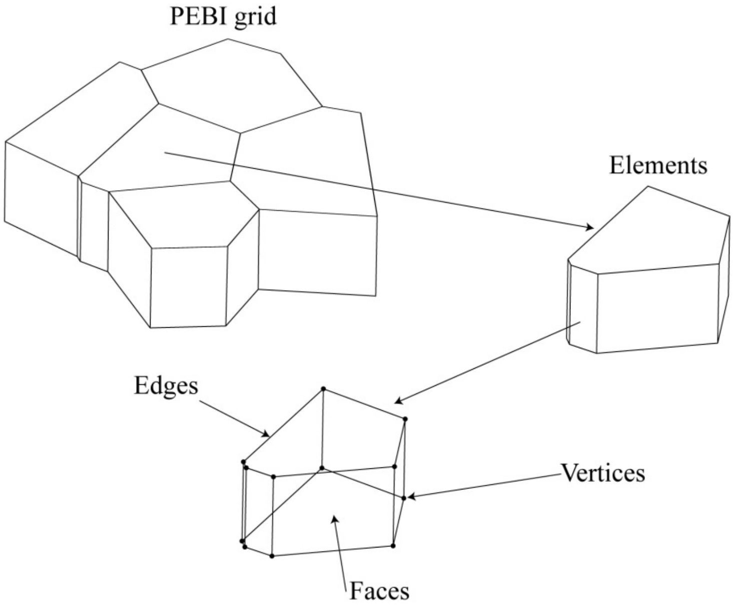


Figure 1-7: Geometric primitives of a grid.

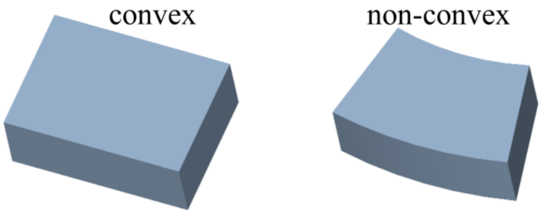


Figure 1-8: convex and non-linear non-convex elements.

Structured grids are used primarily for geostatistical property modeling. Just as in Figure 1-7, they can be decomposed into similar geometric primitives. The grid is referred to as structured because the elements can be traversed in a logical manner. In other words the arrangement of neighbouring elements is constant for all elements, see Figure 1-9. However, there is no constraint on the shape and volume of each element except that it must be a quadrilateral in two dimensions or a hexahedron in three dimensions. Grids may also be regular grids that maintain the same logical structure with all elements having equal volume in the form of rectangles or boxes. A convenient feature of structured grids for geostatistics is that there is always a mapping to a regular grid, see Figure 1-10. The space where the regular grid exists is often called stratigraphic space, depositional space, or chronologic space (Deutsch, 2002; Mallet, 2004; Gringarten et al, 2008) and it is where geostatistical property modeling is carried out. For structured grids, all elements often have the same volume in stratigraphic space; however, in cases with varying sedimentation rates combined with differential compaction, the thickness can vary (Mallet, 2004). Also, distance and covariance computations are not influenced by faulting and folding that has taken place after the time of deposition. For this reason, similar transformations are still required for unstructured grids.

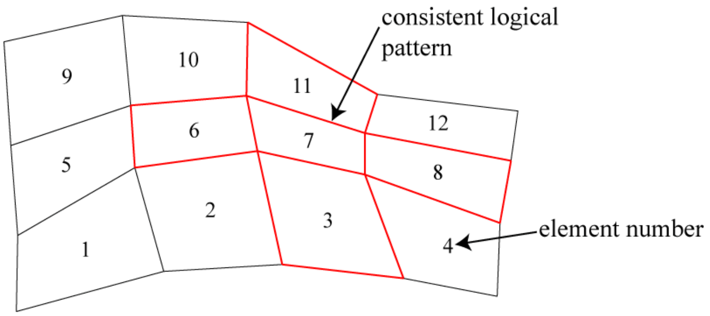


Figure 1-9: Pattern of neighbours for a two dimensional structured grid.

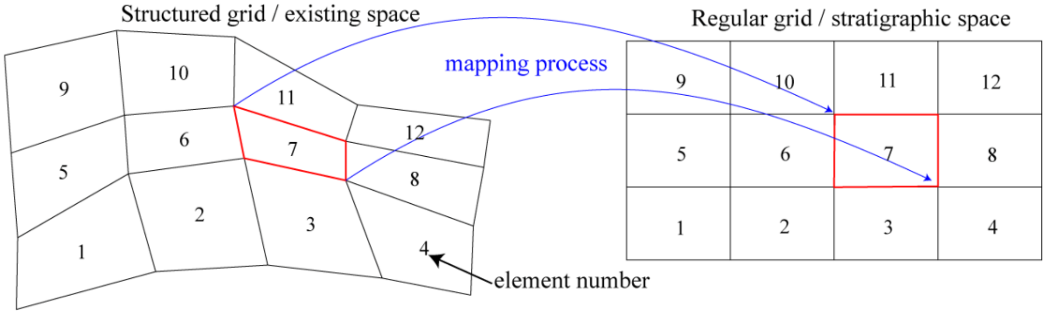


Figure 1-10: Mapping from a structured grid to a regular grid.

There are two common unstructured grid types found in the literature on reservoir modeling and flow simulation: PEBI grids and simplex grids. Both specifications may be used as the coarse scale flow simulation grid or as the refined grid to be populated and upscaled. In cases where a simplex grid is Delaunay (de Berg et al, 2000; Edelsbrunner, 2001), both types of grids are the other's dual as shown by Figure 1-11. The Delaunay condition states that if every edge in a triangulation T is locally Delaunay, then the grid is a Delaunay triangulation, where local Delaunay indicates that the circumscribing circle of any triangular element's vertices contains no other vertex as shown in Figure 1-11 (Edelsbrunner, 2001). The same definitions apply in three dimensions. For these grids, the edges of T form perpendicular bisectors of the edges of its dual Voronoi grid E. This is ideal for the two-point flux approximation; however, it may not be ideal for geostatistical modeling since the vertices of T, denoted p, are not necessarily the centroids of the elements of E, denoted p_c. Estimates of reservoir properties generated at p will not be as representative of those generated at p_c.

When the centroid is required it is computed for simplexes as the average of the element vertices. For Voronoi elements numerical integration is required to evaluate Equation 1.3 for an element v, where $\mathbf{u}$ defines a position in the element and $\rho(\mathbf{u})$ defines a mass or density function.

$$\mathrm{p}_c = \frac{\int_v \mathbf{u}\rho(\mathbf{u})d\mathbf{u}}{\int_v \rho(\mathbf{u})d\mathbf{u}} \qquad 1.3$$

The density function can be assumed constant, but for grid generation purposes it is useful and has been included in the equation. For example, it could reflect the magnitude of permeability or another property and be used to control centroidal Voronoi grid generation (Du et al, 1999). When $\rho(\mathbf{u})$ is constant, the centroid can be computed using geometric decomposition; the element is broken down into a set of simplexes and used in Equation 1.4, where A_k is the area or volume of the k^{th} simplex and p_{ck} is its centroid.

$$\mathrm{p}_c = \frac{\sum_{k=1}^{n} \mathrm{p}_{c,k} A_k}{\sum_{k=1}^{n} A_k} \tag{1.4}$$

Figure 1-12 provides a simple two dimensional example. The volume of Voronoi elements is computed using decomposition as well, where the volume of a tetrahedron is given by Equation 1.5, and p_1, p_2, p_3 and p_4 are its vertices. In two dimensions, decomposition involves triangles and the area is given by Equation 1.6.

$$A = \frac{1}{6}\left|(\mathrm{p}_1 - \mathrm{p}_4)\cdot\left((\mathrm{p}_2 - \mathrm{p}_4)\times(\mathrm{p}_3 - \mathrm{p}_4)\right)\right| \tag{1.5}$$

$$A = \frac{1}{2}\left|(\mathrm{p}_3 - \mathrm{p}_1)_x(\mathrm{p}_2 - \mathrm{p}_1)_y - (\mathrm{p}_2 - \mathrm{p}_1)_x(\mathrm{p}_3 - \mathrm{p}_1)_y\right| \tag{1.6}$$

For the multipoint flux approximation (Aavatsmark et al, 1998), the dual grid used to define interaction regions is not necessarily a Voronoi grid. The concept of interaction regions is discussed in Chapter 4. In Delaunay triangulations, the edges connecting two Voronoi elements do not always physically intersect the face connecting those elements as shown in Figure 1-13. In this work, this is referred to as invalid connections. In multipoint formulations, the interaction region between a set of triangles with a common vertex must be contained within those triangles and this is also shown in Figure 1-13.

Interaction regions for a polygonal grid are not always convex, but their volumes and centroids are computed using decomposition. Performing simulation on the polygonal grid involves triangular interaction regions and reservoir properties are assigned to the polygons; whereas performing simulation on the triangular grid will involve polygonal interaction regions and reservoir properties are assigned to the triangular elements. Both scenarios are shown in Figure 1-14 using the same Delaunay triangulation as in Figure 1-12. As with Delaunay triangulations and Voronoi grids, the vertices of the triangular interaction regions are not necessarily the centroids of the control volumes.

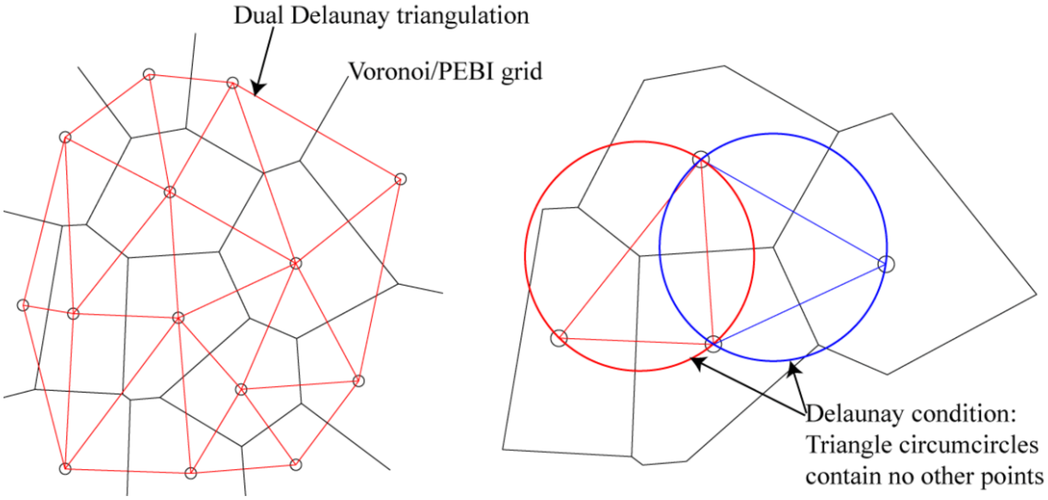


Figure 1-11: Delaunay triangulation and Voronoi grid (left), and local Delaunay condition (right). On the right, no triangle vertices fall inside the circumscribed circles.

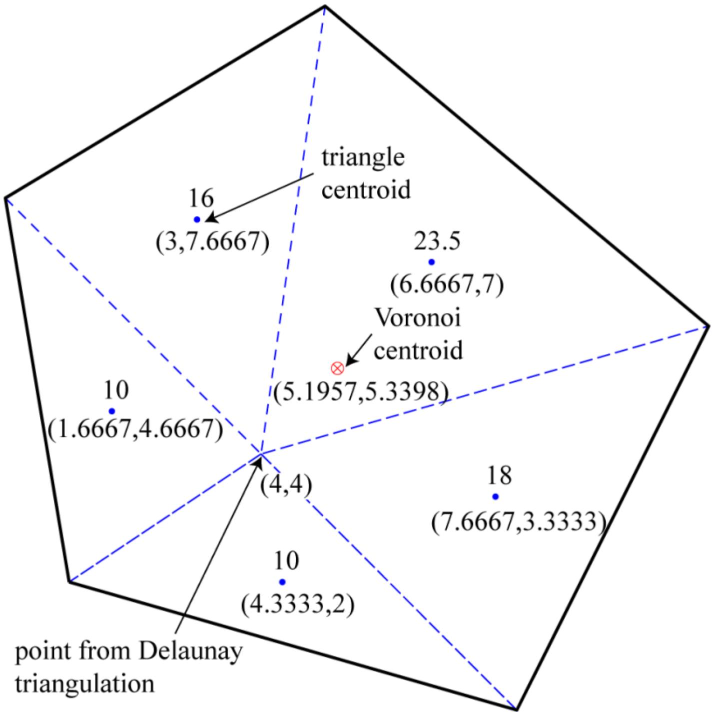


Figure 1-12: Voronoi element centroid by decomposition. Numbers above the triangle centroids are the areas and those in parenthesis are respective *x* and *y* coordinates.

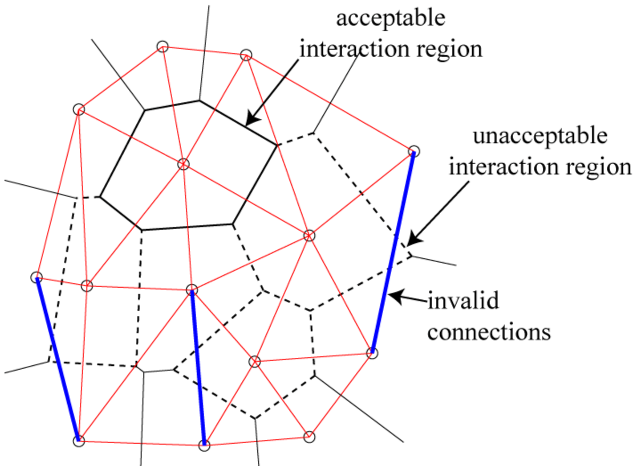


Figure 1-13: Imperfect connections, acceptable and unacceptable interaction regions.

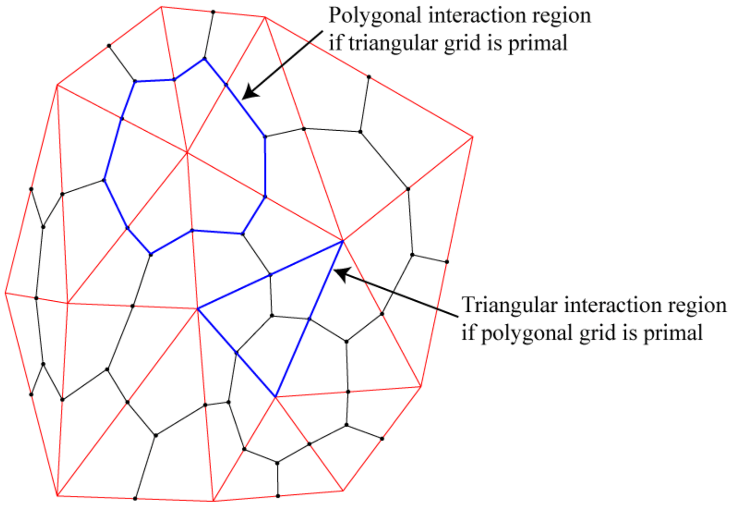


Figure 1-14: Delaunay triangulation and polygonal grid showing interaction regions.

There are many other geometric features to take into consideration for grid generation purposes, for measuring grid quality, or determining how effectively a grid discretizes a differential operator. For the purposes of geologic property modeling, the features of importance include knowing and understanding the type of grid being used, obtaining a set of vertices or centroids to populate, and computing areas or volumes for assessing statistics and upscaling.

1.3.2 Scaling Laws

Unstructured grids introduce a significant amount of geometry into the framework of geostatistical resource characterization. A major difference between unstructured grids and regular grids is variable element scale and shape. This poses a difficulty for geostatistical prediction that is based on statistics at a specific scale. A related problem for geostatistics has been incorporating sample data that are representative of different scales, for example the scale of porosity derived from a neutron log is different than that measured for a core sample, and both of these are different than the scale of grid elements chosen for geological modeling or flow simulation.

Core, log, and seismic data are representative of different scales. Unlike core samples, where the scale is exactly defined by a core plug, log and seismic data do not represent a clear geometric object. The ability to handle these ill-defined geometries with geostatistical scaling laws has been discussed by several authors (Frykman and Deutsch, 1999 and 2002; Frykman, Vejbæk and Rasmussen, 2005). Another scale to consider is that of the geologic model. The scale is the volume of grid elements that is constant for regular grids and generally assumed constant for structured grids after a stratigraphic transformation. Grids designed for flow simulation introduce another scale when regular grids are used, or a variety of scales with unstructured grids. Each grid element having a different volume and geometric specification represents a different scale. Figure 1-15 provides a summary of typical scales encountered in reservoir modeling.

Differences in scale are often accounted for by averaging the available data to the geologic modeling scale. This approach assumes well log and core data adequately represent the geologic grid elements; however, in most cases grid elements are intersected by one vertical well and averages cannot characterize the horizontal direction that is often important for permeability. Unfortunately, there are not many alternatives: geostatistical scaling laws are only applicable to variables that average arithmetically. Change of support and a discussion on the scaling laws of facies and permeability is included below. Porosity is treated as an arithmetically averaging variable and is not discussed further in this section.

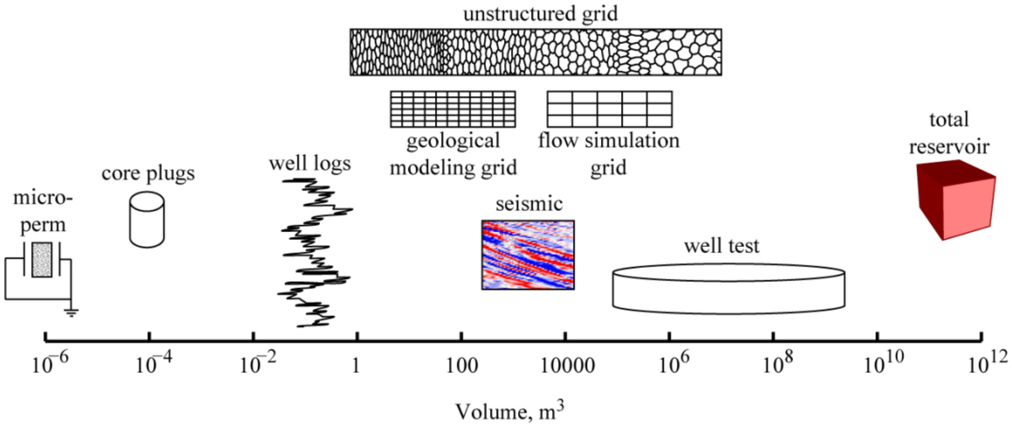


Figure 1-15: Scales for different kinds of reservoir data and grids.

1.3.2.1 Change of Scale

Statistical inference of geological properties depends on the scale or volume that is being inferred. In geostatistics, volume-variance relationships have been defined to explain how the statistics of random variables change with volume, including the distribution, variance, and variogram. The theory pertains to the concept of regularized variables (Journel and Huijbregts, 1978); the arithmetic average of a variable Z_v with volume v to obtain its equivalent Z_V at the larger volume V (Equation 1.7). Determining a set of laws that explain how the statistical and spatial properties for a variable change with volume are the solution to the change of scale problem. The results theoretically permit the incorporation of data with different volumes.

$$Z_V = \frac{1}{|V|}\int_V Z_v dV \tag{1.7}$$

The effect of volume-variance relations or scaling laws on the variogram is discussed. Generally, as the volume that represents a variable increases, the variogram range increases while its variance or sill and nugget effect decrease (Journel and Huijbregts, 1978; Kupfersberger et al, 1998). These effects are expressed mathematically based on the dispersion variance, $D^2(v, \Omega)$, which is the variance of data with volume v within a domain Ω. In the context of this work, the domain could be a hydrocarbon reservoir or an aquifer and v could be a core sample. Letting $v < V < \Omega$, the dispersion variances of v within V and of V within Ω are additive for arithmetically averaging variables via Equation 1.8.

$$D^2(v,\Omega) = D^2(v,V) + D^2(V,\Omega) \tag{1.8}$$

The dispersion variance can be related to the variogram, $\gamma(\mathbf{h})$, by first defining the averaged variogram, $\bar{\gamma}(\mathbf{h})$, where $\mathbf{h}$ is a separation vector between two locations $\mathbf{u}$ and $\mathbf{u}'$, that is, $\mathbf{h} = \mathbf{u} - \mathbf{u}'$. For the volume v, the averaged variogram denoted $\bar{\gamma}(v, v)$ is defined by Equation 1.9, and its relation to dispersion variance by 1.10. Here, $\mathbf{u}$ is a spatial position vector in some coordinate system and $|v|$ denotes the volume of v.

$$\bar{\gamma}(v,v) = \frac{1}{|v|\cdot|v|}\int_v\int_v \gamma(\mathbf{u}-\mathbf{u}')d\mathbf{u}d\mathbf{u}' \tag{1.9}$$

$$D^2(v,\Omega) = \bar{\gamma}(\Omega,\Omega) - \bar{\gamma}(v,v) \tag{1.10}$$

These relations were initially derived for regular grids and therefore only two or three volumes would have to be dealt with depending on the types of sample data available and their sources. Consider V to be the volume of regular grid elements, then the averaged variogram can be derived by the scaling laws. Two approaches are used: 1 – core samples with scale v are averaged or upscaled to V and an experimental variogram calculated and modeled from the result; 2 – the experimental variogram of core samples is calculated and modeled and the result scaled to agree with V. The first is more common in the context of reservoir characterization. In the second approach, the variogram nugget effect, C_0, range, a, and sill, $C_i, i = 1, \ldots, n_i$,with n the number of nested structures, are scaled according to Equations 1.11, 1.12, and 1.13 (Kupfersberger et al, 1998). In Equation 1.12, $\overleftrightarrow{V}$ and $\overleftrightarrow{v}$ describe the extent of the volumes and are direction dependent. These results are based on variables that average arithmetically.

$$C_0^V = C_0^v \cdot \frac{|v|}{|V|} \tag{1.11}$$

$$a^V = a^v + \left(\vec{V} - \vec{v}\right) \qquad 1.12$$

$$C_i^V = C_i^v \frac{1-\bar{\gamma}(V,V)}{1-\bar{\gamma}(v,v)} \qquad 1.13$$

Variogram scaling laws and the additive property of dispersion variance are appropriate for variables that average arithmetically such as porosity. Different change of scale models are required for other variables, such as permeability, that scales according to physical laws rather than statistical averages.

1.3.2.2 Facies

Facies and other categorical variables average arithmetically when they are treated as mutually exclusive indicators. Rather than averaging to a single category, facies average to a distribution of categories for a particular volume (Deutsch and Lan, 2008). In this context, Z_v is a categorical variable that defines nc facies types. The indicator mapping is defined by Equation 1.14, where δ is the Kronecker delta.

$$\begin{aligned} \mathbf{I}_v &= \begin{bmatrix} \delta_{Z_v,1} & \delta_{Z_v,2} & \cdots & \delta_{Z_v,k} & \cdots & \delta_{Z_v,nc} \end{bmatrix} \\ \delta_{Z_v,k} &= \begin{cases} 1, & Z_v = \text{ category } k \\ 0, & \text{otherwise} \end{cases} \end{aligned} \qquad 1.14$$

The result is a vector since the number of possible facies types ranges from 1 to nc. Regionalizing this variable to V as in Equation 1.7 results in a vector of proportions given by Equation 1.15, which also identifies the constraint that all proportions sum to unity.

$$\begin{aligned} \boldsymbol{\varphi}_V &= \left[\frac{1}{|V|} \int_V \delta_{Z_v,k} dV \right], \; k = 1,\ldots,nc \\ 1 &= \sum_{k=1}^{nc} \varphi_{V,k} \end{aligned} \qquad 1.15$$

Upscaled facies form a compositional variable (Pawlowsky-Glahn and Olea, 2004). If Ω was partitioned into a set of elements, as with regular grids, then upscaling results in a distribution of proportions for each category that can be expressed as a probability distribution function.

Results of the above equations are shown experimentally using a hypothetical two dimensional facies model (Figure 1-16) that is upscaled to various regular grids with different scales (Figure 1-17). At the smallest scale each point is defined by a single category and the proportion is either zero or one and the cumulative density function is a step function. As the scale is increased, the density function of each facies proportion converges to the global proportion.

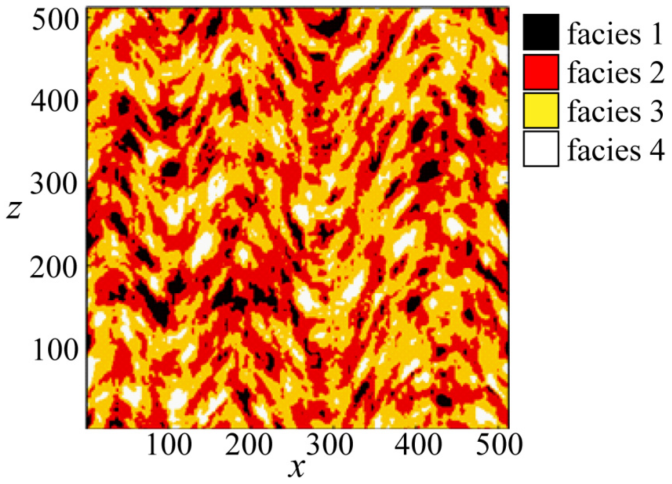


Figure 1-16: Synthetic facies model of four arbitrary facies types.

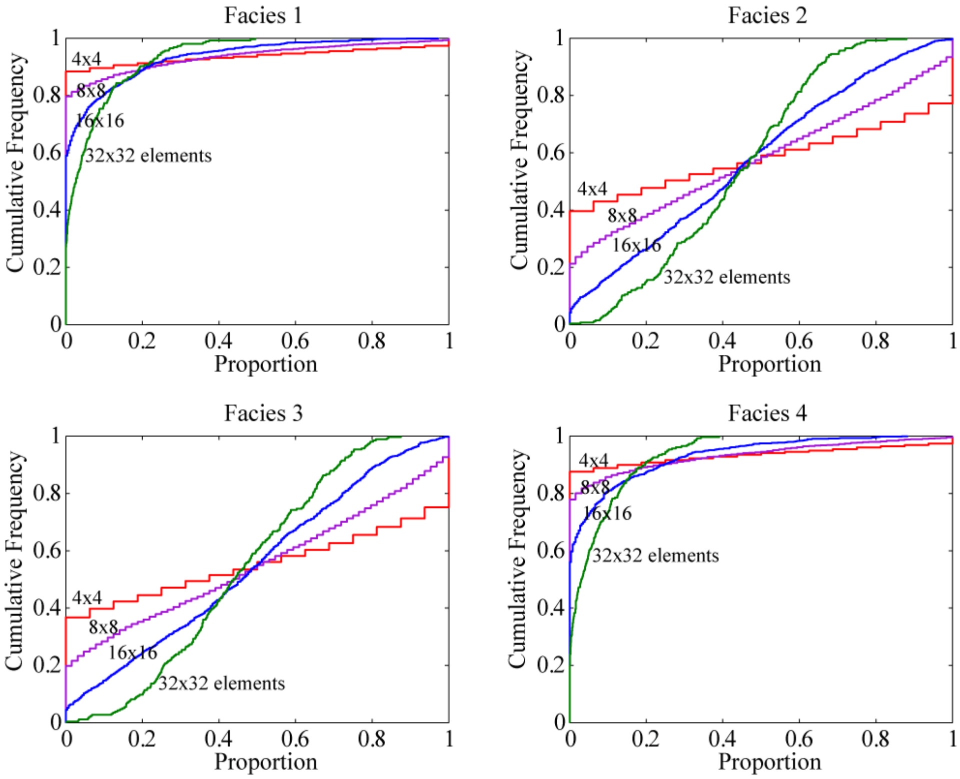


Figure 1-17: Empirical facies proportion distributions at four different scales.

Applying the scaling laws described in the previous section to upscale facies is slightly different due to the compositional character of the result. To ensure that the facies proportions for a given volume sum to one, this constraint must be included in the scaling laws that control the distributions. Compositional scaling of this type can be avoided by dealing strictly with facies at a point scale, where only one facies can exist.

Current practice for geostatistical treatment of facies is to upscale from core and well log scale to the geologic modeling scale using majority vote averaging. The most abundant facies in a grid element is chosen to represent the whole element and global proportions and indicator variograms are calculated based on the result. Considering that the geologic modeling scale is an order of magnitude larger than core and log data, this practice raises concerns. Facies that contribute significantly to the flow response of a reservoir, such as highly permeable sand or impermeable shale, can be neglected completely if they contribute to a small proportion of a grid element, see Figure 1-18.

Two cases are identified in the figure to show that grid positioning can have a drastic effect on flow character using majority vote upscaling. Another issue is that an experimental variogram calculated for a particular facies category is not necessarily representative of the actual variogram for that same category. Consider the vertical variogram calculated for the original layering (the point scale variogram) in Figure 1-18 and for the two upscaled scenarios (Figure 1-19): the upscaled experimental variograms are different from one another and neither is equivalent to the actual vertical variogram. Variograms were computed by discretizing the full thickness into layers with a thickness of 1 unit.

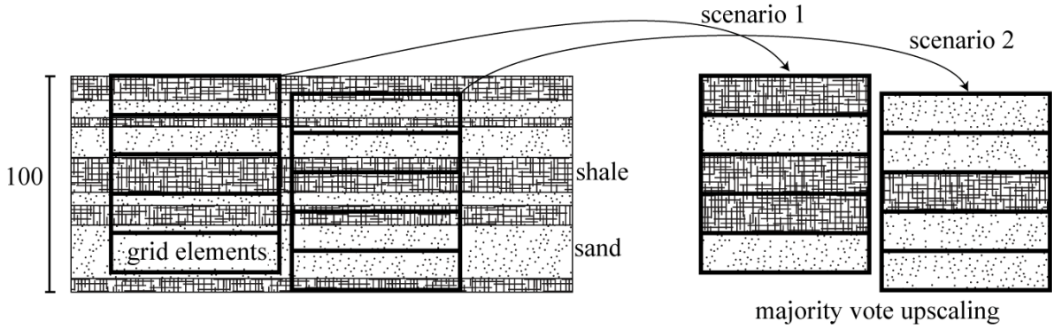


Figure 1-18: Two cases of majority vote upscaling in layered media. The position of the origin of the grid elements leads to different upscaled results. Units are arbitrary.

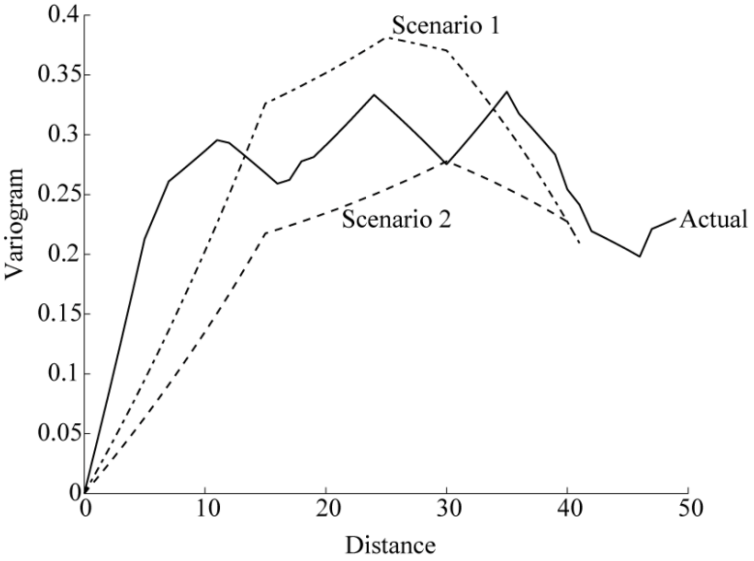


Figure 1-19: Vertical variograms of actual facies and scenarios from Figure 1-18.

1.3.2.3 Permeability and Transmissibility

Only under special conditions does permeability average in a known manner (Renard and de Marsily, 1997). For a discussion of permeability upscaling methods, it is assumed that a permeability field and grid are defined at the fine scale, and that a grid for the coarse scale is also available. For averaging techniques other than arithmetic, the previously discussed scaling laws using average covariance are not applicable. The goal of these upscaling methods is to achieve a similar flow response on all scales.

Arithmetic averaging is appropriate in a layered porous medium when each layer has a constant permeability and flow is parallel to the layers (Figure 1-20). In the same type of media and flow perpendicular to the layers, harmonic averaging is correct (Figure 1-21). Arithmetic and harmonic averages for permeability under these conditions are defined by Equations 1.16 and 1.17, where n is the number of layers and k_A and k_H are respectively the arithmetic and harmonic averages of the layered volume V. Cardwell and Parsons (1945) proved that these two averages define the lower and upper limit of the effective permeability of a block of porous media.

$$k_A = \frac{1}{n}\sum_{i=1}^{n} k_i \qquad 1.16$$

$$k_H = n\left(\sum_{i=1}^{n}\frac{1}{k_i}\right)^{-1} \qquad 1.17$$

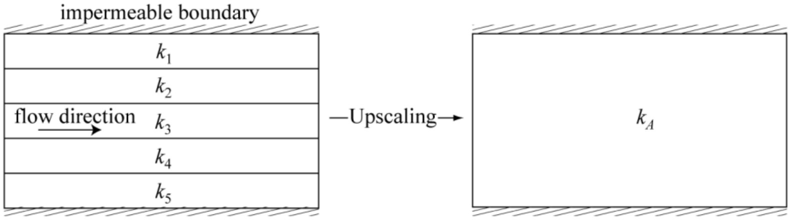


Figure 1-20: Conditions where arithmetic average of permeability is appropriate.

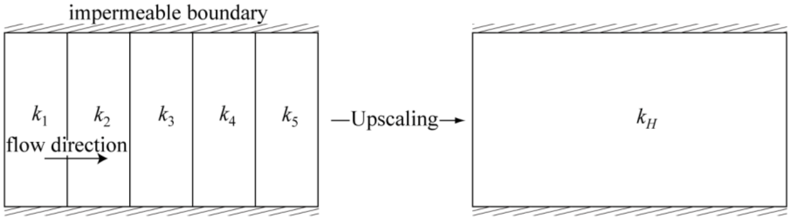


Figure 1-21: Conditions where harmonic average of permeability is appropriate.

In more complex media, simple averages have been evaluated including the geometric average (Jensen, 1991), defined by Equation 1.18, and the power average (Journel et al, 1986; Deutsch, 1987 and 1989) defined by Equation 1.19, where B is the power. Here the sum over n layers has been replaced by a sum over V to indicate the method has been applied to more arbitrary volumes and heterogeneous permeability fields.

$$k_G = \exp\left(\frac{1}{n}\sum_V \ln(k_i)\right) \qquad 1.18$$

$$k_B = \left(\frac{1}{n}\sum_V k_i^B\right)^{\frac{1}{B}} \qquad 1.19$$

The power average is equivalent to the arithmetic average when $B = 1$, to the harmonic average when $B = -1$, and to the geometric average as $B \rightarrow 0$. Some combinations of averages such as harmonic-arithmetic have been advocated to obtain directional permeability for problems in two or three dimensions (Aarnes et al, 2007). For example, in a layered reservoir where flow is primarily in a direction parallel with the layering harmonic-arithmetic averaging proceeds as follows: first, harmonic averaging is applied to evaluate permeability in the direction perpendicular to the layers, and second, the harmonic averaged permeability values are averaged arithmetically in the direction parallel to the layers.

These simple averaging techniques are quite limited, especially in two and three dimensional problems when permeability is a tensor property. Harmonic, arithmetic, power, and combinations can only provide upscaled permeability in specified directions; therefore, they are limited to diagonal tensors and off diagonal or cross flow terms cannot be derived. More complex upscaling models have been developed to accommodate tensor permeability. These methods are referred to as dynamic or flow based upscaling since permeability is used to solve for pressure and flux fields by flow simulators. Resulting pressures and fluxes are used to solve for permeability tensor coefficients or transmissibility.

Several types of flow based upscaling methods exist. The following section describes the setting and methodology behind local upscaling for both permeability and transmissibility. The setting for local upscaling is as follows: there is a fine scale permeability model defined on one grid specification, a coarse scale grid specification, and a series of local problems used to calculate permeability or transmissibility for the coarse grid (Figure 1-22). Two general types of local problems are shown; type 1 uses averaging techniques discussed previously, except pressures and fluxes are used to calculate a permeability tensor, and type 2 involves coarse grid cells that share a common interface where pressures and fluxes are used to calculate the transmissibility between the two coarse cells. For both problem types, boundary conditions must be chosen to induce flow through the element or across the interface.

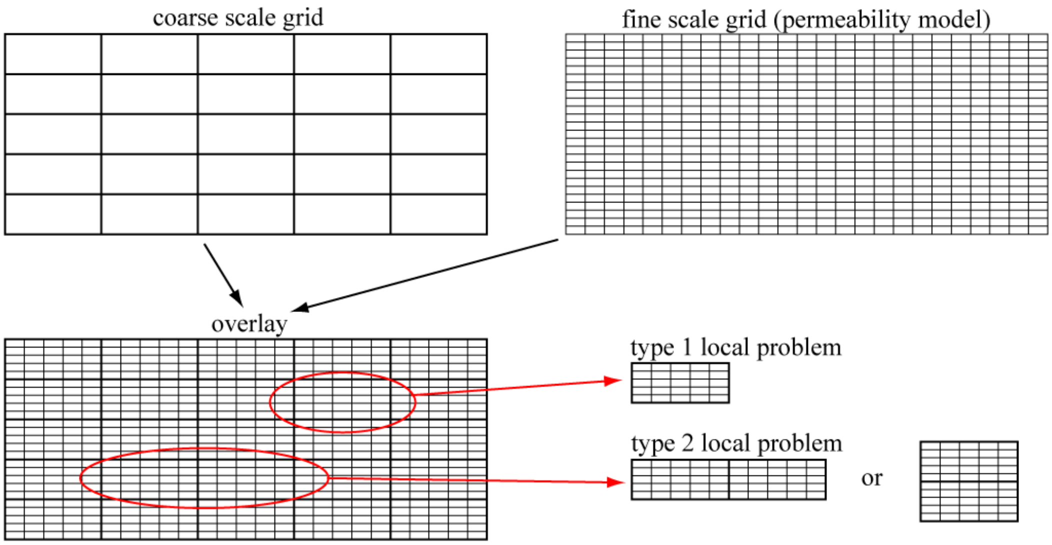


Figure 1-22: general scenarios for upscaling permeability and transmissibility.

Calculating a permeability tensor from known pressure and flux fields at the fine scale is accomplished with Darcy's Law (Equation 1.20).

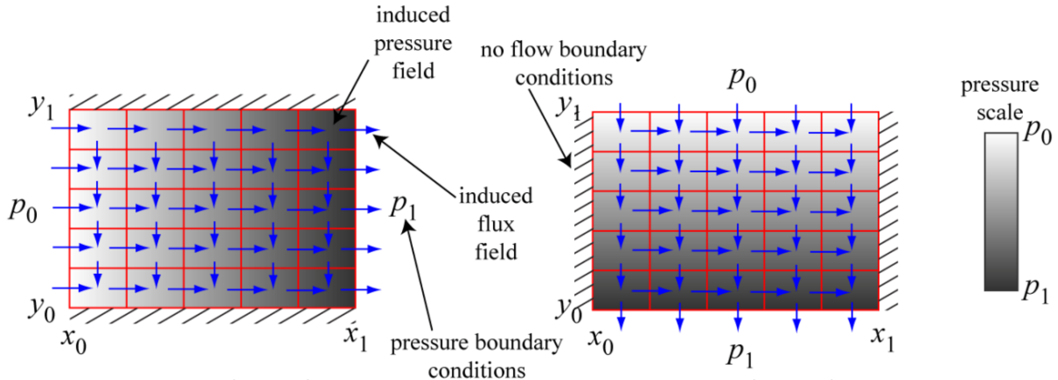
$$\mathbf{q} = -\frac{1}{\mu}\mathbf{k}\nabla p \qquad 1.20$$

Let v denote fine scale grid elements and V denote coarse scale elements ($v < V$), then the upscaled version of Equation 1.20 is given by Equation 1.21, where the upscaled flux is approximated as the average fine scale flux over a coarse grid element and the pressure gradient is approximated by the average fine scale pressure drop across the same region, both are calculated with Equation 1.22.

$$\mathbf{q}_V = -\frac{1}{\mu}\mathbf{k}_V\nabla p_V \qquad 1.21$$

$$\mathbf{q}_V = \frac{1}{|V|}\int_V \mathbf{q}\,dV \qquad \nabla p_V = \frac{1}{|V|}\int_V \nabla p\,dV \qquad 1.22$$

Unless it is assumed that the upscaled permeability tensor is diagonal, Equation 1.21 is underdetermined. For $\mathbf{k}_V$ symmetric, there are six unknowns and only three equations; therefore, multiple sources of $\mathbf{q}$ and ∇p are required. In fact, this is also true for the diagonal tensor case because if the tensor were truly diagonal, $\mathbf{q}$ and ∇p may be exactly zero in one or two of the three possible directions. Multiple sources are obtained by using different boundary conditions and solving for all components of $\mathbf{k}_V$ is a regression problem; the residual between $\mathbf{q}_V$ and $\mathbf{k}_V\nabla p_V$ is minimized. The residual is a measure of the information or detail lost in upscaling the fine scale field to the coarse scale one. Figure 1-23 shows an example that uses two sets of boundary conditions to force flow in different directions such that a diagonal tensor can be derived. In this simplified case, regression is not required and permeabilities are computed directly as shown by the equations in Figure 1-23. The coarse scale flux is computed from the fine scale flux field either at the inlet or outlet faces. The flux field shown is for explanatory purposes only.



$$k_{xx} = -\frac{\mu q (x_1 - x_0)}{(p_1 - p_0)(y_1 - y_0)} \qquad k_{yy} = -\frac{\mu q (y_1 - y_0)}{(p_1 - p_0)(x_1 - x_0)}$$

Figure 1-23: Example type 1 local problems to upscale permeability.

For transmissibility upscaling (type 2 problem in Figure 1-22), where flow is calculated between multiple coarse grid elements, a transmissibility value is to be calculated from known fine scale pressure and flux fields. This type of upscaling is only applicable to two-point flux simulation. Transmissibility is used to express the flux across a cell interface in terms of the pressures in cells adjacent to the interface. This is shown in its simplest form in Equation 1.23 where q_{Vij} is the flux between elements V_i and V_j, T_{ij} is the transmissibility, and p_{Vi} and p_{Vj} are the pressures, see Figure 1-24.

$$q_{Vij} = T_{ij} \left(p_{Vj} - p_{Vi} \right) \tag{1.23}$$

Unlike Equation 1.21 where flux is a vector quantity, flux across an interface is expressed as a magnitude in Equation 1.23. Coarse scale flux is calculated using fine scale fluxes along the interface $\partial V_{ij} = \partial V_i \cap \partial V_j$ in Equation 1.24, where $\mathbf{n}$ is the unit normal vector along ∂V_{ij} pointing from V_i to V_j, and p_{Vi} and p_{Vj} are calculated using the fine scale pressures within elements V_i and V_j respectively by Equation 1.25.

$$q_{Vij} = \int_{\partial V_{ij}} \mathbf{q} \cdot \mathbf{n} dV \tag{1.24}$$

$$p_{Vi} = \frac{1}{|V_i|} \int_{V_i} p dV \qquad p_{Vj} = \frac{1}{|V_j|} \int_{V_j} p dV \tag{1.25}$$

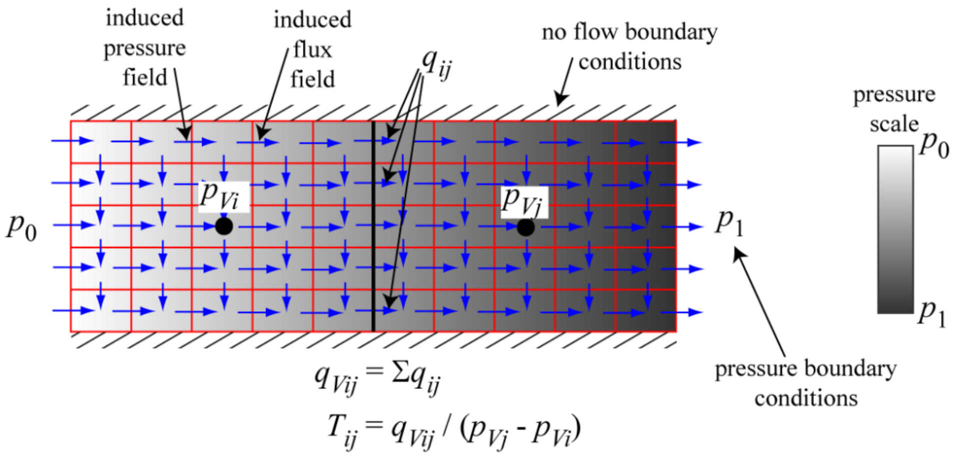


Figure 1-24: Example type 2 transmissibility upscaling.

The type of permeability upscaling used depends on the flow simulation technique, some require permeability while others the transmissibility. In this thesis the focus is upscaling permeability from the fine scale grid to transmissibility on the coarse scale grid. The following section reviews some of the literature related to flow simulators and upscaling.

1.4 Literature Review

Several research topics are involved in this dissertation: unstructured grids, grid design, grid refinement, geostatistics, upscaling, and flow simulation. Each topic is an area of active research with extensive history and much accumulated knowledge. This section reviews previous work on integrating unstructured grids into geostatistical modeling and flow simulation workflows. A brief review of the use of unstructured grids in reservoir modeling is relevant to understand their geometry and properties. More specifically, the review on unstructured grids, which is combined with flow simulation, addresses how they are defined, what features are inherent in good designs, and how they are generated. A literature review on upscaling is also provided.

1.4.1 Unstructured Grids and Flow Simulation

An unstructured grid is a generalized geometric entity that partitions Euclidean space into a set of elements. By definition there are no constraints on the geometry of elements or their configuration pattern apart from being non-intersecting. The design capabilities of unstructured grids are practically unlimited and many forms of elements are possible. Several have been used in the finite element method (FEM) (Hrenikoff, 1941; Zienkiewicz, Taylor and Zhu, 2005), where elements range from tetrahedra to complex curvilinear elements. For reservoir modeling, grids partition the three dimensional space containing the reservoir into a set of elements. The targeted use of these grids is for engineering purposes and the design of elements is done to optimize the quality of flow solutions, or in other words to minimize the error incurred due to gridding. Elements should align with significant geologic features like boundaries, faults, and horizons that are relevant to flow (Kocberber, 1997) and areas with potentially high volumetric flux such as near wells. Grid design is also constrained by the limitations of flow solvers which are discussed later in this section.

Historically, unstructured grids have been in use since the 1940's in various fields of continuum mechanics. They are a recent addition to reservoir simulation with research beginning in the 1980's using Voronoi grids (Pathak et al, 1980). Substantial interest was not seen until the 1990's with PEBI grids (Heinemann et al, 1991; Palagi and Aziz, 1994). Research advancements in both flow simulation on unstructured grids and grid design from the 1990's to date are extensive and a detailed review can be found in Prevost (2004). The majority involve convex elements of varying geometry designed to align with reservoir geology and to capture complex near-well flow or flow through highly permeable conduits. Some recent examples are automatic flow-based grid generation to represent reservoir structure and areas with high volumetric flux (Edwards, 2002; Prevost et al, 2005) and a hybrid grid generation scheme for gaining accuracy near wells (Flandrin, Borouchaki, and Bennis, 2006).

Characterizing reservoir structure and manually constructing a good quality grid is time consuming. With the complexity of reservoir models, automatic grid generation techniques are a necessity. Many techniques exist and several are described by Frey and George (2000). They describe triangulations and tetrahedralizations as being the most flexible for representing complex geometries, which include any combination of faults, horizons, stratigraphic layers, and channels. Three of the main types of unstructured grid generation algorithms are aimed at generating simplex grids and include spatial decomposition, advancing-front, and Delaunay methods.

Several developments have been made in generating coarse grids for reservoirs. Most are based on smaller models for which simulation of the respective fine grid is manageable for comparison purposes. Having the capability to flow simulate the fine grid provides an avenue to assess the effectiveness of the unstructured grid design. One of the earliest automatic unstructured gridding schemes involved PEBI grids and was developed as an advancing front method. Points were distributed through space based on certain criteria, and then connectivity was determined resulting in a Voronoi grid (Palagi and Aziz, 1994). Generating PEBI grids remains an active research topic (Hale, 2002; Mlacnik, Durlofsky, and Heinemann, 2006; Katzmayr and Ganzer, 2009; Yadecuri and Mahani, 2009). Other grids are generated for finite volume simulation methods and multiscale finite element methods and are typically triangulations or tetrahedralizations.

Other methods of generating unstructured grids for reservoirs are hybrid techniques and flow based techniques. In hybrid techniques, reservoirs are initially partitioned by a coarse background grid, which is refined in areas with complex geology and anticipated high saturation and flow gradients (Aziz, 1993). The background grid is commonly regular so does not perfectly align with faults and horizons. Automatic methods are employed to create elements to align with these structures (Kocberber, 1997; Mallet, 2002; Flandrin, Borouchaki, and Bennis, 2006). High saturation gradients are typical near injection and production wells where the background grid is refined to conform to well paths. This refinement is done by removing elements within the vicinity of a well path, then filling the void space with a radial grid.

Flow based grid generation techniques require the solution to a fine scale flow problem from which grid elements are derived. Grid vertices may be distributed based on streamlines (Verma and Aziz, 1996), streamlines and isopotentials (Mlacnik, Durlofsky, and Heinemann, 2006), streamtubes and isobars (Portella and Hewett, 2000), or other flow information. The fine scale problem is usually a basic incompressible, single phase pressure equation, although some research is being done on multiphase upscaling (Christie, 1996 and 2001; Arbogast, 2002; Chen and Li, 2009). Solving the flow equations characterizes the pressure and velocity fields throughout the fine scale grid, which is often a structured regular grid. This is the case since current geologic property modeling techniques are designed for regular grids. The problem of building a permeability model for input to the pressure equation on other discretizations has not been addressed and is a focus of this dissertation.

Both static (Edwards, 2002) and dynamic (Mlacnik, Harrer, and Heinemann, 2003) flow based generation techniques are possible. In the static case, a grid is generated once and remains fixed for the duration of time dependent simulation runs. It may be generated using no wells and arbitrary global boundary conditions so that streamlines are influenced primarily by reservoir structure and heterogeneity, since the element size and density varies according to characteristics of the flow solution. A grid may also be generated for a specific flow scenario, such as a production well configuration where boundary conditions are imposed globally and at well locations. Simulation runs using the resulting grid are only valid if there are no significant changes to the boundary conditions over the time of the run. Shutting in a well or converting a producer into an injector would have a significant impact on flow and thus on the optimal flow based grid. To alleviate the influence of changing reservoir operating conditions over time adaptive grids may be considered. Each alteration to reservoir conditions requires the solution to a fine scale simulation problem, grid generation, grid optimization, and permeability upscaling stages, providing the grid and properties for the next series of time steps on the new grid. Depending on the size of the fine scale model, this series of stages can become time consuming.

For any simulation grid the constraints imposed by discretization of the flow equations must be considered. Flow simulators that work with unstructured grids is an active area of research. Their capabilities dictate how flexible an unstructured grid design can be and the element geometry that is applicable. FEM is an approach applicable for unstructured grids, but is unstable for complex porous flow problems (Farmer, 2005). The simulation technique that has received the most attention due to its use in industry and ability to handle unstructured grids is the finite volume method (FVM).

Most commercial hydrocarbon flow simulators using FVM are limited to corner point grids (Ponting, 1992), which are structured and possibly irregular. Although structured grids do offer some flexibility, there are complications for alignment with well trajectories, overturned surfaces, intersecting faults, and thrust faults (Farmer, 2005).

These simulators are also based on the two-point flux approximation (TPFA); flux from one element to another depends only on the pressures in those two elements. Transmissibility for the face separating the two elements is based on two points typically chosen as the element centers. It is calculated using the permeability tensors for each element, which are usually assumed diagonal. Use of TPFA with full tensors results in error (Edwards and Rogers, 1998; Eigestad and Klausen, 2005; Chen and Mallison, 2007), motivating the development of new simulation techniques.

There are three main types of next generation simulators that are being developed: FVM with multipoint flux approximation (MPFA); hybrid FE-FVM or mixed FEM (MFEM); and multiscale simulators. MPFA methods have been developed for both two dimensions (Aavatsmark et al, 1996; Aavatsmark et al, 1998a,b) and three dimensions (Verma and Aziz, 1997) and are undergoing further development (Eigstad, Aavatsmark, and Espedal, 2002; Mlacnik and Durlofsky, 2006; Chen et al, 2008). MPFA simulators are being developed for triangular, tetrahedral, quadrilateral, and hexahedral grids. Although this appears to be a constraint, practically any unstructured grid consisting of polygonal volumes can be reduced to a set of tetrahedra. Hybrid FE-FVM (Geiger et al, 2004; Paluszny, Matthai, and Hohmeyer, 2007; Maliska, Cordazzo, and Silva, 2007) are very similar to FVM using MPFA with the major difference being a finite element formulation is used to solve the fluid pressure field and finite volumes for the fluid transport phenomena. Tetrahedral grids are the focus, but other possible element configurations are presented.

Multiscale methods (Hou and Wu, 1997; Kippe, Aarnes, and Lie, 2008; Pavliotis and Stuart, 2008) are a recent class of simulation technique that assumes each element is characterized by pressures and fluxes that are functions of position within the element, rather than a constant. Fine scale information is incorporated into the solution of coarse scale equations without having to resolve the fine scale, thus the methods are more efficient than directly simulating at the fine scale. Although still under development, multiscale simulators show promise for complex grid geometries. Aarnes, Krogstad, and Lie (2008) discuss a multiscale mixed/mimetic finite element method that is extensible to general polyhedral cells. They also identify three considerations for generating coarse grids: 1 - Grid geometry should minimize the possibility of bidirectional flow across element interfaces; 2 - Element faces should follow geological layers; 3 - Elements should adapt to flow obstacles. Extending multiscale methods to more complex multiphase flow regimes is ongoing.

A common set of grid attributes for reservoir simulation models are identified in this review on unstructured grids and flow simulation. Attributes are both geological and simulation based. Resulting simulation models should accurately represent large scale geologic structures including the reservoir boundary, faults, and horizons (Kocberber, 1997). Geologic attributes at a smaller scale involve designing elements that do not contain significant permeability discontinuities and do not span structural surfaces like faults. Such occurrences are one source of numerical instability for simulation. Qualities in regards to simulation include the accurate representation of well trajectories, near-well flow, and any other area where complex flow patterns are anticipated, for example through highly permeable conduits or across faults. The overall grid should also be designed to meet the requirements of a particular flow solver and for the most efficient solution of the flow equations. Finally, the process for building a reservoir simulation model should be user-friendly and not overly time consuming. Processes designed to generate unstructured grids should include these attributes.

1.4.2 Geostatistics

There is little previous work involving geostatistics on unstructured grids. The major difficulty in developing geostatistical algorithms for unstructured grids is the variety of grid element volumes. Methods are required to account for the statistics of random variables at different scales; initial research into algorithms for variable scales involved direct sequential simulation (Xu and Journel, 1994). Direct methods were not originally developed for unstructured grids, but for the problem of integrating multiple data types having different scales such as core and log data (Tran, Deutsch, and Xie, 2001).

The application of direct geostatistics on unstructured grids was discussed by Deutsch, Tran, and Pyrcz (2002). Some issues with direct methods were identified including problems with histogram reproduction and the proportional effect inherent in variables having skewed distributions. Solutions to these issues are discussed and more detail is provided in Oz et al (2003). These techniques involve the use of average covariances that reflect the geometry and scale of unstructured grid elements and of the sample data, but their use is restricted to random variables that average arithmetically such as porosity. An approach for permeability is suggested and assumes that arithmetic averaging applies after a power law transformation; however, this technique cannot yield tensor permeability for unstructured grid elements. More information on the use of power averaging can be found in Deutsch (1989) and in the background section. Some follow-up work to address implementation aspects has been done (Manchuk, Leuangthong, and Deutsch, 2004).

Direct methods have rarely been used in practice, which is likely due to it being in too early a stage of development. A series of outstanding issues are discussed by Leuangthong (2004). All other work involving geostatistics on unstructured grids implements an underlying structured grid such that existing geostatistical theory and algorithms are applicable (Caumon, Grosse, and Mallet, 2004). Caumon, Grosse, and Mallet (2004) also highlights three of the potential concerns with implementing geostatistics on unstructured grids including: more complex neighbourhood search, elements being aligned in directions different than the principal directions of geological anisotropy, and computation of average covariance.

1.4.3 Upscaling

Equivalent reservoir properties for unstructured grid elements are computed from values defined on a refined grid. Different properties scale according to different laws and in reservoirs, there are two principle types encountered: arithmetic averaging properties and flow properties. Variables such as porosity, density, fluid saturation, and facies proportions scale arithmetically, while permeability and transmissibility scale according to equations of flow.

Bear (1972) discusses the concept of representative elementary volume in terms of volumetric porosity, which is computed as an arithmetic average. This approach conserves the pore volume across scales (Durlofsky, 2003). Some recent work has been done on scaling of categorical variables such as facies (Deutsch and Lan, 2008). At a small scale, facies are considered categories, but as scale is increased, facies are expressed as proportions that tend to follow a Beta-type distribution function. Proportions are computed arithmetically and within a specific volume, such as an unstructured grid element. The proportions of all facies types must sum to unity, thereby forming a compositional variable (Pawlowsky-Glahn and Olea, 2004).

For permeability, several classes of single phase flow based upscaling exist: local, extended local, global, and coupled local-global. In local methods, grid elements are considered independently, much like a sample of reservoir rock is treated in a lab (Warren and Price, 1961). Flow is hypothetically induced across the element in different coordinate directions by means of prescribed boundary conditions and Darcy's law is used to solve for a permeability tensor. For two-point transmissibility, as used in the TPFA, two elements are considered in the local problem and flow is induced across the interface between them. The average pressure in each element and the flux across the interface are used to compute the transmissibility (White and Horne, 1987). Extended local methods involve a grid element plus an additional region surrounding it (Gomez-Hernandez and Journel, 1990). The intention is to achieve a more accurate representation of flow through the element as it exists within the reservoir. The same principle can be applied to compute transmissibilities. Probably the least used method is global upscaling, which requires solving the flow equations over the entire fine scale geologic model (Farmer, 2002). This is computationally demanding.

Coupled local-global upscaling is more recent (Chen et al, 2003; Chen and Durlofsky, 2006). In this technique, flow equations are solved globally on the coarse grid and used as boundary conditions for local problems, which are solved to update either the permeabilities or transmissibilities for the global problem. The process is iterative. More detailed information on these and other upscaling techniques can be found in several extensive reviews (Wen and Gomez-Hernandez, 1996; Renard and de Marsily, 1997; Farmer, 2002; Durlofsky, 2003; Prevost, 2004; Farmer, 2005; Aarnes et al, 2007).

Advances made towards upscaling on unstructured grids assume the underlying permeability model is defined on a regular grid. He, Edwards, and Durlofsky (2002) develop an upscaling regime for general quadrilateral elements using MFEM. The background grid and overlying quadrilateral element are triangulated. Permeabilities from the regular grid are assigned to the triangulation, which is involved in the flow based upscaling approach. Edwards (2002) describes an upscaling method where unstructured control volumes are enclosed in a bounding box that contains a subset of a regular gridded permeability model, much like the local upscaling methods discussed previously. Local flow based upscaling for regular grids is then used, but only the pressures and fluxes within the enclosed unstructured element and along its boundary are used to compute equivalent permeability. Prevost (2004) describes two upscaling methods for unstructured grids to compute permeability tensors or transmissibilities. Tensors are utilized in the MPFA method and transmissibilities in the TPFA. Again the fine scale permeability model is defined on a regular grid. Development of a technique to upscale permeability from a fine scale unstructured grid to a coarse scale unstructured grid is required.

1.5 Thesis Outline

Chapter 2 discusses many of the pre-processing requirements to undertake reservoir modeling on unstructured grids, including a proposed a workflow. Grid discretization using simplex elements is introduced along with existing tools in two and three dimensions. Tools for visualization are also identified. Spatial transformations from geological space to depositional space are described and their applicability to unstructured grids is demonstrated. A common concern with geostatistics on unstructured grids is data scale; hence a discussion is given along with suggestions for upscaling from sample data scale to a consistent scale for reservoir modeling. This chapter closes with a section on grid design.

Chapter 3 primarily develops a sequential simulation algorithm for populating unstructured grids with categorical and continuous properties. It is based on sequential indicator simulation and sequential Gaussian simulation from GSLIB (Deutsch and Journel, 1998). Input data and the grid specification are discussed. Implementation details covered include random path generation and nearest neighbour search. Model validation for the histogram and variogram are discussed and a tool for computing the experimental variogram of large point sets is developed.

Chapter 4 investigates grid discretization and upscaling. These topics are placed together since the accuracy of upscaled properties depends on the discretization. Theory is developed for arithmetic averaging. A method that links grid element volume and number of discretization points to the expected error in average properties is developed. For permeability, an upscaling approach using the MPFA approach is developed. Permeability that is modeled on the fine scale grid is upscaled to multipoint transmissibilities on the coarse grid for MPFA simulators. Some guidelines on discretization to achieve accurate upscaled transmissibilities are given.

Chapter 5 validates the methods and provides a case study. The discretization approach for arithmetic averaging variables is proven to reproduce the correct variance and covariance at the scale of coarse unstructured grid elements. It is also shown that statistics including the variogram and proportional effect are reproduced through transformations from a Gaussian variable to a reservoir property such as permeability. Potential biases due to irregular discretization are discussed. It is demonstrated that using a low enough error criteria for the grid discretization process removes biases in global statistics. The flow based upscaling approach is validated for homogeneous, layered, and heterogeneous media. The case study goes through all geostatistical modeling steps of the workflow proposed in this dissertation.

Chapter 2 Methodology

Developing geostatistical methods for unstructured grids requires a modified workflow and a novel treatment of data and scale. Firstly, Section 2.1 summarizes an existing workflow for reservoir characterization and flow simulation, and then discusses modifications required for the proposed workflow involving unstructured grids. Section 2.2 explains key pre-processing steps and prerequisites required to do geostatistical modeling including grid discretization; stratigraphic and chronologic transformations; defining a scale for core and geophysical log data; dealing with secondary data such as seismic; and visualization. Lastly, Section 2.3 covers aspects of grid design with examples.

2.1 Workflow

An existing workflow for geostatistical reservoir modeling involves the following (Deutsch, 2002):

1. Data preparation
2. Model large scale features such as horizons and faults
3. Choose a geological modeling grid
4. Preliminary mapping
5. Facies modeling
6. Modeling continuous properties such as porosity and permeability
7. Model validation

Resulting models are used in reservoir flow simulation. If the numbering is continued, these steps involve:

8. Coarse grid design
9. Upscaling to effective properties
10. Flow simulation

Currently, grids designed for geological modeling are structured regardless of the coarse grid designed for flow simulation, i.e. step 3 (choosing a geological modeling grid) rarely considers step 8 (coarse grid design).

The proposed workflow (Figure 2-1) is similar, but ensures the geological modeling grid and flow simulation grid are designed together. Steps 1 and 2 remain unchanged since they are done prior to any choice of grid. Steps 3 to 6 are categorized as preliminary modeling and are carried out on a regular grid for the purpose of coarse grid design. The final coarse grid design is discretized to a fine grid for geological modeling. Results are upscaled to the coarse grid for flow simulation. The proposed workflow is as follows:

1. Data preparation
2. Structural modeling
3. Grid design
    a. Regular gridding and preliminary mapping
    b. Coarse (unstructured) grid design
4. Grid discretization
5. Facies modeling
6. Continuous property modeling
7. Model validation.
8. Upscaling
9. Flow simulation

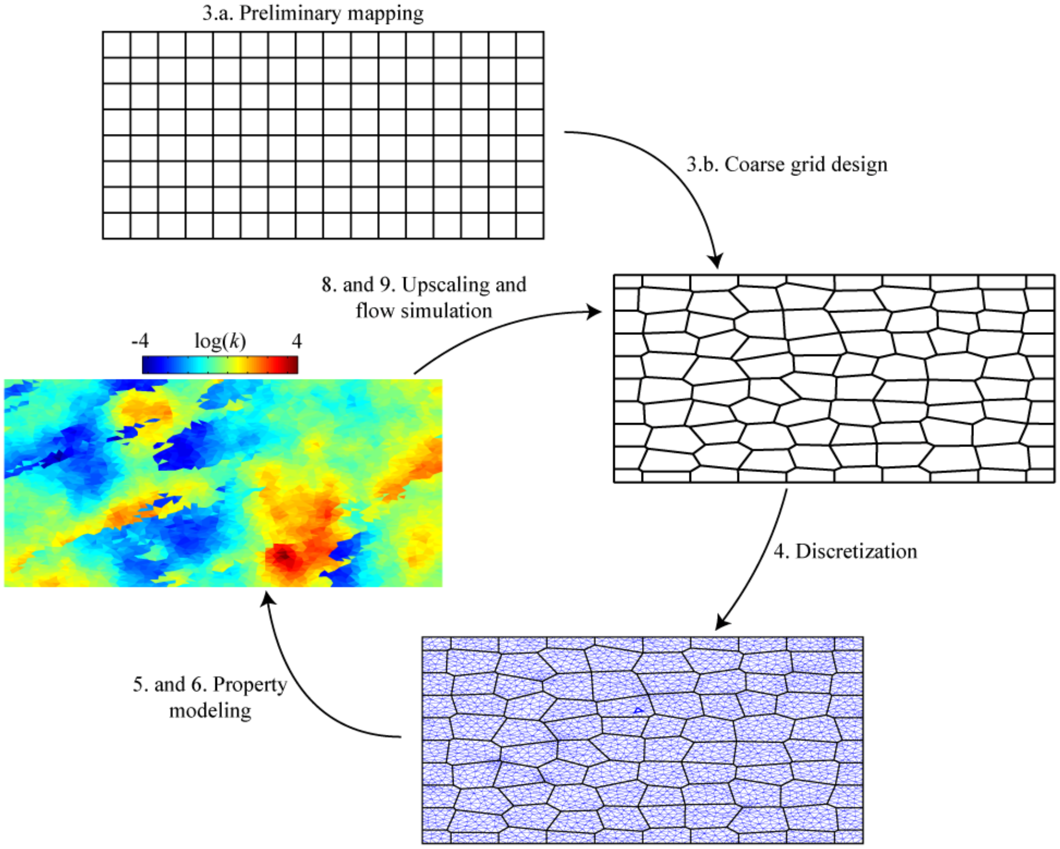


Figure 2-1: Reservoir modeling workflow with unstructured grids. Preliminary mapping is done to provide a background grid of reservoir properties to control grid design. The grid is discretized to a finer scale and then populated with reservoir properties using geostatistics. The resulting models are upscaled to the coarse grid design for flow simulation.

A clear disadvantage of this workflow is geological modeling is carried out twice; once during preliminary mapping, since the coarse grid design depends on geological heterogeneity, and again after grid refinement. However, geostatistical modeling workflows are easily automated once they are parameterized. This involves assessing statistical properties of sample data including probability distribution functions, first and second order moments, multivariate relationships, and spatial covariance (Cressie, 1993).

All of these statistics are dependent on the scale of the sample data from which they are computed. Once all data and parameters are defined for a specific case, geostatistical models are constructed automatically with a computer.

The proposed workflow has several advantages. During grid design, features having an effect on flow are included in the coarse grid design and are therefore included in the discretization. These may include existing or proposed wells for production or injection purposes and multiple-well communication data. To obtain higher accuracy in flow, the scale of unstructured grid elements is much smaller near these features (Fung et al, 1994). For example, radial grids (Figure 2-2) are common around production and injection wells (Pedrosa and Aziz, 1986; Lee et al, 2003; Flandrin et al, 2006). Adjacent to a well, extremely small elements are possible. By including these in an unstructured design and discretization, a higher level of detail in geological heterogeneity is achieved. How important an area is in terms of flow is reflected in the resolution of the geological model. Such information is not incorporated into workflows using structured grids for geological modeling. The resolution of heterogeneity is limited to the structured grid element scale. Another advantage comes from having flexibility in the discretization method as explained in Chapter 1.

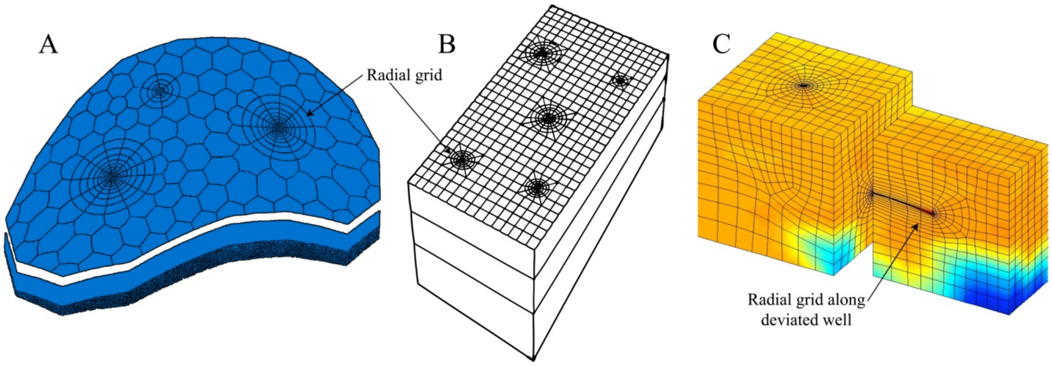


Figure 2-2: Example radial grids: A – 3 wells embedded in a control volume polygonal grid (Fung et al, 1994); B – 5 wells embedded in a regular grid (Pedrosa and Aziz, 1986); C – a deviated well embedded in a distorted hexahedral grid (Lee et al, 2003)

2.2 Preprocessing and Prerequisites

2.2.1 Discretization

Discretization is the process of subdividing space into a set of points or elements. Its purpose in the proposed geological modeling workflow is to provide a grid for describing the geological heterogeneity within coarse unstructured grid elements. In geostatistical modeling and flow simulation, each point represents an area or volume. In this dissertation, unstructured grid elements are discretized into a set of points that are the centers of simplex elements. Each point represents the volume of a simplex. Algorithms for flow simulation and grid generation exist for simplex grids. The discretization of a domain into a set of simplexes is commonly referred to as simplicial mesh generation or meshing (Berg et al, 1998; Frey and George, 2000).

Two existing programs for generating simplex grids are: *Triangle* by Shewchuk (1996) for generating two dimensional triangular grids and *TetGen* by Si (2006) for generating three dimensional tetrahedral grids. Both are constrained generators, meaning the resulting grids conform to boundaries such as faults, horizons, and the faces of unstructured grid elements. Other constraints are also possible; two of the more important ones are simplex size and quality. In two dimensions, triangle area is constrained to a user specified maximum either globally or locally within segment bounded regions. The same concept is used in TetGen to constrain the volume of tetrahedral elements, except local control is within facet bounded regions.

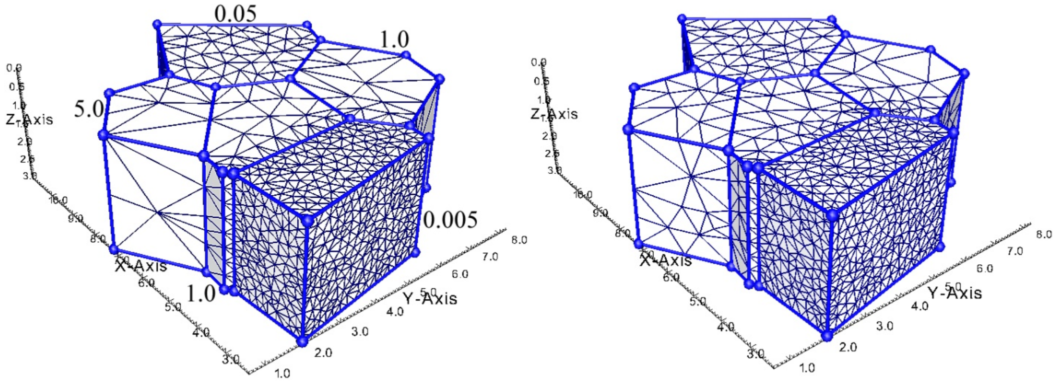


Figure 2-3: Example of volume constraints using TetGen. Numbers adjacent to elements indicate the assigned volume constraint. Radius-edge ratio constraints of 2 (left) and 1.4 (right) are shown. Units are arbitrary.

Using Triangle, quality is measured by a lower bound angle constraint for triangles. For example, specifying an angle constraint of 20 degrees results in a triangular grid where all interior angles are no less than 20 degrees. Using TetGen, quality is measured using an upper bound radius-edge ratio measure for tetrahedra (Miller et al, 1995). The radius refers to a tetrahedron's circumsphere and the edge refers to the length of the tetrahedron's shortest edge. Quality measures are an important link between grid geometry and the accuracy of solving flow equations using the grid (Aavatsmark et al, 1998b; Knabner and Angermann, 2003). A three dimensional example of volume constraints (Figure 2-3) is shown using coarse scale PEBI grid elements as the bounded regions. Two radius-edge ratio constraints show the difference in tetrahedron shape achieved.

Discretization is important when the grid designed for flow simulation cannot take into account all geological features for reasons of computational feasibility and time. The discretization process is used to include these details such that their effect on flow is captured through upscaling. As mentioned, capturing the variation of reservoir properties within coarse unstructured grid elements is the purpose of discretization, but this may involve more than facies, porosity, and permeability. For example, a stratigraphic unit may consist of several thin beds that are not included in the coarse grid. The beds may not correlate with the stratigraphic unit either in the case of cross-bedding (Figure 2-4).

As long as such features can be modeled, they can be included in the discretization and upscaling process.

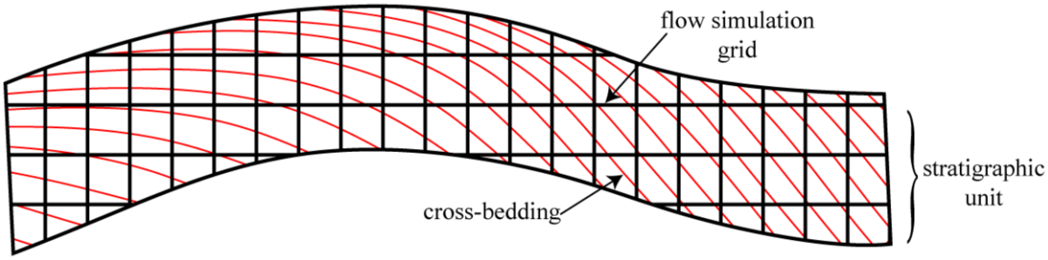


Figure 2-4: Example of cross bedding not captured by a flow simulation grid.

2.2.2 Spatial Transformation

Most reservoirs in their existing configuration, referred to as geological space or physical space, are complex geometric environments that have been affected by post-depositional events including folding, faulting, compaction, and erosion. This results in a distorted image of the reservoir and poses challenges for geostatistical modeling. Determining how reservoir properties correlate is confounded by post-depositional events. In geostatistics, spatial correlation between two points is defined as a function of the Euclidean distance between them. Using the Euclidean distance in geological space directly will not account for geometric deformations. An example involving a fold is shown in Figure 2-5. A common assumption is that the spatial correlation of reservoir properties is entrained in the data prior to any post-depositional alterations (Deutsch, 2002; Mallet, 2004; Gringarten et al, 2008). The space of a reservoir at the time of deposition is referred to as depositional space. Mapping a reservoir from geological space to depositional space involves reversing all post-depositional events. It requires a structural model of the reservoir that includes major stratigraphic surfaces and fault surfaces. Types of erosional events and differential compaction that occurred are also required. Refer to Deutsch (2002) and Mallet (2002) for a review of structural modeling.

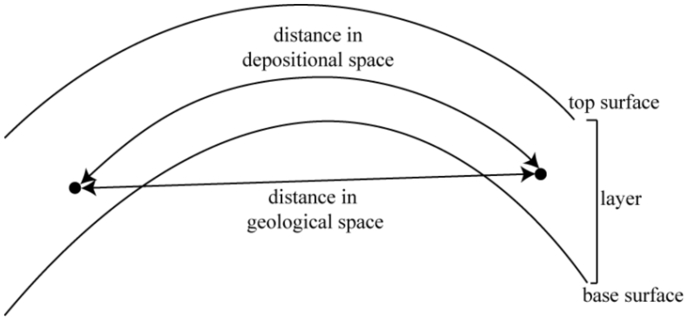


Figure 2-5: Euclidean distances in geological and depositional space. The distance in geological space is shorter than in depositional space leading to a covariance that is artificially high. Not to scale

Two mapping techniques are stratigraphic coordinates (Jones et al, 1986; Deutsch, 2002) and the geo-chronological model (Mallet, 2004). Both techniques assume that geometric deformations affect only the spatial arrangement of a reservoir, not the reservoir properties.

Stratigraphic coordinates define a vertical coordinate mapping within a stratigraphic layer. Layers are defined by two surfaces from the structural model that are referred to as the existing top and existing base of the layer. Due to erosion or onlap, these are not necessarily the correlation top and correlation base of the layer. The vertical coordinates of these surfaces are used to define the stratigraphic coordinates in Equation 2.1, where z_s is the stratigraphic depth, z is the depth of any point within the layer, and z_{ct} and z_{cb} are the depths of the correlation top and correlation base (Figure 2-6). Transformation results will differ depending on the geological interpretations of the reservoir and events resulting in the existing top and base.

$$z_s = \frac{z - z_{cb}}{z_{ct} - z_{cb}} \tag{2.1}$$

Stratigraphic coordinates can be applied to unstructured grids as well. The structural model provides the information required to compute the geological coordinates of any point in the reservoir. If the stratigraphic layer in Figure 2-6 was discretized using a triangulation, the vertices and centers of the triangles could be mapped to stratigraphic space (Figure 2-7).

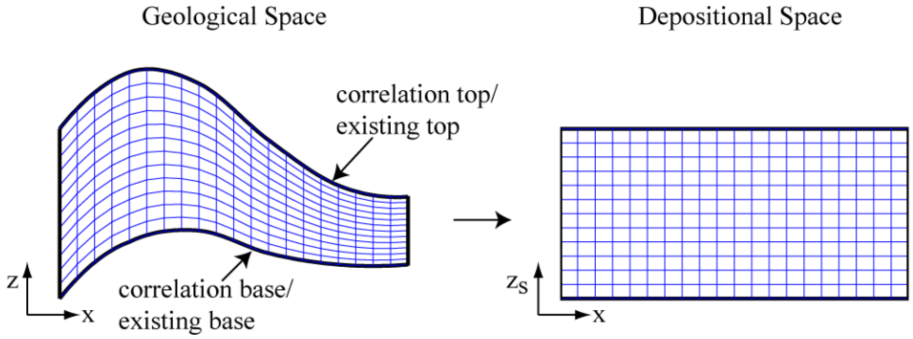


Figure 2-6: Example of stratigraphic coordinates. Not to scale.

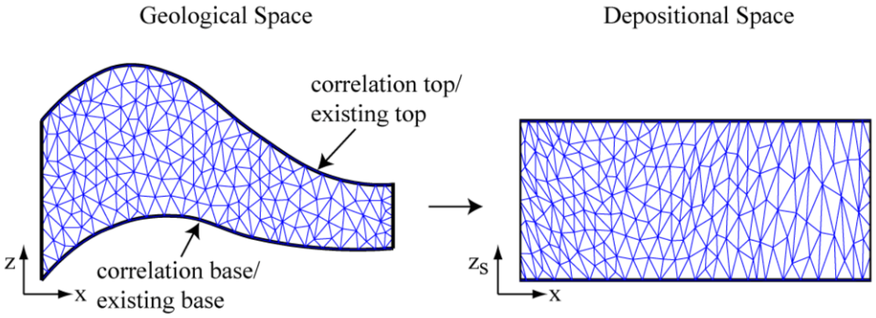


Figure 2-7: Example of stratigraphic coordinates with a triangulation. The triangulation was made using the program Triangle (Shewchuk, 1996). Not to scale.

The geo-chronological model (Mallet, 2004) defines a curvilinear coordinate system in geological space that can be mapped to depositional space while minimizing distortions in distance and volume. One of the differences between stratigraphic and geo-chronological coordinates is how the time axis is defined. Stratigraphic coordinates assume time is vertical in both geological space and depositional space. Geo-chronological coordinates assume time is vertical only in depositional space. It is derived by tracing a curve through particles that were deposited earliest to those deposited latest (Figure 2-8). This transformation is also possible for unstructured grids (Figure 2-9). For the stratigraphic coordinate case, the triangulation is distorted only in the vertical direction; whereas for geo-chronological coordinate, the triangulation is distorted in all directions.

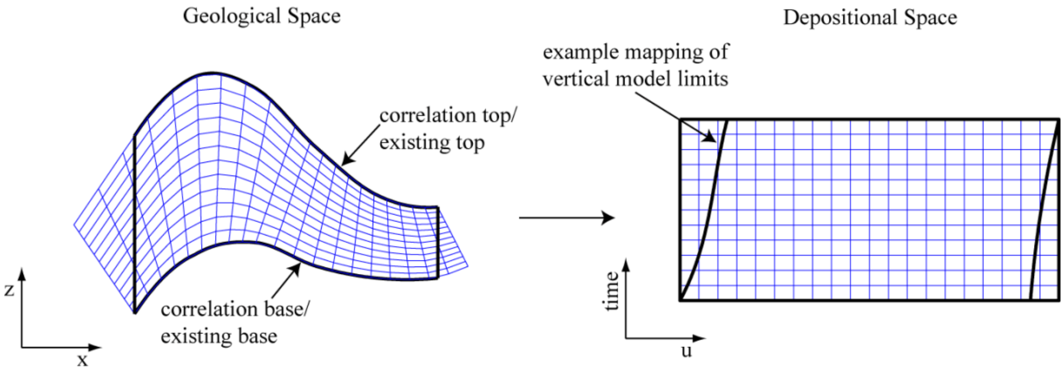


Figure 2-8: Example of geo-chronologic coordinates. Not to scale.

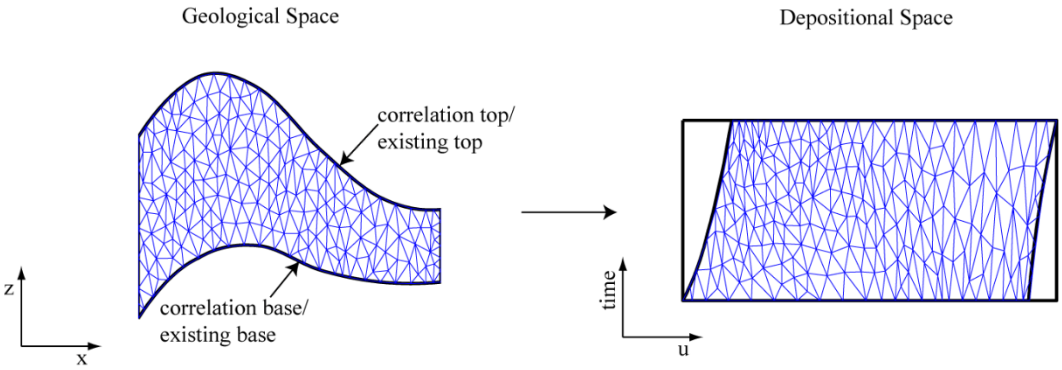


Figure 2-9: Example of geo-chronological coordinates with a triangulation. The triangulation was made using the program Triangle (Shewchuk, 1996). Not to scale.

2.2.3 Data Scale

Geostatistical property modeling requires an association between conditioning data and scale because statistics depend on scale. In geological modeling workflows involving structured grids, the scale is equal to the structured grid element volume in depositional space. It is common to assume that all elements are of equal volume, which is useful in assessing and modeling various statistics including probability distribution functions, multivariate relationships, and variograms. However, sample data is not initially available at this consistent scale. It may include core photos, formation micro image

(FMI) logs, core samples, high resolution and standard wireline logs, seismic data, and well test data (Figure 2-10). Sample data at a smaller scale than structured grid elements (hard data) is upscaled while sample data at a larger scale (soft data) is integrated using other methods that are discussed in the following section on seismic data.

Upscaling from the scale of samples to the geological modeling scale has received little attention as compared to upscaling from the geological model scale to the flow simulation grid scale (Dubrule and Damsleth, 2001). The static averaging methods described in Chapter 1 are often used to upscale core and logs to the geological modeling scale (Figure 2-11). A variety of other averaging and regression approaches are used for permeability as summarized by Deutsch (2009) along with a micro-modeling and flow based upscaling technique using FMI logs. The use of micro-models for upscaling has also been discussed by Hosseini et al (2008). These techniques are needed to unify the scales of sample data for geostatistics on unstructured grids.

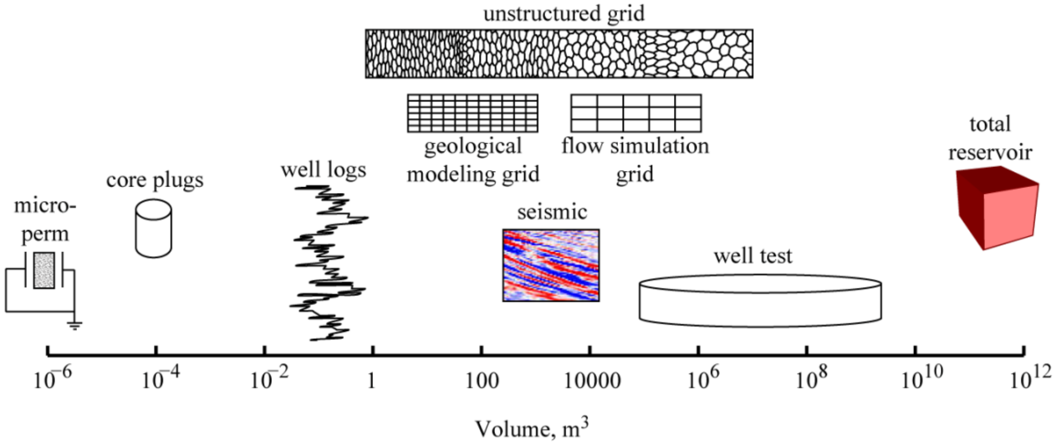


Figure 2-10: Scales for different kinds of reservoir data

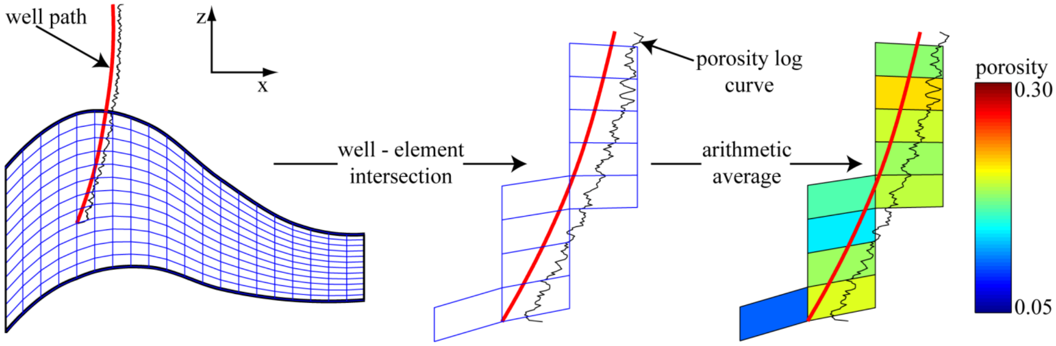


Figure 2-11: Example of arithmetic averaging of a synthetic porosity log to structured grid elements. The section of log samples that fall within each grid element are averaged and assigned to the element. The arithmetic average is used for porosity to preserve pore volume across scale. Not to scale.

Coarse unstructured grids and their discretizations cover a variety of scales in both geological and depositional space. Upscaling hard data to the discretized grid would lead to a set of scale-inconsistent samples, each with a scale equal to their respective grid element volume (Figure 2-12). In cases where a grid is designed to include wells, samples along the well may not fall within an element but on its edges. However, some degree of scaling is necessary to achieve a consistent scale for integrating various types of data and so that variances and covariances between variables are consistent for geostatistical modeling purposes.

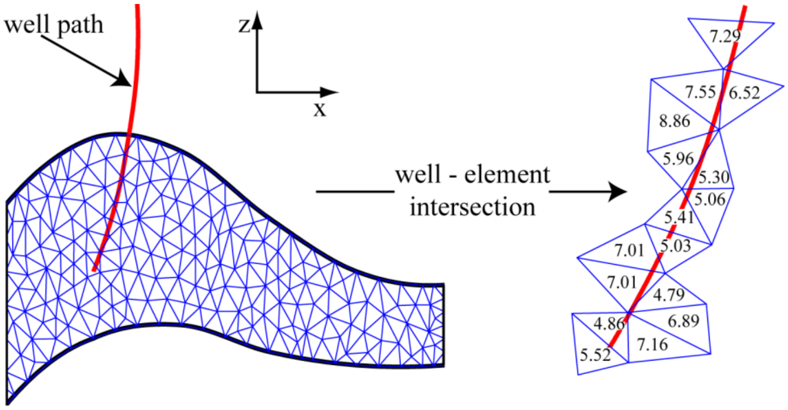


Figure 2-12: Triangular elements intersected by a well. Areas of intersected triangles are shown. Not to scale and units are arbitrary.

This dissertation proposes that a reasonable scale is larger than the small sample data and on the order of the smallest grid element scale in geological space. The scale must be larger than the hard sample data for several reasons. One is for practicality. Each type of data measures a different scale of heterogeneity. For example, a full core sample measuring 5 cm in diameter and height is roughly 100 cm^3; therefore it is possible to take roughly 10,000 core samples form a cubic meter of rock. Given that reservoirs are on the order of millions of cubic meters, assessing this level of heterogeneity throughout is impractical. Another reason is that hard data samples are sparse. Attempting to derive their spatial correlation is made challenging by short scale variability, which is smoothed through upscaling. This concept is shown in Figure 2-13 using an interval from an actual gamma ray log in a carbonate reservoir and arithmetic averaging. By averaging out short scale variability, a higher spatial correlation is observed over a change in depth of 1.8 meters.

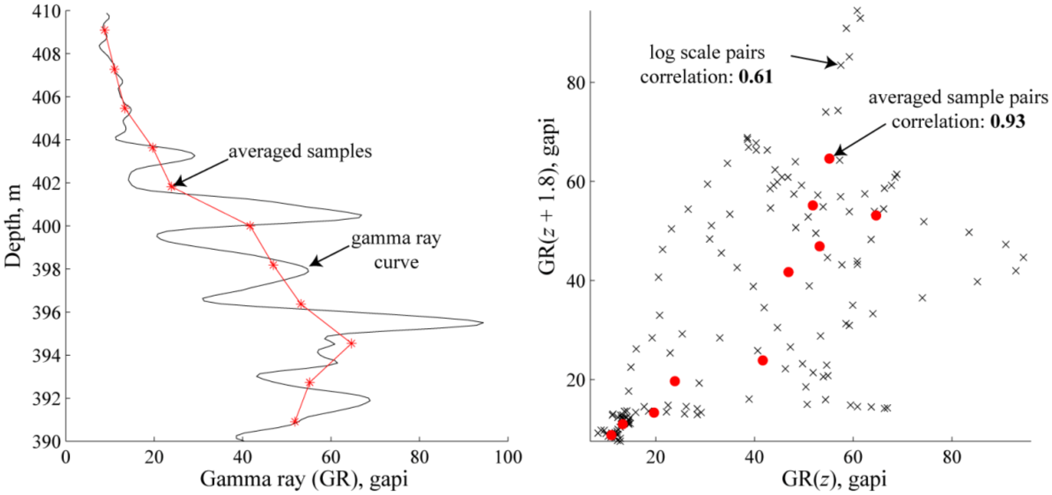


Figure 2-13: Effect of short scale variability on spatial correlation.

Figure 2-13 also highlights a loss of detail as one consequence of upscaling. To limit the loss of detail, a reasonable scale should be no larger than the smallest grid elements used in geological modeling. This choice relates to unstructured grid design, which involves smaller elements in areas of increased importance. In many cases, the smallest elements will occur adjacent to wells. By limiting the reasonable scale to be no larger than these elements, the amount of smoothing or loss of detail in areas of importance is bounded. From here on, the chosen reasonable scale is referred to as point scale. To summarize, define the reasonable point scale as v_p, the scale of hard data as v_ε and the scale of the smallest grid element in a discretization as V_{min} in depositional space. The point scale is chosen according to Equation 2.2 with two goals: to unify hard data to a single scale and avoid excessive loss of detail.

$$v_\varepsilon < v_p \le V_{min} \tag{2.2}$$

To obtain point scale samples, the upscaling approach shown in Figure 2-11 is not advocated. There are inconsistencies in both the position of samples within grid elements and the number of samples within grid elements. A more robust approach is to upscale samples along wells into equal volume intervals independent of the gridding scheme. A concern with this approach is aliasing, that is, the effect of phase shifts on the upscaled properties. Equal interval upscaling was used to compute average values along the gamma ray log in Figure 2-13. It is equivalent to the case where wells are vertical and intersect a column of regular grid elements through the element centers (Figure 2-14). Differences occur when the well does not intersect the grid element centers since it is typical to assign the upscaled samples to the center locations. The proposed method assigns upscaled samples to the center of the averaging volume.

As with upscaling to a structured grid, the intervals for upscaling are not necessarily constant through a reservoir. The chosen scale, v_p, is dependent on post-depositional events and is affected by the transformation from geological space to depositional space. If the chosen point scale defines a length for upscaling log data, then the length must correspond to depositional space. Upscaling to equal intervals in this space accounts for deformation like folding and differential compaction.

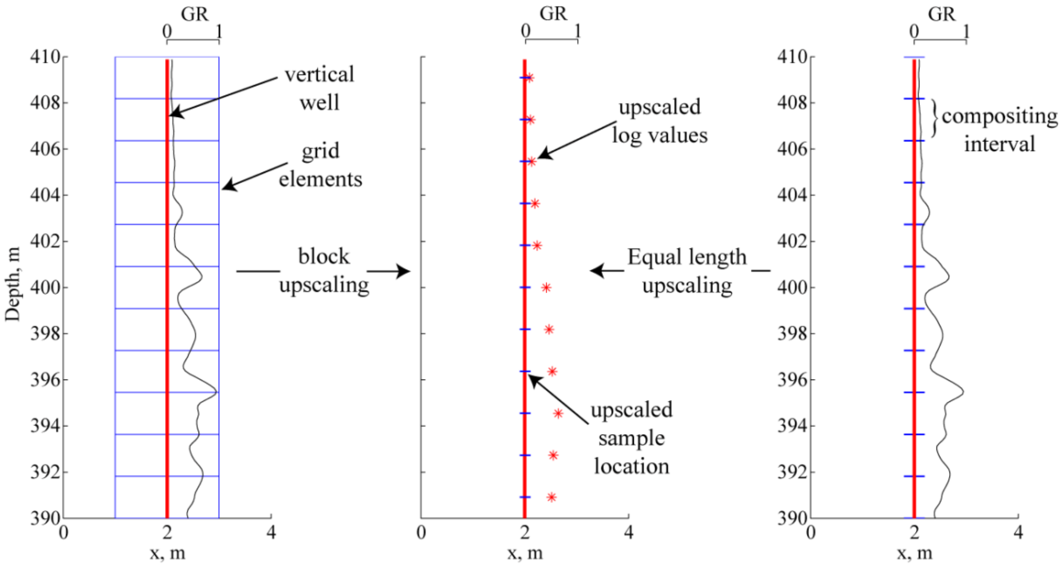


Figure 2-14: Similarity of averaging with vertical wells and a regular grid (left) to upscaling with equal intervals (right). Identical upscaled values are achieved with arithmetical averaging (middle).

A side effect of using point scale data is the scale of larger elements is not precisely accounted for. In geostatistical modeling, the error in assigning point scale estimates to larger scale elements can be measured. For arithmetically averaging variables such as porosity, the error is expressed as the decrease in variance incurred through an increase in scale. It is computed using dispersion variance or average variograms in Equation 2.3, where f_σ is the reduction in variance going from scale v to V, where v, V, and Ω denote the point, grid element, and reservoir volume scales respectively. A variance reduction parameter can be specified, which is translated into an upper bound on element volume. Equation 2.4 provides the average variogram value from which the volume is derived.

$$f_\sigma = \frac{D^2(v,V)}{D^2(v,\Omega)} = \frac{\bar{\gamma}(V,V) - \bar{\gamma}(v,v)}{\bar{\gamma}(\Omega,\Omega) - \bar{\gamma}(v,v)} \qquad 2.3$$

$$\bar{\gamma}(V,V) = f_\sigma \bar{\gamma}(\Omega,\Omega) + (1 - f_\sigma)\bar{\gamma}(v,v) \qquad 2.4$$

Deriving a volume is non-trivial. Several reasons for this are: 1 – each reservoir property has a different variogram, giving different average variograms; 2 – variograms depend on orientation and using simplex discretization results in grid elements with many different orientations; 3 – distances and variogram computations are defined in depositional space whereas the discretization is computed in geological space. The last point leads to a spatially dependent volume in cases where post-depositional and syn-depositional events have altered the volume of reservoir rock as with differential compaction and variable sedimentation rate.

These three problems are handled using the following approaches and assumptions. To deal with multiple variograms, specify a single variogram for grid design purposes so that for any separation vector, $\mathbf{h}$, the variogram for grid design, $\gamma_g(\mathbf{h})$ is not less than all other variogram models, $\gamma_k(\mathbf{h}), k = 1, \dots, m$ variables, that are involved (Equation 2.5). Average variogram values using $\gamma_g(\mathbf{h})$ will always be larger, leading to a volume that in the worst case results in the specified variance reduction parameter.

$$\gamma_g(\mathbf{h}) \geq \gamma_k(\mathbf{h}) \;\forall \mathbf{h}, k = 1, \dots, m \qquad 2.5$$

Deriving the volume from $\gamma_g(\mathbf{h})$ for the second problem and handling orientation effects is possible using an iterative approach like that implemented in Frykman and Deutsch (1999). This is done in depositional space where volumes have not been affected by post-depositional events. In the third problem with spatially dependent volumes, spatial dependence is a consequence of the transformation. In depositional space, the time-space volume is constant. Generating the fine scale grid for geological modeling is still carried out in geological space, since it depends on the geometry of the coarse grid for flow simulation. Element volume calculations are done in depositional space to impose volume constraints on grid generation; however, this requires mapping the elements from geological space. The validity of this process assumes that through the transformation from geological to depositional space, distortions in a grid element's geometry are minimal. This is a reasonable assumption since the scale at which the spatial transformation operates is significantly larger than the scale of grid elements. For example, in Figure 2-7 a triangle with straight edges in geological space remains a triangle with straight edges in depositional space. Using a triangle in the region that has undergone a substantial amount of compaction, a comparison between the actual triangle geometry through the transformation is shown as compared to the straight edge approximation (Figure 2-15). The difference in area is 1.785%, and this is for a relatively extreme synthetic case of compaction and deformation.

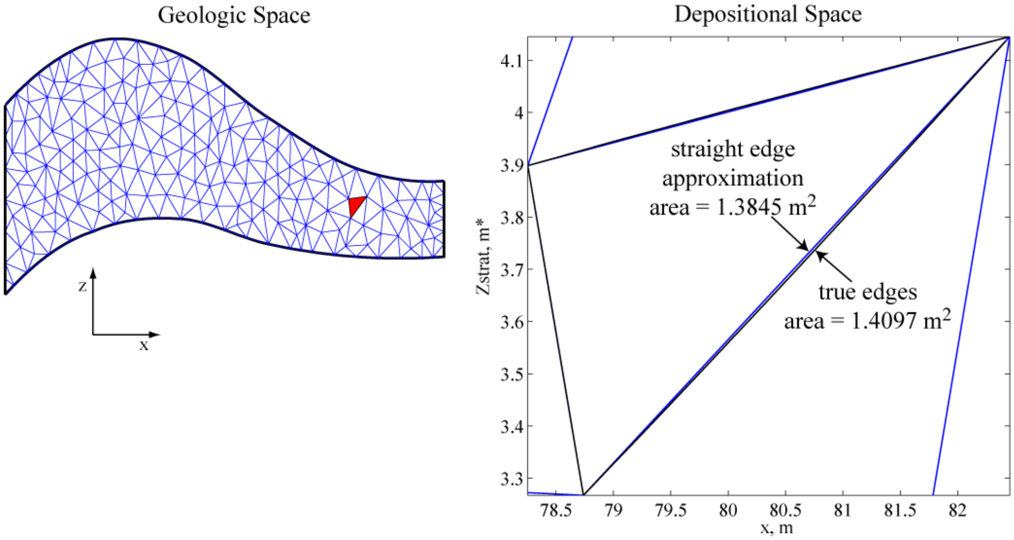


Figure 2-15: Difference in triangle area for the actual geometry through the mapping process and assuming straight edges. The triangle of interest is shown on the left.

More detail on deriving a volume and controlling grid discretization is covered in Chapter 4 due to its relation to upscaling. It is shown that the error incurred in going from point scale to a discretized grid element scale is mitigated through upscaling. Discretization volume constraints are derived from the error in upscaled properties rather than the error from point scale to element scale.

2.2.4 Seismic Data

Other forms of sample data are large relative to geological modeling grid elements. Some typical examples are seismic, well test data, and historical production data. This section focuses on incorporating seismic data into geostatistical modeling on unstructured grids. It is often referred to as secondary data because it is used to guide modeling of reservoir properties (primary data) directly used in flow simulation. Several techniques to account for seismic data in geostatistical reservoir modeling include regression, external drift, locally varying mean, block cokriging, collocated cokriging, and simulated annealing (Xu et al, 1992; Deutsch et al, 1996; Deutsch and Journel, 1998). Most techniques are discussed in terms of enhancing lithology and porosity modeling. For the remainder of this section, it is assumed that seismic has been processed or inverted into a useful attribute such as acoustic impedance (AI) that is correlated with a primary variable (Haas and Dubrule, 1994).

Apart from block cokriging, all methods require the seismic attribute, typically AI, be available at all locations. Existing modeling practices using structured geological grids involve two situations with seismic data: 1 – the seismic attributes are available at a higher resolution laterally and a coarser resolution vertically than the geological grid; 2 – similar to 1, but the seismic attributes are available at the same vertical resolution as the geological grid. The lateral resolution of seismic is typically high with 12.5 by 12.5 meter sample spacing (Liner, 2004). Structured geological models often involve elements that are 16 to 64 times larger in area, requiring upscaling of the seismic attributes to obtain collocated samples. Vertical resolution of seismic is typically larger than the geological grid and is interpolated at points matching the grid element centers.

For unstructured grids, several situations are possible. The resolution of seismic both laterally and vertically may be higher or lower than the size of the unstructured grid elements. However, the point scale chosen for hard data as discussed in the previous section is constant. It is the point scale data that is involved in seismic calibration processes (Deutsch et al, 1996), which uses seismic at its available resolution. Therefore, making point scale estimates on unstructured grids should not involve any change in scale for seismic data, such as block averaging to unstructured elements. Rather, a nearest neighbour re-sampling approach could be used to obtain data for collocated cokriging. Full cokriging could also be used.

2.2.5 Visualization

Visual display of reservoir models is important to their analysis. Geostatistical models should be checked numerically and visually for consistency and correctness. Visual checks can reveal issues with specific components of the workflow, for example: unrealistic intersection of geological structures; coincident samples, which cause a singularity in kriging; and unreasonable grid element geometry. They are also useful for determining if changes to a grid design are necessary, for example if surfaces and faults are represented accurately or if a well has been integrated into the design properly. Tools to accomplish these tasks using unstructured grids are becoming available. To list a few, Paradigm GOCAD has two plugins available for unstructured grids and Eclipse FloGrid from Schlumberger handles unstructured grids as well. Developments are also being made in other areas of research. This dissertation utilizes an open-source visualization application called Paraview (Squillacote, 2008).

Paraview accepts many types of geometric primitives and grid element types (Schroeder et al, 2002). Four built in element types useful for reservoir modeling are quadrilateral, hexahedron, triangle, and tetrahedral elements. Voronoi elements can be defined in two dimensions using the polygon element type; however, in Paraview there is no equivalent in three dimensions. Either a set of polygons can be used to represent the faces of a three dimensional Voronoi element, or a set of tetrahedral elements can be used. One method of defining structured and unstructured grid designs and element properties to visualize in Paraview is through keyword based text files (Kitware, 2006). Table 2-1 provides a sample file that defines the four element types listed previously (Figure 2-16).

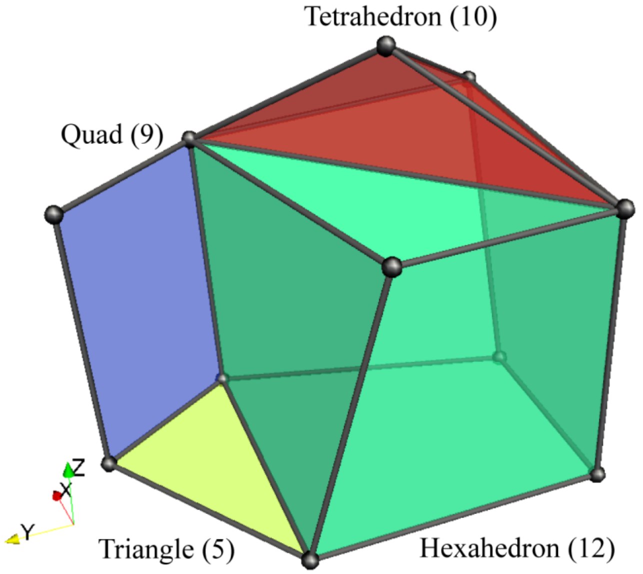


Figure 2-16: Illustrative example of four element types visualized in Paraview. Elements are numbered according to their type.

Table 2-1: Example Paraview input file for Figure 2-16.

Line	Text	Line	Text
1	`# vtk DataFile Version 3.0`	18	`CELLS 4 23`
2	`Figure 2-11`	19	`8 0 1 2 3 4 5 6 7`
3	`ASCII`	20	`4 4 5 6 8`
4	`DATASET UNSTRUCTURED_GRID`	21	`4 9 2 6 10`
5	`POINTS  11 float`	22	`3 9 2 3`
6	`0.000  0.000  0.000`	23	
7	`7.000  0.000  0.000`	24	`CELL_TYPES 4`
8	`10.000 10.000  0.000`	25	`12`
9	`0.000 10.000  0.000`	26	`10`
10	`0.000  0.000 10.000`	27	`9`
11	`10.000  0.000 10.000`	28	`5`
12	`10.000 10.000 10.000`	29	
13	`0.000  7.000 10.000`	30	`CELL_DATA  4`
14	`5.000  5.000 14.000`	31	`SCALARS z float 1`
15	`7.000 15.000  0.000`	32	`LOOKUP_TABLE default`
16	`7.000 15.000 10.000`	33	`2`
17		34	`4`
		35	`1`
		36	`3`

2.3 Grid Design

Two grids are involved in reservoir analysis; a fine grid for geostatistical modeling and a coarse grid for flow simulation. Consistency between the two grids is important for up and down scaling between grids. Fine grids are designed to honour major geological features such as faults and surfaces and provide adequate detail for upscaling. Flow simulation grids or coarse grids have the following characteristics: they are solvable by existing simulators; they honour major geological features, wells and other boundary conditions; and minimize error that occurs due to approximating the flow with a discrete system of elements. Additional goals are needed for geological grids to ensure a seamless transition to flow simulation grids. These goals were identified in Chapter 1 and are reiterated here:

1. The coarse and fine grids are both amenable to existing flow simulation algorithms.
2. The accuracy of upscaling processes is consistent for all coarse elements.
3. Fine elements do not overlap one another and are fully embedded in coarse elements.
4. The interfaces of fine elements align with coarse element interfaces.

These are not part of current geological grid design that involves structured grids in geological space. Structured grids often have non-orthogonal features (Figure 2-17), which is contradictory to the first goal, that is, such features are not ideal for the TPFA. Non-orthogonal geometry and partial element connections cause grid orientation effects with existing commercial flow simulators that use the TPFA method (Aziz, 1993; Aarnes et al, 2007). Other problems with structured grids for flow simulators are summarized by Farmer (2005). Problems are avoided by using different geological grid designs such as simplex grids, and targeting different flow simulation algorithms for upscaling and full field simulation.

This section covers types of inputs that contribute to grid design ranging from choice of flow simulator to seismic, log, and core data. The influence each input has on the design is discussed with some examples for demonstration purposes.

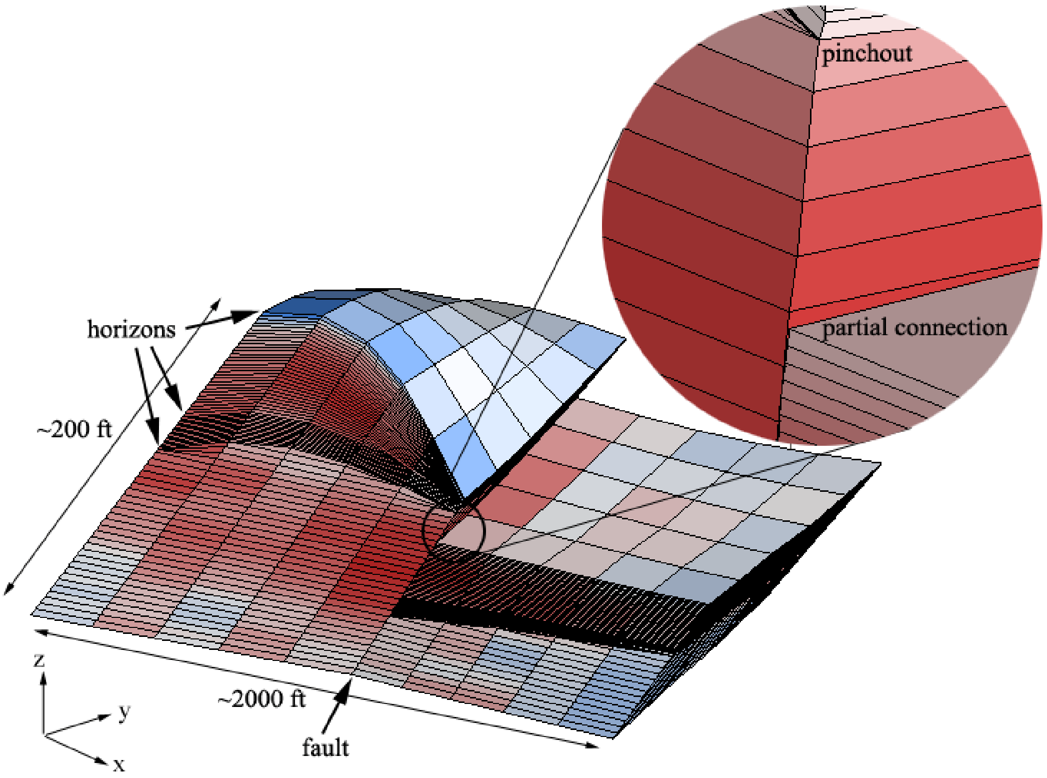


Figure 2-17: Structured grid aligned with horizons and faults. Not to scale.

2.3.1 Inputs

An abundance of data sources exist for reservoir analysis and can be used for grid design. Several variables provide information about the structure, heterogeneity, and conditions of a reservoir, the most common ones are listed in Table 2-2 and their impact on grid design is summarized. Variables in Table 2-2 are organized in an order they might be incorporated for grid generation:

1. Determine the grid type.
2. Define boundaries.
3. Define interior geometry and detail.

The following sections cover grid design with simulator choice, geological structure, wells, and hard and soft data.

Table 2-2: Variables and their influence on grid design

Variable	Influence
Simulator choice	Determines how the flow equations and grid are discretized to achieve convergence to the correct solution. For example, use of the TPFA technique limits grids to regular or locally orthogonal, whereas use of the MPFA method permits general polygonal or tetrahedral grids.
Grid purpose	The type of production the grid is intended to simulate will influence grid parameters such as the type of elements to use and their volume and resolution. For example, a large scale application may use a fairly coarse PEBI grid, whereas a smaller SAGD application may use a high resolution structured grid.
Geological structure	Horizons, faults, geologic objects, and boundaries that are expected to influence flow should be incorporated into grid design. Flow across or along these structures is captured more accurately and it is possible to incorporate additional information about the surfaces. For example, faults may be conductive, transmissive, or sealed and impermeable. Structural modeling is also important for mapping from geological space to depositional space.
Existing wells	Wells involved in the production process, such as those used for injection or production, influence the grid. Unstructured grids commonly incorporate dense radial grids around wells to capture radial flow there. Well trajectories and perforated intervals can also be used to constrain near well discretization.
Future wells	If future well sites for production are known in advance, they can be incorporated into the grid in a similar manner to existing wells.
Single well test data	Drawdown and buildup tests provide information about the effective permeability within the vicinity of a well, which can be used to condition the near-well permeability field (Wen et al, 2005). They also provide information about the area of influence of the test, which can be used for controlling radial grid design.
Multiple well test data	Interference tests and tracer data can provide information about the occurrence and flow character of faults, effective permeability, connectivity between wells, and pressure distributions. This data can be used to characterize surfaces with flow parameters such as transmissibility multipliers, to condition permeability fields between wells based on connectivity, and to provide a rough idea of pressure gradients and streamlines for flow based gridding.
Historical production data	Provides similar information as multiple well test data. Additional information includes well drainage volume and well communication.
Seismic data	Helps identification and modeling of structures and influences geological property modeling, which can control grid resolution.
Well log and core data	Provide information about distributions and variography of facies, porosity, and permeability. Preliminary mapping of these variables helps to identify the spatial distribution and level of uncertainty which can be tied to grid resolution. Permeability models can also be used in flow based gridding.

2.3.2 Simulator Choice

Selecting the type of grid to use is dependent on the simulator, the application, the capability of the model, and on knowledge about the reservoir including the depositional environment and fluids involved. In Chapter 1, several simulation techniques were discussed along with grids they can handle; however, flow simulators have additional limitations that impact grid design. Simulators using TPFA should be limited to regular or locally orthogonal grids and diagonal tensor permeability (Chen et al, 2007). More flexibility in design is possible with MPFA simulators. Hexahedral, polygonal, and simplex grids with full tensor permeability are possible. However, finding a solution for the pressure field is sensitive to the magnitude and anisotropy of permeability. In extreme cases, it can be shown that the system matrix for MPFA is not an M-matrix, resulting in non-physical pressure oscillations (Aavatsmark et al, 1998; Eigstad and Klausen, 2005; Mlacnik and Durlofsky, 2006). Oscillations that reduce pressure can lead to dissolution of gasses where this would normally not occur. If it is possible that the environment and fluids will lead to these conditions the MPFA method may not be appropriate, or care must be taken in subsequent grid design steps. For example, grid optimization can be applied to improve the quality of the solution (Mlacnik and Durlofsky, 2006).

2.3.3 Grid Purpose

Grid selection is also influenced by the targeted application and scale of the problem. Conventional production processes may be effectively discretized using coarse hybrid grids. If little structure is involved, a regular grid may suffice. Enhanced recovery processes (Donaldson et al, 1989) such as immiscible and miscible displacement may require more accurate representation of structure and increased resolution to better describe the geometry of the frontal and interface regions (Sammon, 2003). Different discretizations may be required for thermal applications such as cyclic steam injection and SAGD. Describing the distribution of temperature and viscosity and simulating the gravity drainage process demands a higher resolution model than conventional production for example (Christensen et al, 2004; Skoreyko, 2003).

2.3.4 Geological Structure

In existing workflows and the proposed workflow suggested in Section 2.1, structural modeling is carried out prior to any gridding. Structures can be used to constrain resulting discretization. In many cases, incorporating faults and stratigraphic surfaces into a structured grid cannot be accomplished while maintaining orthogonality resulting in distorted grid designs and partial element connections (Figure 2-17). If orthogonality is a condition of grid design, then a more appropriate discretization is a PEBI grid (Figure 2-18).

Other grid choices like tetrahedral grids can make the inclusion of structure more straightforward. Specifying a point distribution defines the vertices and faces of tetrahedral elements directly, so vertices can be placed directly on the surfaces thereby reproducing them (Figure 2-18). Since the set of element faces along surfaces form a triangulation, it would be logical to use triangulations for initial structural modeling. They can be reproduced exactly by the final tetrahedralization. Constrained tetrahedral grid generation is well developed for a variety of problems, with the majority of research targeting grid optimization (Du and Wang, 2006).

Structural surfaces are demonstrated in an example involving four horizons, which form three stratigraphic layers, and two faults (Figure 2-19). The area of interest covers a 1,000 by 500 meter area and is roughly 25 meters thick. Upper and lower layers are permeable while the middle layer is assumed to be a very low permeability shale barrier. It is assumed that only folding and faulting of the layers has occurred, in that order, and stratigraphic coordinates are satisfactory. Mapping is a two stage process: first the fault block is translated to align the horizons, then the layers are flattened using stratigraphic coordinates.

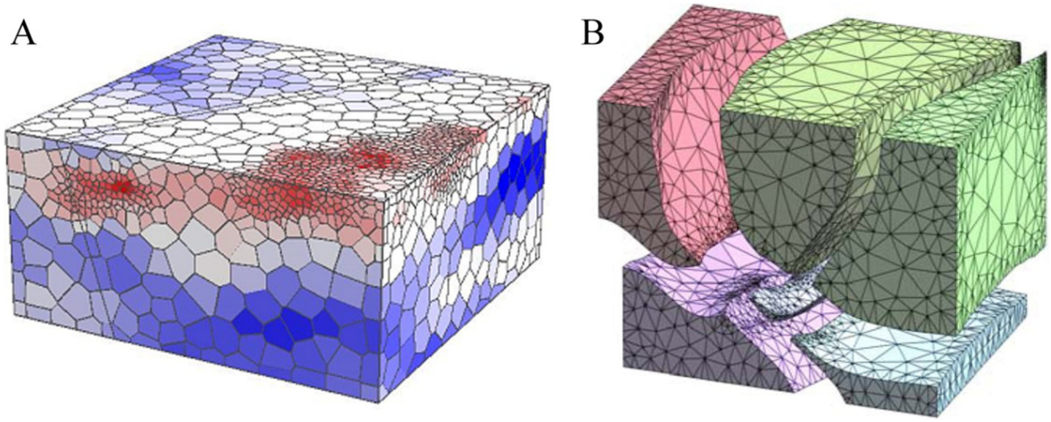


Figure 2-18: Example grids from Chapter 1: A – PEBI grid with fault, blue to red is low to high velocity (Prevost et al, 2005); B – Tetrahedral grid with multiple intersecting faults (Prevost et al, 2005).

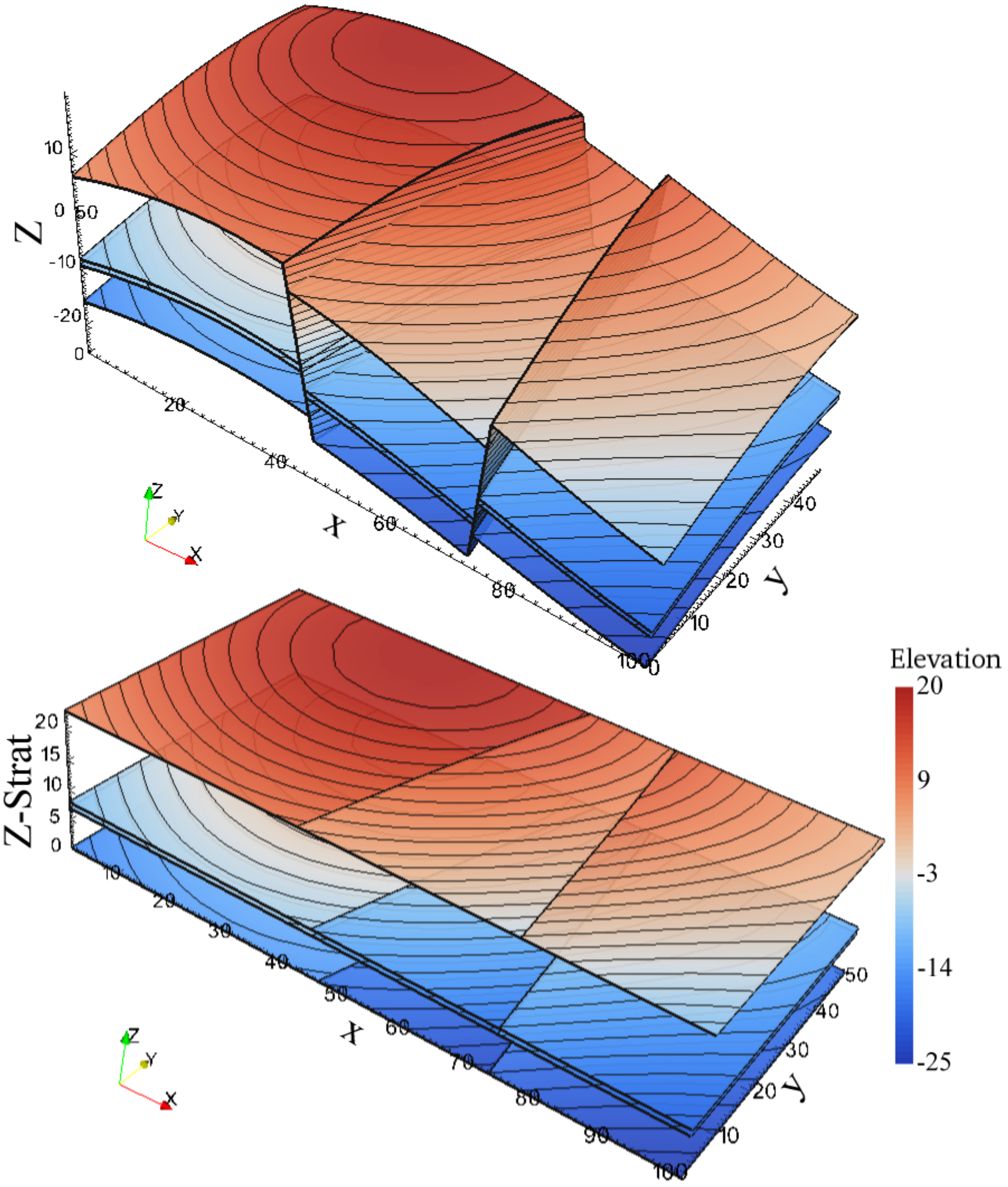


Figure 2-19: Isometric view of structures for grid design example (top); unfaulted and flattened surfaces in depositional space (bottom). The Z-Strat axis has been scaled to reflect the average thickness of each layer. Lateral coordinate are in 10's of meters and elevation is relative to sea level in meters.

2.3.5 Wells

Existing wells and proposed future wells that are involved in the production process should be included in the grid design. This is often accomplished using high resolution radial grids centered along the well path. Increased resolution can provide higher accuracy of resulting pressure and flow gradients. Several methods have been developed to incorporate wells into a grid design (Palagi and Aziz, 1994; Flandrin et al, 2006). They work under the assumption that a grid design for a reservoir already exists without any wells: first elements are deleted within the vicinity of wells; then new elements are reconstructed to form radial or other grid types and stitched to the existing grid. Some work has also been done on generating tetrahedral grids down to the scale of well perforations (Karimi-Fard and Durlofsky, 2009).

Continuing with the example from the previous section, the model is also going to test an injection-production well pair. It is also assumed that the available simulator uses MPFA so tetrahedral grids are possible. Using tetrahedral grids for both the coarse and fine scale achieves the four goals listed at the beginning of Section 2.3. The locations of the production and injection wells for the example are shown in Figure 2-20. Injection is proposed into the lower layer and production from the upper layer. Using simplex grids around wells is not exactly like the radial grids shown in Figure 2-2; however, for the purpose of capturing the radial flow behaviour near wells, simplex grids with increasing resolution towards wells provide the necessary accuracy. Addition of a radial grid is shown for the producing well (Figure 2-21 and Figure 2-22). The grid was given a radius of 70 meters and discretized using 7 radial increments that grow quadratically (the spacing from one radius to the next is a quadratic function as opposed to linear) and 11 tangential increments.

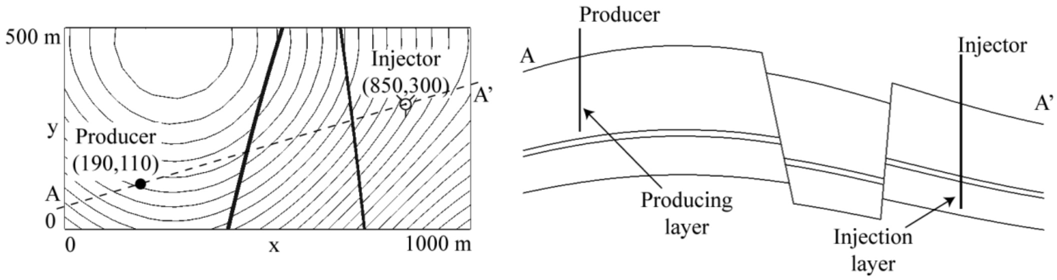


Figure 2-20: Locations of injection and production wells.

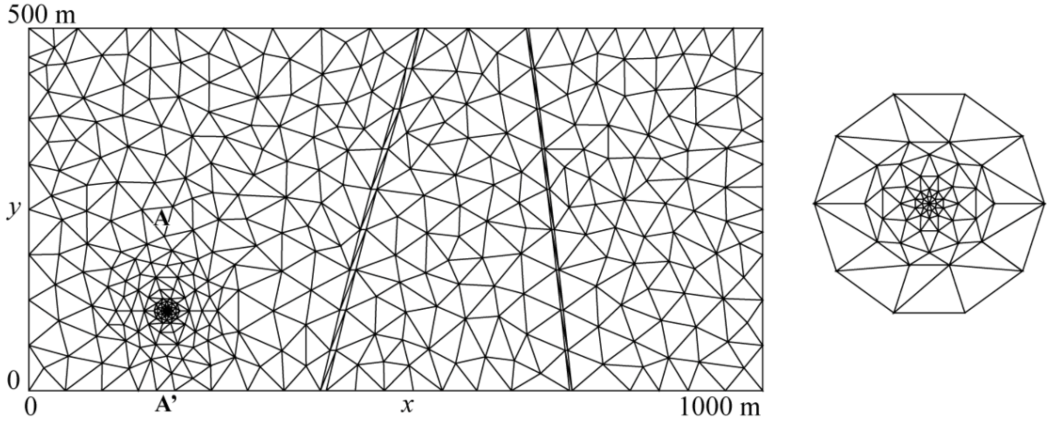


Figure 2-21: Triangular radial grid at the producer location (left) and close-up (right). A Cross section from A to A' is shown in Figure 2-22.

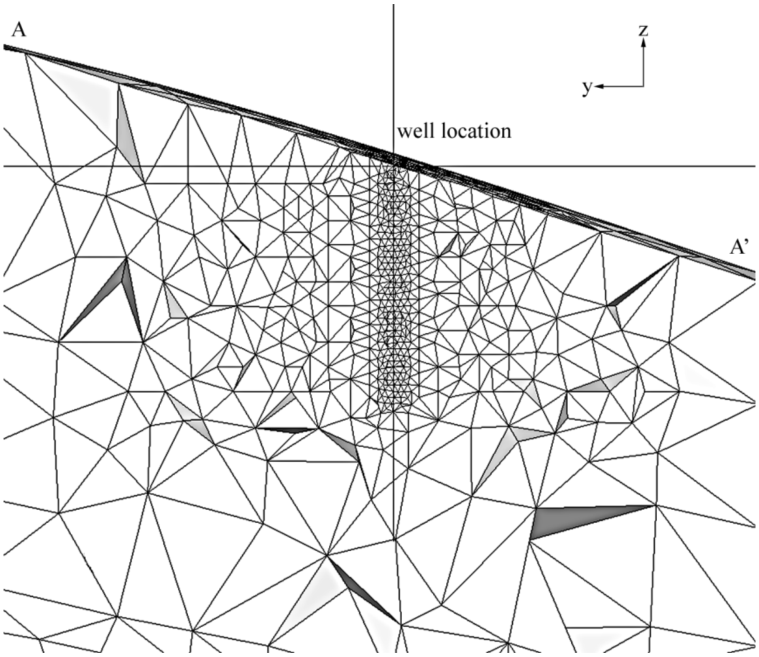


Figure 2-22: Cross section showing tetrahedral radial grid along a well.

2.3.6 Seismic, Well Log and Core data

The spatial distribution of properties can be used to control the local resolution of grid design. Reservoir properties are used as soft constraints on grid design, unlike reservoir structures that are used as discrete constraints. Reservoir structures specify exactly where grid element interfaces must lie; whereas spatial distributions can provide control in several ways: to specify limits on the variability within unstructured grid elements; be converted to a spatial distribution of element volumes to control grid generation; and in the case of permeability, be used in flow based grid design.

Use of property distributions requires some degree of preliminary mapping. It provides an idea of the spatial distribution of variables that correlate with flow such as facies, porosity, and permeability. In geostatistics, mapping is accomplished with one of several forms of kriging, which provides an estimate of the expected value of a property, as well as a homoscedastic estimation variance that serves as a measure of local uncertainty. Integrating the estimation variance into grid generation is not appropriate for highly skewed variables such as permeability. It is possible to grossly underestimate the variance in areas where high permeability is expected, and overestimate it in areas where low permeability is expected, leading to a suboptimal grid. The phenomenon for such distributions is known as the proportional effect (David, 1977; Manchuk et al, 2009), and various techniques are available to obtain a more appropriate estimation variance including distribution mapping (Oz et al, 2003), interpolation variance calculation (Yamamoto, 2000), and indicator kriging methods (Goovaerts, 1994).

Another important aspect to consider during preliminary mapping is data dependence. Relationships may exist between permeability, porosity, and facies for example. Such information is analysed and incorporated into modeling during data preparation. It is best to use the same modeling techniques and parameters for preliminary mapping that will be used in generating the final models. Using numerical models of reservoir properties to control unstructured grid design is discussed by several authors (Pirzadeh, 1993; Yadecuri and Mahani, 2009). Other references to flow based grid design are found in Chapter 1.

Preliminary mapping using geostatistics has two possible outcomes: 1 – a single model that gives local distributions of uncertainty for each property (kriging); 2 – a set of models or realizations of each property that are equally probable (simulation). Using kriging results, or similarly the expectation of a set of realizations, to control grid design gives one design. Design of this type is referred to as designing in expectation, and the objective is to generate a grid to account for a set of realizations, rather than for a specific realization. Using simulation results for grid design gives a different design for each realization. This could be very tedious unless a fully automatic grid generation system is available.

Using the same stratigraphic layout and well locations as in Figure 2-19 and Figure 2-20, preliminary mapping on a regular grid is done to provide a synthetic permeability model that is used for grid generation. The regular grid was arbitrarily chosen in depositional space with the number of elements in *x*, *y*, and *z* of 39, 19, and 49 respectively (Figure 2-23).

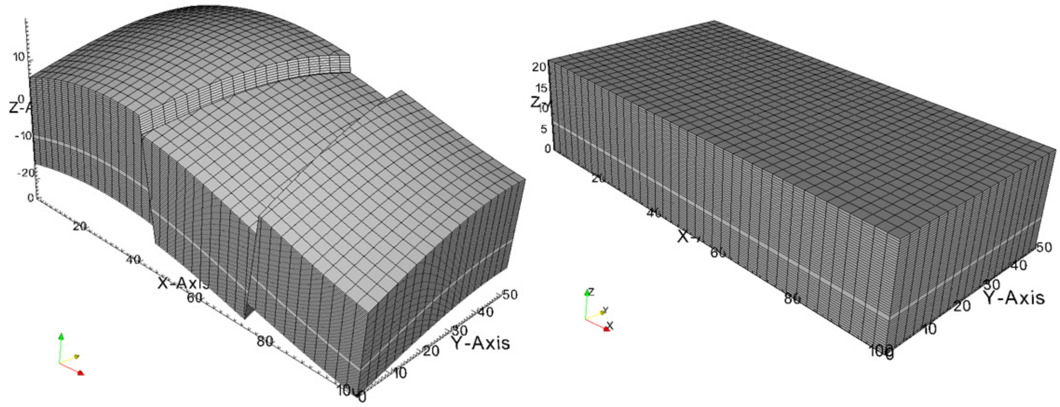


Figure 2-23: Regular grid for preliminary mapping in geological space (left) and depositional space (right). Units are in 10's of meters.

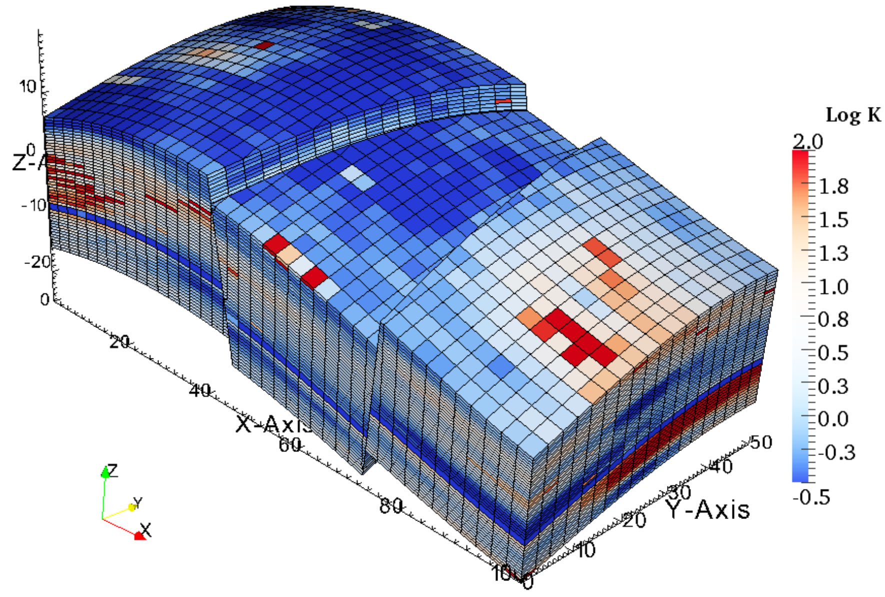


Figure 2-24: Synthetic permeability model from preliminary mapping. Units are in 10's of meters.

Typical geostatistical methods were used to generate several realizations that were averaged to provide an expected permeability model (Figure 2-24). The full range of log(k) is -9 to 4. This is used to control grid generation in two ways: 1 – element volume is related inversely to the magnitude of permeability; 2 – element volume is controlled by flow velocity in a flow-based gridding approach.

Generating a tetrahedral grid using TetGen with permeability based volume control is done in an iterative fashion. One of the command line arguments for TetGen is the 'r' switch that reconstructs or refines a grid that has already been generated. Several refinement modes are possible, but the one of interest involves creating a text file of volume constraints that defines the maximum volume for each tetrahedron. Using the 'r' switch along with 'a' tells TetGen to read in the volume file and apply the constraints. This capability is used as follows:

1. Convert the permeability model into a volume model. A high contrast volume model is used in the example (Equation 2.6). The large volume is set high so that no refinement is applied for low permeability.

$$Volume(\mathbf{k}) = \begin{cases} 1\times10^{5}\,m^{3} & \log(\mathbf{k}) < 0 \\ 500\,m^{3} & \log(\mathbf{k}) \geq 0 \end{cases} \qquad 2.6$$

2. Generate an initial coarse tetrahedral grid with the 'p' switch, possibly with a quality constraint or global maximum volume constraint ('pq' or 'pqa'). For the example, 'pq1.4' was used that imposes the quality radius-edge ratio of 1.4.
3. Find the nearest point in the volume model to the tetrahedra barycenters and generate a volume file.

4. Refine the mesh from 2 using ‘ra’ and repeat step 3. Iterate until the grid is desirable or does not violate any volume constraints. In the example, ‘rq1.4a’ is used to maintain the quality.

In the model, there is also the very low permeability layer where no flow is expected to occur. This can be eliminated from the tetrahedral grid by identifying the layer as a hole for TetGen. Holes are specified as three dimensional points falling inside an enclosed portion of the piecewise linear complex (Miller et al, 1996; Si, 2006) used to define the reservoir model. All tetrahedra within the enclosed region specified as a hole are eliminated from the grid.

After three iterations the grid has 96,915 tetrahedral elements (Figure 2-25). Volumes range from less than 0.02 m^3 in the central portion of the radial grids to roughly 80 m^3 in regions with low permeability. In comparison with Figure 2-24, the grid has small elements where log(**k**) is above one as well as near wells and large elements elsewhere. Also notice the low permeability layer is not included in the grid.

One of the motivations for going to flow based gridding is shown in the permeability based grid. There are areas of the reservoir that have been gridded with a high resolution even though no flow is expected to occur there. The cross section in Figure 2-26, which is oriented exactly as that in Figure 2-20 through the well locations, indicates areas where low flow is expected to occur based on the injection and production sites. It may be possible to grid these regions very coarsely or omit them from the model completely resulting in fewer elements and more efficient flow simulation. One possible workflow for triangular or tetrahedral flow based gridding involves the following steps:

1. Generate an initial coarse tetrahedral grid for the reservoir. It may include surfaces, faults, and radial grids around wells.
2. Use nearest neighbour interpolation to assign permeability from preliminary mapping to the unstructured grid element centers.
3. Build and solve the flow equations for pressure and flux for a given set of boundary conditions.
4. Refine the grid based on one of several measures such as pressure gradient, flow velocity, or streamline density. Reiterate from step 2.

Refining a grid with this approach leads to a grid with smaller elements only where flow occurs. However, the grid is only valid or optimal for the boundary conditions used for its design. Adding a well or changing boundary conditions will result in a different grid.

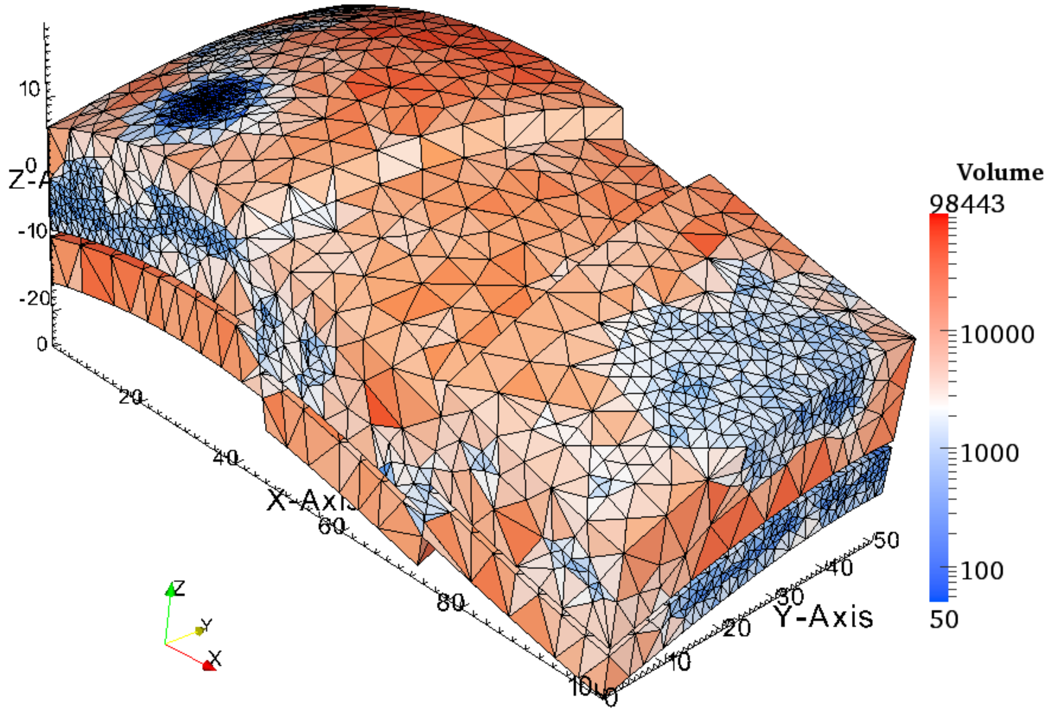


Figure 2-25: Tetrahedral grid with permeability based element volume control. Axes in 10's of meters; volume in m^3.

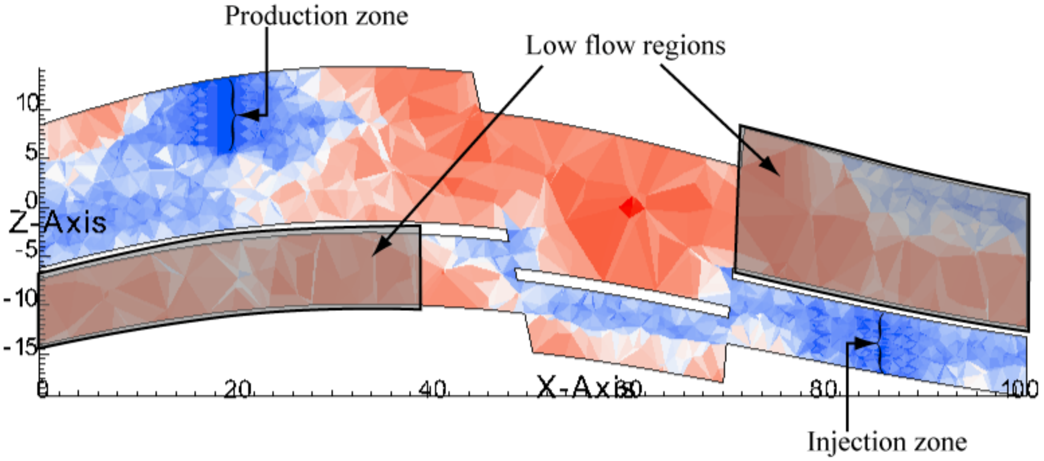


Figure 2-26: Cross section through wells showing areas of low expected flow. Axes in 10's of meters.

2.4 Summary

This chapter discussed a workflow that incorporates unstructured grids into reservoir modeling and flow simulation. It included the following steps:

1. Data preparation
2. Structural modeling
3. Grid design
4. Grid discretization
5. Facies modeling
6. Continuous property modeling
7. Model validation
8. Upscaling
9. Flow simulation

Some of the key components were covered. Discretization is used to increase the resolution of coarse unstructured flow simulation grids for geostatistical property modeling and upscaling. Spaces involved in reservoir modeling including geological and depositional space were covered and it was shown that existing mapping techniques that apply to regular grids also apply to unstructured grids. A discussion on data scale highlighted the scales involved in reservoir modeling and criteria to choose a reasonable scale to use. A reasonable scale is chosen to be practical and minimize loss of detail. Some issues with existing methods of upscaling sample data from core and log scale to the geological modeling scale were pointed out and a different method was suggested whereby upscaled values are not assigned to grid elements, but to the center of the upscaling volume. Incorporating larger scale data such as seismic was also discussed. Existing techniques for this are applicable to unstructured grids.

Visualization is a very important component to reservoir modeling. It facilitates the detection of errors, workflow issues, and areas for improvement or adjustment. Some available software for this was identified. Visualization is also important for grid design that was covered in moderate detail.

Chapter 3 Sequential Simulation

This chapter focuses on using geostatistics to populate a discretized unstructured grid with reservoir properties. It develops a sequential simulation algorithm that populates grids with categorical variables such as facies and continuous variables such as porosity and permeability. The general workflow of the algorithms is described in Section 3.1. Some of the input parameters required for the algorithm are covered in Section 3.2. Section 3.3 develops components of the workflow that require a different approach than regular grids including defining a random path and searching for data. Section 3.4 discusses model validation including histogram reproduction and variogram reproduction.

3.1 Sequential Simulation

Sequential indicator simulation (SIS) and sequential Gaussian simulation (SGS) are two widely used geostatistical algorithms for generating reservoir models (Deutsch and Journel, 1998; Deutsch, 2002). They are typically used to generate a set of equally probable models called realizations that reproduce input distribution functions, heterogeneity, and patterns of spatial continuity. Each realization defines the spatial distribution of all properties of interest such as facies, porosity, and permeability. The set of realizations are used to quantify the uncertainty in other analyses such as flow simulation.

Simple kriging is the estimation method used in SIS and SGS algorithms. Given a set of conditioning data, the kriging estimate and estimation variance, also referred to as the conditional mean and conditional variance respectively, define a distribution function. Assuming the random variable is Gaussian, the conditional mean and variance from kriging define a Gaussian distribution function. The distributions are used with Monte Carlo simulation to generate realizations. Several variants of kriging exist and are discussed in Chiles and Delfiner (1999), Isaaks and Srivastava (1989), and others.

Simple kriging is a linear estimator that minimizes the error variance between the true value and the estimate based on the covariance function of the property. The estimate, $\hat{y}(\mathbf{u})$, is a linear combination of n conditioning data, $y(\mathbf{u}_i)$, given by Equation 3.1, where λ_i are the kriging weights.

$$\hat{y}(\mathbf{u}) = \sum_{i=1}^{n} \lambda_i y(\mathbf{u}_i) = \boldsymbol{\lambda}^T \mathbf{y} \qquad 3.1$$

Minimizing the estimation variance defined by Equation 3.2 leads to the system of equations in Equation 3.3, where σ_y^2 is the variance of y and $C(\mathbf{u}_i, \mathbf{u}_j)$ is the covariance between locations $\mathbf{u}_i$ and $\mathbf{u}_j$ calculated based on the variogram of y. This forms a linear system of equations in Equation 3.4, where $\mathbf{C}$ is the covariance matrix between conditioning data and **c** is vector of covariance between conditioning data and the estimate.

$$\sigma_{SK}^2(\mathbf{u}) = \sigma_y^2 + \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i C(\mathbf{u}_i, \mathbf{u}_j)\lambda_j - 2\sum_{i=1}^{n}\lambda_i C(\mathbf{u}, \mathbf{u}_i) \quad 3.2$$

$$\sum_{i=1}^{n} C(\mathbf{u}_i, \mathbf{u}_j)\lambda_j = C(\mathbf{u}, \mathbf{u}_j),\ j = 1, \ldots, n \quad 3.3$$

$$\mathbf{C}\boldsymbol{\lambda} = \mathbf{c} \quad 3.4$$

The kriging equations are applicable to any data configuration. Kriging and simulation algorithms using kriging are therefore applicable to discretized unstructured grids.

A general workflow for building reservoir models using SIS and SGS is described in the following list. It is executed within a domain that is assumed to be first and second order stationary (Journel and Huijbregts, 1978; McLennan, 2007). These are typically the individual stratigraphic layers within a reservoir. Sample data, referred to as conditioning data for geostatistical modeling, is used to derive input parameters for the geostatistical simulation algorithms. A grid specification is also required to define the spatial locations where reservoir properties are modeled. Once all parameters are defined, sequential simulation proceeds as follows:

1. Data transformation. Facies are transformed to indicators for SIS. Porosity and permeability are transformed to a standard normal distribution for SGS.
2. Define a random path to traverse through the grid elements.
3. For each element:
    a. Search for conditioning data and previously simulated values.
    b. Perform kriging of facies proportions or the conditional mean and variance.
    c. Draw from the conditional distribution and add to the pool of previously simulated values.
4. Back transformation. Indicators are converted to categories and normal values to their corresponding input distributions.

Resulting models are validated to ensure input statistics are reproduced. Two common checks are histogram reproduction and variogram reproduction. Histogram reproduction compares the histogram of the realizations with the histogram of the conditioning data. Variogram reproduction compares the realization variograms with the input variogram. When multivariate data are involved, the correlations between properties are also checked to ensure the correlation of the realization match those of the conditioning data.

Applying the sequential simulation algorithm to discretized unstructured grids is done using the centers of the grid elements. Centers form an irregular configuration of points throughout the reservoir in geological space. The configuration remains irregular in depositional space. When structured grids are used, the mapping results in a regular gridded set of points. There are no changes to the algorithm for irregular sets of points, but components requiring some attention are: the input histogram; generating a random path; searching for nearby conditioning data; and validating models. A sequential simulation program called psgsim was developed for this and is described in the Appendix.

3.2 Input Parameters

Geostatistical simulation involves several parameters that are derived from hard and soft conditioning data including: trends and other non-stationary behaviour such as locally varying moments; univariate and multivariate distribution functions; correlation coefficients; and variograms. Several general books on geostatistics describe these parameters (Goovaerts, 1997; Deutsch, 2002; Wackernagel, 2003; Webster and Oliver, 2007). Statistical parameters explained in this section include the histogram and variogram. Another input parameter for simulation is a grid. In existing sequential simulation algorithms, the grid is regular and is fully defined by 9 parameters: in each coordinate direction, the number of grid elements, the origin of the grid, and the element extent are defined. The algorithm developed in this chapter is intended to work with any spatial configuration of points. When the algorithm is applied to an unstructured grid discretization, the points are the centers of the grid elements. The algorithm can be applied to any configuration of points.

3.2.1 Histogram and Variogram

The histogram and variogram derived from conditioning data depend on the scale of the data and on the space where the data resides. The scale of data for modeling a discretized unstructured grid is equal to the scale of conditioning data as discussed in Chapter 2, where the scale was referred to as point scale. The histogram and variogram represent the same point scale. These parameters also depend on the space in cases where the mapping from geological space to depositional space is non-linear. In a non-linear transformation, either the relative position of points change affecting the variogram, or the volume each point represents changes affecting the histogram. The variogram is affected by changes in relative position because this changes the distance. An example where relative position is changed, but not volume, is through reversing the displacement caused by a fault (Figure 3-1). Experimental variograms and variogram modeling makes sense in depositional space where Euclidean distances provide the correct correlation. An example where volume is changed but not relative position is in undoing the differential compaction of Figure 3-2. If the point configuration were not central in the z-axis as shown, the relative point position would also change. Transformations will usually change both relative position and volume.

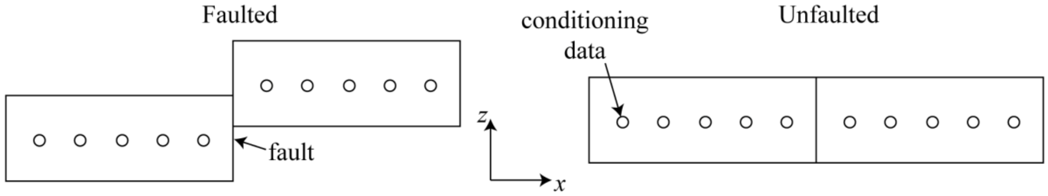


Figure 3-1: Change in relative point position from undoing a fault.

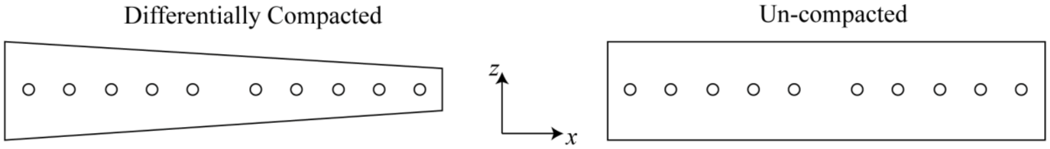


Figure 3-2: Change in volume from undoing differential compaction.

Changing the volume through a spatial transformation can lead to a different histogram in each space. Consider the example in Figure 3-3 involving a hypothetical porosity model with a trend in the horizontal direction. The hypothetical porosity model was created by linearly varying porosity along the *x* axis and adding some uniform random noise. Values were scaled to an arbitrary range of 0.07 to 0.44. In geological space, the histogram has a lower mean and is negatively skewed whereas in depositional space the histogram is roughly uniform with zero skew. If the trend is removed from the porosity field, the distribution of the uniform random values in each space is nearly identical; however, the trend may not be so evident in all cases. Also, the effects of spatial transformation on the histogram may not be clear because we rarely have exhaustive knowledge of the reservoir properties.

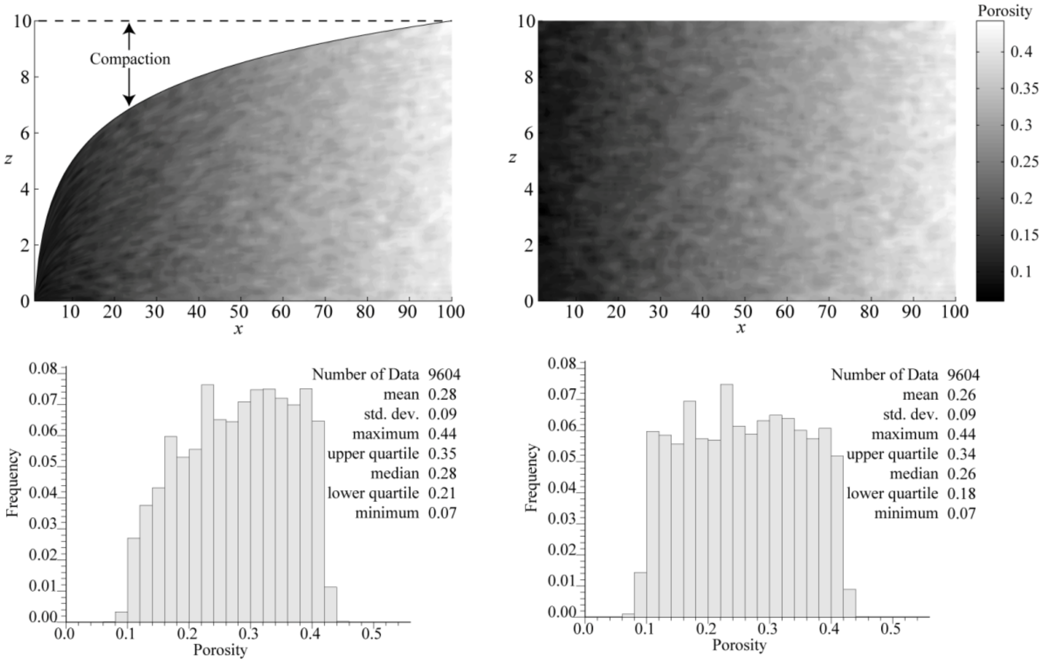


Figure 3-3: Effect on the histogram of transforming differential compaction (left) to depositional space (right).

The final histogram of interest is that of the reservoir properties in geological space. Computations such as hydrocarbon volume require the distribution of porosity and oil saturation in geological space. The distribution of permeability is required in geological space for flow simulation. From Figure 3-3, the histogram on the left is an example of the resulting porosity distribution that a realization from SGS should reproduce; however, using this distribution for input will not give these results using structured or unstructured grids. In the structured grid case, the grid is mapped to depositional space where it is regular and defines the scale and variogram space for SGS (Figure 3-4). Even though the realization produces the input distribution, it is at the scale of the grid in depositional space. Computing the distribution in geological space requires the values of the realization be weighted by the grid element volumes in geological space. Weighting was already applied to obtain the input distribution. Applying weighting again will lead to the wrong porosity distribution and in this case inflation of the mean due to the negative skew. For SGS to reproduce the correct porosity histogram in geological space, the input histogram should be the depositional (Figure 3-3, right).

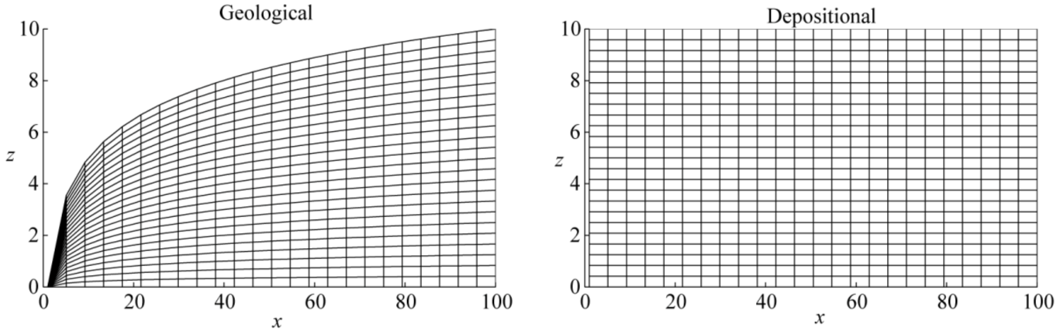


Figure 3-4: Structured grid associated with Figure 3-3 mapped to depositional space.

In almost all cases the histogram of a reservoir property is not known and it must be estimated from available conditioning data. When a reservoir is not sampled uniformly, the data are often declustered to estimate the true distribution (Journal, 1983; Deutsch, 1989). Declustering should be applied in depositional space where the distribution is required. Moreover, this has the advantage that post-depositional effects do not influence the declustering process.

3.2.2 Grid Specification

Grids define a set of spatial locations that are populated using geostatistical methods. In existing techniques, reservoirs are modeled using structured grids and the set of locations is the grid element centers. The transformation from geological space to depositional space results in a regular grid for modeling. Using a discretized unstructured grid is similar. The centers of the elements define a set of locations that are mapped to depositional space; however, the configuration of points is irregular. In this thesis, unstructured grids are always discretized using a simplex grid, but other discretizations are possible so this section is kept general.

Not all unstructured grids are defined using element centers, for example, in the finite element method properties can be assigned to the element vertices. Another example involves PEBI grids, where the set of points that define its dual grid are not always PEBI element centers (Figure 3-5). This dissertation focuses on populating points that are inside the corresponding grid elements. In the following, it is found that using element centroids is optimal.

Each grid element also has a volume that is different than the scale of conditioning data. Kriging estimates based on the scale of conditioning data have a variance with the same scale, not a variance representative of the grid element scale. The associated error was discussed in Chapter 2 and for an unstructured grid discretization will not exceed the variance reduction parameter f_σ. Another error involves the magnitude of the estimate that depends on the location of the estimate within the grid element. Minimizing both errors is based on simple kriging.

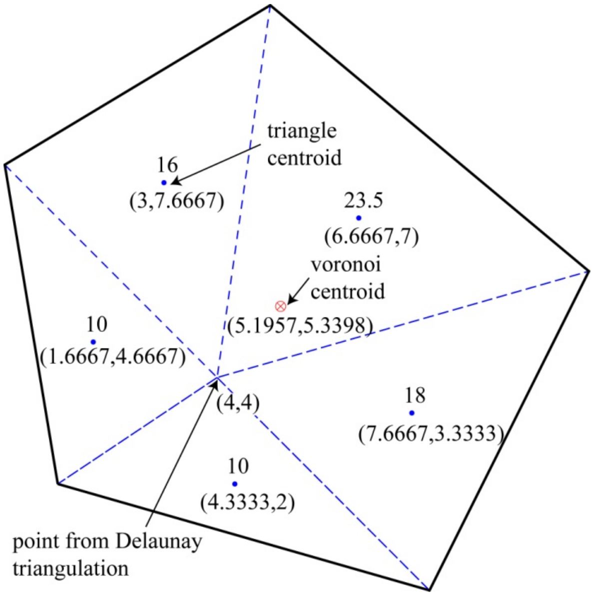


Figure 3-5: Voronoi element centroid by decomposition from Chapter 1. Numbers above the triangle centroids are the areas and those in parenthesis are x and y coordinates

Consider a particular element v from an unstructured grid: the point scale estimate using simple kriging is defined by Equation 3.1 and the v-scale estimate by Equation 3.5, where an overbar is used to indicate kriging weights are derived using the average covariance between conditioning data and the grid element (Equation 3.6). Also, the estimate $\hat{y}_v$ is no longer associated with a location, **u**, but with the whole element.

$$\hat{y}_v = \sum_{i=1}^{n} \bar{\lambda}_i y(\mathbf{u}_i) = \bar{\boldsymbol{\lambda}}^T \mathbf{y} \tag{3.5}$$

$$\mathbf{C}\bar{\boldsymbol{\lambda}} = \bar{\mathbf{c}} \tag{3.6}$$

Average covariance between $y(\mathbf{u}_i)$ and v is defined by Equation 3.7, where $|v|$ is the element volume (Equation 3.8).

$$\bar{C}(\mathbf{u}_i, v) = \frac{1}{|v|}\int_v C(\mathbf{u}_i, \mathbf{w})d\mathbf{w} \tag{3.7}$$

$$|v| = \int_v d\mathbf{w} \tag{3.8}$$

The point with position $\mathbf{u}$ should be located to minimize the error between the point scale estimate made without average covariance and the v-scale estimate made with average covariance.

The assumptions involved include:

1. Properties scale arithmetically at least up to the scale of the grid element permitting the use of average covariance.
2. The element v is small relative to the scale of variation of the covariance function, that is, the covariance function is nearly linear for any distance that spans v.
3. The element v is convex.

The minimization problem is defined by Equation 3.9 and is simplified after substituting in the equations for the kriging weights.

$$\begin{aligned} \min_{\mathbf{u}} : & \bar{\boldsymbol{\lambda}}^T\mathbf{y} - \boldsymbol{\lambda}^T\mathbf{y} \\ & = \mathbf{y}^T\mathbf{C}^{-1}\bar{\mathbf{c}} - \mathbf{y}^T\mathbf{C}^{-1}\mathbf{c} \\ & = \mathbf{y}^T\mathbf{C}^{-1}\left(\bar{\mathbf{c}} - \mathbf{c}\right) \end{aligned} \tag{3.9}$$

Under the assumptions specified, Equation 3.9 is minimized by setting the value in parentheses to zero. Covariance functions in geostatistics are functions of distance; therefore under assumption 2 listed previously the covariance is proportional to the distance and a proxy for the average covariance of Equation 3.7 is given by Equation 3.10, where $r(\mathbf{u}_i, \mathbf{w})$ is the Euclidean distance between $\mathbf{u}_i$ and $\mathbf{w}$.

$$\begin{gathered} \bar{c}(\mathbf{u}_i, v) = \frac{1}{|v|}\int_v C(\mathbf{u}_i, \mathbf{w})d\mathbf{w} = \frac{1}{|v|}\int_v r(\mathbf{u}_i, \mathbf{w})d\mathbf{w} \\ r(\mathbf{u}_i, \mathbf{w}) = \left\|\mathbf{u}_i - \mathbf{w}\right\| \end{gathered} \tag{3.10}$$

Consider the case with one conditioning datum, then Equation 3.9 is written as Equation 3.11.

$$\min_{\mathbf{u}} : \mathbf{y}^T\mathbf{C}^{-1}\left[\bar{C}(\mathbf{u}_i, v) - C(\mathbf{u}_i, \mathbf{u})\right] \tag{3.11}$$

Setting $\bar{C}(\mathbf{u}_i, v) - C(\mathbf{u}_i, \mathbf{u})$ to zero and rewriting using the distance function results in Equation 3.12. Reorganized into Equation 3.13, the left hand side defines the centroid of v in terms of the distance coordinate r. This is a contour of constant distance around $\mathbf{u}_i$.

Considering a second conditioning datum $\mathbf{u}_{i+1}$, the resulting contours intersect near the centroid of v.

$$\frac{1}{|v|}\int_{v} r(\mathbf{u}_i,\mathbf{w})d\mathbf{w} = r(\mathbf{u}_i,\mathbf{u}) \qquad 3.12$$

$$\frac{\int_{v} r(\mathbf{u}_i,\mathbf{w})d\mathbf{w}}{\int_{v} d\mathbf{w}} = r(\mathbf{u}_i,\mathbf{u}) \qquad 3.13$$

Equation 3.13 is evaluated for three conditioning data and a triangular element with an area of 1.5 m^2 in Figure 3-6. As v gets larger, especially in a direction orthogonal to the vector from $\mathbf{u}_i$ to $\mathbf{u}$, the contour departs from the centroid because the distance surface over v is cylindrical. However, when several conditioning data are present, the average of the intersections near the center of v will approximate the center of v and minimize the error between the estimate and estimation variance made with and without the use of average covariance.

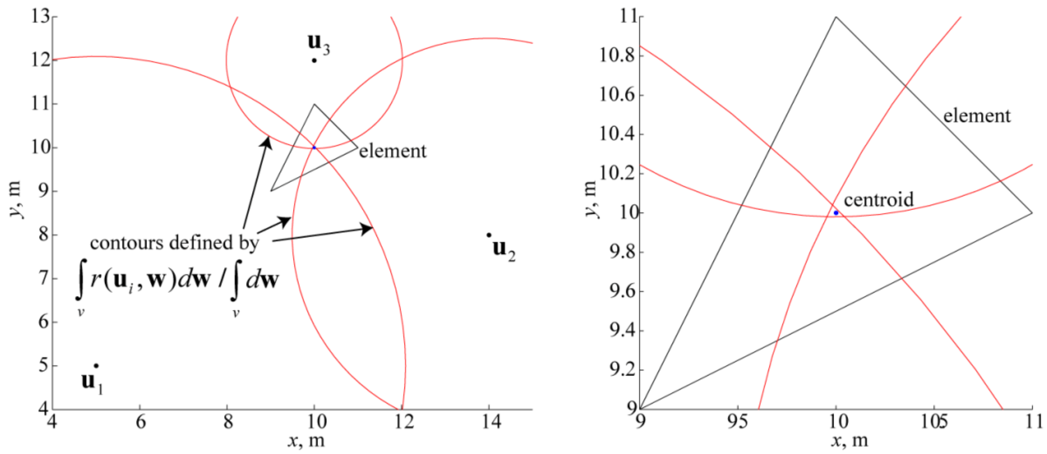


Figure 3-6: Contours equal to the average distance between points and a triangular element intersecting at the element centroid.

When geostatistical simulation is used to populate a grid and average covariance is not used, the points should be the centroids of the grid elements. This holds true for all grids including those where the centroids are not used in other processes such as flow simulation. For example, flow simulation on PEBI grids uses the vertexes of the dual triangular grid, which do not always coincide with the centroids of the PEBI grid elements. In another scenario, if a PEBI grid were used to discretize an unstructured grid, it would be populated using geostatistics and the element centroids, then flow based upscaling would use the vertices of the dual triangular grid.

The consequences of using points other than the centroid to model reservoir properties are demonstrated with a configuration of points that are not evenly spaced, but that result in a Voronoi grid that is also a regular grid (Figure 3-7). Random Gaussian fields are generated with no conditioning data (unconditional) using the centroids and using the

Voronoi grid points. Fields are generated with LU simulation (Davis, 1987) that is a technique for generating Gaussian processes using Cholesky factorization and involves the following steps:

1. Compute the covariance matrix between all points, $\mathbf{C} = C(\mathbf{u}_i, \mathbf{u}_j),\ i, j = 1, \ldots, n$, where n is the number of points.
2. Compute the Cholesky factorization of the covariance matrix, $\mathbf{C} = \mathbf{L}\mathbf{L}^T$
3. Generate a vector of Gaussian random numbers with length n, $\mathbf{y}_0$.
4. Correlate the random vector with the lower triangular matrix, $\mathbf{y} = \mathbf{L} \cdot \mathbf{y}_0$

Resulting Gaussian fields have the correct covariance structure or variogram. Multiple realizations are generated by repeating steps 3 and 4 with different Gaussian random vectors, $\mathbf{y}_0$. With this approach, the same vector $\mathbf{y}_0$ is used to generate a random field at the centroids and at the Voronoi points; therefore they can be compared directly since the differences originate strictly from the covariance matrix that depends on the position of the points.

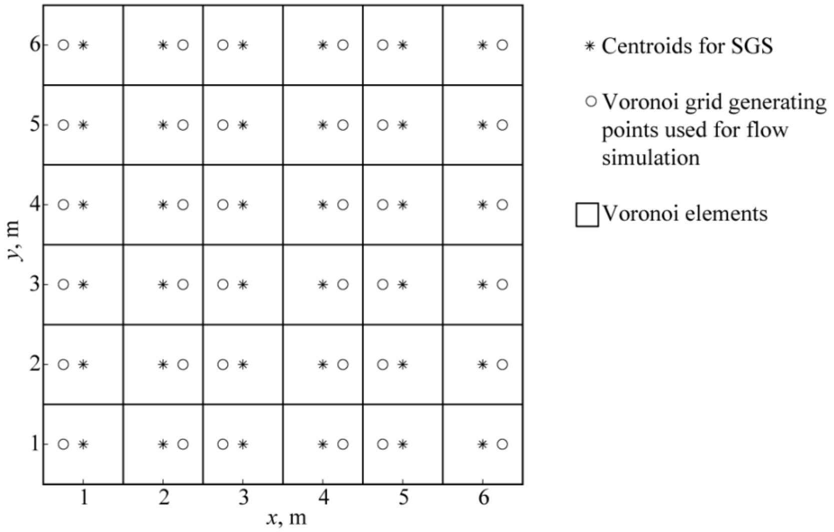


Figure 3-7: Voronoi grid with points that do not coincide with the element centroids.

In this example, the Gaussian fields from LU simulation are converted to synthetic permeability fields using $\mathbf{z} = \exp(\mathbf{y})$. An isotropic spherical variogram with a range of 5 meters and no nugget effect is considered for this example. Several items are compared visually and with sum squared error (SSE) measures between using the centroids, which is correct, and using the Voronoi points including the differences in:

1. Synthetic permeability fields.
2. Two-point transmissibilities computed using the Voronoi points.
3. The upscaled permeability using the static averages covered in Chapter 1 including arithmetic, harmonic, and geometric.

SSE is measured for a single realization and its distribution over a set of 500 realizations is also computed. For upscaled permeability, the distribution of the difference over the set of realizations is used for comparison.

Two-point transmissibility is used to relate the flux across an interface in the grid to the pressure in the elements that share the interface (Chen et al, 2007). Transmissibility across a vertical interface in the Voronoi grid is computed using Equation 3.14, where Δy is the length of the y-interface that is intersected orthogonally by the vector from point $\mathbf{u}_i$ to $\mathbf{u}_j$, Δx is the length of the vector from $\mathbf{u}_i$ to $\mathbf{u}_j$, and k_x is the permeability in the x direction (Figure 3-8). The equation for the transmissibility across the horizontal interfaces is similar with Δy and Δx flipped and k_y instead of k_x. For this example, permeability is isotropic, that is, $k_x = k_y$.

$$T_{ij} = 2\frac{\Delta y}{\Delta x}\frac{k_{xi} \cdot k_{xj}}{k_{xi} + k_{xj}} \qquad 3.14$$

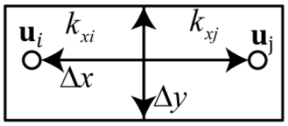


Figure 3-8: Components for evaluating Equation 3.14.

A realization of permeability using Voronoi points, centroids, and the difference is shown in Figure 3-9. The SSE between the fields is 12.53. The difference between adjacent simulated values caused by the difference in distance between centroids and Voronoi points can be seen, for example, in the elements at $(x = 2, y = 1)$ and $(x = 3, y = 1)$. For the Voronoi point pairs across vertical element interfaces that are closer than centroid points, the values are more similar as expected since the covariance between them is higher. The reverse is true for Voronoi point pairs that are further apart.

Differences in the permeability fields lead to different transmissibility fields in x and y (Figure 3-10 and Figure 3-11). In flow simulation, this will result in a different pressure and velocity field and therefore in a different flow response. Histograms of the permeability and transmissibility errors for 500 realizations were also generated (Figure 3-12). Errors tend to follow a lognormal distribution since the synthetic permeability realizations are lognormal.

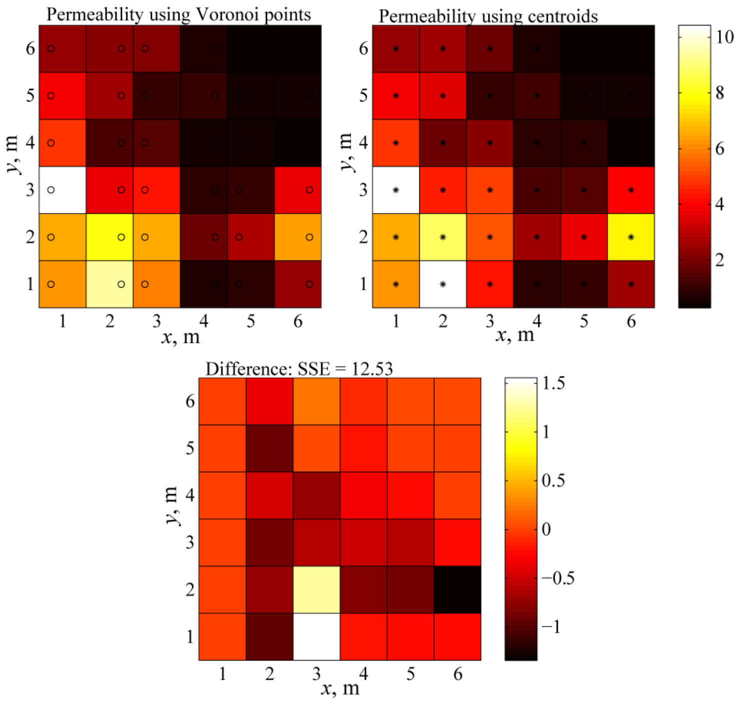


Figure 3-9: Permeability fields using Voronoi points, centroids and the difference.

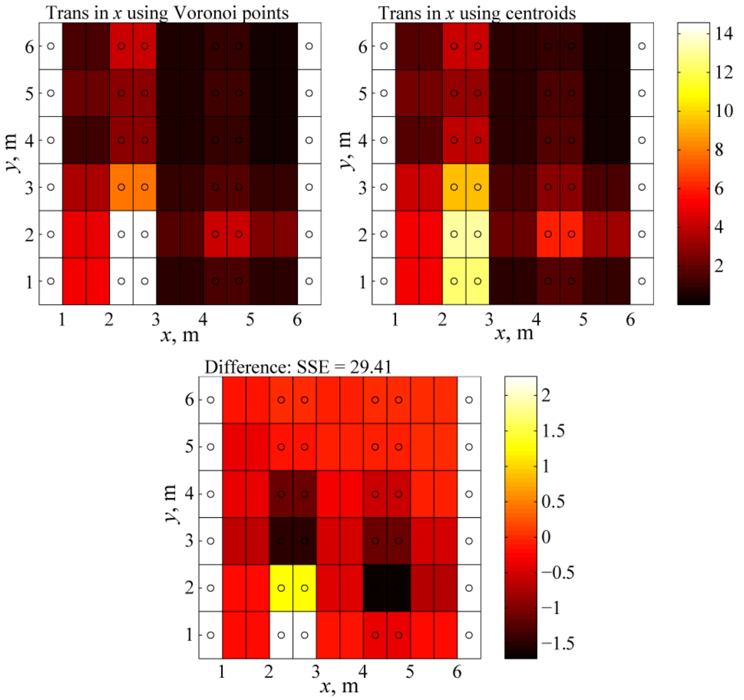


Figure 3-10: x-Transmissibility fields using Voronoi points and the permeabilty fields from Figure 3-9.

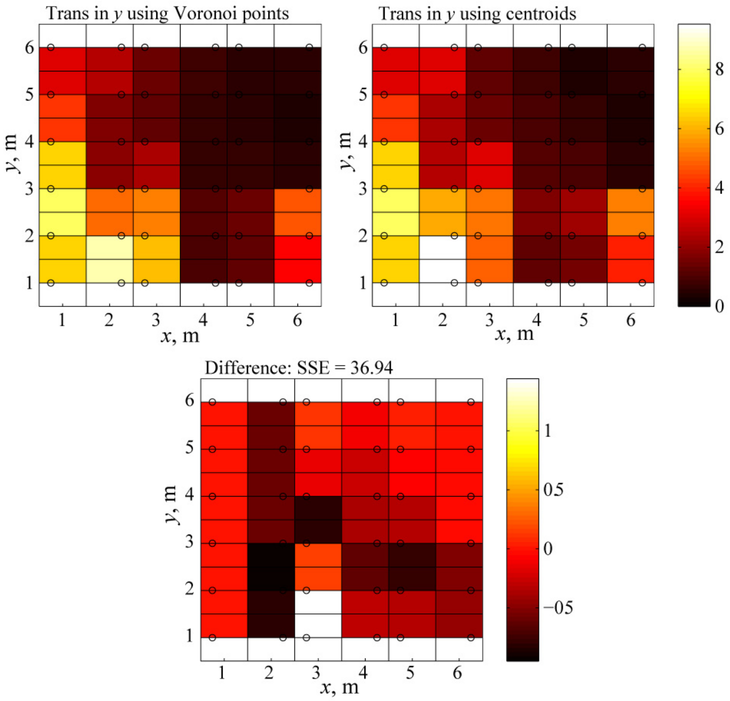


Figure 3-11: *y*-Transmissibility fields using Voronoi points and the permeability fields from Figure 3-9.

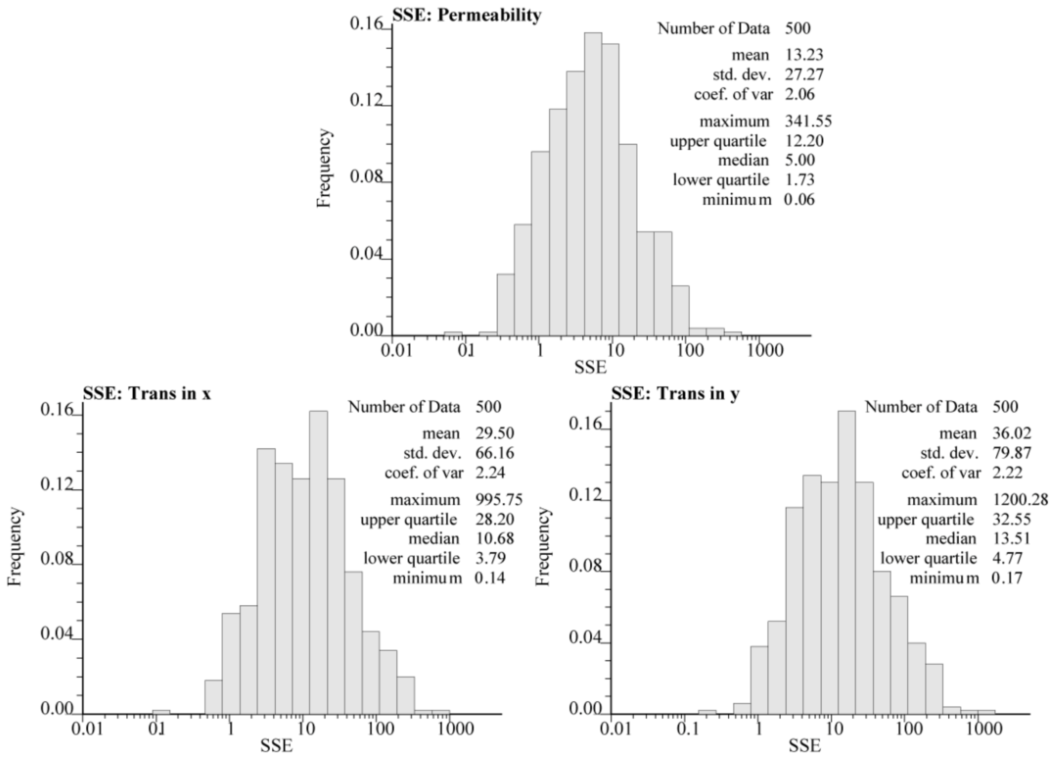


Figure 3-12: Histograms of SSE measures for 500 realizations.

Three averages were used to approximate upscaled permeability of the 6 m by 6 m region: arithmetic, geometric, and harmonic. Histograms of the error incurred when computing the upscaled permeability using the field generated with Voronoi points as opposed to element centroids are shown in Figure 3-13. The errors involved are similar to the average SSE divided by the number of element in the underlying permeability field and transmissibility fields. Although the errors do not appear high in magnitude, the minimum and maximum permeability values observed in the 500 realizations is 0.016 and 46.35 mD respectively. Permeability models can involve much larger values, and the error scales with magnitude. Consider a permeability field with a higher mean, so instead of creating a synthetic field using $\mathbf{z} = \exp(\mathbf{y})$, we use $\mathbf{z} = \exp(\mathbf{y} + 4)$. The range of permeability observed for 500 realizations using this equation is 0.85 to 2,530.7 mD. Errors in the average permeabilities are significantly higher (Figure 3-14).

These errors in average permeability are for one coarse grid element. For a grid of such elements, the errors in average permeability translate into errors in transmissibility as they did in the fine scale model, resulting in a different pressure and flow response in flow simulation.

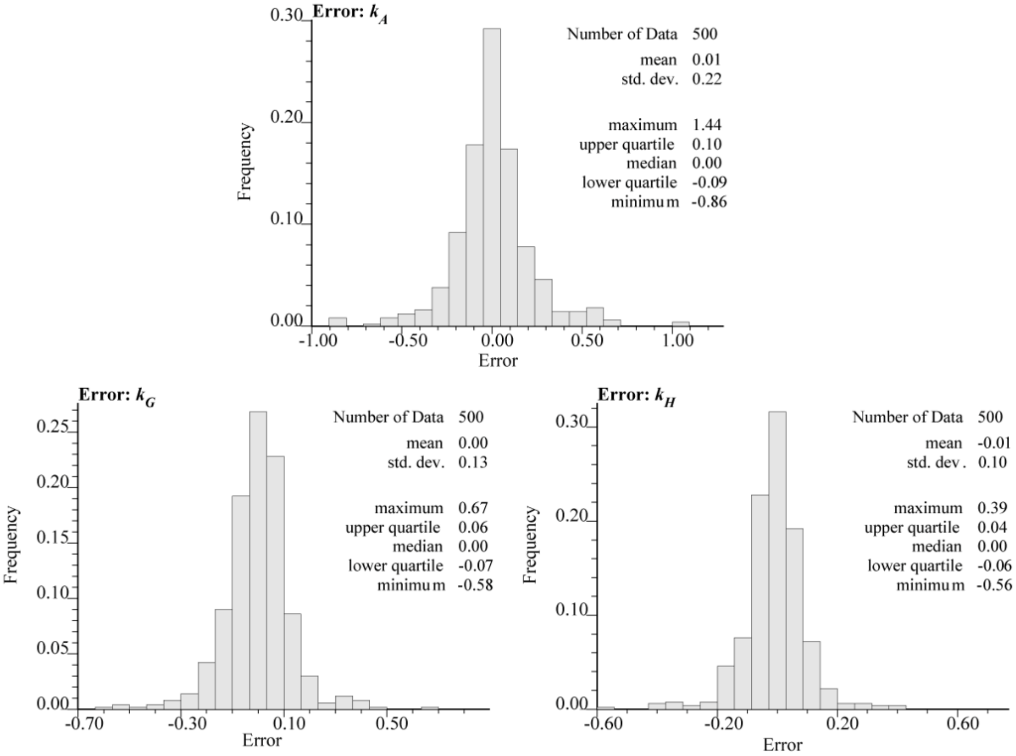


Figure 3-13: Error histograms for three static permeability averages of 500 realizations.

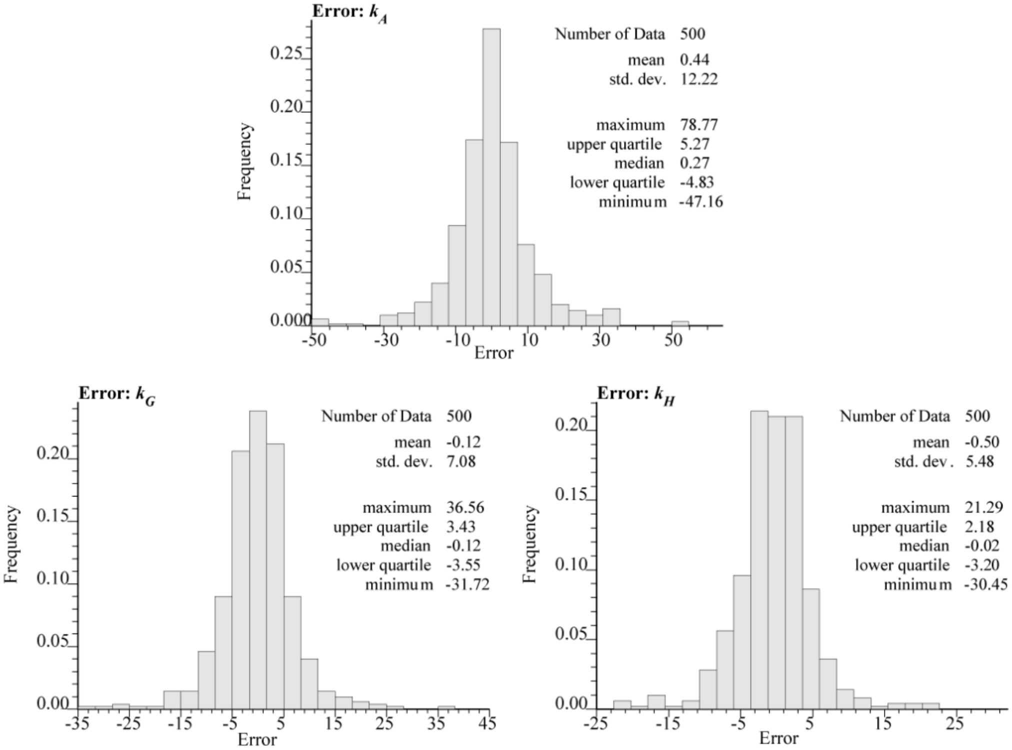


Figure 3-14: Increase in error for three static permeability averages of 500 realizations using permeability with a larger magnitude.

3.3 Algorithm Details

Designing a sequential simulation algorithm to work on irregular sets of points involves a few differences from existing approaches designed for regular grids. Differences include: how point locations are defined; computation of covariance between points; generation of a random path for sequential simulation to follow; and searching for nearby data for kriging.

In geostatistical modeling algorithms that use regular grids, the location of points is defined implicitly. Cartesian regular grids defined by center points require nine parameters in three dimensions:

- Number of elements along each Cartesian axis: n_x, n_y, n_z
- Origin of the first element center: $\mathbf{u}_0 = (u_x, u_y, u_z)$. In this definition, the origin or first block is the lower, left, back (minimum z, y, x respectively), element.
- Size of each element: $d\mathbf{u}_0 = (du_x, du_y, du_z)$

The center of any grid element based on its index, **I**, using the above nine parameters is done with Equation 3.15, where $\mathbf{I} = \left(I_x, I_y, I_z\right), I_x \in 1, \dots, n_x, I_y \in 1, \dots, n_y, I_z \in 1, \dots, n_z$ and $\mathbf{1} = (1,1,1)$.

$$\mathbf{u}_{\mathbf{I}} = \mathbf{u}_0 + (\mathbf{I} - \mathbf{1}) \cdot d\mathbf{u}_0 \qquad 3.15$$

With irregular sets of points, the locations are specified explicitly. There are two computational disadvantages to defining points explicitly: 1 – the points are usually stored on disk (hard drive) and must be loaded into main memory (random access memory or RAM) and this is not required with regular grids; 2 – storing point explicitly requires more memory, leaving less for other processes. Both disadvantages are not significant because in most cases, points will only be read into main memory once and there is usually an adequate amount of RAM available, especially with 64 bit computers.

Computing covariance between two points is another difference. When the relative position of two points in a regular grid is identical, the covariance does not change regardless of the location in the grid, assuming second order stationarity. If the covariance between two points is computed once, it is known for all other points with the same configuration in the regular grid and does not have to be recomputed, but the covariance must be stored. Covariance values are stored in a lookup table, denoted C_T, based on the relative index between the two points involved. For two points in a regular grid with indexes $\mathbf{I}_1$ and $\mathbf{I}_2$, the relative index is $d\mathbf{I} = \mathbf{I}_1 - \mathbf{I}_2$, and the covariance is accessed from or assigned to the lookup table: $C_{12} = C_T(d\mathbf{I})$. A two dimensional example is given in Figure 3-15.

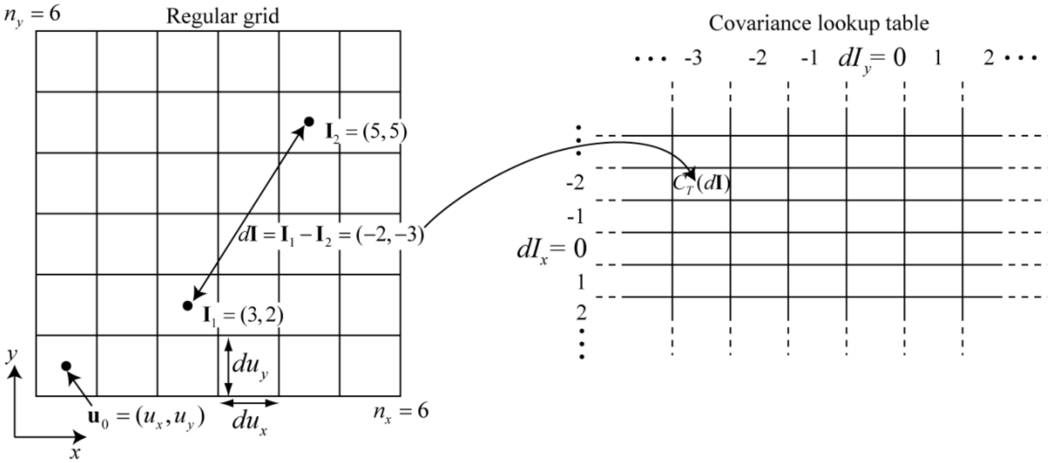


Figure 3-15: Covariance lookup table for a regular grid.

When the points are irregularly spaced, no two configurations of points may be the same and the type of covariance lookup table for regular grids is not possible. It is also not feasible to store all covariances as $n(n-1)/2$ pairs are involved, where n is the number of points in the irregular point set. The number of pairs grows quadratically with n, exceeding one billion with n as low as 45,000. The disadvantage of no lookup table is the covariance between any pair of points encountered must be computed, rather than referenced from a lookup table.

Two other differences involving random path generation and searching for data are more extensive and are covered in the following sections.

3.3.1 Random Path

Sequential simulation involves estimating conditional distributions for all center points of a grid. It is based on a factorization of the joint probability of grid points and conditioning data defined by Equation 3.16, where P is probability, n is the number of conditioning data and N is the number of grid points. For Gaussian random fields, conditional distributions are given by simple kriging and used to generate Gaussian random values, $y_n, \dots, y_{n+N}$.

$$\begin{aligned} &\mathrm{P}\left\{y_{n+1} \le Y_{n+1}, \dots, y_{n+N} \le Y_{n+N} \mid y_1, \dots y_n\right\} = \\ &\qquad \mathrm{P}\left\{y_{n+1} \le Y_{n+1} \mid y_1, \dots y_n\right\} \times \\ &\qquad \mathrm{P}\left\{y_{n+2} \le Y_{n+2} \mid y_1, \dots y_n, y_{n+1}\right\} \times \\ &\qquad \vdots \\ &\qquad \mathrm{P}\left\{y_{n+N} \le Y_{n+N} \mid y_1, \dots y_n, y_{n+1}, \dots, y_{n+N-1}\right\} \end{aligned} \tag{3.16}$$

For very small problems ($n + N \le 10{,}000$) when all conditioning data and previously simulated values can be included in kriging, the path taken through the grid points has no effect on results; however, N is typically large ($N \ge 10^6$) for realistic problems and the conditional probabilities are approximated using a set of nearest neighbours. In this case, the path has an effect on results. Several approaches to avoid path-related artifacts in realizations are based on random or quasi-random paths. For regular grids, the multigrid approach is common (Gomez-Hernandez and Cassiraga, 1994; Chiles and Delfiner, 1999). It is a stratified sampling method (Asmussen and Glynn, 2007) that provides improved reproduction of the variogram over pure random sampling, especially when anisotropy and more complex nested variogram models are involved.

A multigrid approach for unstructured grids is accomplished using a spatial decomposition data structure called a quadtree for two dimensional problems and an octree for three dimensions (Klinger, 1972; Berg et al, 2000). In two dimensional problems, the set of grid points is decomposed into four quadrants that form the coarsest layer. Each quadrant is further divided into four more quadrants, forming the second layer, and so on. This recursive decomposition is continued until a quadrant only contains one grid or data point (Figure 3-16). Each quadrant or octant is called a node; if nodes contain more than one point, they are split into more quadrants constituting a new level; nodes with only one point are called leaf nodes.

A path is generated from a quadtree or octree by randomly traversing from the highest level to a leaf node. Each time a level is accessed, a different node is chosen than the last time it was accessed. Each time a leaf node is reached, the datum is added to the path and the cycle is repeated from the top level. Leaf nodes are visited only once. A small unconditional simulation example is provided to show variogram reproduction results using a purely random path (Figure 3-17), the multigrid approach (Figure 3-18), and the quadtree method (Figure 3-19). A 200 by 200 node regular grid of 1 m elements was used. Sgsim (Deutsch and Journel, 1998) with the number of previously simulated nodes limited to 30 was used to generate the conditional distributions. Average variograms were calculated by taking the average of the experimental variograms of each individual realization.

In all cases, variogram reproduction is close visually. The SSE between the input model and the average variogram of the 50 realizations was computed as well (Table 3-1). Errors were calculated up to two different distances: to the variogram range in each direction and to a distance of 100 m in each direction. Samples for the SSE calculation were spaced at the element size of 1 m. For SSE up to 100 m, errors are highest when using the fully random path. The multigrid and quadtree are close, with the quadtree performing best in the y-direction. For SSE up to the variogram range, the quadtree approach performs best followed by the multigrid.

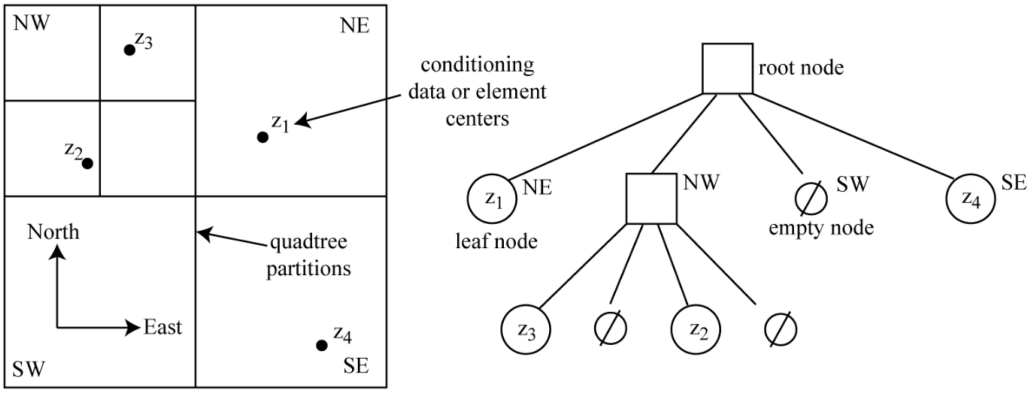


Figure 3-16: Schematic of a quadtree for 4 data points.

Table 3-1: SSE of variogram reproduction for different random paths.

Random Path Type	SSE in x 100 m	SSE in y 100 m	SSE in x 15 m	SSE in y 50 m
Random	0.0775	0.0444	0.0203	0.0241
Multigrid	0.0234	0.0382	0.0123	0.0132
Quadtree	0.0243	0.0120	0.0089	0.0085

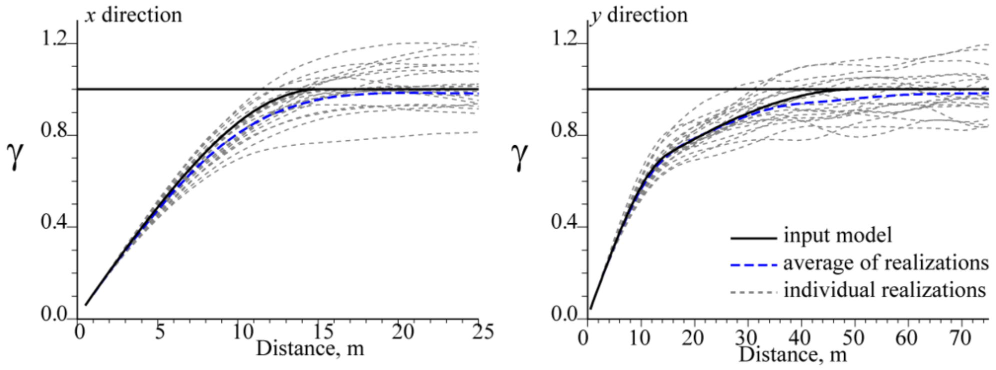


Figure 3-17: Variogram reproduction using a pure random path.

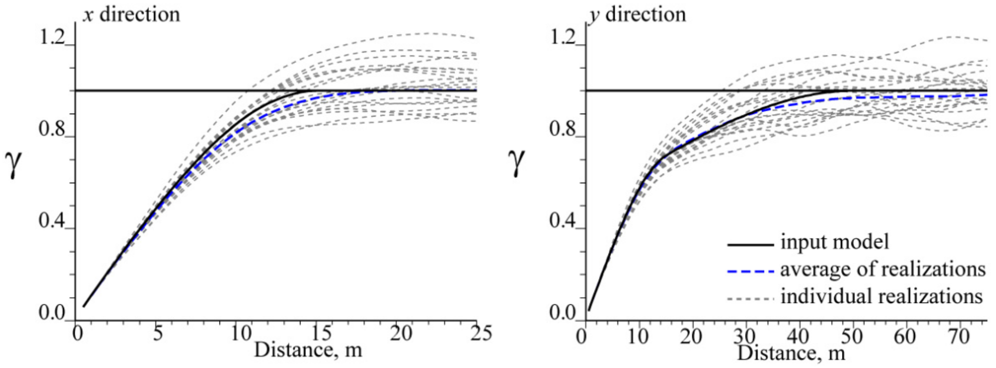


Figure 3-18: Variogram reproduction using the multigrid approach.

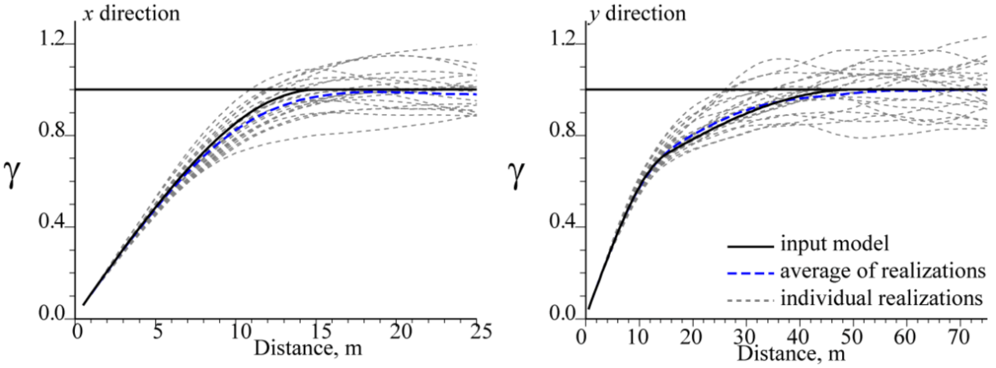


Figure 3-19: Variogram reproduction using the quadtree approach.

Quadtrees and octrees are well suited for unstructured grids where the element volume, and hence the point density, varies in space. The number of levels depends on the point density; higher density requires more levels. Since the dimension of quadrants from a particular level is tied to data density, multigrid sampling is achieved for all data densities encountered in an unstructured grid. Two tests are done to show that variogram reproduction is achieved for the quadtree approach on non-uniformly distributed points. Unconditional realizations are generated using psgsim (see Appendix).

Test 1: a random point set with 10,000 points is generated using a Gaussian mixture model distribution within a 100 by 100 meter domain. The mixture model consists of three Gaussian kernels centered at $(x = 50, y = 20)$, $(80,70)$, and $(25,75)$ with standard deviations of 15, 22 and 8 respectively. The concentration of points in these three areas could represent discretization within the vicinity of three wells for example. An anisotropic variogram model defined in Table 3-2 is used with the 30 nearest previously simulated points for kriging. A single Gaussian realization is shown in Figure 3-20, and variogram reproduction results for fifty realizations are shown in Figure 3-22. Visually, the match between the input model and average experimental variogram is good. The SSE computed at a spacing of 1 m in the x-direction and 3 m in the y-direction are 0.0053 and 0.0087 respectively. Errors are accumulated up to distances of 20 m and 80 m in x and y respectively.

Test 2: the same variogram, nearest neighbours, and domain size is used, but the points are distributed along randomly positioned lines that could represent faults or fractures. 10,500 points are used. There is a background set of 3,000 uniform random points, and each line is surrounded by 1,500 random points with distances from the line that are Gaussian distributed. A single Gaussian realization is shown in Figure 3-21, and variogram reproduction results for fifty realizations are shown in Figure 3-23. Variogram reproduction results are good as with the last example by visual inspection. The SSE computed at a spacing of 1 m in the x-direction and 3 m in the y-direction are 0.0057 and 0.0127 respectively. Errors are accumulated up to distances of 20 m and 80 m in x and y respectively.

Table 3-2: Variogram model for Test 1 and 2.

Structure	Structure Type	Nugget Effect	Variance	Range in y (m)	Range in x (m)
1	Spherical	0	0.5	30	10
2	Exponential	0	0.5	100	10

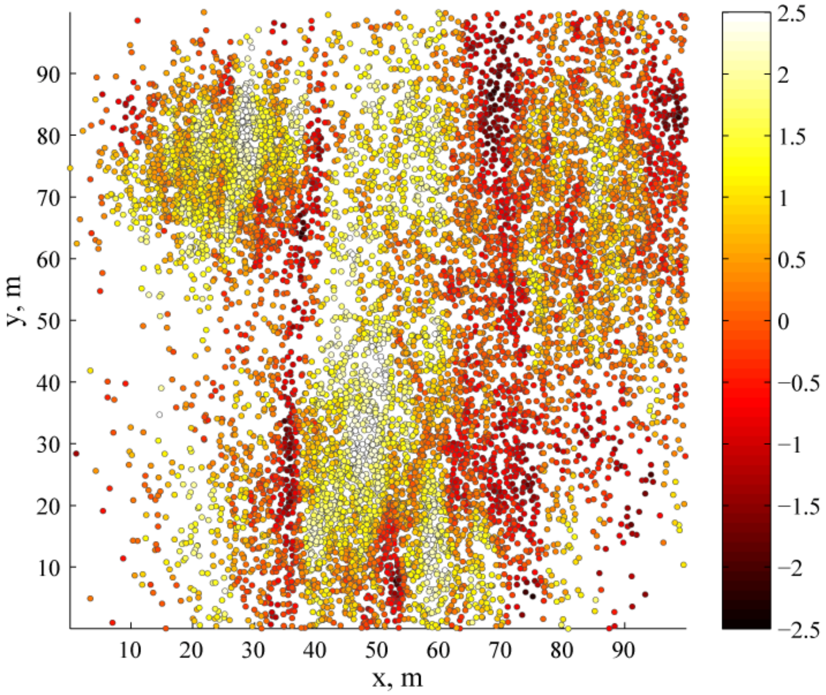


Figure 3-20: Point realization for Test 1. Point values are Gaussian distributed.

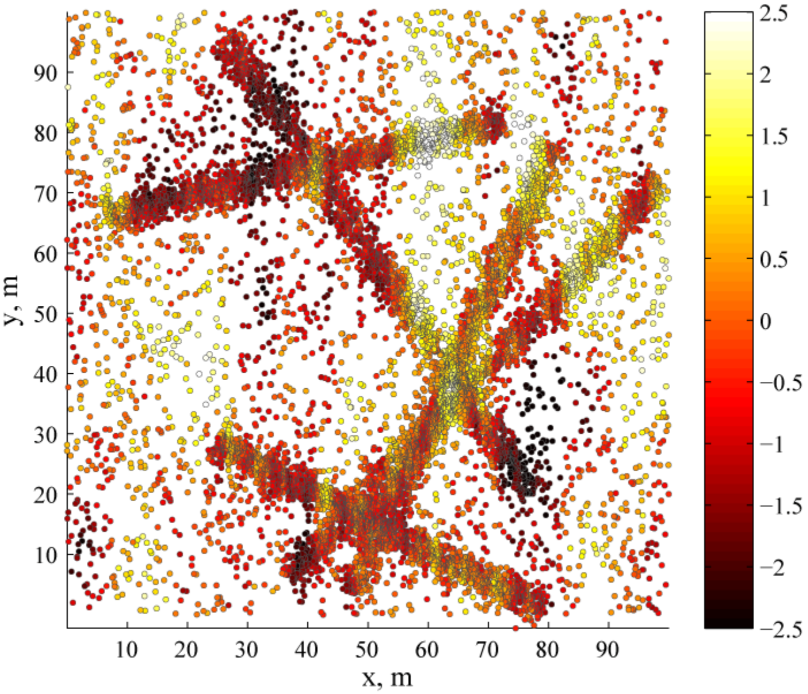


Figure 3-21: Point realization for Test 2. Point values are Gaussian distributed.

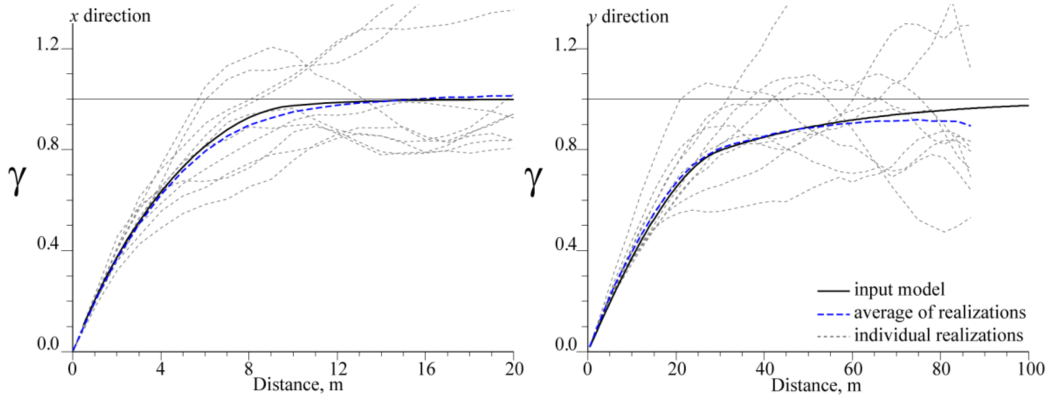


Figure 3-22: Variogram reproduction for Test 1.

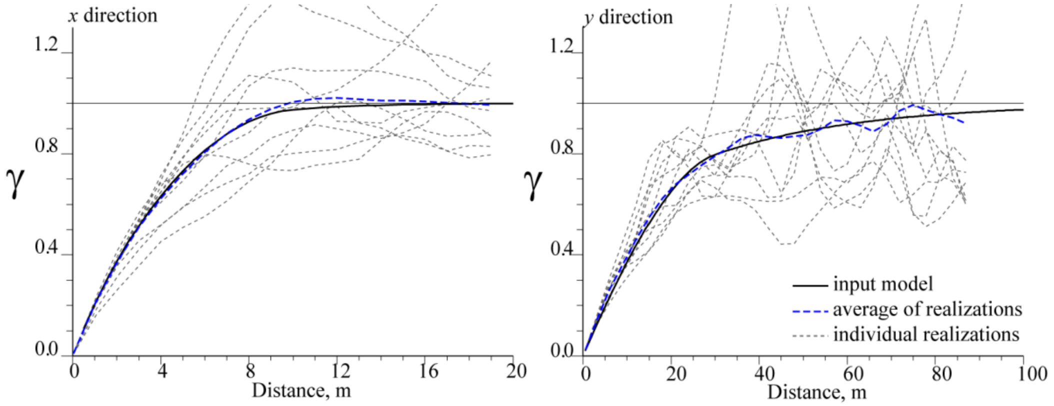


Figure 3-23: Variogram reproduction for Test 2.

3.3.2 Search

Estimating conditional distributions using a set of nearest neighbours requires finding the nearest points to the estimation location. Two search strategies utilized in GSLIB are the superblock search and spiral search (Deutsch and Journel, 1998). The superblock search is applicable to unstructured grids and irregularly spaced sample data. Many other search strategies exist (Skiena, 1998; Zezula et al, 2005), and they are often designed for very large and high dimensional datasets. The superblock search and another effective search structure called k-dimensional trees or kd-trees (Skiena, 1998; Samet, 1990a; Samet, 1990b) are explored. Other search strategies are not explored as they will likely offer no significant improvement; however, additional study in this area could be undertaken.

For geostatistical modeling the number of dimensions does not exceed three and the search space is Euclidean in depositional space. Executing the search in depositional space will yield the nearest neighbours with highest covariance for a specific variable, but this may not occur if the search is done in geological space based on Euclidean distance. When multiple variables are involved, different covariance functions are possible and the nearest neighbours of one variable may not be the same as those for another variable. This may warrant a different search for each variable in a multivariate problem.

3.3.2.1 Superblock Search

The superblock search decomposes space into a regular grid of elements. Conditioning data and grid points are stored in each element (Figure 3-24). For example, superblock (3,3) in the figure references the point marked with a cross. Searching for the nearest neighbours of a point, **u**, is done by retrieving the containing superblock (the query block) using an index calculation (Equation 3.17), then spiralling out through superblocks until enough neighbours are located. Equation 3.17 is Equation 3.15 rearranged for **I**, where the floor function rounds real numbers down to the nearest integer and 1.5 is added instead of 1 to account for points **u** that do not align exactly with the center of a superblock.

$$\mathbf{I} = \text{floor}\left(\frac{\mathbf{u}-\mathbf{u}_0}{d\mathbf{u}_0}+1.5\right) \qquad 3.17$$

Grid parameters are required to setup a superblock search. All conditioning data and grid points are added to the grid prior to simulation by determining the superblock index via Equation 3.17 and storing the index of the data and points. The index of a conditioning datum or point from an irregular point set refers to a position in a vector containing the locations of the conditioning data and points. The path taken when spiralling out from the query block is optimized using a search template (Figure 3-25) that defines a set of superblock index offsets to follow. Superblocks in the template are sorted by increasing distance from the center of the query block, where distance is measured to the nearest corner of other superblocks. Since the grid is regular, several superblocks are equidistant from the query block and they are all searched to ensure closer points within them are not omitted.

As a superblock search is done, the set of k-nearest neighbours is maintained using a maximum heap (Cormen et al, 2002) keyed by distance. Once k data have been found, the size of the ball containing them is the maximum distance in the heap (Figure 3-26). If additional points are found that have a smaller distance, a replace-maximum operation is done on the heap and the ball size is updated (Figure 3-27). This process continues until the distance to the closest corner of the next superblock is greater than the ball size. The resulting set of k points contains the exact k-nearest neighbours.

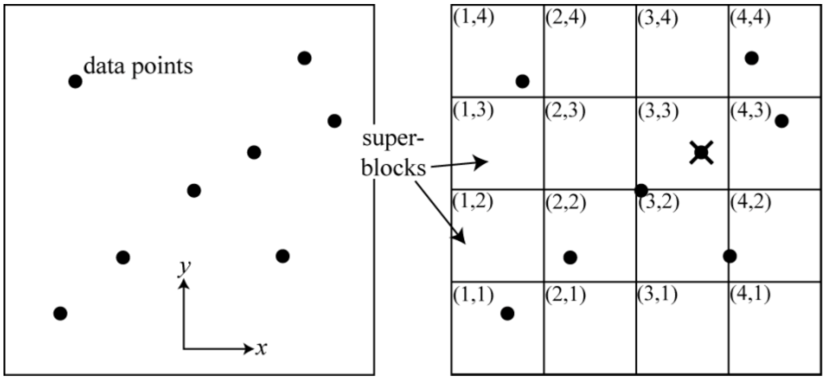


Figure 3-24: A two dimensional superblock structure. Block indexes are in parenthesis.

14	13	12	10	9	9	9	10	12	13	14
13	11	8	7	6	6	6	7	8	11	13
12	8	6	5	4	4	4	5	6	8	12
10	7	5	3	2	2	2	3	5	7	10
9	6	4	2	1	1	1	2	4	6	9
9	6	4	2	1	0	1	2	4	6	9
9	6	4	2	1	1	1	2	4	6	9
10	7	5	3	2	2	2	3	5	7	10
12	8	6	5	4	4	4	5	6	8	12
13	11	8	7	6	6	6	7	8	11	13
14	13	12	10	9	9	9	10	12	13	14

27	26	25	23	21	18	15	13	13	14	17
26	24	22	20	16	12	10	9	9	11	14
25	22	19	13	9	7	6	6	7	9	13
23	20	13	8	5	3	3	4	6	9	13
21	16	9	5	2	1	1	3	6	10	15
18	12	7	3	1	0	1	3	7	12	18
15	10	6	3	1	1	2	5	9	16	21
13	9	6	4	3	3	5	8	13	20	23
13	9	7	6	6	7	9	13	19	22	25
14	11	9	9	10	12	16	20	22	24	26
17	14	13	13	15	18	21	23	25	26	27

Figure 3-25: Superblock search template for isotropic case (left) and anisotropic case (right) with a 2:1 anisotropy ratio at an azimuth of 45 degrees. Numbers indicate search order from the query block numbered 0.

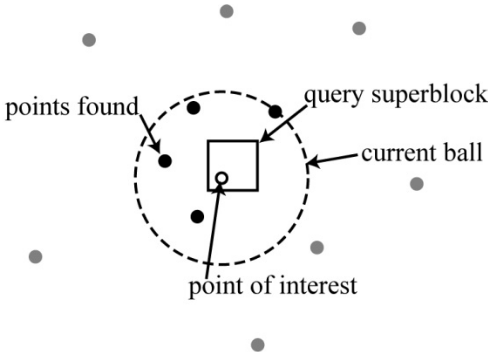


Figure 3-26: Ball size in a maximum heap.

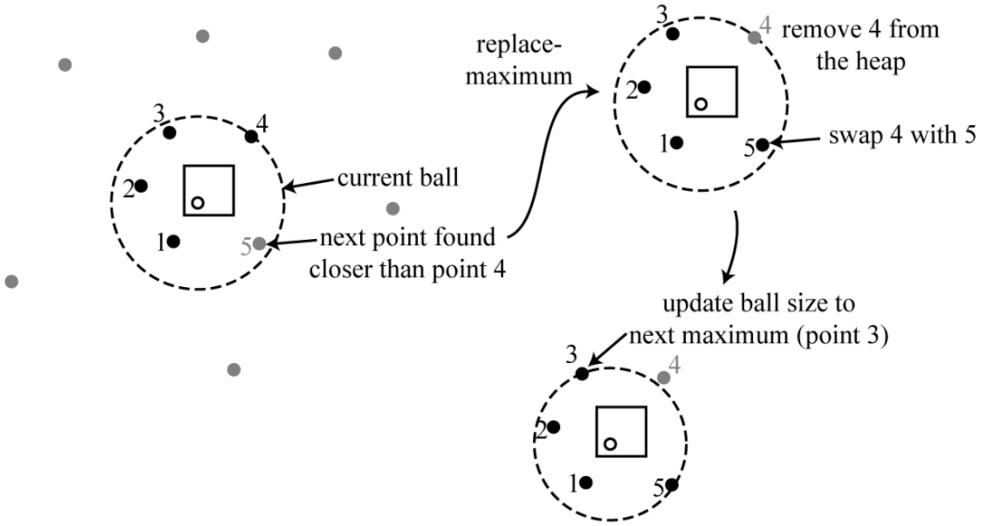


Figure 3-27: Replace-maximum operation for $k = 4$.

3.3.2.2 kd-trees

The kd-tree decomposes space using half-planes so that conditioning data and grid points can be accessed efficiently by location (Figure 3-28). For kriging, a kd-tree is constructed using the conditioning data and is static for all grid points. For sequential simulation, two approaches are possible: 1 – the tree is constructed once using all grid points and nearest neighbour queries return both previously simulated and unvisited points; 2 – the tree is dynamic, growing as grid points are visited.

Building the tree once using all grid points has the advantage that a balanced kd-tree is possible, which yields faster queries. The downside is when the ratio of un-simulated to simulated points is high; many points are accessed to find a set of nearest neighbours. With a dynamic kd-tree no time is taken accessing un-simulated points since they are not stored in the tree; however, some additional time is used for inserting points, and some efficiency is lost because the resulting tree is unbalanced. A tree that is balanced has the cutting plane positioned at the median of the data for all levels, which minimizes the number of levels required and the query time. With unbalanced trees, the number of levels depends on the location of the query points. This degrades the efficiency of kd-trees for nearest neighbour queries in geostatistical modeling applications.

To insert a point, the kd-tree is traversed until the containing node is found and the point's index is added to the kd-tree node. Traversing the tree involves making logical comparisons between the point location and the half-plane location for a particular node. When the number within a node reaches a predefined maximum, the node is split along the dimension with maximum extent. It is possible to ensure a kd-tree is balanced by the end of generating a realization since the points used to construct it are static. All half-planes can be pre-computed based on all points and used during the incremental construction of the tree. The tree is balanced at the end of generating a realization, but may not be balanced at intermediate steps.

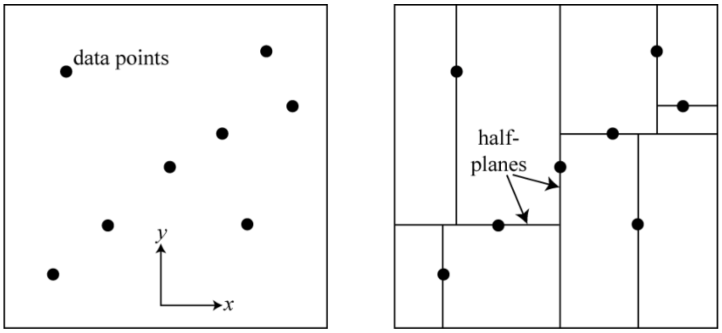


Figure 3-28: Example two dimensional kd-tree

The kd-tree has three advantages over the superblock search:

1. A search radius and grid specification is not required.
2. There is no overhead to store empty blocks.
3. There is no additional overhead to construct and store a search template

The problem of having to specify a search radius is that conditioning data or grid points beyond it are ignored. This may be important in the early stages of simulating a realization when few previously simulated values are available and also for locations that are far from conditioning data. Moreover, having to specify or determine optimal grid settings with a superblock search is an iterative process. Large superblocks cause a slow search because each block may contain many points and more distances are computed than necessary to find the k-nearest neighbours. Small superblocks require excessive memory because the grid can contain excessive numbers of superblocks.

Using superblock grids superimposed over unstructured grids will invariably lead to empty blocks, but these still require memory and are still visited during the search process to test if data is contained within them. Lastly, the search template can involve a substantial amount of memory, especially for cases where small superblocks are used and the search radius is large. It is possible to use larger superblocks; however, this can result in blocks containing more than the required k-nearest neighbours; in this case, many points are checked against the heap for the initial superblocks searched. This is the cause of the slower time trials in Figure 3-29.

Both search strategies are tested in two and three dimensions for a series of uniformly distributed random point sets ranging from 100 to 1 million. A computer with a 3.2 GHz processor was used. Unconditional realizations are generated for each point set using a maximum of 20 previously simulated points for kriging. The average time per realization is calculated based on ten realizations. Results are shown in Figure 3-29. The kd-tree performs best in all cases. In the two dimensional problems, two versions of the superblock search are tested. One was of fixed size ($n_x = n_y = 31$); therefore, as the size of the point set grows, the number of data in each superblock increases. The consequence is a quadratic increase in time with the number of points. The second superblock search was optimized to contain on average 10 points per superblock giving results similar, but slower than the kd-tree. Two superblock grids were also used for the three dimensional sets. One was optimized to have an average of 10 points per superblock and another to have an average of 5 points per superblock.

For comparison, an unconditional realization on a 1 million element regular grid takes approximately 35 seconds, which is 3.2 times faster than the equivalent unstructured problem using psgsim. The principal reason for the difference in run times is the covariance lookup table. Distances and covariance do not have to be computed for every pair of points encountered.

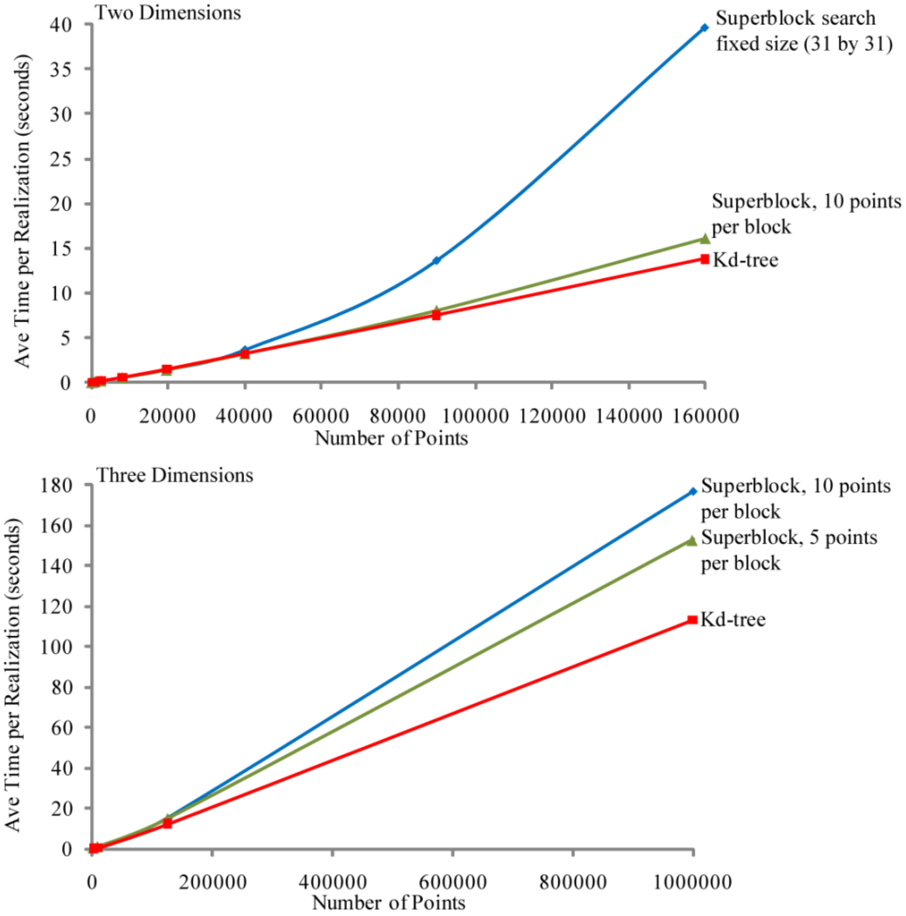


Figure 3-29: Time trials for two dimensions (top) and three dimensions (bottom).

3.4 Model Validation

For geostatistical models it is useful to assess statistics such as the histogram and variogram for three reasons: 1 – to ensure that the random fields are consistent with the input statistics; 2 – to check that other input parameters are satisfactory, such as the number of data to use for kriging; 3 – to confirm that models are consistent with the assumptions made regarding first and second order stationarity. Both the histogram and variogram are checked in depositional space where they are defined as input to the sequential simulation algorithms. The following sections cover the validation of these statistics for psgsim. Variogram reproduction was confirmed in Section 3.3.1 for two arbitrary cases with lots of scatter; however, more detail on computing the experimental variogram of large irregular point sets is provided.

3.4.1 Histogram Reproduction

One test is done to show that psgsim generates fields with a Gaussian distribution in depositional space. The importance of this test is based on the normal score transformation that is applied to reservoir properties prior to modeling them. In using the normal score transform, we assume the resulting property is Gaussian distributed. By generating a field with a Gaussian distribution, the back transformation yields the original input distribution of the reservoir property.

The point set for the test is generated differently than in Section 3.3.1 since grid elements are needed. Element areas or volumes are used as weights when computing the empirical distribution function of the realizations generated by psgsim. Points are generated randomly and the Delaunay triangulation provides grid elements. The triangular element centers are used as input to psgsim. A dataset similar to Test 2 in Section 3.3.1 involving random points generated along lines that could represent faults or fractures is used. The variogram from Table 3-2 is used and 50 realizations are generated to check for Gaussianity (Figure 3-30).

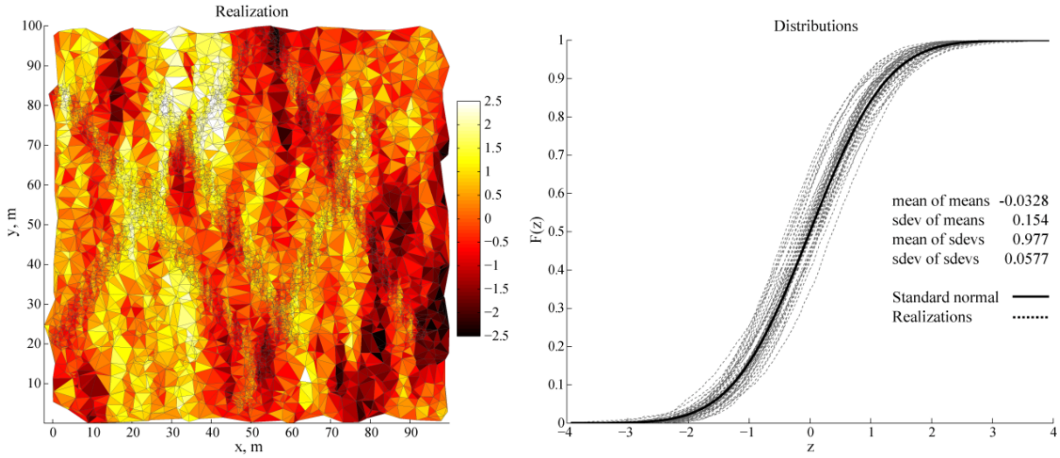


Figure 3-30: A single realization and empirical distributions of 50 realizations.

Normality of the realizations is tested using the Chi Square test with the null hypothesis that the realizations follow a normal distribution. Since a finite domain is used to generate realizations, there are fluctuations in the sample mean and variance from one realization to the next; therefore the reference distribution for each realization is determined based on the sample mean and sample variance. The sample distribution is non-standard normal; however, the average mean and average standard deviation across multiple realizations reflect a standard normal distribution. Empirical distributions of each realization are fit using 500 quantiles.

The mean of the realization means (-0.0328) and the mean of the realization standard deviations (0.977) are close to 0 and 1 respectively. At a significance level of 0.05, the Chi Square tests rejected the null hypothesis once out of 50 realizations. Chi Square tests used 20 bins to compute observed frequencies and since sample mean and variance were used to compute expected frequencies, 17 degrees of freedom remain.

3.4.2 Variogram Reproduction

Confirming variogram reproduction is more challenging because of implementation issues. Points are irregularly spaced and there are potentially millions of them. Computing experimental variograms involves point pairs; a grid with N points has N^2 pairs, which cannot feasibly be searched during the computation. This is still infeasible when considering symmetry, where $N(N+1)/2$ pairs are considered. Moreover, multiple directions are typically checked, with the directions aligning with principal directions of correlation.

Gamv from GSLIB (Deutsch and Jounel, 1998) computes the variogram of irregularly spaced data; however, it searches all $N(N+1)/2$ pairs. This is feasible for the intended purpose of gamv, which is for a limited number of sample data. Often sample data are sparse and in small number. A different approach is developed for very large irregular point sets. Some of the enhancements are: data organization to eliminate array indexing calculations; the number of required square roots is minimized; and the outermost or slowest cycling loop is by direction rather than over the point pairs. Several parameters required for computing the experimental variogram are (Figure 3-31): azimuth angle, θ, azimuth tolerance, $d\theta$, horizontal bandwidth, b_h, dip angle, α, dip tolerance, $d\alpha$, vertical bandwidth, b_v, lag separation distance, h, and lag tolerance, dh.

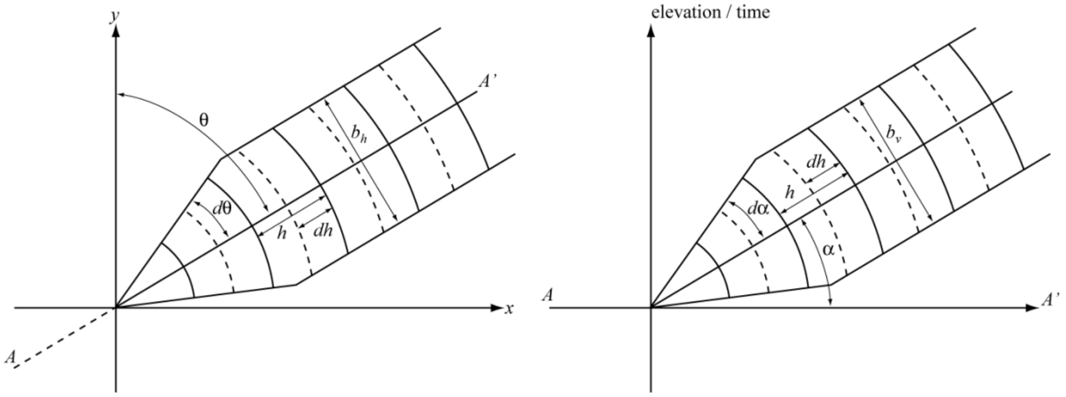


Figure 3-31: Experimental variogram parameters for irregularly spaced data.

For each direction the coordinate system is rotated so that the new x-coordinate, x', aligns with the direction. Locating pairs into their proper lag windows in the rotated coordinate system is a simple index calculation. For two dimensional problems, points are sorted by y' (the rotated y-coordinate) and makes clipping by the bandwidth parameter straightforward. Points within the bandwidth will always be contiguous in memory and defined by a lower and upper index. From a given index i, the range is easily determined by searching in $+i$ and $-i$ until $y'(\pm i)$ exceeds the bandwidth. Rotation also simplifies clipping points outside the angle tolerance window. For two points with indexes i and j, the pair is clipped for the condition defined by Equation 3.18, where tan $(90-\theta)$ only needs to be evaluated once for each direction.

$$\frac{|x_i' - x_j'|}{|y_i' - y_j'|\tan(90 - \theta)} < 1 \qquad 3.18$$

Enhancements for three dimensional problems are similar. Rather than sorting, points are added to a two dimensional kd-tree based on their y' and z' coordinates. Points within the bandwidth specifications are found using an orthogonal range query with extents $y' \pm b_h/2$ and $z' \pm b_v/2$. Clipping by the angle tolerance window is identical to the 2D case for the azimuth tolerance, but requires an additional check using the dip tolerance (Equation 3.19), where $\tan(90 - \alpha)$ only needs to be evaluated once for each direction.

$$\frac{|x_i' - x_j'|}{|z_i' - z_j'|\tan(90 - \alpha)} < 1 \qquad 3.19$$

The enhancements discussed are only necessary for large data sets when bandwidths that are small relative to the domain of the problem are used. Computing omnidirectional experimental variograms with large bandwidths still involves all $N(N + 1)/2$ pairs and no appreciable difference in performance between gamv and gamvf is expected. Three time trials were done to show the feasibility of using gamvf. Times for gamv are included to illustrate the time required for all $N(N + 1)/2$ pairs. The first trial dataset is two dimensional with 10,000 random samples and experimental variograms for two directions are calculated. The same dataset is also used with large bandwidth and 90 degree direction tolerances so that gamvf searches all $N(N + 1)/2$ pairs. The second example is three dimensional with 125,000 random samples and three directions are explored (Table 3-3). The domain for both examples is 100 m along each axis. All azimuth and dip tolerances were set to 10.0 degrees and horizontal and vertical bandwidths to 1 except for the omnidirectional case.

Table 3-3: Run time for gamv and gamvf

Dimensions	n	Time, seconds (gamv)	Time, seconds (gamvf)
2 (omni)	10,000	17.875	8.766
2	10,000	17.875	0.172
3	125,000	10,495.255	2.641

More comprehensive time trials are done with gamvf. Table 3-4 gives results for several values of N and different bandwidths and angle tolerances. The domain in all cases is a 100 m cube and points are uniform random distributed throughout. Each time is for a single direction and single realization. An isotropic spherical variogram with a range of 30 m was used.

Experimental variograms are plotted in Figure 3-32. They are not labelled since most overlap significantly. Variograms do not match the input model exactly because this is one realization, not the average variogram of many realizations. Variograms for the

8,000 point case show the largest separation, where curves with increased variability are associated with smaller bandwidths. Resulting times are insensitive to the angle tolerance for the range of bandwidths explored in each case. For the 125,000 and 1 million point cases, the experimental variograms are smooth and a smaller bandwidth could be applied resulting in reduced execution time. Three smaller bandwidths were used for this and results are summarized in Table 3-5 and Figure 3-33. The number of pairs involved in computing each variogram value is also shown. Smooth results are obtained using a bandwidth of 0.1 and each point is supported by 10,000 to 20,000 pairs. The number of pairs decreases as distance increases due to the finite domain.

Using a bandwidth of 0.1 for 100 realizations of 1 million points in three directions, variogram reproduction checks would take approximately 20 minutes. Since times appear to increase linearly with N, this approach is feasible for applications involving multimillion element models.

Table 3-4: Time trials for gamvf.

N	Bandwidth (b_h=b_v)	Angle Tolerance ($d\theta$=$d\alpha$)	Time (seconds)
8,000	2.0	10, 30, 50	0.0156, 0.0156, 0.0156
	4.0	10, 30, 50	0.0469, 0.0469, 0.0469
	8.0	10, 30, 50	0.0781, 0.0781, 0.0781
125,000	1.0	10, 30, 50	0.6719, 0.6875, 0.6718
	2.0	10, 30, 50	1.5156, 1.5156, 1.5000
	4.0	10, 30, 50	4.7188, 4.6719, 4.6719
1,000,000	0.2	10, 30, 50	4.4688, 4.6875, 4.7344
	0.4	10, 30, 50	7.7188, 7.7813, 7.7188
	0.8	10, 30, 50	19.563, 19.609, 19.500

Table 3-5: Reduced bandwidth time trials for N = 1,000,000.

N	Bandwidth ($b_h = b_v$)	Angle Tolerance ($d\theta = d\alpha$)	Time (seconds)
1,000,000	0.01	10	2.906
	0.05	10	3.438
	0.10	10	3.813

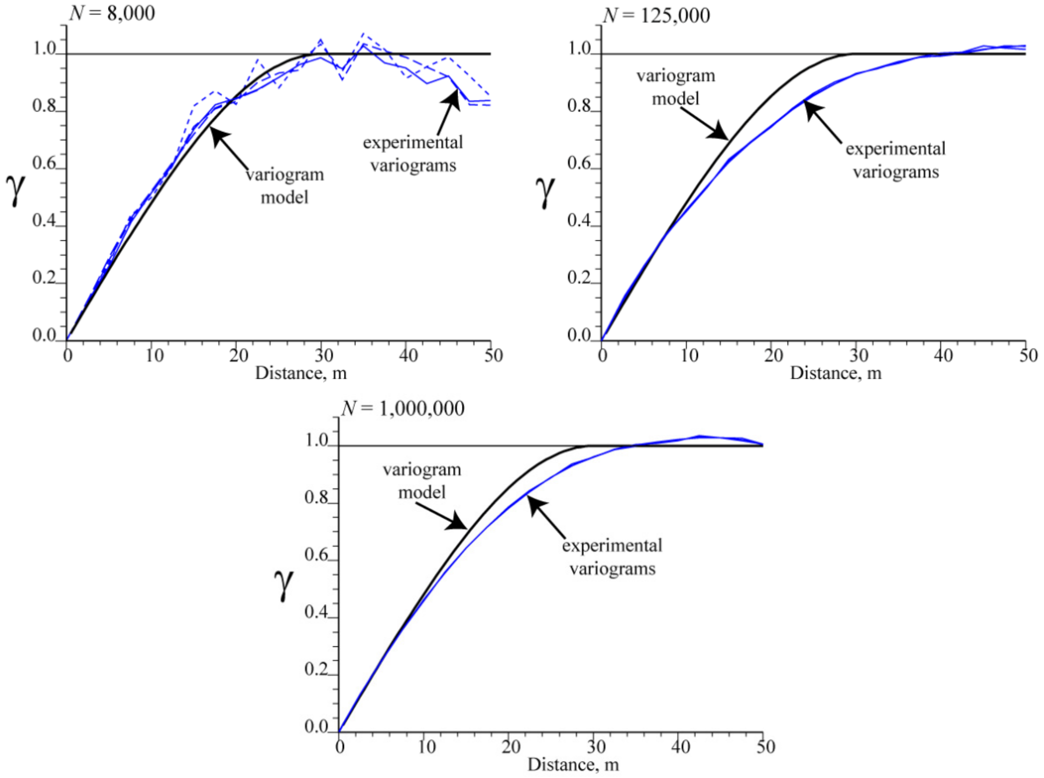


Figure 3-32: Experimental variogram results from gamvf for various *N*, bandwidths, and angle tolerances.

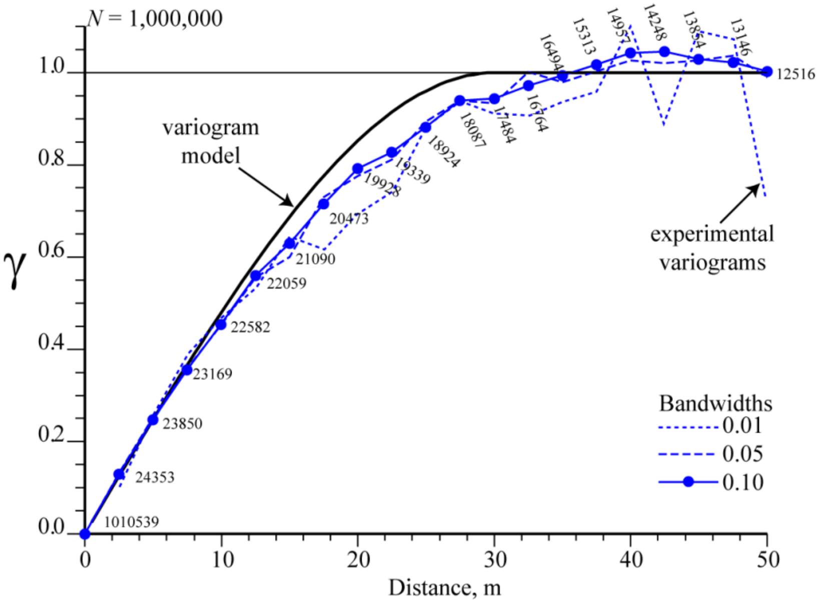


Figure 3-33: Experimental variograms for small bandwidths and 1 million points. Number of pairs for each point is shown for $b_h = b_v = 0.1$.

3.5 Summary

This chapter covered many of the necessary components to populate unstructured grids with sequential simulation. Of particular importance is the space where statistics are derived. The distribution and covariance function of a random variable are applied in depositional space; therefore, these parameters are derived in depositional space. An algorithm was developed to populate unstructured grids with point scale or pseudo-point scale realizations. Since points are used instead of attempting to populate values that represent the scale of grid elements, all of the geometry required to specify a grid is not required. Psgsim requires only a set of points. To minimize the error due to scale differences, the optimal points are the element centers. Other underlying implementation details including specifying point locations explicitly instead of implicitly, the inability to use covariance lookup tables, generating random paths, and searching for the k-nearest neighbours were covered.

Model validation is an important part of geostatistics. Histogram and variogram reproduction for unstructured grids was discussed. It was shown that psgsim gives Gaussian random fields with the correct spatial correlation based on an example with a point set having a high degree of scatter. The histogram is weighted by the unstructured grid element volumes, operating in the same way as declustering weights in polygonal declustering. This Chapter focused primarily on two dimensional examples to demonstrate correctness. Further demonstration of geostatistics on unstructured grids in three dimensions is provided in Chapter 5 in a case study.

Chapter 4 Upscaling

Grids populated with reservoir properties are at a higher resolution than grids used for flow simulation. Upscaling is used to compute effective properties for the coarse grid using properties from the fine grid. Accuracy of upscaled properties depends on the geometry of the fine grid that is controlled by discretization; therefore the geometry of the discretization should be influenced by the upscaling procedure. This chapter develops upscaling of arithmetic variables such as porosity and facies in Section 4.1. A flow based upscaling approach for transmissibility using the multipoint flux approximation is developed in Section 4.2. Within this is a section on the convergence of upscaled transmissibility as the resolution of the discretization is increased. Section 4.3 discusses aspects of discretization element quality including angle and aspect-ratio for triangles and tetrahedra respectively.

4.1 Arithmetic Averaging

Several variables in reservoir modeling applications scale arithmetically including porosity, facies, and fluid saturations. Upscaling from a fine scale simplex grid to a coarse scale element is accomplished using a weighted arithmetic average of the values given by Equation 4.1, where Z_{Vi} is the upscaled value assigned to coarse element i with volume V, n is the number of simplexes within element i with volumes w_k, $Z(\mathbf{u}_k)$ are the point scale simulated values in each simplex, and N is the number of coarse elements.

$$Z_{Vi} = \frac{1}{|V_i|} \sum_{k=1}^{n} w_k \cdot Z(\mathbf{u}_k), i = 1,...,N \qquad 4.1$$

This holds for continuous and indicator variables; however, the result for indicators is a vector of proportions, $\boldsymbol{\varphi}_{Vi}$, rather than a single average (Equation 4.2), where δ is the indicator transform function defined by Equation 4.3 and nc is the number of categories.

$$\boldsymbol{\varphi}_{Vi} = \frac{1}{|V_i|} \sum_{k=1}^{n} w_k \cdot \delta_{Z,j}(\mathbf{u}_k), j = 1,...,nc \qquad 4.2$$

$$\delta_{Z,j} = \begin{cases} 1, Z = j \\ 0, \text{otherwise} \end{cases} \qquad 4.3$$

The volume of element i is given by Equation 4.4. Note that the locations, $\mathbf{u}_k$, depend on i and the number of discretization points may be different for each i as well.

$$|V_i| = \sum_{k=1}^{n} w_k \qquad 4.4$$

Equation 4.1 can be evaluated in either geological space or depositional space as long as the simplex volumes and coarse element volumes are computed in the same space. The resulting values are equal unless non-uniform deformation of the elements has taken place from post-depositional events. It is recommended that upscaling is done in geological space, since flow simulation with the coarse elements is done in geological space.

Computing the arithmetic average is similar to infill asymptotics (Cressie, 1993). With infill asymtotics, the domain is kept constant and the convergence of statistics with increasing sample size is analyzed. In this case, the domain is a coarse unstructured grid element and the sample size is controlled through the discretization process. The statistic is the mean of an underlying random field. In the literature, infill asymptotics assess statistical parameters of random processes such as the mean of the associated random variable. Arithmetic averaging is similar in that for a given realization of a random field, each average should represent the associated realization, rather than all possible realizations of the random process.

One question that arises is if the mean of a realization, whether for a whole reservoir or within a single unstructured grid element, converges as the number of points increases to infinity. From a sequential simulation perspective and considering a finite domain, the mean of a realization should converge to a constant value, but it may not be attainable. Letting the number of points, n, go to infinity, the estimate, $\hat{Z}$, and estimation variance, $Var(\hat{Z})$, is expressed as a recursion in Equation 4.5, where $\mathbf{c}$ is a vector of covariances between $\mathbf{u}_n$ and $\mathbf{u}_{n-1}$, $\mathbf{C}$ is the covariance matrix between all $\mathbf{u}_i, i = 1, \dots, n-1$, and z is the set of $n-1$ simulated values. Since the distance between points $h \rightarrow 0$ as $n \rightarrow \infty$, $C(\mathbf{u}_n, \mathbf{u}_j) \rightarrow 1$ for some $j \in 1 \dots n-1$. In this case $\mathbf{c}^T\mathbf{C}^{-1} = \boldsymbol{\lambda} = [0_1\ 0_2\ \cdots 0_{j-1}\ 1_j\ 0_{j+1} \cdots 0_{n-1}]$, which yields the result in Equation 4.6; therefore, the realization ultimately converges.

$$\begin{aligned} \hat{Z}(\mathbf{u}_n) &= \mathbf{c}^T\mathbf{C}^{-1}z(\mathbf{u}_{n-1}) \\ Var\left[\hat{Z}(\mathbf{u}_n)\right] &= \sigma_Z^2 - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c} \end{aligned} \qquad 4.5$$

$$\begin{aligned} \hat{Z}(\mathbf{u}_n) &\rightarrow Z(\mathbf{u}_j) \;\; n \rightarrow \infty \\ Var\left[\hat{Z}(\mathbf{u}_n)\right] &\rightarrow 0 \end{aligned} \qquad 4.6$$

A problem for this analysis is the mean to which a realization converges is unknown. The mean could be observed as n is increased, but until the domain is saturated the mean will fluctuate, i.e. it has non-zero variance. An alternative measure is needed to assess the quality of an average. The problem is similar to numerical integration, for example using the midpoint rule (Kythe and Schaferkotter, 2004). The goal is the same: to estimate the true integral defined by Equation 4.7. The difference is that for triangulations the intervals are not necessarily of equal width.

$$Z_V = \frac{1}{V}\int_V Z(\mathbf{u})d\mathbf{u} \qquad 4.7$$

Numerical integration results in a mean squared error measure that is summarized by Equation 4.8.

$$\varepsilon_Z^2(n,V)=E\left\{(Z_V-\hat{Z}_V)^2\right\} \tag{4.8}$$

The right hand side is equal to the estimation variance assuming the expected value of Z_V is zero (Equation 4.9).

$$\begin{aligned} Var\left\{Z_V-\hat{Z}_V\right\} &= E\left\{(Z_V-\hat{Z}_V)^2\right\}-E\{Z_V\}E\left\{\hat{Z}_V\right\} \\ &= E\left\{(Z_V-\hat{Z}_V)^2\right\} \end{aligned} \tag{4.9}$$

Expanding the square and substituting Equation 4.7 in for Z_V and Equation 4.1 for $\hat{Z}_V$:

$$\begin{aligned} \varepsilon_Z^2(n,V) &= E\left\{Z_V^2\right\}-2E\left\{Z_V\hat{Z}_V\right\}+E\left\{\hat{Z}_V^2\right\} \\ &= \frac{1}{|V|\cdot|V|}\int_V\int_V E\{Z(\mathbf{u})Z(\mathbf{v})\}d\mathbf{u}d\mathbf{v}-\frac{2}{|V|\cdot|V|}\sum_{k=1}^{n}w_k\int_V E\{Z(\mathbf{u}_k)Z(\mathbf{v})\}d\mathbf{v} \\ &+\frac{1}{|V|\cdot|V|}\sum_{k=1}^{n}\sum_{l=1}^{n}w_kE\{Z(\mathbf{u}_k)Z(\mathbf{v}_l)\}w_l \end{aligned} \tag{4.10}$$

The expected values of the products is equal to the covariance, that is, $E\{Z(\mathbf{u})Z(\mathbf{v})\}=C\big(Z(\mathbf{u}),Z(\mathbf{v})\big)+E\{Z(\mathbf{u})\}E\{Z(\mathbf{v})\}$ with the assumption $E\{Z(\mathbf{u})\}=E\{Z(\mathbf{v})\}=0$. Covariance is expressed as a function that is independent of the values of Z, so replacing $E\{Z(\mathbf{u})Z(\mathbf{v})\}$ in Equation 4.10 by $C(\mathbf{u},\mathbf{v})$:

$$\begin{aligned} \varepsilon_Z^2(n,V) &= \frac{1}{|V|\cdot|V|}\int_V\int_V C(\mathbf{u},\mathbf{v})d\mathbf{u}d\mathbf{v}-\frac{2}{|V|\cdot|V|}\sum_{k=1}^{n}w_k\int_V C(\mathbf{u}_k,\mathbf{v})d\mathbf{v} \\ &+\frac{1}{|V|\cdot|V|}\sum_{k=1}^{n}\sum_{l=1}^{n}w_kC(\mathbf{u}_k,\mathbf{v}_l)w_l \end{aligned} \tag{4.11}$$

All terms are average covariance expressions. Defining $\bar{C}(V,V)$ as the exact average covariance of V (first term), $\bar{C}(\mathbf{u},V)$ as the exact average covariance between point $\mathbf{u}$ and volume V (integral component of the second term), and $\hat{C}(V,V)$ as the approximate average covariance computed from the discretization (third term), Equation 4.11 is written as Equation 4.12 and is similar to the estimation variance in geostatistics (Chiles and Delfiner, 1999), except that weights are used.

$$\varepsilon_Z^2(n,V)=\bar{C}_Z(V,V)-2\sum_{k=1}^{n}\bar{C}_Z(\mathbf{u}_k,V)+\hat{C}_Z(V,V) \tag{4.12}$$

Equation 4.12 is written in terms of the covariance function of Z, C_Z; however, the covariance used in constructing models using sequential simulation discussed in Chapter 3 is typically that of the Gaussian transform, $Y=G(Z)$, denoted C_Y. Since the transformation is rank preserving and finite, the error using either covariance function is representative of the other, but the convergence rate to zero of $\varepsilon^2(n, V)$ with increasing n may differ (Equation 4.13). Note that the association of V, n and w to element i is implicit in Equations 4.10 and 4.11.

$$\frac{\varepsilon_Z^2(n+1,V)}{\varepsilon_Z^2(n,V)} \neq \frac{\varepsilon_Y^2(n+1,V)}{\varepsilon_Y^2(n,V)} \qquad 4.13$$

In the case of indicator modeling of categorical variables, the covariance is that of the indicator variable for evaluating Equation 4.12, which becomes a sum over the number of categories. Each category may have a different covariance function and weights are based on element volumes in geological space.

Equation 4.12 assumes the random field within V is stationary; although this is not a requirement. If V contains multiple stratigraphic units or facies, it is partitioned into stationary sub-volumes to evaluate the error in a piecewise manner. Evaluating the error is made more complex by the exact integrals that often require numerical methods to be evaluated (Journel and Huijbregts, 1978), and because the covariance function is defined in depositional space. As is shown in the following section, Equation 4.12 converges to 0 as $n \to \infty$.

4.1.1 Number of Points

Determining how many points to use for the discretization process depends on the desired accuracy to achieve through arithmetic averaging. The result depends on several parameters including the grid element geometry, variogram structure, and distribution function of the variable. All of these are problem specific; however, a general guide on selecting the number of points is developed. The approach is based on evaluating Equation 4.12 and understanding how it responds to different parameters. Two responses indicated in the literature are the error decreases as n increases and the error decreases as the variogram becomes more regular (Chiles and Delfiner, 1999, page 130). As n increases and assuming the discretization maintains regularity (i.e. points remain uniformly distributed in the element), the second and third terms in Equation 4.8 approach the first term, thereby reducing the error. Increased regularity of the variogram increases the accuracy of the second and third terms, which reduces the error. For example, if the regularity of the covariance function within V approaches a constant covariance of one, the error would be zero for any choice of $n \geq 1$.

The variance of Z scales the error linearly. In most cases, the covariance function is standardized to have unit variance. Expressing the error in the same variance units as Z is accomplished by multiplying the error using the standardized covariance of Z by the variance of Z. Variables such as porosity, facies, and water saturation range between zero and one and the variance cannot exceed 0.25. Other variables such as grain density and permeability may have different variance depending on the units (kg/m^3, lbs/ft^3, m^2, and others). Because of the variety of units, choosing the number of points is based on standardized variables.

Increasing the nugget effect, C_0, reduces the average covariance by a factor of $\sigma^2 - C_0 = 1 - C_0$, which reduces the error by the same factor. As in the example where the covariance function approaches a constant of one, in the same manner as the covariance approaches zero ($C_0 = \sigma^2$), the error would be zero for any choice of $n \geq 1$.

The average covariance is related to the volume and geometry of V. As the average covariance of V decreases for a given covariance function and constant n, the error increases. A reduction in average covariance generally indicates an increase in V, although two elements with equal volume and different orientations have different average covariance when there is anisotropy in the covariance function.

Determining the number n that achieves a given accuracy is done by evaluating Equation 4.12 and extracting a contour of constant error to guide the discretization process applied to the coarse unstructured grid. A disadvantage of this approach is the average covariance can have a range of values for all elements in an unstructured grid with the same volume since the elements can have different shapes. However, if the orientation and geometry of all elements in an unstructured grid with a given volume V is consistent, then the average covariance of those volumes will also be consistent. In some cases, this is a valid assumption because grids designed for flow simulation often follow some consistency limitations to achieve stable and convergent systems of equations.

Because of the wide variety of variogram models and grid geometry, the error function is developed empirically. Data are created by randomly generating volumes and choosing different n to evaluate Equation 4.12. Results are fit with a surface of the form $\varepsilon_{\bar{Z}}^2(n, V)$. The resulting surface is used to choose n to control grid discretization. It is also possible to choose a contour of constant error that provides a function of n, which is dependent on the average covariance or volume. A FORTRAN program called cbarerror was developed to do this. It generates grid elements and discretizes them using a range of n values to compute the error. The program uses a parameter file so users control the size and shape of volumes to explore, the range of n, and the variogram model (Table 4-1).

Table 4-1: Parameters for cbarerror.

Line	Parameters	Description
1	`1 0.0`	nested structures, nugget effect
2	`0 1.0 0.0  0.0  0.0`	structure, variance, angle 1 2 3
3	`10.0 10.0 10.0`	range 1 2 3
4	`0.1 20.0 10`	min max volume to explore, number
5	`1`	type of shape (1-tri/tet, 2-quad/hex)
6	`30 2`	max points to use and step
7	`3`	point infill method: 1-deterministic, 2-simplex
8	`2`	dimensionality of the problem
9	`errcurve.out`	file for output

The minimum and maximum volumes on line 4 are divided uniformly (for the parameters shown, 10 volumes are explored ranging between 0.1 and 20). Two types of elements are randomly generated: triangles or tetrahedra and quadrilaterals or hexahedra. The same number of random elements is generated for each volume as there are multiple volumes to explore (10 random elements per volume for the parameters shown, 100 in total). The error is evaluated for several values of n (line 6). The number always starts at 1 and increments by the step size up to the maximum. Two different point infill methods are possible. For the deterministic option, elements are gridded with a barycentric coordinate system for simplexes and a regular grid for quadrilaterals or hexahedra. Triangulation is done with the program Triangle (Shewchuk, 1996) and tetrahedralization with the program TetGen (Si, 2006).

Examples in two and three dimensions are developed for three common variogram models: the spherical, exponential, and Gaussian functions. All are isotropic with a range of 10 meters. Triangular and quadrilateral volumes range from 0.1 to 30 m^2 discretized by triangulation with the number of points ranging from 1 to 29 in steps of 2. Triangles are constrained to have interior angles no less than 20 degrees resulting in some variation in n from the targeted range. In three dimensions, tetrahedral and hexahedral volumes range from 0.1 to 90 m^3 discretized by tetrahedralization for the same point range. Tetrahedral quality was constrained to a radius-edge ratio of 1.5. An approach to model the error function is developed below in reference to the two dimensional examples. This is followed by three dimensional results. True average covariances were approximated by using an excessive number of points per block: 500 per element in two dimensions and 700 per element in three dimensions. The average covariance for different volumes based on the two and three dimensional examples described converges rapidly to a stable value (Figure 4-1).

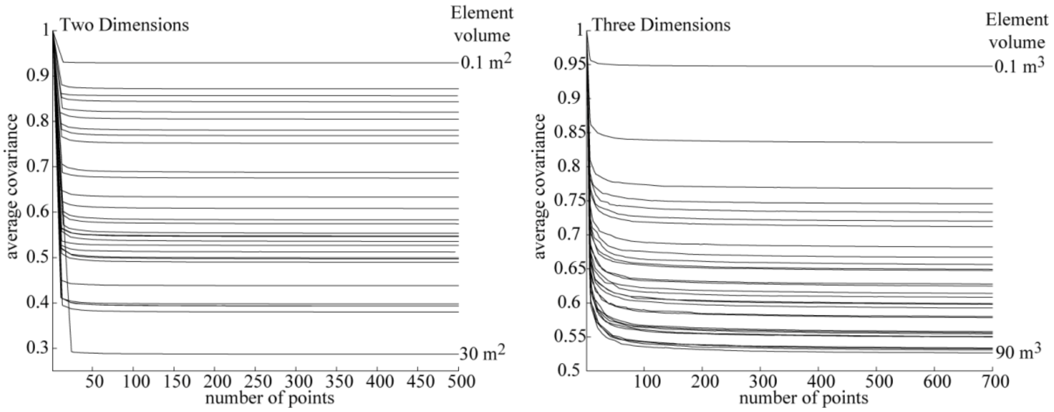


Figure 4-1: Convergence of average covariance with increasing number of points for different triangular and tetrahedral volumes in two and three dimensions.

An initial look at the resulting data for the spherical variogram (Figure 4-2, results are similar for the other variogram models) shows the error varies like n^{-1} and $V^{1/2}$. The error is also exponentially distributed and taking the logarithm reveals more information. The function is linearized by replacing n with its inverse, $1/(n+\tau)$, and raising the volume to a power, V^{β}. Through visual inspection for the data shown in Figure 4-1, τ and β were chosen to be 10 and 0.15 respectively (Figure 4-3). Alternatively, a simple minimization process to find τ and β involves the following iteration:

1. Choose initial parameter, τ or β.
2. Compute the independent variables, $(n+\tau)^{-1}$ or V^{β}, and standardize it.
3. Perform linear regression and compute the error.
4. Increase or decrease the parameter based on the change in error.
5. Iterate until the error in the residuals is small (the stopping criteria).

With this approach and using an error stopping criteria of 0.001, τ and β were found to be 6.67 and 0.2 respectively (Figure 4-3). Although the difference between the visually chosen parameters and optimized ones is not immediate, the effect is observed in the regression error of the final model. After linearization, the function is a plane with $(n+\tau)^{-1}$ and V^{β} as the independent variables and $\log(\varepsilon_Z^2(n,V))$ as the dependent variable (Equation 4.14). The mean squared error between the predicted error via regression and the actual error is 2.1×10^{-4} using the hand picked parameters and 5.1×10^{-5} using the optimized ones for the data shown in Figure 4-2.

$$\log \varepsilon_Z^2(n,V) = A_0 + A_1 V^{\beta} + A_2 \left(n+\tau\right)^{-1} \tag{4.14}$$

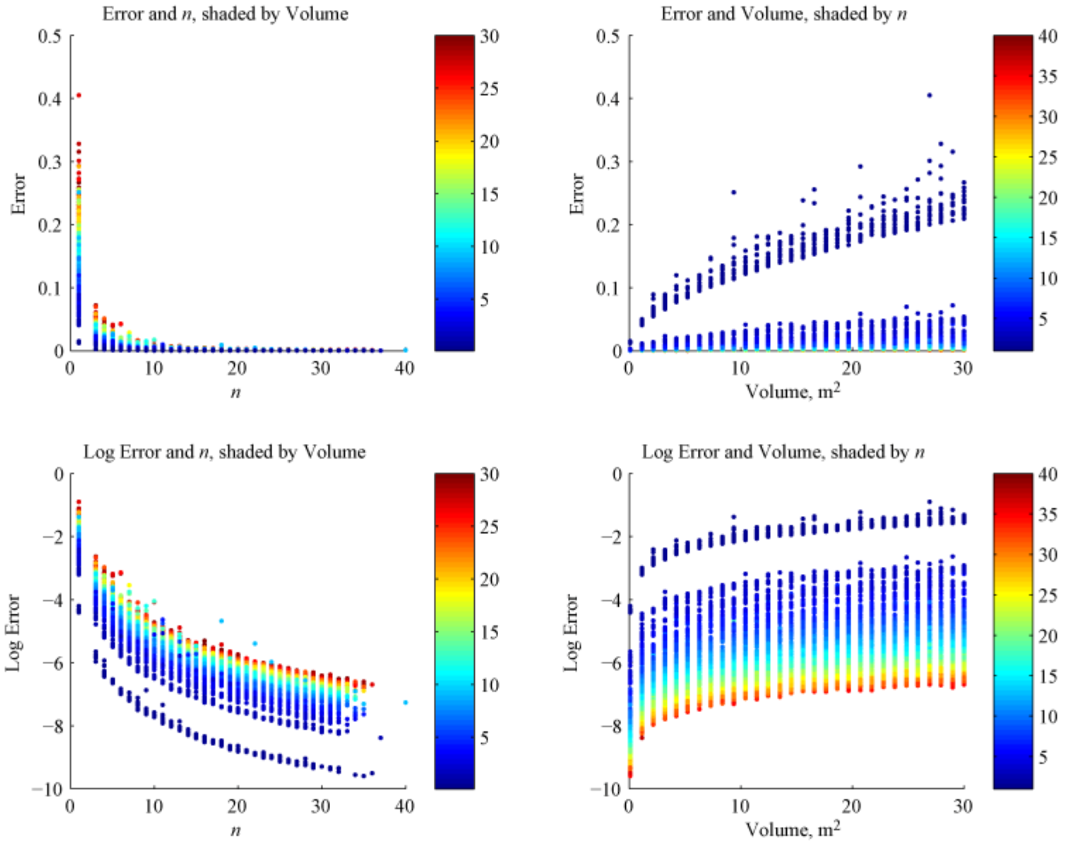


Figure 4-2: Variation of mean squared error with n and volume. The large separation in the upper right plot is due to the large reduction in error from $n = 1$ to $n = 2$.

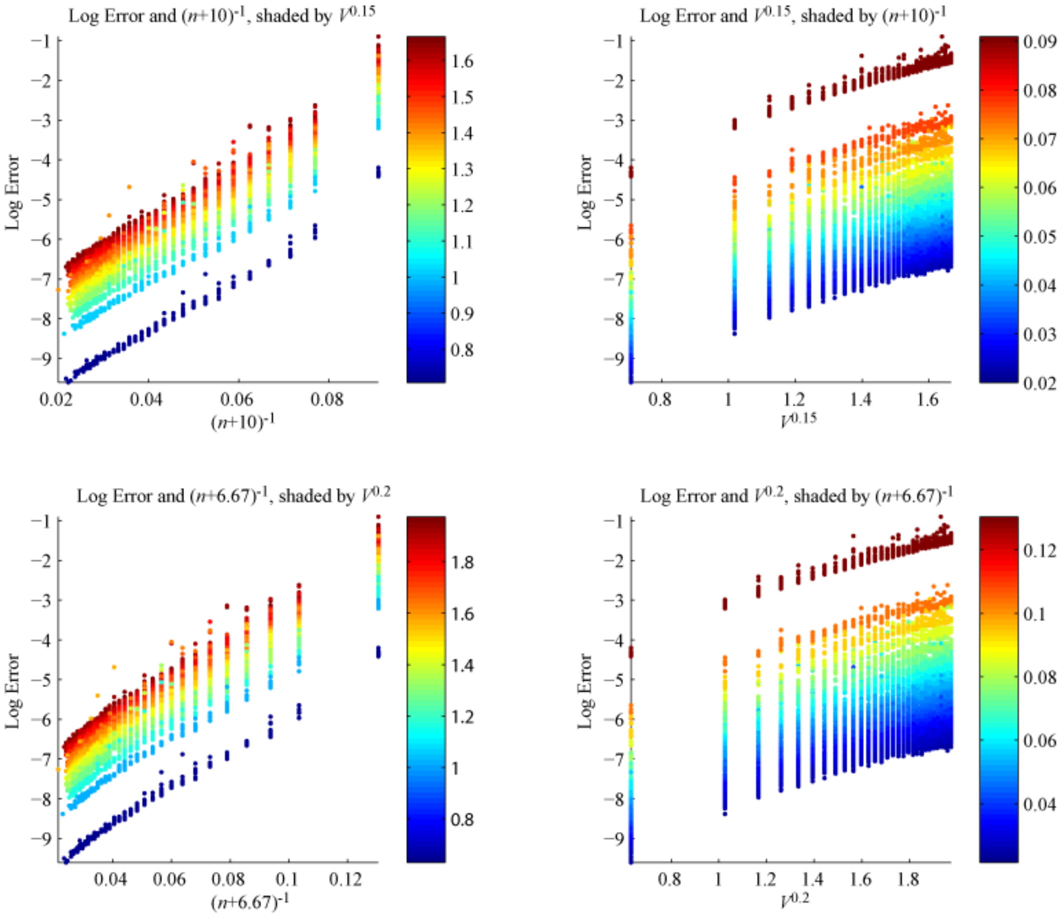


Figure 4-3: Linearized plots of mean square error with functions of *n* and volume.

Results for the other variogram models are summarized in Table 4-2 for triangular elements and Table 4-3 for quadrilaterals. Error surfaces are displayed as contours for spherical, exponential, and Gaussian variograms in Figure 4-4 to Figure 4-9. The presence of a nugget effect leads to a lower error as identified previously. In Figure 4-4 through Figure 4-9, this appears as a change in the error scale by a factor of approximately $70\% = 1 - C_0$.

Different variogram models yield different error surfaces as well. The occurrence of decreasing error from exponential to spherical to Gaussian variograms is due to their shape. The exponential covariance function has the largest gradient near the origin and more points are necessary to approximate the average covariance. The Gaussian variogram has a gradient of zero at the origin and fewer points are required to achieve the same accuracy for a given volume. The Spherical model falls in-between.

Volume shape has some effect on the quality of the regression. The residual error using triangular elements was larger than quadrilateral elements indicating more variability in the average covariances for given volumes. This can be attributed to the occurrence of randomly generated slivers or low quality triangles that cover a large extent of the covariance function in one direction and a small extent in the other. This leads to smaller average covariance and larger error for smaller volumes. Quadrilaterals tend to be more symmetric in their coverage of the covariance function for a given volume leading to more consistent average covariance and results with lower error.

Any variogram model can be used to generate the data required to perform regression and obtain parameters τ, β, A_0, A_1 and A_2. Results are used to choose an acceptable mean squared error, which is used to extract a contour that relates element volume and n to use in the discretization process. An error contour of Equation 4.14 is defined by Equation 4.15, with the resulting n being rounded to the nearest positive integer. Rounding is not necessary if the result is converted into a volume constraint, V/n, for the discretization process.

$$n\left(V,\varepsilon_Z^2\right)=\left[\frac{-A_0-A_1V^{\beta}+\log\varepsilon_Z^2}{A_2}\right]^{-1}-\tau \tag{4.15}$$

In practice, several variables are involved and each may have a different variogram leading to different error surfaces. Choosing one variogram model that contains all others involved allows a single error surface to be modeled and utilized for choosing n. This concept was discussed in Chapter 2, but in terms of deriving a volume for which the error is small when restricted to $n = 1$. In this case, the volume curve for constant n is defined by Equation 4.16. Curves are shown for $n = 1$ and the spherical, exponential and Gaussian cases with no nugget effect and quadrilateral elements in Figure 4-10. An interesting point is the error incurred in representing the discretizing elements with one point is mitigated in the averaging process. Using the exponential, quadrilateral, zero nugget effect case (Figure 4-6) for an example and assuming an error of 0.005 is targeted, a volume of 30 m^2 would be discretized by roughly 15 points. Each point represents an approximate volume of $V/n = 2$ m^2. These volumes are represented by 1 point, which involves an error of roughly 0.08 when scaled off the curve in Figure 4-10.

$$V\left(n,\varepsilon_Z^2\right)=\left[\frac{-A_0-A_2\left(n+\tau\right)^{-1}+\log\varepsilon_Z^2}{A_1}\right]^{\frac{1}{\beta}} \tag{4.16}$$

Findings are similar in three dimensions and the same error surface model is applicable. For the same set of example variograms, the resulting error function surfaces are summarized in Table 4-4 for tetrahedral elements and Table 4-5 for hexahedral elements. Error contour surfaces for spherical, exponential, and Gaussian variograms are shown in Figure 4-11 through Figure 4-16. There is not a significant increase in the number n required to obtain a low error with the added dimension. The growth of n is based on point spacing. Using the spherical variogram, zero nugget case for an example, a triangle with an area of 16 m^2 requires approximately 10 points (approximate point spacing of $(16/10)^{1/2} = 1.26$) to achieve an error of 0.003, while a tetrahedron having similar dimensions with a volume of 64 m^3 ($16^{3/2}$) requires roughly 25 points (approximate point spacing of $(64/25)^{1/3} = 1.36$). These numbers are based on points being distributed using the programs Triangle and TetGen as explained previously.

Choosing the error is dependent on the reservoir modeling application, practicality, and user discretion. One should attempt to minimize the error while maintaining a reasonable number of fine scale grid elements. Small problems with few coarse grid elements could use a high degree of refinement and achieve very low error. Larger models may have to be limited to fewer n and relaxed error criteria. Arithmetic averaging is not the only upscaling process involved in reservoirs and the choice of n cannot be limited to them. An upscaling method for transmissibility using the MPFA method is developed in Section 4.2. This process is different than an arithmetic average and results in different criteria for choosing n.

Table 4-2: Error function regression surfaces for triangular elements.

Model	Nugget	τ	β	A_0	A_1	A_2	Fit Error
Spherical	0.0	6.67	0.100	-13.91	4.51	48.62	1.41E-3
	0.3	6.67	0.100	-14.26	4.51	48.62	6.90E-4
Exponential	0.0	6.67	0.091	-13.52	4.89	47.18	1.01E-3
	0.3	6.67	0.091	-13.88	4.89	47.18	5.00E-4
Gaussian	0.0	4.67	0.083	-34.69	17.41	54.77	1.51E-3
	0.3	4.67	0.083	-34.69	17.13	54.66	7.40E-4

Table 4-3: Error function regression surfaces for quadrilateral elements.

Model	Nugget	τ	β	A_0	A_1	A_2	Fit Error
Spherical	0.0	6.67	0.200	-11.29	1.89	46.83	5.07E-5
	0.3	6.67	0.200	-11.65	1.89	46.83	2.49E-5
Exponential	0.0	7.67	0.091	-13.69	4.83	52.91	1.67E-4
	0.3	7.67	0.091	-14.05	4.83	52.91	8.19E-5
Gaussian	0.0	6.67	0.083	-35.80	18.11	68.74	2.06E-5
	0.3	6.67	0.083	-35.87	17.90	68.59	1.03E-5

Table 4-4: Error function regression surfaces for tetrahedral elements.

Model	Nugget	τ	β	A_0	A_1	A_2	Fit Error
Spherical	0.0	6.67	0.286	-8.60	0.61	38.51	3.42E-5
	0.3	6.67	0.286	-8.95	0.61	38.51	1.68E-5
Exponential	0.0	6.67	0.091	-10.33	2.98	38.01	5.58E-5
	0.3	6.67	0.091	-10.69	2.98	38.01	2.73E-5
Gaussian	0.0	7.67	0.083	-26.78	12.38	51.02	4.79E-6
	0.3	7.67	0.083	-27.13	12.38	51.02	2.35E-6

Table 4-5: Error function regression surfaces for hexahedral elements.

Model	Nugget	τ	β	A_0	A_1	A_2	Fit Error
Spherical	0.0	12.67	0.286	-9.43	0.61	78.99	2.64E-5
	0.3	12.67	0.286	-9.79	0.61	78.99	1.29E-5
Exponential	0.0	12.67	0.100	-10.87	2.67	78.43	3.66E-5
	0.3	12.67	0.100	-11.22	2.67	78.43	1.79E-5
Gaussian	0.0	18.67	0.083	-28.58	12.62	139.36	2.48E-6
	0.3	18.67	0.083	-28.94	12.62	139.36	1.21E-6

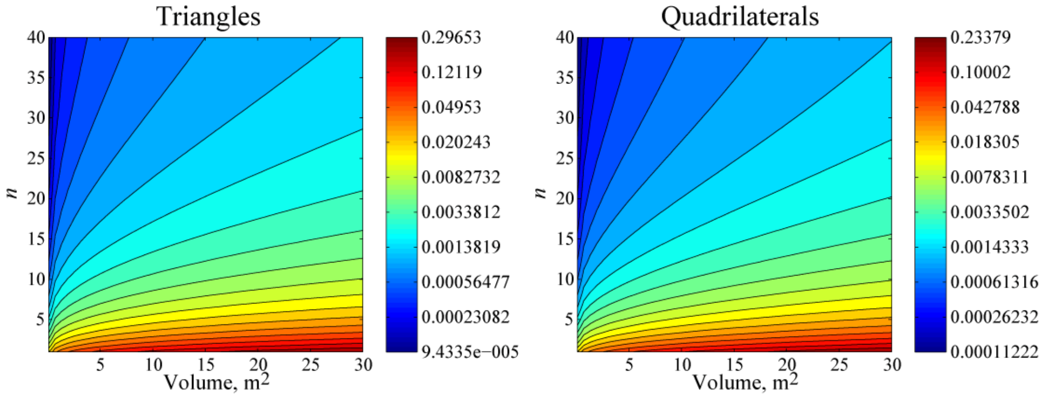


Figure 4-4: Error surfaces for a spherical variogram: range 10 m; nugget effect 0.

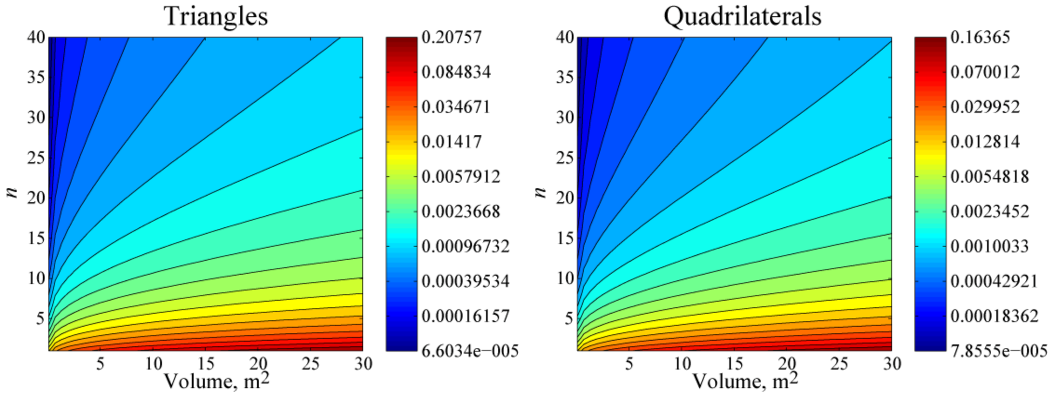


Figure 4-5: Error surfaces for a spherical variogram: range 10 m; nugget effect 0.3.

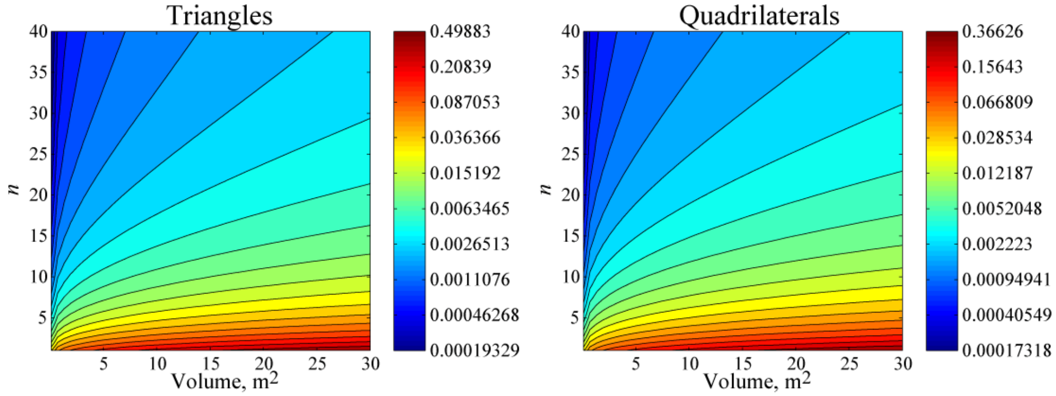


Figure 4-6: Error surfaces for an exponential variogram: range 10 m; nugget effect 0.

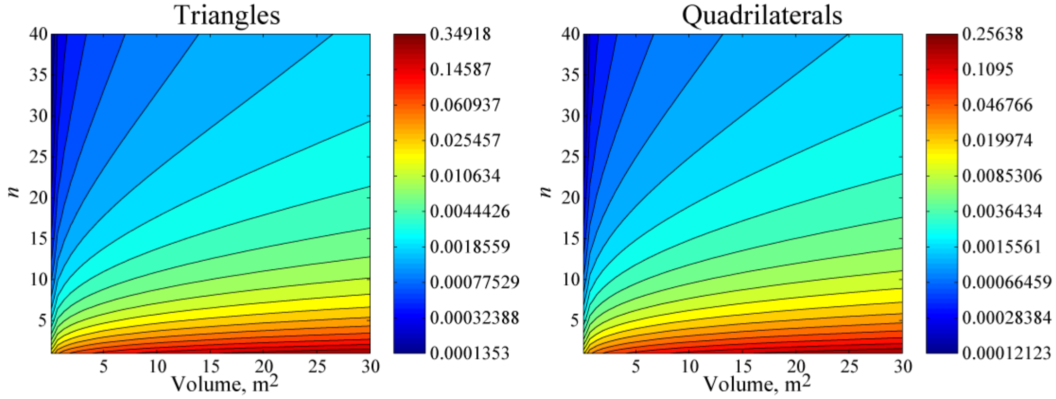


Figure 4-7: Error surfaces for an exponential variogram: range 10 m; nugget effect 0.3.

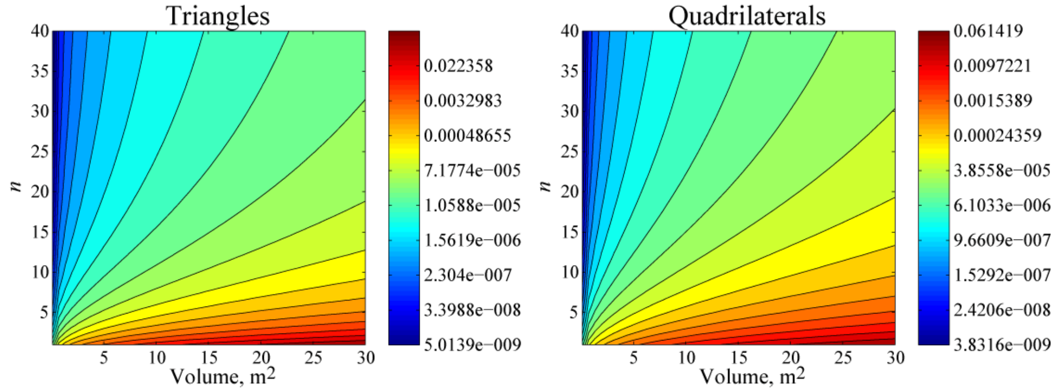


Figure 4-8: Error surfaces for a Gaussian variogram: range 10 m; nugget effect 0.

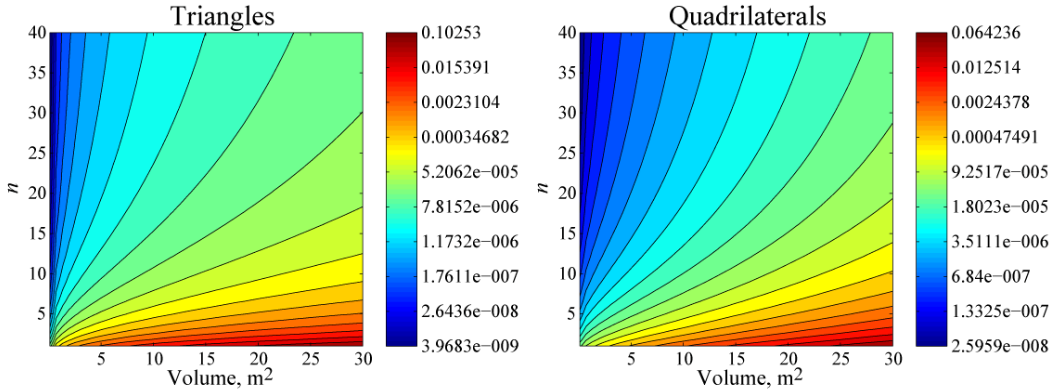


Figure 4-9: Error surfaces for a Gaussian variogram: range 10 m; nugget effect 0.3.

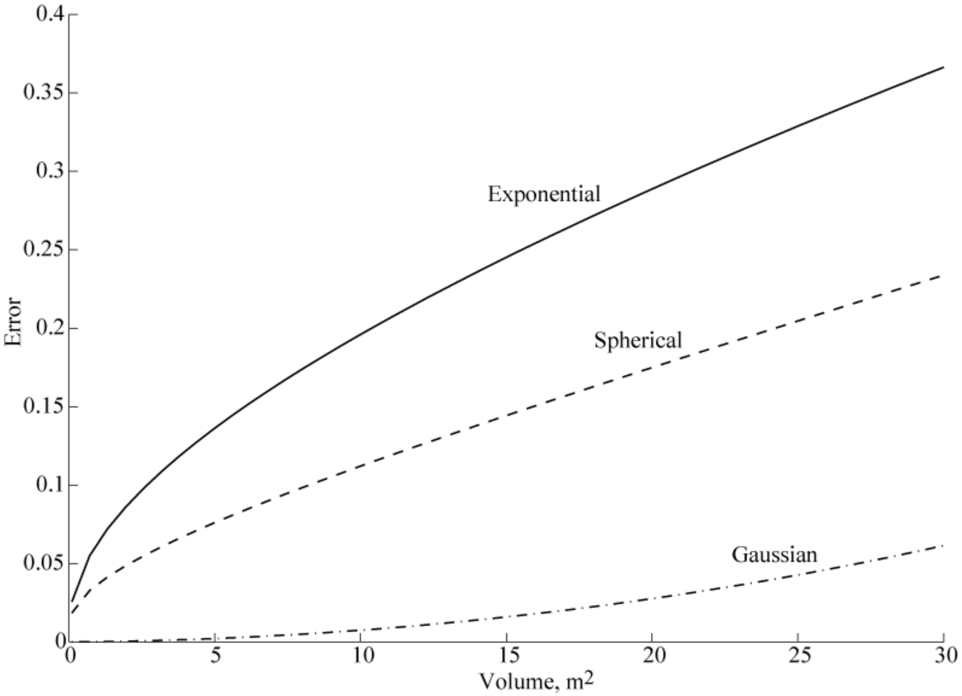


Figure 4-10: Error curves for *n* = 1.

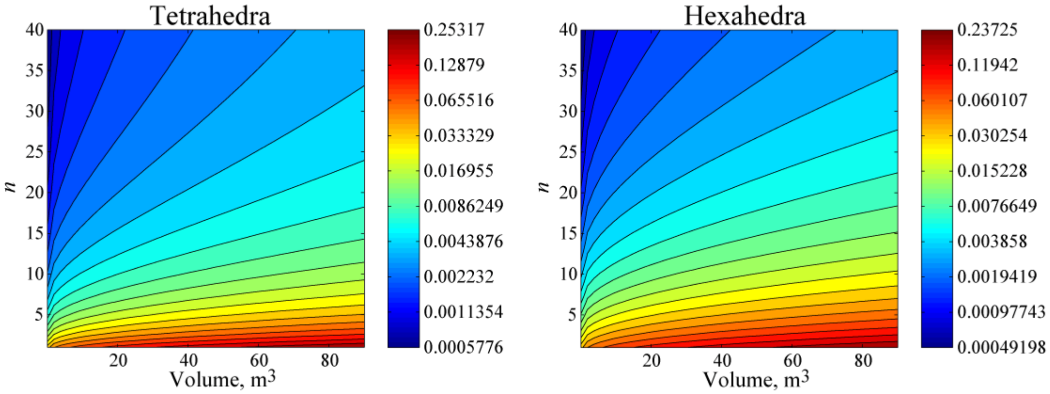


Figure 4-11: Error surfaces for a spherical variogram: range 10 m; nugget effect 0.

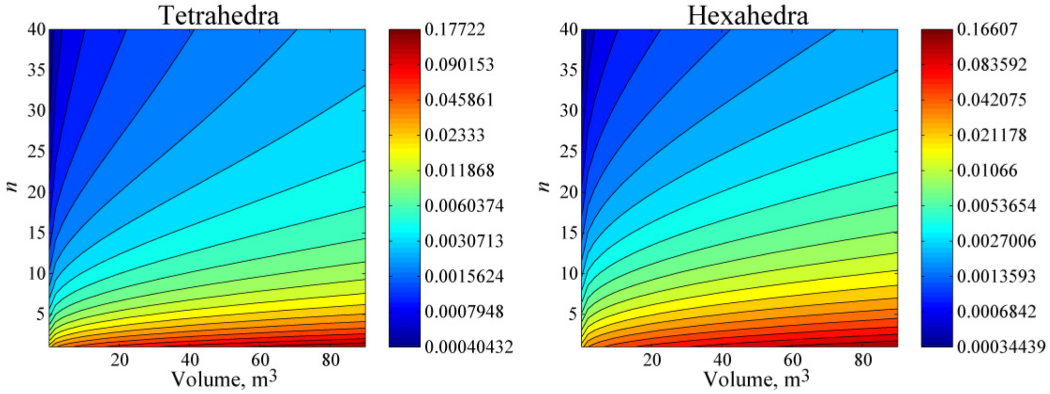


Figure 4-12: Error surfaces for a spherical variogram: range 10 m; nugget effect 0.3.

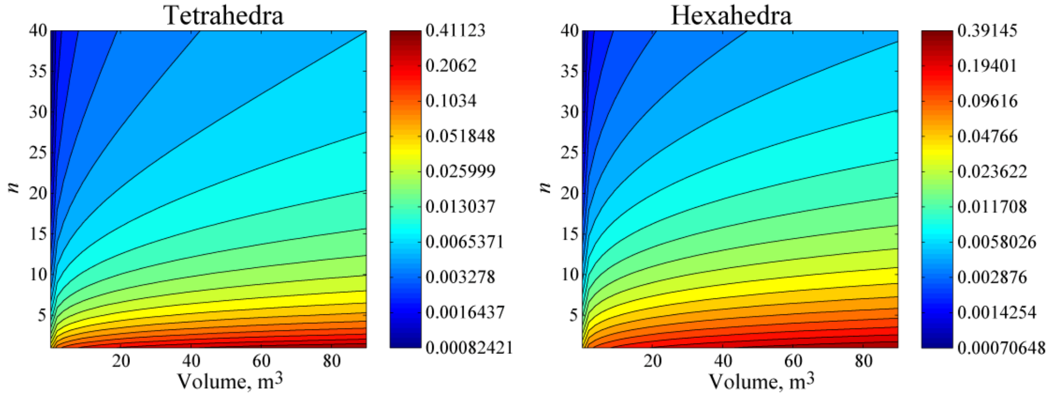


Figure 4-13: Error surfaces for an exponential variogram: range 10 m; nugget effect 0.

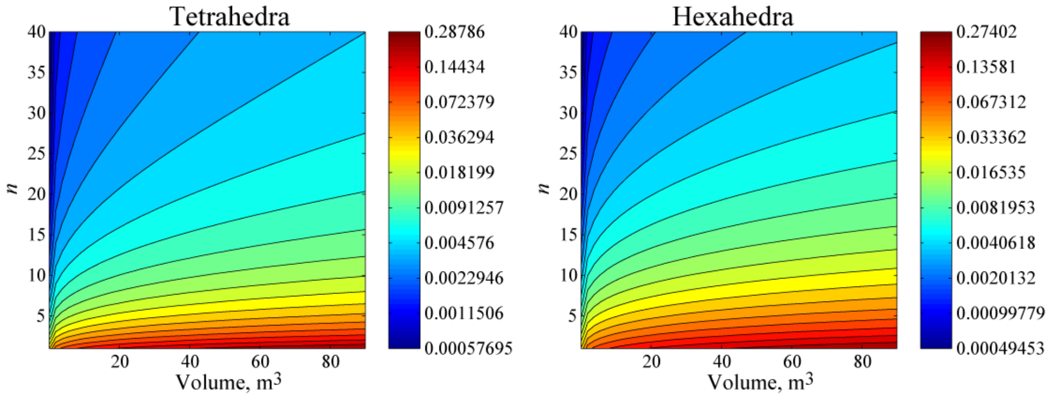


Figure 4-14: Error surfaces for an exponential variogram: range 10 m; nugget effect 0.3.

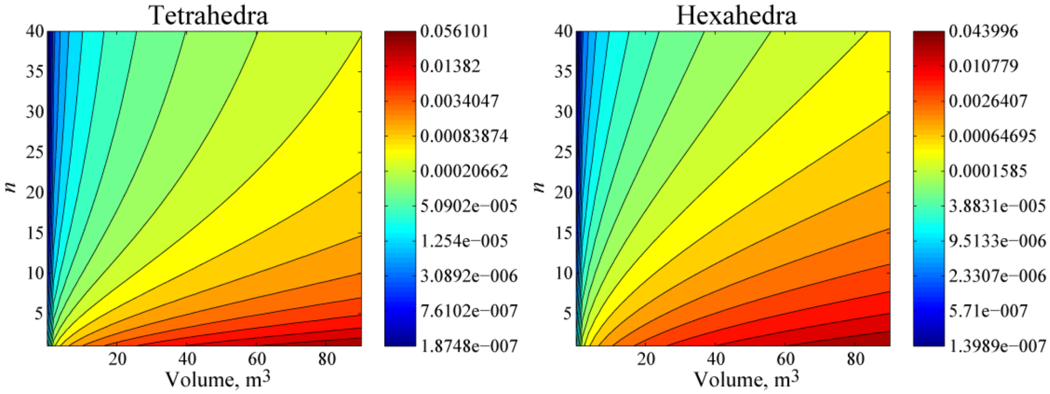


Figure 4-15: Error surfaces for a Gaussian variogram: range 10 m; nugget effect 0.

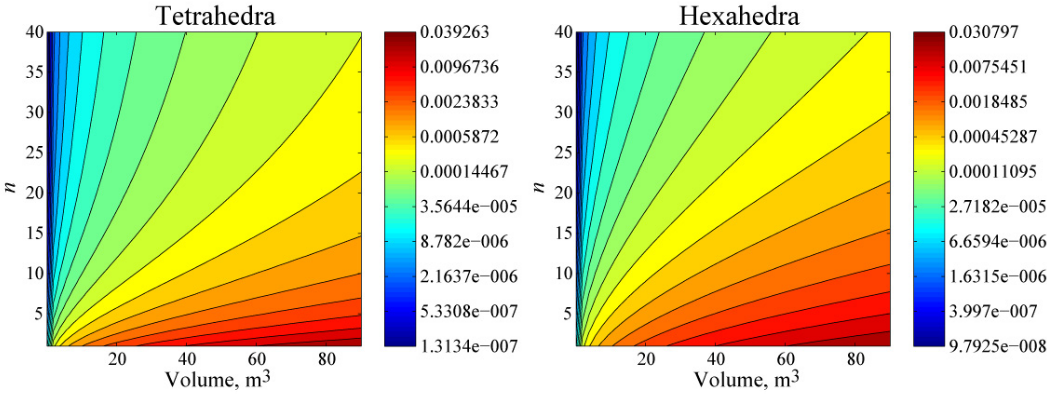


Figure 4-16: Error surfaces for a Gaussian variogram: range 10 m; nugget effect 0.3.

4.2 Flow-based Upscaling on Simplex Grids

In this section, a flow based upscaling approach for transmissibility is developed using the multipoint flux approximation. The fine grid is assumed to be a simplex grid and the coarse grid a PEBI grid or polygonal grid; although the methods developed are theoretically applicable to other grids that the MPFA method is applicable to including triangular, tetrahedral, quadrilateral, and hexahedral grids. Details of the MPFA method for triangular and polygonal grids in two dimensions are presented by Aavastmark et al (1998a,b) and in three dimensions by Verma and Aziz (1997). The focus of this dissertation is the single phase pressure equation given by Equation 4.17, where ∇ is the gradient operator, ρ and μ are the fluid density and viscosity, $\mathbf{k}(\mathbf{u})$ is a symmetric positive definite permeability tensor dependent on spatial location $\mathbf{u}$, p is the pressure, and q are the sources and sinks that typically model production and injection wells. The component in parenthesis is the Darcy velocity.

$$-\nabla \cdot \left(\frac{1}{\mu} \mathbf{k}(\mathbf{u}) \nabla p \right) = \frac{q}{\rho} \tag{4.17}$$

$$\mathbf{k} = \begin{bmatrix} k_{xx} & k_{xy} & k_{xz} \\ k_{yx} & k_{yy} & k_{yz} \\ k_{zx} & k_{zy} & k_{zz} \end{bmatrix} \tag{4.18}$$

Permeability tensors given by Equation 4.18 are populated on the discretized grid using geostatistics. They are often assumed diagonal and isotropic ($k_{xx} = k_{yy} = k_{zz}$ and $k_{xy} = k_{xz} = k_{yz} = 0$) or transversely isotropic ($k_{xx} = k_{yy} \neq k_{zz}$ and $k_{xy} = k_{xz} = k_{yz} = 0$). Full symmetric tensors are obtained through upscaling from the discretized grid to the unstructured flow simulation grid; however, most flow simulation methods use transmissibilities that are derived from permeability, rather than the permeability tensors directly.

The MPFA method is a control volume scheme to solve Equation 4.17, where each element of the coarse grid forms a control volume. The divergence theorem, that is, conservation of mass (Equation 4.19) is used to construct a system of equations to solve for the pressure, where V is a control volume with bounding surface ∂V, $\mathbf{n}$ is the unit normal pointing out of the surface, S (Figure 4-17), and the viscosity and density terms have been dropped for simplicity. Approximations such as the MPFA are invoked to simplify the integrals. Transmissibility across grid elements is derived from permeability to discretize the flow equations. Upscaling is used to derive transmissibility at a larger scale much like permeability is upscaled to full tensors.

$$\int_V q\,dV = -\int_{\partial V} \left(\mathbf{k} \nabla p \right) \cdot \mathbf{n}\,dS \tag{4.19}$$

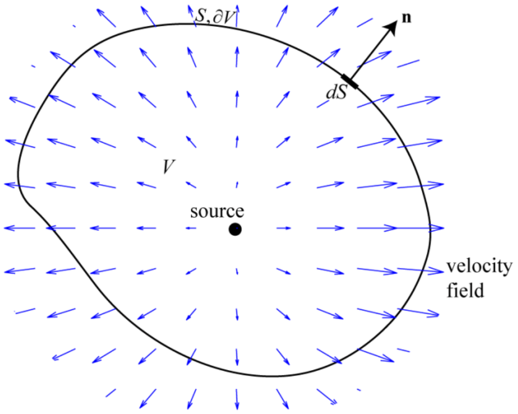


Figure 4-17: Illustration of the divergence theorem for a two dimensional volume. The mass of fluid originating from the source must equal the mass of fluid crossing the surface indicated by the solid line.

4.2.1 MFPA Background

The MPFA method is defined for both triangular grids and polygonal grids that were discussed in Chapter 1. Typically, the dual grid of a triangulation is a Voronoi grid or a PEBI grid; however, this is not always the case for the MPFA formulation even when the triangulation is Delaunay. Reversing the argument and considering a PEBI grid as primal, the dual grid for the MPFA formulation is not necessarily a triangulation. Both cases are shown in Figure 4-18.

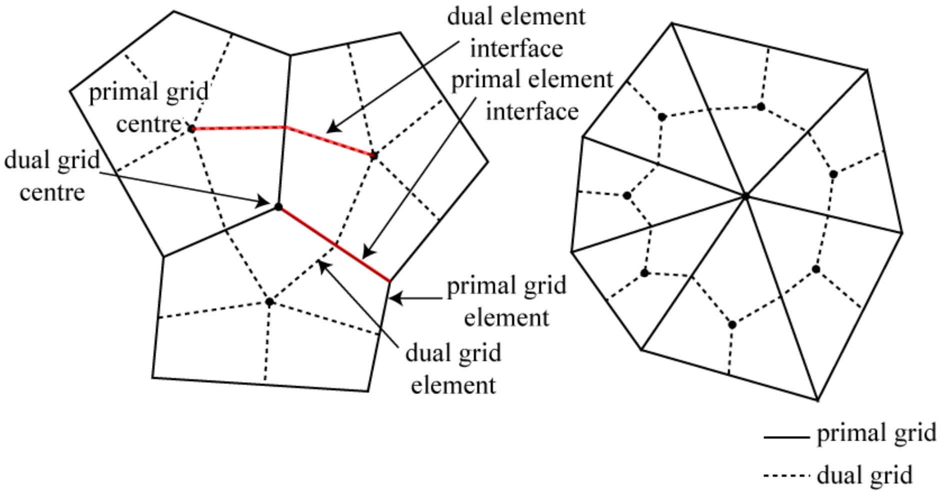


Figure 4-18: Primal PEBI grid (left) and primal triangular grid (right) in the MPFA formulation.

Dual grids for both cases are generated by connecting the primal grid element centers through the midpoints of the interfaces between neighbouring elements. If this is relaxed so that the dual grid element edges are permitted to intersect the primal grid interfaces at any point, it is possible to design PEBI grids so that the dual is a triangulation. The criterion for this is that the PEBI grid vertices fall inside the triangles or tetrahedra that generate them.

Elements of the dual grid are termed interaction regions as they describe how fluids interact between the control volumes that generate them. Components of interaction regions are defined with reference to a triangular region, but they apply to polygonal interaction regions as well. In Figure 4-19, points labelled p_i are the centers of three adjacent control volumes, V_i, b is the interaction region barycenter, m_i are the midpoints along edges $\overline{\mathrm{p}_i\mathrm{p}_{i+1}}$, $\mathbf{n}_i$ are the normal vectors pointing from p_i to p_{i+1}, and $\mathbf{v}_i$ are normal vectors to the face opposite of p_i. The normal vectors are area vectors in that their length is equal to the area of the associated face. Vector $\mathbf{v}_i^+$ and $\mathbf{v}_i^-$ have lengths equal to edges $\overline{\mathrm{p}_i\mathrm{m}_i}$ and $\overline{\mathrm{p}_i\mathrm{m}_{i-1}}$ respectively. Interfaces defined by $\overline{b\mathrm{m}_i}$ form the flux interfaces between control volumes V_i and V_{i+1}. The triangles formed by $\Delta\mathrm{p}_i\mathrm{m}_i\mathrm{m}_{i-1}$ are variation triangles with areas A_i.

Interaction region parameters in three dimensions for tetrahedral regions (Figure 4-20) are a straightforward extension of the triangular case. There are five barycenters, one for each triangle, b_{ijk}, and one for the tetrahedron, b. Six midpoints, m_{ij}, along the tetrahedron edges are involved. Normal vectors, $\mathbf{n}_{ij}$ and $\mathbf{v}_i$ are associated with faces and their lengths are equal to the face areas.

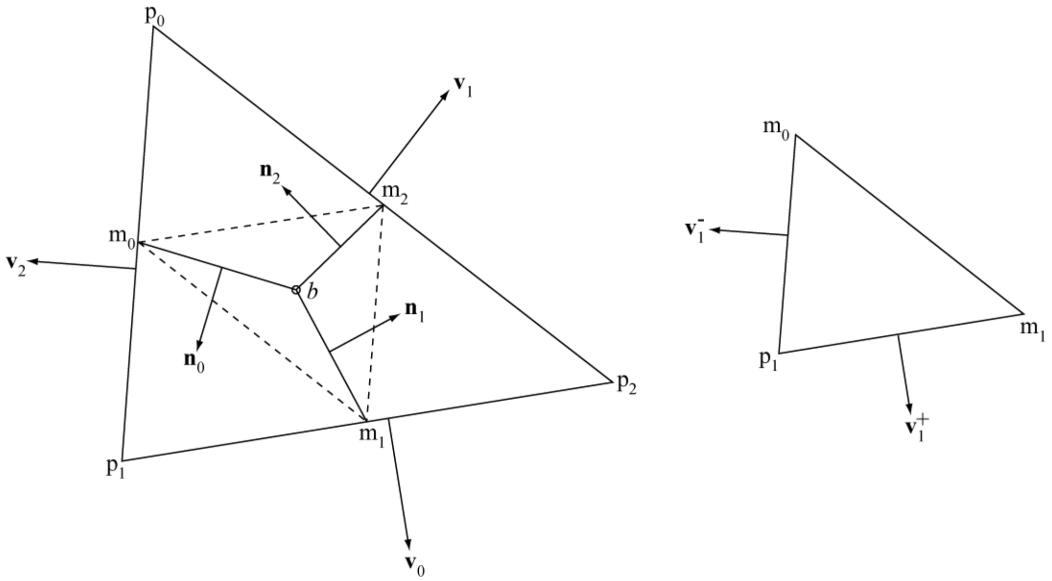


Figure 4-19: Triangular interaction region (left) and variation triangle (right). Vectors are not drawn to scale.

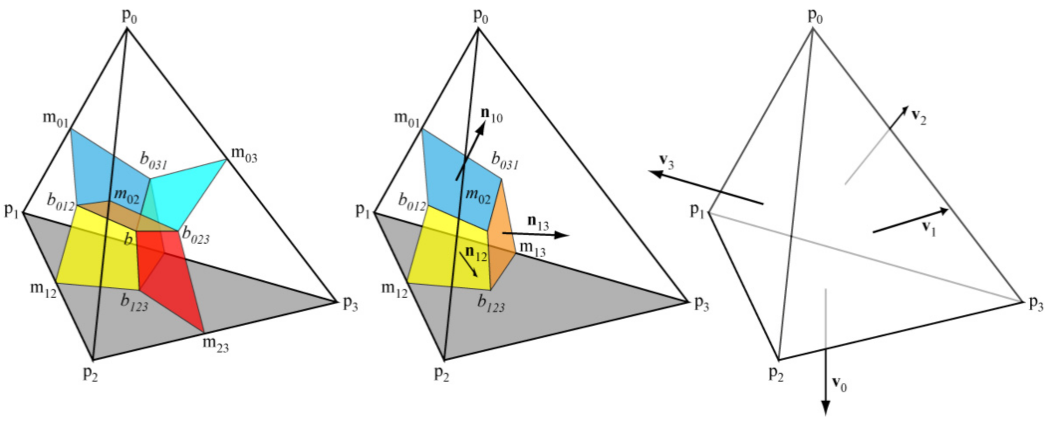


Figure 4-20: Three views of a tetrahedral interaction region. Vectors not drawn to scale.

Interfaces within the interaction regions are used to discretize Equation 4.19. The two dimensional case is used to demonstrate the MPFA methodology. In the MPFA method, multiple points are used to express the flux, q_i, across each of the interfaces ($\overline{b\mathrm{m}_i}, i = 0,1,2$) in terms of the pressures in the adjacent control volumes ($p_i, i = 0,1,2$) given by Equation 4.20, where t_{ij} are weights. If all pressures are identical, that is, if there is no pressure gradient within the interaction region, the flux across all interfaces must be zero; therefore the weights must sum to zero (Equation 4.21). Otherwise, there would be flow across an interface with no force driving it. Using matrix notation, fluxes across all interfaces within an interaction region are expressed in Equation 4.22, where **T** is the transmissibility matrix for the interaction region.

$$q_i = \sum_{j=0}^{2} t_{ij} p_j \qquad 4.20$$

$$\sum_{j=0}^{2} t_{ij} = 0, \quad i = 0,1,2 \qquad 4.21$$

$$\mathbf{q} = \mathbf{T}\mathbf{p} = \begin{bmatrix} t_{00} & t_{01} & t_{02} \\ t_{10} & t_{11} & t_{12} \\ t_{20} & t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ p_2 \end{bmatrix} \qquad 4.22$$

The weights are transmissibility coefficients, where t_{ij} specifies the contribution of pressure p_j to flux q_i across flux interface i. Derivation of t_{ij} depends on the geometry of the problem, the underlying permeability field, and assumptions regarding the continuity of pressure and flux within the interaction regions. It is assumed that pressure varies linearly within the variation triangles and that pressure is continuous across the edge midpoints. Pressure is permitted to be discontinuous along the flux interfaces at all other points. Flux is assumed continuous over the flux interfaces.

Assuming pressure varies linearly within a variation triangle, the pressure gradient is constant and Equation 4.19 is simplified to Equation 4.23 for interface i and control volume j. The index j is superscripted for $\mathbf{k}$ to indicate it is the permeability tensor of control volume j. The area of the interface is equal to the length of the normal vector.

$$q_{ij} = -(\mathbf{k}^j \nabla p_j) \cdot \mathbf{n}_i \tag{4.23}$$

The pressure gradient, ∇p_j, is dependent on the pressures at the vertices of the variation triangle $\Delta \mathrm{p}_i \mathrm{m}_i \mathrm{m}_{i-1}$, with the pressures at midpoints m_i and m_{i-1} being unknown. The purpose of the flux continuity assumption previously mentioned is to provide sufficient conditions to account for the unknown midpoint pressures. Flux continuity is achieved by setting the flux across interface i due to pressure gradient in variation triangle j equal to the flux across interface i due to pressure gradient in variation triangle $j+1$ (Equation 4.24). The pressure gradient within each variation triangle is calculated from Equation 4.25, where $\overline{p_i}$ is the pressure at midpoint m_i. Substituting Equation 4.25 into 4.24 and solving results in Equation 4.26 for the transmissibilities, where $\mathbf{T}$ is the transmissibility matrix used in Equation 4.22, ω_{ij} is defined by Equation 4.27 and each matrix is defined afterwards in Equation 4.28.

$$\begin{aligned}(\mathbf{k}^0 \nabla p_0) \cdot \mathbf{n}_0 &= (\mathbf{k}^1 \nabla p_1) \cdot \mathbf{n}_0 \\ (\mathbf{k}^1 \nabla p_1) \cdot \mathbf{n}_1 &= (\mathbf{k}^2 \nabla p_2) \cdot \mathbf{n}_1 \\ (\mathbf{k}^2 \nabla p_2) \cdot \mathbf{n}_2 &= (\mathbf{k}^0 \nabla p_0) \cdot \mathbf{n}_2\end{aligned} \tag{4.24}$$

$$\nabla p_i = -\mathbf{v}_i^+ \cdot \frac{\overline{p}_{i-1} - p_i}{2A_i} - \mathbf{v}_i^- \cdot \frac{\overline{p}_i - p_i}{2A_i} \tag{4.25}$$

$$\mathbf{T} = \mathbf{A}(\mathbf{A}-\mathbf{C})^{-1}(\mathbf{D}-\mathbf{B}) + \mathbf{B} \tag{4.26}$$

$$\omega_{ij}^{\pm} = \frac{\left(\mathbf{k}^j \cdot \mathbf{n}_i\right)^T \mathbf{v}_j^{\pm}}{2A_j} \tag{4.27}$$

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} \omega_{00}^- & 0 & \omega_{00}^+ \\ \omega_{11}^+ & \omega_{11}^- & 0 \\ 0 & \omega_{22}^+ & \omega_{22}^- \end{bmatrix} \\ \mathbf{B} &= \begin{bmatrix} -\omega_{00}^+ - \omega_{00}^- & 0 & 0 \\ 0 & -\omega_{11}^+ - \omega_{11}^- & 0 \\ 0 & 0 & -\omega_{22}^- - \omega_{22}^+ \end{bmatrix} \\ \mathbf{C} &= \begin{bmatrix} \omega_{01}^+ & \omega_{01}^- & 0 \\ 0 & \omega_{12}^+ & \omega_{12}^- \\ \omega_{20}^- & 0 & \omega_{20}^+ \end{bmatrix} \\ \mathbf{D} &= \begin{bmatrix} 0 & -\omega_{01}^+ - \omega_{01}^- & 0 \\ 0 & 0 & -\omega_{12}^+ - \omega_{12}^- \\ -\omega_{20}^+ - \omega_{20}^- & 0 & 0 \end{bmatrix}\end{aligned} \tag{4.28}$$

Transmissibilities are determined similarly for polygonal interaction regions (Verma and Aziz, 1997; Verma, 1998; Aavastmark, 1998a,b). This approach is used to compute transmissibilities for all interaction regions in a grid such that the expression for the flux across all control volume interfaces in terms of the pressures within the control volumes is defined. The result is a discretized version of Equation 4.19 where the right hand side is the total flux out of a control volume. For single phase incompressible flow, this flux is zero unless the control volume acts as a source or sink (injector or producer). The unknowns to be solved for are the pressures in each control volume. For N control volumes, this forms a sparse system of N equations (Equation 4.29), where $\underline{\mathbf{T}}$ is the matrix of transmissibilities between all control volumes, **p** is the vector of pressures at the control volume centers, and **q** is the vector of out-fluxes for all control volumes. $\underline{\mathbf{T}}$ is constructed using **T** matrices from Equation 4.26 for each interaction region.

$$\underline{\mathbf{T}}\mathbf{p} = \mathbf{q} \tag{4.29}$$

After solving for **p**, transmissibilities are used to compute the flux across the interfaces between control volumes so that the transfer of fluid from sources to sinks can be assessed, in other words, flow simulation.

4.2.2 Local Transmissibility Upscaling

In this dissertation, there are two grids involved: a coarse grid for flow simulation, which is assumed to be a PEBI grid, polygonal grid, or a simplex grid, and a fine simplex grid for geological modeling. When using the MPFA method on the coarse scale grid, upscaling involves deriving the transmissibility matrices of the coarse scale interaction regions from the permeability field defined on the fine scale simplex grid. This approach constrains the geometry of the grid discretization, since one of the goals of grid design is conforming coarse and fine scale grids.

Equation 4.22 is used to derive the transmissibility matrix for coarse scale interaction regions using a flow based upscaling approach. In the local approach to flow based upscaling discussed in Chapter 1, the flow equations are solved on the fine scale using arbitrary boundary conditions. Resulting pressures and fluxes are determined for the coarse scale grid and either coarse scale permeability or transmissibility is solved for. Since Equation 4.22 involves 9 unknowns for triangular interaction regions, two sets of boundary conditions are required along with the constraint in Equation 4.21 to solve for the unknown transmissibility matrix.

Pressure boundary conditions are chosen to induce flow across the flux interfaces as well as allow flow in and out of the interaction region, i.e. the pressure gradient is non-zero (Equation 4.30). To agree with the derivation of MPFA transmissibility, the pressure along the edges $\overline{p_i p_j}$ varies linearly, but the pressure on the interior of the interaction region is unconstrained.

$$p_i > p_j \geq p_k; \quad i,j,k \in 0,1,2; \quad i \neq j \neq k \tag{4.30}$$

The two boundary condition scenarios and sum constraint form the system in Equation 4.31, where q_j^k is the j^{th} flux due to the k^{th} boundary condition and p_j^k is pressure at vertex j of boundary condition k. Choosing boundary conditions that forms the 3×2 identity matrix (Equation 4.32) provides a trivial solution defined in Equation 4.33.

$$\begin{bmatrix} q_0^1 & q_0^2 & 0 \\ q_1^1 & q_1^2 & 0 \\ q_2^1 & q_2^2 & 0 \end{bmatrix} = \begin{bmatrix} t_{00} & t_{01} & t_{02} \\ t_{10} & t_{11} & t_{12} \\ t_{20} & t_{21} & t_{22} \end{bmatrix} \times \begin{bmatrix} p_0^1 & p_0^2 & 1 \\ p_1^1 & p_1^2 & 1 \\ p_2^1 & p_2^2 & 1 \end{bmatrix} \tag{4.31}$$

$$\begin{bmatrix} q_0^1 & q_0^2 & 0 \\ q_1^1 & q_1^2 & 0 \\ q_2^1 & q_2^2 & 0 \end{bmatrix} = \begin{bmatrix} t_{00} & t_{01} & t_{02} \\ t_{10} & t_{11} & t_{12} \\ t_{20} & t_{21} & t_{22} \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.32}$$

$$\begin{aligned} t_{00} &= q_0^1 \quad t_{01} = q_0^2 \quad t_{02} = -(t_{00} + t_{01}) \\ t_{10} &= q_1^1 \quad t_{11} = q_1^2 \quad t_{12} = -(t_{10} + t_{11}) \\ t_{20} &= q_2^1 \quad t_{21} = q_2^2 \quad t_{22} = -(t_{20} + t_{21}) \end{aligned} \tag{4.33}$$

Obtaining q_j^k based on p_j^k involves solving for the pressure using Equation 4.29 within the interaction region and reconstructing the fluxes across the flux interfaces. More detail on the discretization follows.

Grid discretization for upscaling permeability to MPFA transmissibility matrices requires the interaction regions to act as boundary constraints. For a coarse scale PEBI grid and triangular interaction regions, the fine scale simplex grid conforms to the triangle edges (tetrahedron faces in 3D) and flux interfaces. Because flux interfaces form the PEBI grid elements, the simplex grid conforms to both the control volumes and the interaction regions. This situation is identical for coarse simplex grids, only control volumes refer to simplexes and interaction regions to polygons or polyhedra (Figure 4-21).

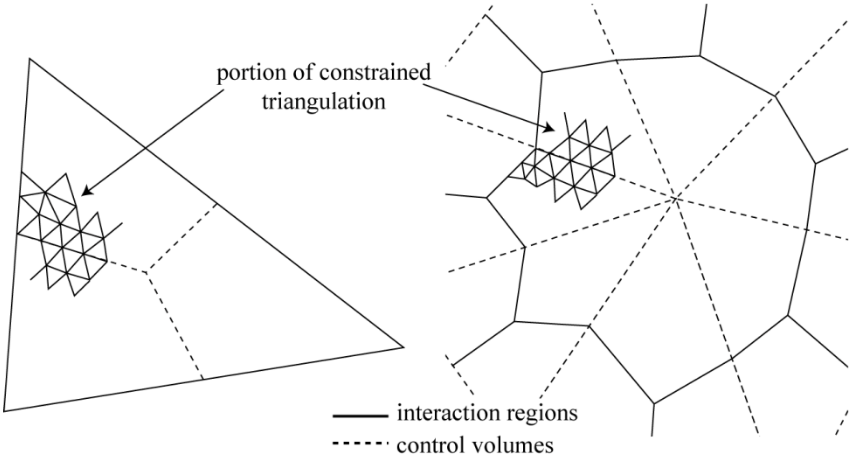
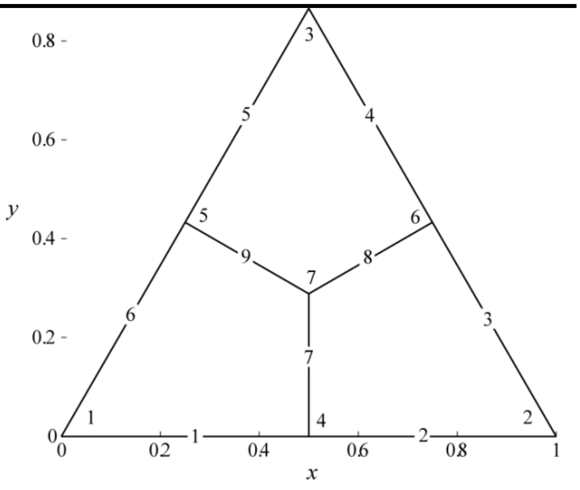


Figure 4-21: Constrained triangulation of triangular interaction regions (left) and polygonal interaction regions (right).

An advantage of constraining the discretization by the interaction regions is that for a given problem, upscaling to the MPFA transmissibility matrices is possible for both interaction region types using the same discretization. The coarse grid may be used as either a grid of control volumes, or a grid of interaction regions in setting up Equation 4.29. Development of geometry for upscaling is done in terms of triangular interaction regions.

Using the program Triangle (Shewchuk, 1996) to discretize an interaction region requires the boundaries and flux interfaces are input as segments. A single triangular interaction region involves seven points, which are the corner points, midpoints and barycenter. It also involves nine segments including six connecting the corner points to the midpoints and three connecting the barycenter to the midpoints. A sample text file for input into Triangle is given in Table 4-6 along with the interaction region geometry in Figure 4-22. Line 1 defines the number of vertices ($n_v = 7$), dimensionality (always 2), the number of attributes, and an option to use boundary markers (0 or 1). The next n_v lines include the vertex number, x and y coordinates, attributes if any, and boundary markers that are all set to 1 in this example. Line 9 is the start of the segment list that specifies the number of segments ($n_s = 9$) and number of boundary markers. The next n_s lines identify the segment number, vertices involved in the segment (always 2) as indexes into the vertex list, and the boundary markers. Ordering of the vertices and segments is arbitrary.

Table 4-6: Triangle input file example for triangular interaction region.

Line	Text	Result
1	`7 2 0 1`	 **Figure 4-22: Geometry and boundary markers corresponding to Triangle input in this table.**
2	`1 0.0000 0.0000 1`	
3	`2 1.0000 0.0000 1`	
4	`3 0.5000 0.8660 1`	
5	`4 0.5000 0.0000 1`	
6	`5 0.2500 0.4330 1`	
7	`6 0.7500 0.4330 1`	
8	`7 0.5000 0.2887 1`	
9	`9 1`	
10	`1 1 4 1`	
11	`2 4 2 2`	
12	`3 2 6 3`	
13	`4 6 3 4`	
14	`5 3 5 5`	
15	`6 5 1 6`	
16	`7 4 7 7`	
17	`8 6 7 8`	
18	`9 5 7 9`	
19	`0`	

Generating the discretization using Triangle involves several command line switches as well. Chapter 2 mentions user control of triangle area and quality. It is also possible to ensure the triangulation is conforming Delaunay (all Voronoi vertices fall within the triangulation) and to generate various output files including a segment file. The following is a sample command executed in Cygwin to triangulate the interaction region of Figure 4-22, with the result shown in Figure 4-23:

Program Switches Input file

```
$ ./triangle -pDq30a0.005e g
```

Specifying the input file requires only the name and not the extension, which is .poly by default. Command line switches used include:

- p: triangulate a planar straight line graph
- D: generate a conforming Delaunay triangulation
- q30: constrain interior angles to greater than 30 degrees
- a0.005: constrain the maximum area of triangles to be less than 0.005 square meters
- e: generate an edge list file.

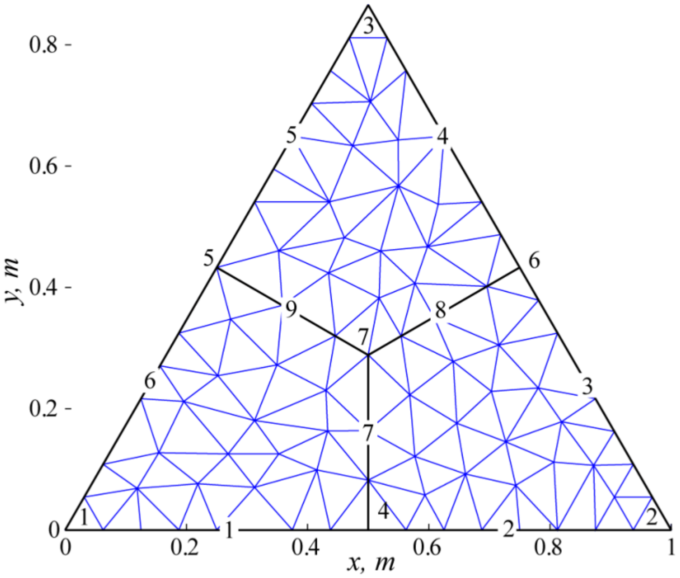


Figure 4-23: Interaction region discretized using Triangle.

Using segment boundary markers and generating the segment file is useful for identifying the edges that fine scale triangles are adjacent to. Segments in the file that coincide with an input segment are assigned the same boundary marker, making it straightforward to determine which triangles fall on the boundaries to recover the fluxes. Solving for interface fluxes on the triangulation using MPFA involves modeling the permeability in each triangle (Chapter 3), building the polygonal interaction regions, computing transmissibilities, specifying boundary conditions, and building and solving Equation 4.29. These steps are developed using the triangulation of Figure 4-23. Permeability is homogeneous with a value of 1 for all simplexes to validate the approach.

Polygonal interaction regions are built exactly like the triangular ones discussed in Section 4.2.1, the barycenters of triangles are connected to the midpoints of the triangle edges (Figure 4-24). Transmissibilities are compiled into $\underline{\mathbf{T}}$ from Equation 4.29 that is a sparse $N \times N$ matrix, where N is the number of fine scale triangles. N is equal to 139 in this case. Pressure boundary conditions are specified along the whole boundary of the coarse triangular interaction region based on the arbitrary pressures chosen for its vertices via Equation 4.30. Enforcing linear pressure variation along the outer edges of the coarse variation triangles and pressure continuity at the midpoints is done by assuming linear pressure variation between the assigned vertex pressures; however, only pressures along the bounding fine scale triangle barycenters are kept.

Assigning linearly varying pressures to the barycenters of the fine scale triangles is done as follows:

1. Identify the set of bounding triangles, S, which are those having an edge that borders only one triangle.
2. Compute the barycentric coordinates, $\boldsymbol{L}$, of all triangles in S relative to the coarse simplicial interaction region.
3. Interpolate the corner point pressures from the coarse interaction region, $\mathbf{p}$, to the fine scale triangles, $\mathbf{p}(S) = \boldsymbol{L}\mathbf{p}$.

Figure 4-25 shows the resulting pressure boundary conditions for one case. Triangles are shaded fully based on their pressures for visibility; however, the pressures exist only at the barycenters.

Pressure boundary conditions are applied to $\underline{\mathbf{T}}\mathbf{p} = \mathbf{q}$ for each boundary triangle, $S(k)$, by zeroing row $S(k)$ of $\underline{\mathbf{T}}$ and setting the same row of $\mathbf{q}$ equal to the pressure $p(S(k))$. Specified pressures are recovered exactly in the solution vector $\mathbf{p}$. Fluxes across the fine scale triangle interfaces are determined from the corresponding interaction region transmissibilities and the pressure solution. For the constant permeability case, the resulting pressure field has a constant gradient. Reconstructed fluxes are shown using velocity vectors, which are calculated as flux divided by the interface area (Figure 4-26).

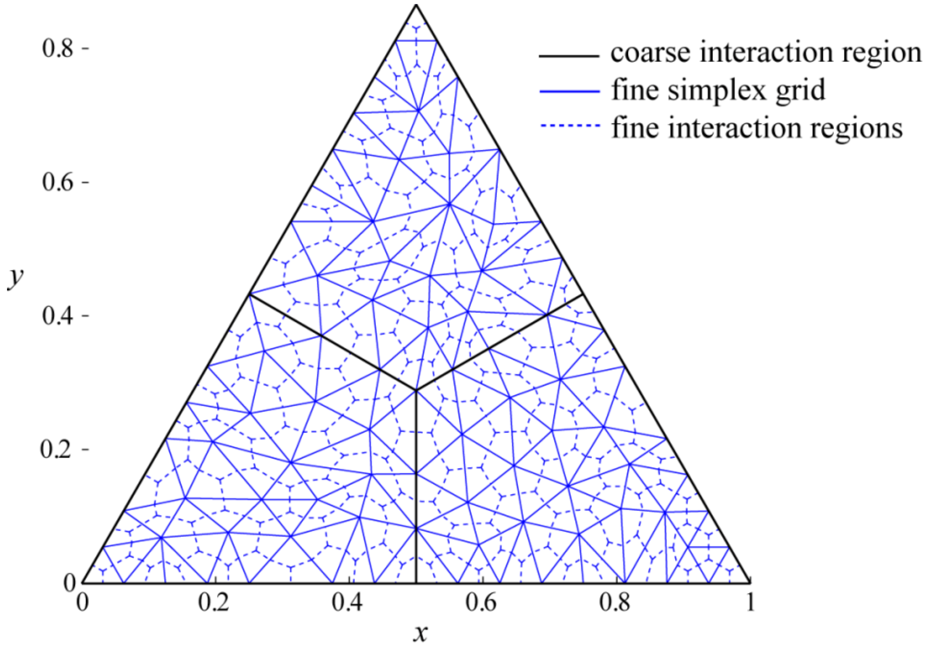


Figure 4-24: Fine scale polygonal interaction regions for MPFA upscaling.

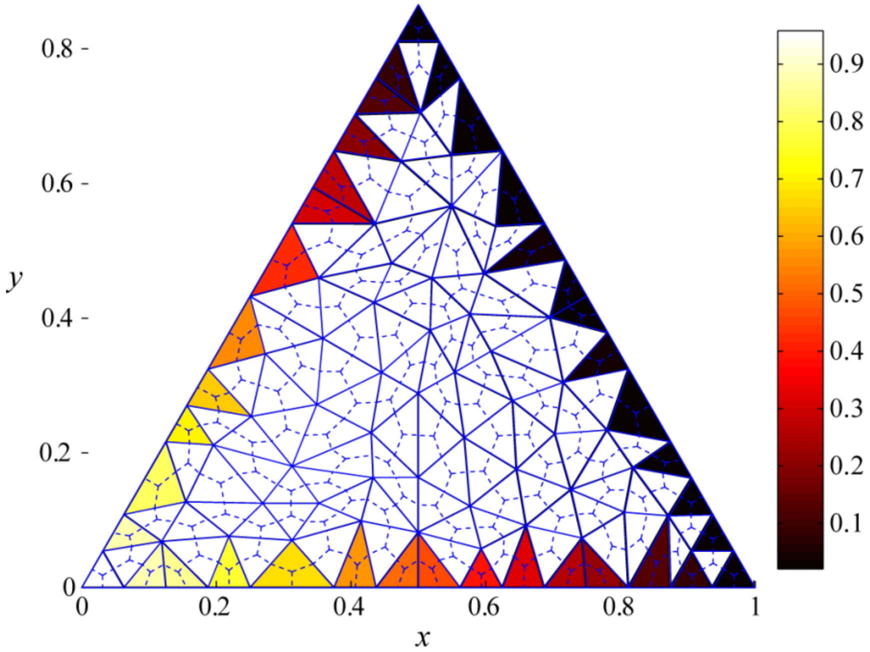


Figure 4-25: Linear pressure boundary conditions on a coarse triangular interaction region for the local fine scale pressure solution. Pressures are stored at the fine scale triangle barycenters.

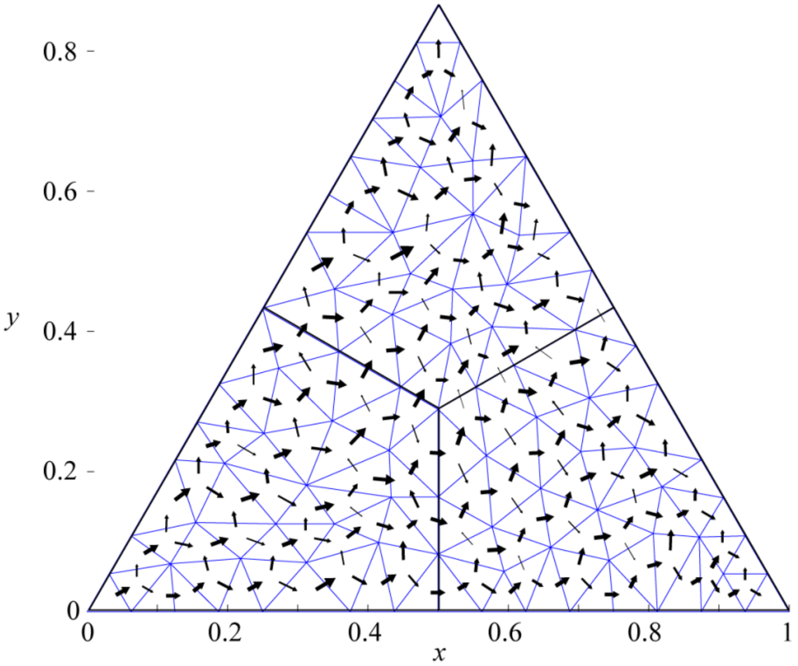


Figure 4-26: Velocity field with magnitude indicated by arrow width.

$$\begin{bmatrix} 0.289 & -0.288 & 0 \\ 0.002 & 0.288 & 0 \\ -0.289 & -0.001 & 0 \end{bmatrix} = \begin{bmatrix} t_{00} & t_{01} & t_{02} \\ t_{10} & t_{11} & t_{12} \\ t_{20} & t_{21} & t_{22} \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} t_{00} & t_{01} & t_{02} \\ t_{10} & t_{11} & t_{12} \\ t_{20} & t_{21} & t_{22} \end{bmatrix} = \begin{bmatrix} 0.289 & 0 & -0.289 \\ -0.289 & 0.289 & 0 \\ 0 & -0.289 & 0.289 \end{bmatrix} \qquad 4.34$$

The fluxes to compute the transmissibility matrix of the coarse scale interaction region are the sum of all fine scale fluxes along the coarse scale flux interfaces. The result is the first column of the left hand side matrix in Equation 4.32; however, the sign given to the fluxes must agree with the MPFA approach. Flux q_0^1 is the flux across interface 0 from p_0 to p_1, flux q_1^1 is the flux across interface 1 from p_1 to p_2, and flux q_2^1 is the flux across interface 2 from p_2 to p_0, with outflux being positive. The final set of equations and coarse transmissibility matrix is given by Equation 4.34.

The previous development of MPFA upscaling using constant permeability and an equilateral triangle confirms two results:

1. The pressure gradient across the coarse flux interface that is along the zero pressure edge is zero; therefore the flux must be zero there as indicated by the results.
2. Across the coarse flux interfaces with a pressure gradient of 1 and with permeability a constant 1 Darcy, the flux is equal to the area of the interface. For the equilateral triangle of edge length 1, the flux interfaces have a length of $3^{1/2}/6 \cong 0.28868$ that was recovered in Equation 4.34 with some error involved

due to numerical round-off and a triangulation consisting of triangles with different shapes and sizes.

The upscaling methodology presented in this section is new and is a contribution of this thesis. It is needed so that permeability models on high resolution simplex grids can be upscaled to a coarse scale polygonal grid. Methods shown apply to triangular interaction regions, but they can be generalized to other interaction regions including: polygonal interaction regions when the grid of control volumes is triangular and hexahedral interaction regions when the grid is structured. For the case with polygonal interaction regions, the geometric components are identical to those presented for triangular interaction regions. Pressures are defined at the triangular control volume centers. If the triangular and polygonal grids are dual, then the edge midpoints are coincident as are the edges connecting them. Differences include: the definition of flux interfaces and interaction region edges are interchanged; the number of equations is equal to the number of control volumes in an interaction region. If this number is n, then $n-1$ boundary conditions are required along with the constraint that transmissibilities sum to zero to solve for the transmissibility matrix of the polygonal interaction region.

4.2.3 Boundary Conditions

Using pressure boundary conditions with a constant gradient is limiting because regardless of the point pressures selected for p_i, p_j, and p_k, the coarse scale transmissibilities do not change. Incorporating this approach into any local-global upscaling regime (Chen et al, 2003; Chen and Durlofsky, 2006) is ineffective because the global pressures that are used as local boundary conditions have no effect on the resulting upscaled transmissibility matrix. Other approaches to construct boundary conditions for local upscaling problems are presented to achieve an improved description of the flow that actually occurs through an interaction region. One option found in the literature uses one dimensional pressure solves around the perimeter of a volume to be upscaled rather than a linear pressure variation as in Section 4.2.2 (Chen et al, 2003). The method is applicable to local MPFA upscaling.

Solving for perimeter pressures for a triangular interaction region is a one dimensional problem that is solved using two-point transmissibilities. Consider edge $\overline{\mathrm{p}_1\mathrm{p}_2}$ of the triangle used in Figure 4-23. Pressures assigned at points p_1 and p_2 are boundary conditions used to determine pressure at the midpoints of all fine scale triangles along the edge (Figure 4-27). Each triangle has a permeability $\mathbf{k}_i$. Two-point transmissibilities, T_{ij}, defined by Equation 4.35 are used to setup the system of equations, where $\mathbf{k}_i$ and Δx_i are respectively the permeability and width of element i along the edge and $T_i = \mathbf{k}_i/\Delta x_i$ is the transmissibility of element i.

$$T_{ij} = 2\left(\frac{1}{T_i}+\frac{1}{T_j}\right)^{-1} = 2\left(\frac{\mathbf{k}_i\mathbf{k}_j}{\Delta x_i\mathbf{k}_j+\Delta x_j\mathbf{k}_i}\right) \qquad 4.35$$

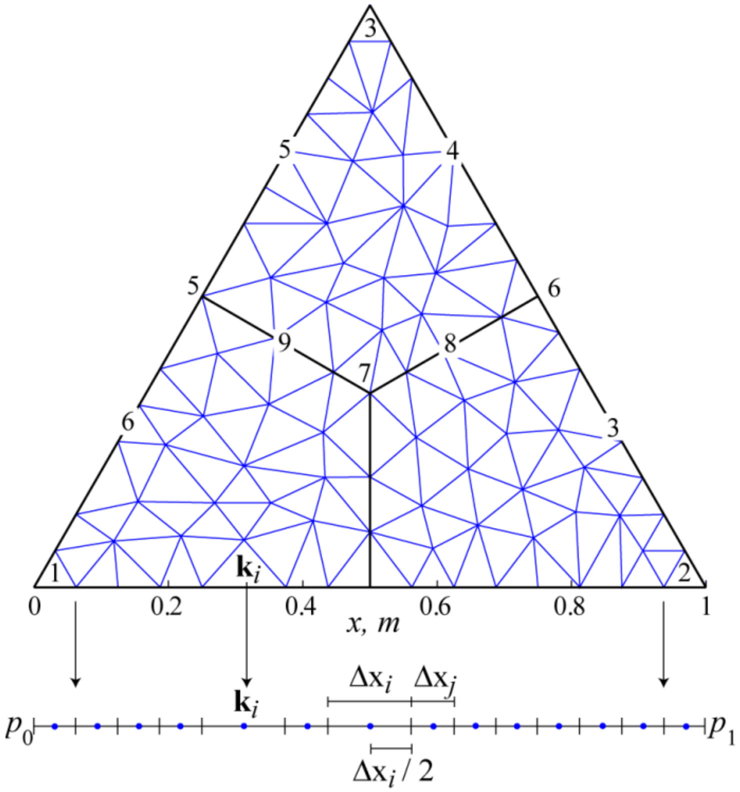


Figure 4-27: One dimensional pressure problem shown for lower edge.

Applying conservation of mass for each element via Equation 4.36 provides a system of n equations having the same structure as Equation 4.29, where n is the number of elements along the edge and $\underline{\mathbf{T}}$ is tridiagonal with diagonal entries $T_{i-1,i} + T_{i,i+1}$, lower diagonal entries $T_{i-1,i}$ and upper diagonal entries $T_{i+1,i}$.

$$\begin{aligned} q_i &= -T_{i-1,i}(p_{i-1} - p_i) + T_{i,i+1}(p_i - p_{i+1}) \qquad i = 1,\ldots n \\ &= -T_{i-1,i}p_{i-1} + (T_{i-1,i} + T_{i,i+1})p_i - T_{i,i+1}p_{i+1} \end{aligned} \tag{4.36}$$

Three such problems are setup, one for each side of the triangular interaction region to characterize the pressure distribution around its perimeter. Since two sets of boundary conditions are required to determine the transmissibilities in Equation 4.31, six one-dimensional problems are involved. Using boundary conditions of (1,0,0) and (0,1,0), the solution to edges with zeros at the vertices is zero and only four problems are involved.

Solutions to the one dimensional problems can be used in two ways: 1 – only the pressures at the coarse interaction region midpoints are retained and used to impart different pressure gradients in each variation triangle; 2 – all pressures are used as boundary conditions for the interior problem. Method 2 is not explored; however, it is expected that results are better than linearly varying boundary conditions since more permeability information is incorporated into the upscaling process. This is an area of future development. The first approach is similar to that in Section 4.2.2 in that a linear pressure variation is applied to each variation triangle; however, the pressure gradient is different along edges $\overline{p_\iota m_\iota}$ and $\overline{m_\iota p_{\iota+1}}$ if the permeability is heterogeneous. For constant permeability, the same solution is obtained as Equation 4.34 for the equilateral triangle example.

An example with constant permeability in each control volume connected by the equilateral interaction region previously used demonstrates the result (Figure 4-28). Different pressure gradients in each variation triangle are clear. They were derived based on the corner point and midpoint pressures only. The apparent gradient in the direction from point 1 to 5 of the interaction region is greater than that from corner 1 to 4. This is as expected since permeability increases from point 1 to 3 through 5 and decreases from point 1 to 2 through 4. Using these pressures along the boundary of the interaction region to solve for the interior pressures gives a smooth pressure field (Figure 4-29).

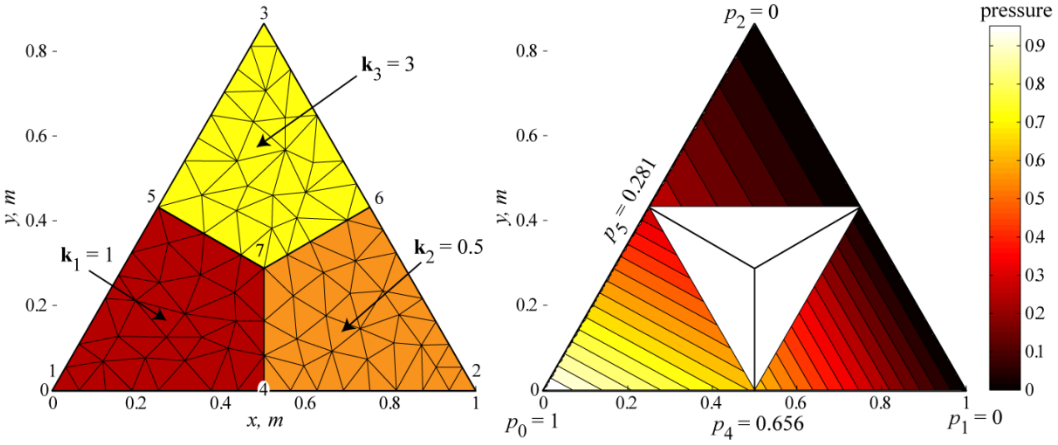


Figure 4-28: Heterogeneous permeability (left) and linear varying pressure in the variation triangles (right) based on midpoint pressures from one dimensional boundary solutions.

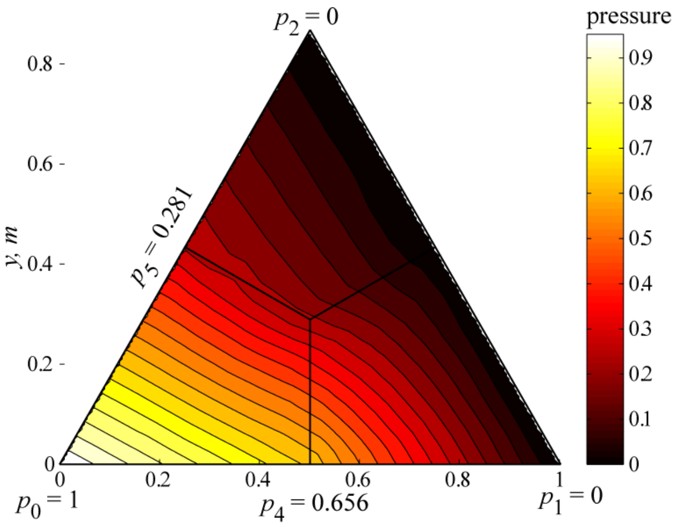


Figure 4-29: Interior pressure solution for three permeabilities case.

The resulting transmissibility matrix is summarized in Equation 4.37. Using the same boundary conditions but without the one dimensional solves for the midpoint pressures results in the transmissibility matrix in Equation 4.38. Knowing the permeability within each variation triangle makes it possible to compute the transmissibility matrix analytically using Equation 4.26 with the solution given by Equation 4.39. Interestingly, the matrix obtained without the midpoint pressures is closer to the analytical result. Evaluating Equation 4.27 always results in a zero pressure gradient contribution from p_{ι} across the opposing flux interface $\overline{b\mathrm{m}_{\iota+1}}$ when the permeability tensor is diagonal and isotropic. Using linear boundary conditions, the computed pressure gradient across flux interfaces along the zero pressure edge is approximately zero as well, leading to a solution that is closer to the analytical solution.

However, the assumptions and simplifications made by the MPFA approach do not consider flow at a higher resolution within the interaction regions. The fine scale pressure solution gives an improved understanding of flow on the interior of the interaction region beyond the assumption of a constant pressure gradient in the variation triangles. By arguing that fluid takes the path of least resistance, flow must occur across the interface on the zero pressure edge for the example with three permeabilities. This is found using boundary conditions that involve the midpoint pressures and is observed in Figure 4-29. Pressure contours not perpendicular to a flux interface induces flow across that interface.

$$\mathbf{T}_{midpoint} = \begin{bmatrix} 0.148 & 0.109 & -0.517 \\ -0.120 & 0.212 & -0.079 \\ -0.027 & -0.321 & 0.596 \end{bmatrix} \tag{4.37}$$

$$\mathbf{T}_{cornerpoint} = \begin{bmatrix} 0.183 & 0.029 & -0.456 \\ -0.177 & 0.238 & -0.023 \\ -0.006 & -0.267 & 0.479 \end{bmatrix} \tag{4.38}$$

$$\mathbf{T}_{analytical} = \begin{bmatrix} 0.192 & 0 & -0.433 \\ -0.192 & 0.247 & 0 \\ 0 & -0.247 & 0.433 \end{bmatrix} \tag{4.39}$$

An issue with the one dimensional pressure solution is a possible discontinuity in pressure along the edge and at the midpoint when computed for the neighbouring interaction region if the fine scale permeability is different on opposite sides of the edge (Figure 4-30). In this case, the permeability from two neighbouring triangles sharing the edge involved in the one dimensional pressure problem is averaged arithmetically. This is equivalent to assessing flow through layered media parallel to the layers and the arithmetic average is correct (Figure 4-31). Resulting pressures along the edge are guaranteed to be continuous between interaction regions.

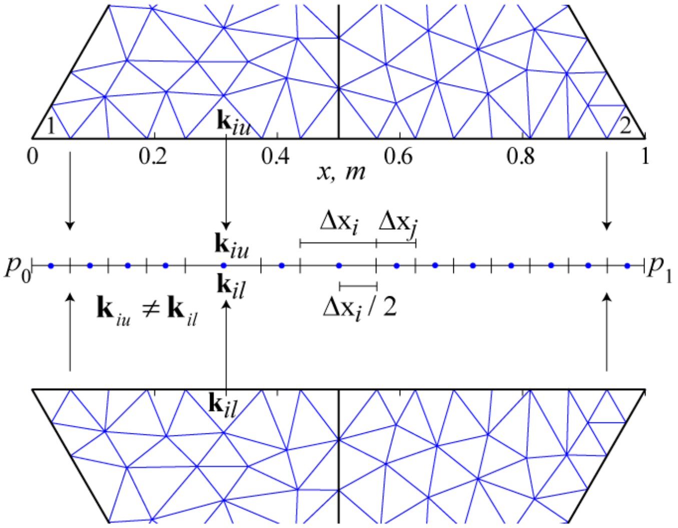


Figure 4-30: One dimensional pressure problem leading to discontinuous pressure between two triangular interaction region edges.

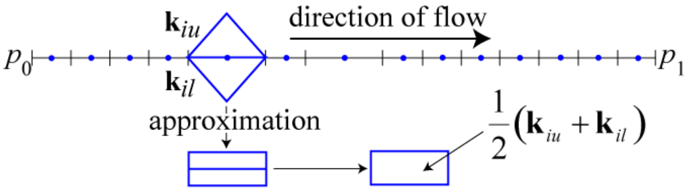


Figure 4-31: Average permeability to ensure continuous pressure between interaction regions when using one dimensional pressure solves.

Using the midpoint pressures to define boundary conditions provides a unique solution in that regardless of the two sets of pressure boundary conditions chosen the transmissibility matrix is unique. Although more representative transmissibilities are obtained than without the midpoint pressures, they are independent of the boundary conditions. This may not be the case when the interaction regions involved in the upscaling process contain heterogeneous permeability and are embedded in a larger grid.

All variations on local boundary conditions that are based solely on the pressures at the corners of the interaction region lead to a unique transmissibility matrix. This occurs because the flux across the interaction region boundary is directly linked to the corner point pressures. Only when the boundary conditions are controlled by more than the local information is dependence between pressure and transmissibility obtained. For example, if the midpoint pressure was dependent on more information than the corner point pressures, a different transmissibility matrix would result. Such information might include pressures around the region being upscaled or a better approximation of the flux across its boundary; the limit being the global pressure and flux fields on the fine scale grid that is typically unattainable. As mentioned in Chapter 1, additional information is brought into the equations using extended local upscaling (Gomez-Hernandez and Journel, 1990). The objective is to obtain better estimates of flow through the upscaling region. Extended local upscaling methods are not explored in this dissertation; however, additional study in this area could be undertaken.

4.2.4 Convergence

Upscaled transmissibilities are used to approximate the flow that is happening through the fine scale grid. As with arithmetic averaging, there is dependence between the number of fine scale elements involved and the accuracy of the upscaled values. In the arithmetic averaging case, it is possible to compute the expected value of the error since the true value can be approximated. There is no straightforward approximation for the upscaled transmissibility based on the fine scale permeability field. There are several additional dependencies including the permeability field, boundary conditions, geometry of the fine scale discretization, and geometry of the interaction region. Reconciling these requires numerical studies.

Choosing the number of fine scale elements, n, to discretize a coarse element will primarily be chosen for practicality: grids must be kept to some manageable size to meet time constraints and computational limitations. A few convergence studies are done to give a general idea of potential error when choosing n. It is assumed that the variation of permeability within interaction regions is not extreme in the presence of different facies and/or geological units (Figure 4-32), that is, the accuracy of transmissibility assuming a constant facies and geology is representative. This is applicable for small grid elements. Larger elements in an unstructured grid design are typically in areas of the reservoir where low flow rates occur and the accuracy is not as important as areas adjacent to wells, where grid elements are small. Further convergence study is required to relax the assumptions made for the studies presented.

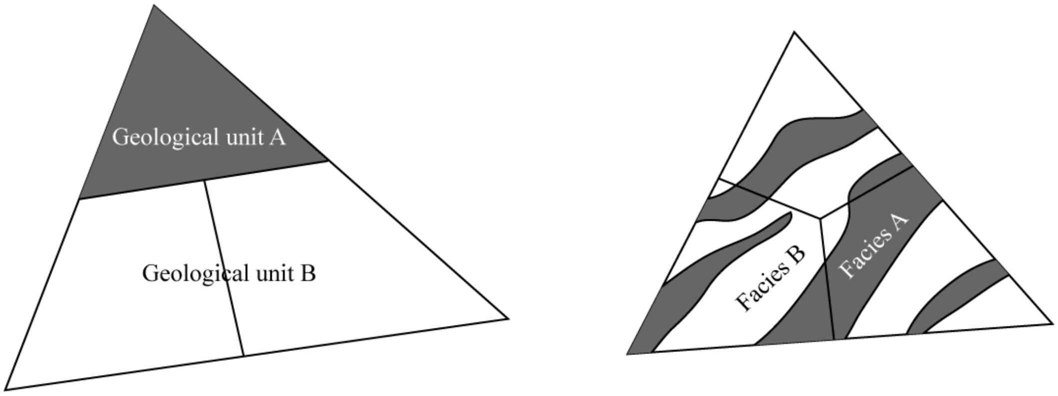


Figure 4-32: Interaction region intersecting different geological units (left) and facies (right).

Computing the error in transmissibility matrices requires an estimate of the true values and approximations for different n. True values are estimated numerically with a high resolution discretization consisting of $N \gg n$ elements; however, accessing both transmissibility matrices for the same realization is problematic as a realization filling n elements can lead to many different realizations of size N. The approach used to observe the convergence of transmissibility for a few simple examples is:

1. Generate a set of high resolution permeability realizations.
2. For some arbitrary interaction region covered by the realizations and for a variety of discretization volume constraints
    a. Resample the realizations using nearest neighbour interpolation.
    b. Upscale the transmissibilities.
3. Observe the changes in transmissibility with increasing n.

Nearest neighbour interpolation is used in step 2.a. for two reasons: 1 – it is similar to generating a point scale realization on the different discretizations using geostatistics, but rather than simulating a value at the centre of an element, the value from the already known realization from step 1 is selected; 2 – it prevents the permeability field from being smoothed out by using other methods such as kriging and linear interpolation.

In two dimensions, the coarsest possible discretization that is constrained by the boundary and the flux interfaces has $n = 6$ triangles (Figure 4-33). This is used as the starting point for all examples. A difference between these studies and those involving the arithmetic average is an interaction region is used here and in the arithmetic average cases a grid element is used. Defining n_A as the number chosen for arithmetic averaging and n as the number chosen for transmissibility upscaling, the relationship between them for simplex interaction regions is approximated by Equation 4.40, where O is a cluster of interaction regions connected to the center points of a grid element, $\bar{\bar{O}}$ is the number of interaction regions in the cluster, and d is the dimensionality.

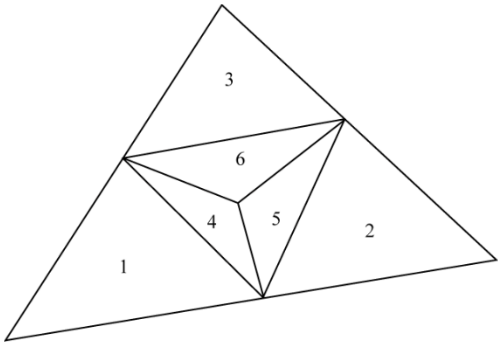


Figure 4-33: Coarsest discretization possible for a triangular interaction region.

$$n_A \cong \bar{\bar{O}} \cdot \frac{n}{d+1} \qquad 4.40$$

The first two examples covered involve the equilateral triangle used previously and observe the computed transmissibility with increasing n. Example 1 uses a constant permeability of 1 Darcy. In this case, the coarsest triangulation yields the exact transmissibility matrix. Increasing n results in less symmetry in the discretization because the program Triangle (Shewchuk, 1996) inserts new points for the triangulation without considering global grid quality; however, there is very little deviation from the analytical transmissibilities (Figure 4-34). In the second example, permeability is 1 Darcy in two control volumes and 10 Darcy in the third. The transmissibilities quickly converge to a consistent value (Figure 4-35).

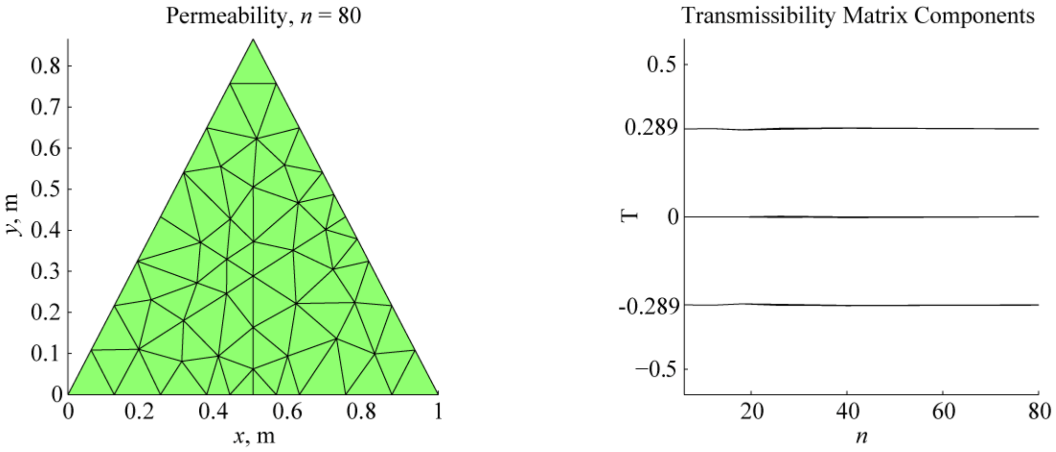


Figure 4-34: Convergence of transmissibility for constant permeability.

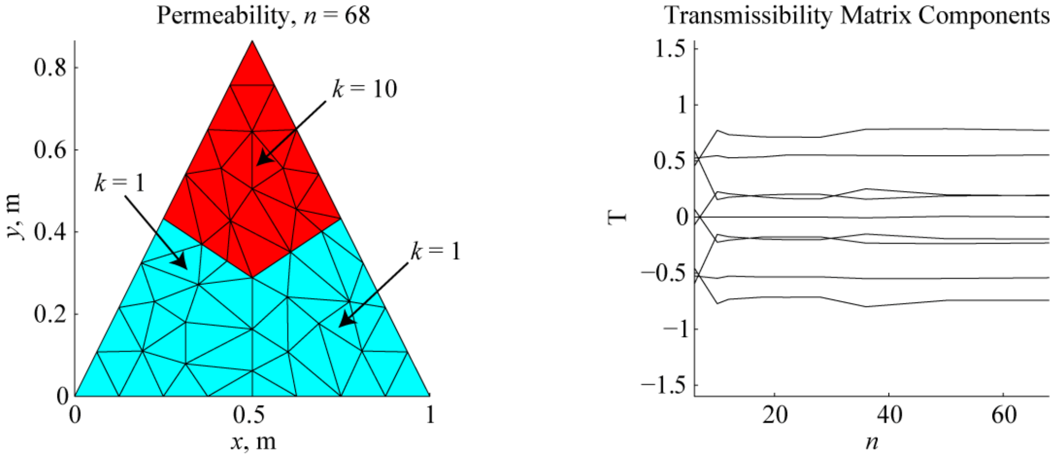


Figure 4-35: Convergence of transmissibility for different permeability in each control volume joined by the interaction region.

Further examples involve unconditional permeability fields and multiple realizations. Fields are generated on a 51 by 51 node regular grid that covers the interaction region using Cholesky factorization or LU simulation (Davis, 1987). The realizations from LU simulation are Gaussian and are transformed using $Z = \exp(Y)$ to reflect permeability. Convergence of transmissibilities is expressed as a sum of squares of the change in value with increasing n (Equation 4.41), where $\hat{T}_{ij}(n)$ is a component of the transmissibility matrix computed using a discretization with n elements. Initial transmissibilities without discretization, $n = 0$, are approximated using the geometric mean of the permeability in each variation triangle and Equation 4.26. Triangulation was done using the program Triangle (Shewchuk, 1996) with a minimum interior angle constraint of 30 degrees.

$$SSE(\mathbf{T}) = \sum_{i=1}^{3}\sum_{j=1}^{3}\left(\hat{T}_{ij}(n) - \hat{T}_{ij}(n+1)\right)^2 \qquad 4.41$$

Three SSE curves averaged over 25 realizations are shown for the equilateral triangle case in Figure 4-36 corresponding to a spherical, exponential, and Gaussian variogram model. All models were isotropic with a range 5 times the extent of the interaction region. Results are as expected for the effect of variogram structure on error: variograms with less variation show less error. Another observation is the error from the geometric average approach, $n = 0$, to using $n = 6$ is less than the error from $n = 6$ to $n = 12$. Using $n = 6$ is not much different than the analytical approach and the error is small; hence the resulting transmissibilities are more similar. This indicates that the geometric average and too few n are unsatisfactory. For the three variograms shown, n should be at least 12.

Convergence of error with variogram range is shown in Figure 4-37 using an isotropic spherical model and ranges 1, 2, 3, 5 and 8 times the interaction region extent. As range decreases relative to the interaction region, which is equivalent to more variation in the variogram, the error increases for a constant n. For ranges 1 and 2, $n = 18$ discretization points may be necessary for a spherical variogram. However, in most applications variation in the variogram is small relative to the grid element volume. A similar convergence test using a spherical variogram with range of 5 meters and a nugget effect varying between 0 and 0.2 is shown in Figure 4-38.

A final convergence test (Figure 4-39) using a spherical variogram with a range of 5 meters shows the change in error as the interior angle of the interaction region is changed. To accommodate the changing angle, the quality constraint for the program Triangle (Shewchuk, 1996) was decreased along with the quality of the interaction region. Results are somewhat mixed, but generally the error seems to increase as the angle decreases. The variation is not substantial. Studies to observe the change in error with permeability contrast were not done since the transmissibility scales linearly with permeability. Higher permeability translates to larger transmissibilities that yield larger squared error. When scaled by the magnitude of the permeability contrast, $\max(\mathbf{k}) / \min(\mathbf{k})$, the error curves would show no difference. There is some concern in all the convergence plots (Figure 4-36 to Figure 4-39) because the error is not monotonically decreasing. This is caused by the variation in the geometry of the underlying triangulation when increasing n.

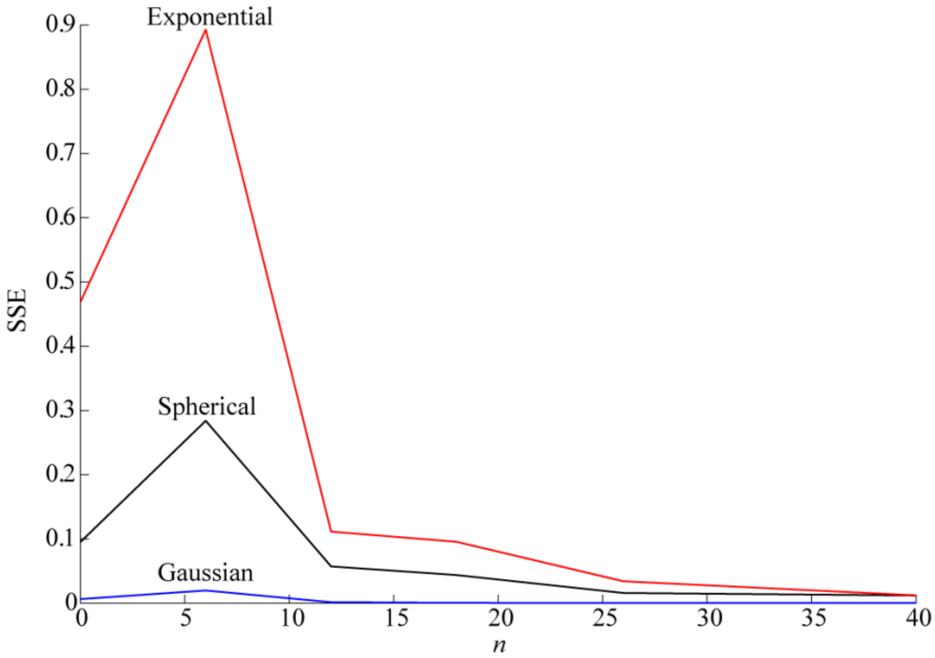


Figure 4-36: SSE convergence with n for different variograms.

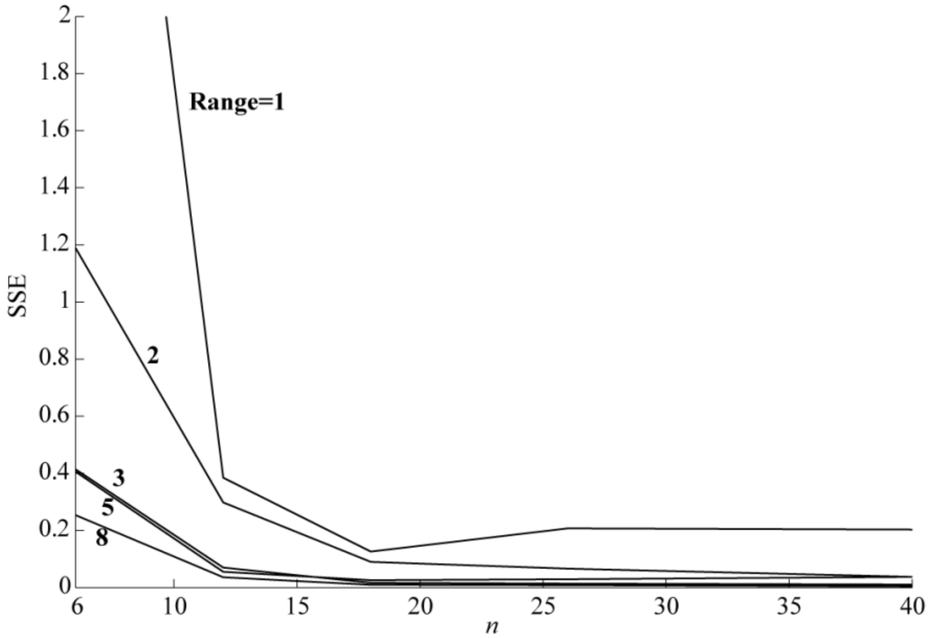


Figure 4-37: SSE convergence with n for a spherical variogram with different ranges.

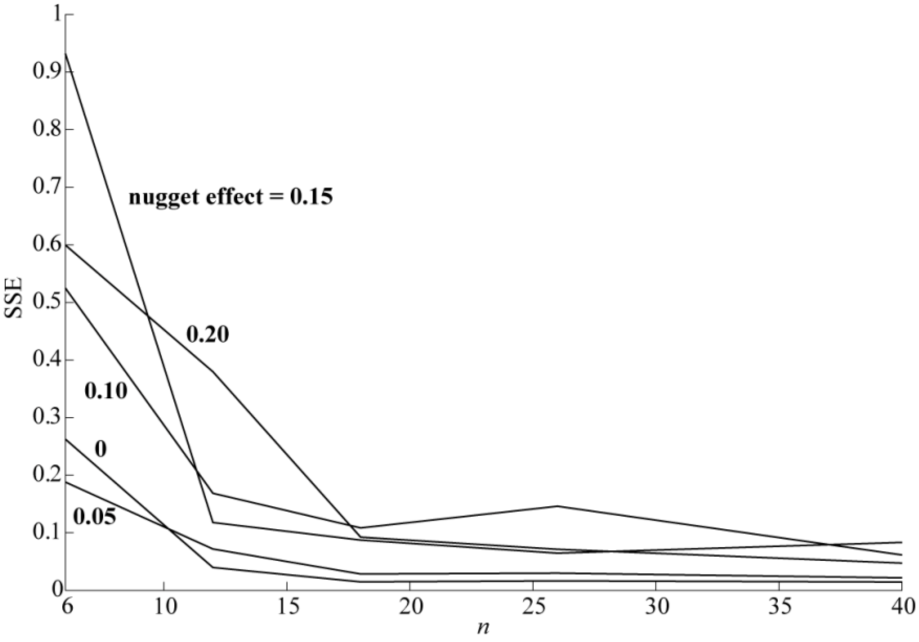


Figure 4-38: SSE convergence with n for a spherical variogram with a range of 5 meters and different nugget effects.

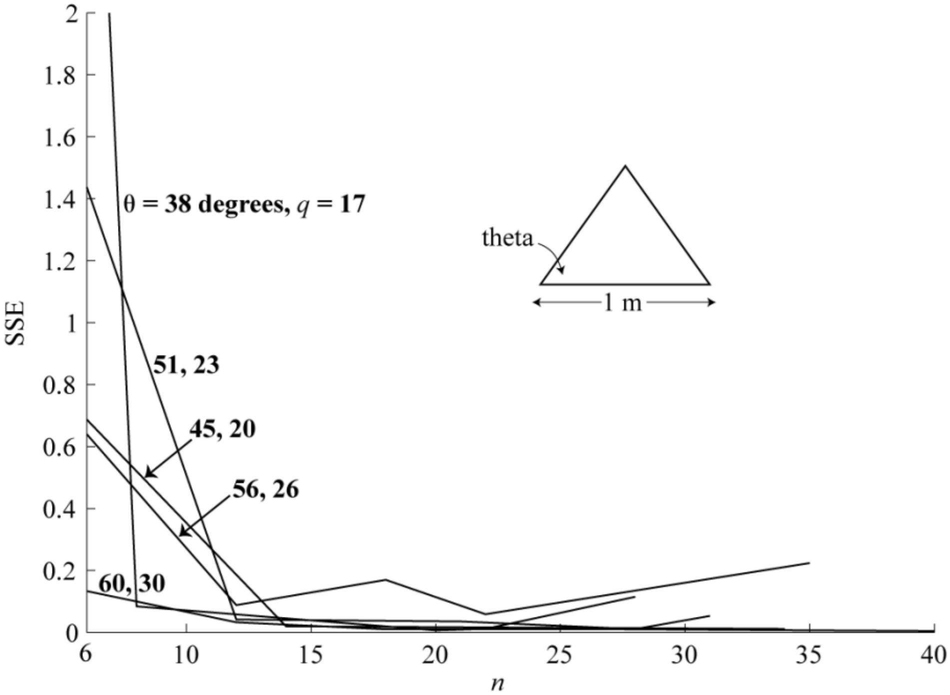


Figure 4-39: SSE convergence with n for a spherical variogram with a range 5 meters and different interior angles for the interaction region.

These basic convergence examples show that choosing n for upscaling transmissibility is not as exact as for arithmetic averaging. Because of the high number of dependencies, no explicit approach presents itself; however, some general guidelines are identified:

- As variation in the variogram relative to the size of interaction regions decreases, n can be reduced. Variation decreases from exponential to spherical to Gaussian variogram models. The presence of a nugget effect increases the variation and reduces the convergence rate.
- As the smallest interior angle of triangular interaction regions decreases, the error generally increases. Discretization also becomes more challenging since this angle may violate the minimum angle constraint specified for Triangle.
- Specifying an angle constraint of 30 degrees for Triangle typically leads to at least 12 fine scale triangles regardless of the area constraint. This is due to Triangle having to match both the triangle, barycenter, and flux interfaces.
- In general for two dimensional problems, n should be chosen to be 12 to 18 for triangular interaction regions. The upscaling methods were not extended to three dimensions; therefore these guidelines are not necessarily applicable in three dimensions.

Realistically, the volume of interaction regions is on average small relative to the variation in the variogram. The majority of grid elements are concentrated around wells, significant geological features, and known flow paths and are small for obtaining accurate flow simulation results.

4.3 Comments on Discretization Quality

Another constraint for generating simplex discretizations is element quality. For approximating the arithmetic average of a stationary random field with a triangular discretization, it is clear that a more uniformly distributed set of triangles yields a better approximation. This is achieved with a high angle quality constraint with Triangle and low radius-edge ratio with TetGen. For the MPFA, quality also depends on the underlying permeability tensor field that acts as a local coordinate system. Some geometric and permeability configurations are known to cause non-physical solutions to the pressure equation using MPFA simulation (Aavatsmark, 2002; Mlacnik and Durlofsky, 2006). This situation is not encountered if the inverse of the transmissibility matrix, $\underline{\mathbf{T}}$ in Equation 4.29, is positive, $\underline{\mathbf{T}}^{-1} \geq 0$. A sufficient condition is that $\underline{\mathbf{T}}$ is an M-matrix (Fujimoto and Ranade, 2004), which is a matrix with properties defined in Equation 4.42.

$$\begin{aligned} T_{ii} &> 0 \quad \forall i \\ T_{ij} &< 0 \quad \forall ij, i \neq j \\ \sum_j T_{ij} &\geq 0 \quad \forall i \end{aligned} \qquad 4.42$$

For isotropic permeability in two dimensions, an M-matrix is achieved using a triangular grid when the sum of the interior angles opposite a shared edge is less than π (Figure

4-40) (Aavatsmark et al, 1998b). This is achieved using a Delaunay triangulation, which is an option in Triangle (command line switch 'D'). For example, using an angle constraint of 30 degrees, the worst possible case for the angle sum is $120 + 120 = 240$ degrees, but to ensure the triangulation is Delaunay the shared edge is flipped and the new sum is $60 + 60 = 120$ degrees (Figure 4-41). For tetrahedral grids, a sufficient condition for an M-matrix is the dihedral angle of any two faces of a tetrahedron is less than $\pi/2$; however this is not a necessary condition (Mlacnik and Durlofsky, 2006). An M-matrix is achieved when the sum of the contributions from all tetrahedra sharing an edge is less than or equal to zero.

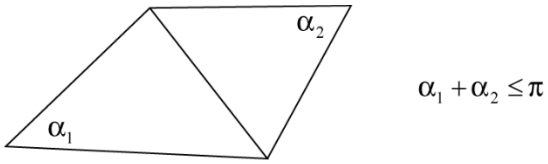


Figure 4-40: Interior angle constraint for an M-matrix in two dimensions and for isotropic homogeneous permeability.

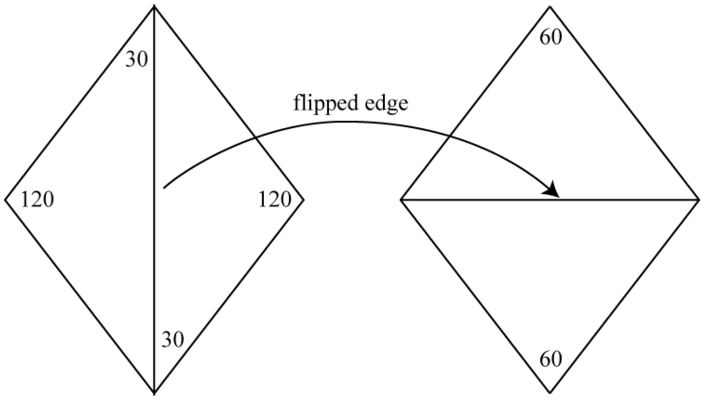


Figure 4-41: Example edge flip in a Delaunay triangulation.

For heterogeneous and anisotropic permeability, the conditions for an M-matrix do not necessarily hold. It becomes difficult to define quality constraints for discretization to ensure a good quality solution to the pressure equation. Achieving these conditions is a grid optimization problem discussed by Mlacnik and Durlofsky (2006). In this thesis, it is assumed that the coarse unstructured grid was designed with such considerations. Also, it is assumed that discretization to achieve Delaunay triangulations and good quality elements is sufficient to achieve stable upscaled transmissibilities. This implies that any M-matrix issues that occur on the fine scale grid are mitigated through upscaling and will not degrade flow simulation results on the coarse grid.

4.4 Summary

This chapter introduced two new components for upscaling: 1 – a general way to measure error for arithmetic averaging that can be used to control grid discretization; 2 – a new method to upscale transmissibility for PEBI and polygonal grids using the MPFA method. For random variables that scale with an arithmetic average, the expected error is dependent on the variogram and geometry. It is possible to derive a problem-specific function that relates the expected error to the number of discretization points used for the average. For cases when the variation in the variogram is low within the extent of grid elements, which is often the case, low error is possible with few discretization points. It was also found that when going from two dimensions to three dimensions, the number of points required to maintain a similar accuracy is achieved by maintaining the same point spacing.

Upscaling transmissibilities using the MPFA method was developed for triangular interaction regions. It was shown to yield exact results for homogeneous permeability. Extending the results to tetrahedral interaction regions is possible and additional study could be undertaken for this. Local upscaling methods were developed with two types of boundary conditions: 1 – linearly varying pressure around the entire interaction region; 2 – linearly varying pressure in each variation triangle, with midpoint pressures derived from one dimensional pressure solves. Regardless of the type of boundary conditions chosen, as long as they are local the resulting transmissibility matrices for triangular interaction regions are independent of the boundary conditions; i.e. any two sets of boundary conditions along with the constant sum constraint yields the same transmissibility matrix.

Determining an approach to choose the discretization points for transmissibility upscaling is more complex than for arithmetic averaging. Some general guidelines were identified to acquire more stable transmissibility matrices in two dimensions. Generally, more points are required when there is increased variation in the variogram and lower quality interaction regions. In two dimensions, 12 to 18 points are adequate. When a low enough error is chosen for the arithmetic average discretization criteria, the number of points should be adequate for transmissibility upscaling as well, although further study is required.

Averaging and flow based upscaling is one stage of the proposed reservoir modeling workflow of this dissertation. Starting with a coarse unstructured grid designed based on some a priori knowledge of the permeability distribution and boundary conditions, the approach is to discretize the coarse grid with a higher resolution simplex grid. Discretization is controlled by the desired accuracy of averaging and flow based upscaling and also by practicality, i.e. the time required to do the modeling depends on the number of fine and coarse grid elements. Final upscaled properties are used in flow simulation studies. The upscaled properties reflect the underlying fine scale models for all scales in the coarse unstructured grid.

Chapter 5 Demonstration and Case Study

This chapter focuses on demonstrating aspects of the workflow proposed in Chapter 2. Section 5.1 shows that the correct local distributions of uncertainty are obtained using simplex discretization. Recovery of statistics through the Gaussian transformation is shown for a lognormal variable including the covariance function and the proportional effect. A comparison of global statistics computed using regular grids and conventional approaches to using discretized unstructured grids is also done.

Section 5.2 demonstrates that the transmissibility upscaling approach in homogeneous, layered, and heterogeneous media gives good results in comparison to using traditional flow simulation methods on high resolution regular grids. The last section goes through a reservoir modeling case study covering the geostatistical modeling components of the proposed workflow in Chapter 2. It utilizes programs developed in Chapter 3 for modeling and checking variogram reproduction and the error function from Chapter 4 for choosing discretization parameters.

5.1 Simplex Discretization

5.1.1 Covariance Reproduction

In Chapter 4 it was shown for simplex discretization that as discretization density increases, the average covariance converges to the correct value. This section shows that geostatistical simulation on the discretization followed by upscaling is consistent with scaling theory for arithmetically averaging random variables. It is shown that averaged simulated values within an element are distributed with a mean and variance equal to those obtained by kriging with the average covariance of the same element. This is first done assuming the underlying random field is Gaussian, which removes any non-linear effect imposed by a Gaussian transformation.

For Gaussian fields, simple kriging using the average covariance of an element yields the same distribution as simulating random values in the element and averaging them. Consider a case where a Gaussian value is conditionally simulated based on N conditioning data to represent an unstructured grid element using average covariance that is approximated by discretizing the element into n points. In Equation 5.1, $\mathbf{y}$ is the vector of conditioning data, $\overline{y}_s$ is the simulated value, $\boldsymbol{\lambda}$ is the vector of kriging weights, $\overline{\mathbf{c}}$ is the vector of average covariances between data samples and the grid element, $\mathbf{C}_{11}$ is the covariance matrix between the conditioning data, and r is a zero mean Gaussian random deviate with variance given by Equation 5.2. Here, $\overline{C}(V,V)$ is the average covariance of the element that describes the variance at scale V.

$$\begin{aligned}\overline{y}_s &= \boldsymbol{\lambda}^T \mathbf{y} + r \\ &= \overline{\mathbf{c}}^T C_{11}^{-1} \mathbf{y} + r\end{aligned} \qquad 5.1$$

$$\sigma_r^2 = \bar{C}(V,V) - \boldsymbol{\lambda}^T \bar{\mathbf{c}} \qquad 5.2$$

Now consider representing the same element by simulating point scale values at each of the n points and averaging them arithmetically. For simplicity, assume all values represent an equal fraction of V, that is, $w_k = 1/n, k = 1, \ldots, n$. To account for all covariances the LU formalism (Davis, 1987) is used. Covariances between all conditioning data and discretization points is expressed in Equation 5.3 along with its Cholesky decomposition, where $\mathbf{C}_{12}$ is the covariance between conditioning data and discretization points and $\mathbf{C}_{22}$ is between discretization points.

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}^T & \mathbf{C}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{11} & 0 \\ \mathbf{A}_{12} & \mathbf{L}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{11} & \mathbf{B}_{12} \\ 0 & \mathbf{U}_{22} \end{bmatrix} \qquad 5.3$$

A conditional realization for all n points is generated using Equation 5.4, where $\mathbf{y}_s$ is the vector of simulated values and $\mathbf{r}$ is a vector of zero mean Gaussian random deviates with unit variance.

$$\mathbf{y}_s = \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y} + \mathbf{L}_{22}\mathbf{r} \qquad 5.4$$

Computing the mean and variance of the mean of Equation 5.4 recovers exactly the result in Equations 5.1 and 5.2 as shown in Equations 5.5 and 5.6. The first step replaces $\mathbf{C}_{11}^{-1}\mathbf{y}$ by $\boldsymbol{\alpha}$ as in dual kriging and the expected value of $\mathbf{L}_{22}\mathbf{r}$ is a zero vector since $E\{\mathbf{L}_{22}\mathbf{r}\} = \mathbf{L}_{22}E\{\mathbf{r}\} = 0$. Summation of $\mathbf{C}_{12}$ over i and multiplying by $1/n$ gives the same average covariance vector in Equation 5.1 and replacing $\boldsymbol{\alpha}$ with its original matrix vector product recovers the same kriging estimate.

$$\begin{aligned} \mathrm{E}\{\mathbf{y}_s\} &= \mathrm{E}\{\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}\} + \mathrm{E}\{\mathbf{L}_{22}\mathbf{r}\} \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbf{C}_{12,ij}\boldsymbol{\alpha}_j \\ &= \bar{\mathbf{c}}^T\boldsymbol{\alpha} = \bar{\mathbf{c}}^T\mathbf{C}_{11}^{-1}\mathbf{y} \qquad \square \end{aligned} \qquad 5.5$$

$$\begin{aligned} \mathrm{Var}\{\mathrm{E}\{\mathbf{y}_s\}\} &= \mathbf{d}^T\left[\mathbf{C}_{22} - \mathbf{C}_{12}^T\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\right]\mathbf{d} \\ &= \mathbf{d}^T\mathbf{C}_{22}\mathbf{d} - \mathbf{d}^T\mathbf{C}_{12}^T\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{d} \\ &= \bar{C}(V,V) - \bar{\mathbf{c}}^T\mathbf{C}_{11}^{-1}\bar{\mathbf{c}} \\ &= \bar{C}(V,V) - \boldsymbol{\lambda}^T\bar{\mathbf{c}} \qquad \square \end{aligned} \qquad 5.6$$

Computing the variance of the mean is a quadratic form expression of the variance of $\mathbf{y}_s$ shown in Equation 5.6, where $\mathbf{d} = [1/n \ldots 1/n]^T$. These results are applicable when no transformation is used and the variable being estimated averages arithmetically.

Using point scale estimates is therefore consistent with geostatistical scaling for Gaussian random fields. Averaging a set of point scale values accounts for scale and shape as the average covariance approach does. It is also shown in Equations 5.10 and 5.11 that the estimate and estimation variance are correct when there are multiple grid elements involved. In this case, the average covariance between conditioning data and grid elements and between the elements must be recovered. Consider two grid elements, a and b, and LU simulation using average covariance, then Equation 5.1 is written as Equation 5.7, where $\bar{\mathbf{c}}_a$ and $\bar{\mathbf{c}}_b$ are vectors of average covariance between conditioning data and grid element a and b respectively, and $\mathbf{C}_{11}$ is unchanged. $\mathbf{L}_{22}$ is from the Cholesky decomposition of the covariance matrix in Equation 5.8 and the estimation variance is defined by Equation 5.9.

$$\bar{\mathbf{y}}_s = \left[\bar{\mathbf{c}}_a \quad \bar{\mathbf{c}}_b\right]^T \mathbf{C}_{11}^{-1}\mathbf{y} + \mathbf{L}_{22}\mathbf{r} \tag{5.7}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}^T & \mathbf{C}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{11} & \bar{\mathbf{c}}_a & \bar{\mathbf{c}}_b \\ \bar{\mathbf{c}}_a^T & \bar{C}_{aa} & \bar{C}_{ab} \\ \bar{\mathbf{c}}_b^T & \bar{C}_{ba} & \bar{C}_{bb} \end{bmatrix} \tag{5.8}$$

$$\Sigma = \begin{bmatrix} \bar{C}_{aa} & \bar{C}_{ab} \\ \bar{C}_{ba} & \bar{C}_{bb} \end{bmatrix} - \begin{bmatrix} \lambda_a \\ \lambda_b \end{bmatrix} \left[\bar{\mathbf{c}}_a \quad \bar{\mathbf{c}}_b\right] \tag{5.9}$$

Discretizing the elements and taking the expected value in Equation 5.10 yields the same result. Equation 5.11 shows the same estimation variance matrix is recovered as well.

$$\begin{aligned} \mathrm{E}\{\mathbf{y}_s\} &= E\{\mathbf{y}_{sa} \quad \mathbf{y}_{sb}\} \\ &= \mathrm{E}\left\{[C_{1a} \quad C_{1b}]^T \mathbf{C}_{11}^{-1}\mathbf{y}\right\} + \mathrm{E}\{\mathbf{L}_{22}\mathbf{r}\} \\ &= \left[\bar{\mathbf{c}}_a \quad \bar{\mathbf{c}}_b\right]^T \mathbf{C}_{11}^{-1}\mathbf{y} \qquad \square \end{aligned} \tag{5.10}$$

$$\begin{aligned} \mathrm{Var}\left\{\mathrm{E}\{\mathbf{y}_s\}\right\} &= \mathbf{D}^T\left[\mathbf{C}_{22} - \mathbf{C}_{12}^T\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\right]\mathbf{D} \qquad \mathbf{D} = \begin{bmatrix} \mathbf{d} & 0 \\ 0 & \mathbf{d} \end{bmatrix} \\ &= \mathbf{D}^T \begin{bmatrix} C_{aa} & C_{ab} \\ C_{ba} & C_{bb} \end{bmatrix} \mathbf{D} - \mathbf{D}^T \begin{bmatrix} C_{1a} \\ C_{1b} \end{bmatrix} \mathbf{C}_{11}^{-1}[C_{1a} \quad C_{1b}]\mathbf{D} \\ &= \begin{bmatrix} \bar{C}_{aa} & \bar{C}_{ab} \\ \bar{C}_{ba} & \bar{C}_{bb} \end{bmatrix} - \begin{bmatrix} \lambda_a \\ \lambda_b \end{bmatrix} \left[\bar{\mathbf{c}}_a \quad \bar{\mathbf{c}}_b\right] \\ &= \Sigma \qquad \square \end{aligned} \tag{5.11}$$

These equations show that in the Gaussian space of geostatistical simulation, the proposed approach using discretization is consistent with arithmetic scaling laws. The advantage of using the discretization approach is realized when transformations and non-linear averaging are involved. Reservoir properties are typically not Gaussian. Quantile transforms are often used to obtain a Gaussian random variable so that geostatistical techniques apply; however, this is a non-linear transform and the average covariance approach is not applicable.

5.1.2 Statistics through a transformation

Statistics of point scale values through a transformation are reproduced since the geostatistical simulation algorithm developed in Chapter 3 results in a Gaussian field. The correct covariance is recovered and other statistics such as the proportional effect (David, 1977; Journel and Huijbregts, 1978; Manchuk et al, 2009) are recovered as well.

These results are found in the literature (Journel and Kyriakidis, 2004); however they are demonstrated with a single two dimensional example involving a lognormally distributed random variable, Z. There is an analytical relationship (Equation 5.12) between the covariance functions of Z and $Y = \ln(Z)$. Y is normally distributed. In Equation 5.12, σ_Y^2 is the variance of Y (1 in this case), m is the mean of Z defined by Equation 5.13, and m_Y is the mean of Y (0 in this case).

$$C_Z(\mathbf{h}) = m^2\left[\exp\left(\sigma_Y^2 C_Y(\mathbf{h})\right) - 1\right] \qquad 5.12$$

$$m = \ln\left(m_Y + \sigma_Y^2 / 2\right) \qquad 5.13$$

Checking that the variogram is reproduced at the coarse grid element scale requires computing the experimental variogram; however, computing the experimental variogram of an unstructured grid with a variety of element shapes and volumes is non-trivial. For simplicity, a coarse regular grid is used so that the experimental variogram can be computed as only a function of $\mathbf{h}$, rather than of $\mathbf{h}$, volume, and shape. Elements are discretized using a triangulation as would be the case with a coarse unstructured grid.

An arbitrary regular grid with 60 by 30 elements in x and y of size 10 by 10 m is used. The program Triangle is used to discretize the grid with a quality constraint of 20 degrees and an area constraint of 10 m^2. There are nominally 15 triangles per regular grid element and a total of 27,883 triangles. On the triangle barycenters, 50 unconditional Gaussian realizations are generated using psgsim with a spherical variogram with a range of 50 meters in x and 100 meters in y. Realizations are converted to a lognormal distribution and both variables are averaged to the coarse regular grid (Figure 5-1).

Checking reproduction of the covariance at the coarse scale is done by computing the experimental variogram of each coarse scale realization and averaging the set of 50. Analytical values of the sill were calculated for the input variogram model. The sill of the average of Y is 0.88 and of Z is 3.84. Input point scale variograms are plotted with the average variogram of the coarse realizations in Figure 5-2. The experimental and analytical sills are visually similar. Other changes agree with variogram scaling law theory (Frykman and Deutsch, 1999): the range is larger at the coarser scale; the variograms are shifted to the right due to the smoothing effect of averaging; the overall shape of the variogram has been preserved. Note the average variogram is not computable at a spacing less than the coarse grid element size.

The presence of a proportional effect is also confirmed for this example with a cross-plot of the mean and standard deviation of each element for a single realization (Figure 5-3). Checking against the analytical relationship requires evaluating Equation 5.12 at $\mathbf{h} = 0$, written as Equation 5.14.

$$\sigma_Z^2 = m^2\left[\exp\left(\sigma_Y^2\right) - 1\right] \qquad 5.14$$

Assuming second order stationarity, the variance of Y is constant throughout the field and is determined using the dispersion variance covered in Chapter 1 (Equation 5.15). The dispersion variance, discussed in Chapter 1, between the coarse elements and the domain was already calculated as $D^2(V,\Omega) = 0.88$.

$$D^2(v,\Omega) = D^2(v,V) + D^2(V,\Omega) \qquad 5.15$$

For large domains and Gaussian field, the dispersion variance between point scale (v in this case) and the domain is assumed 1; however, for the domain in this example, $D^2(v,\Omega) = 0.98$. The variance, σ_Y^2, used in Equation 5.14 is then equal to $0.98 - 0.88 = 0.10$. This value is approximated from the lognormal realizations using linear regression of the form of Equation 5.16, where k is the slope.

$$\sigma_Z^2 = km^2,\ \sigma_Y^2 = \ln(k+1) \qquad 5.16$$

Based on this approach using the 50 realizations, $\sigma_Y^2 = 0.096$. Resulting lines are shown on Figure 5-3 and are in close agreement.

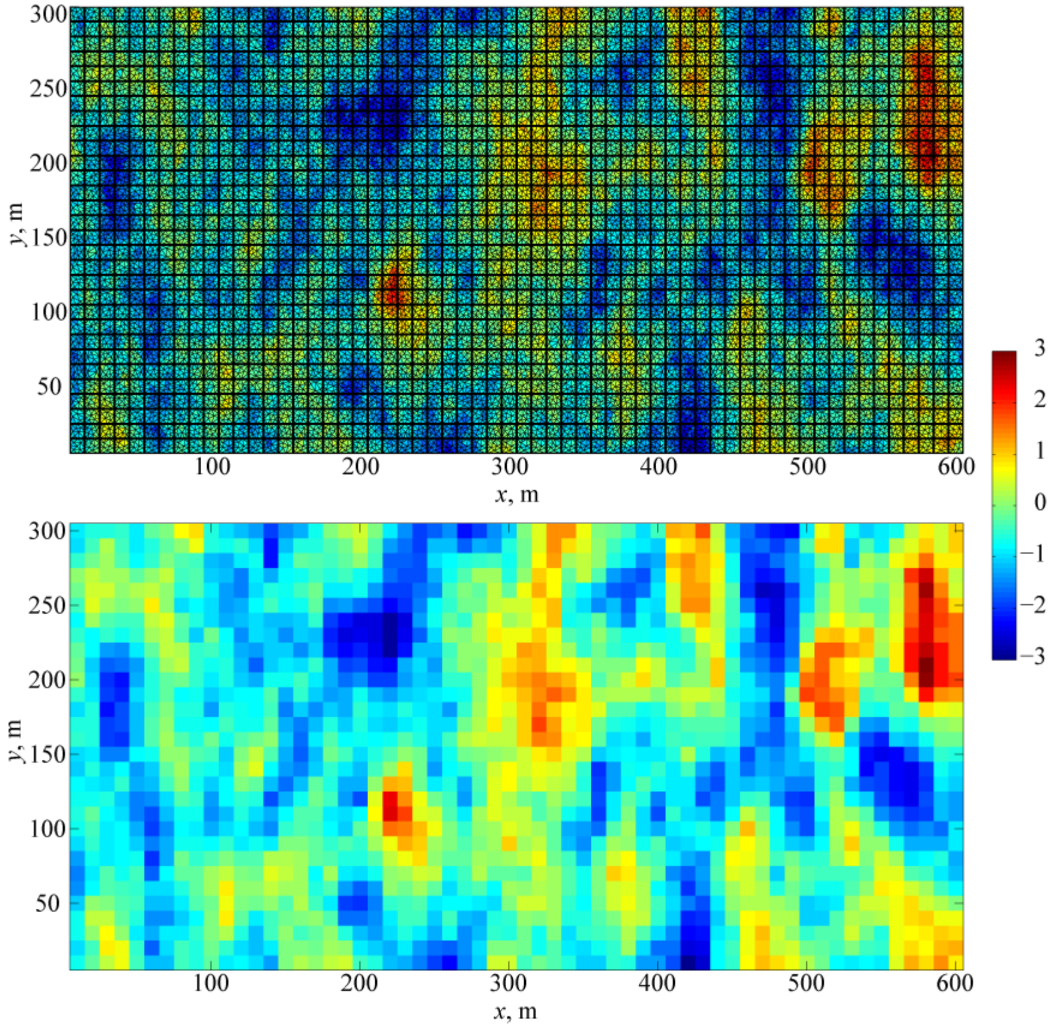


Figure 5-1: Gaussian realization on a fine scale triangular grid (top) and arithmetically averaged to a coarse regular grid (bottom).

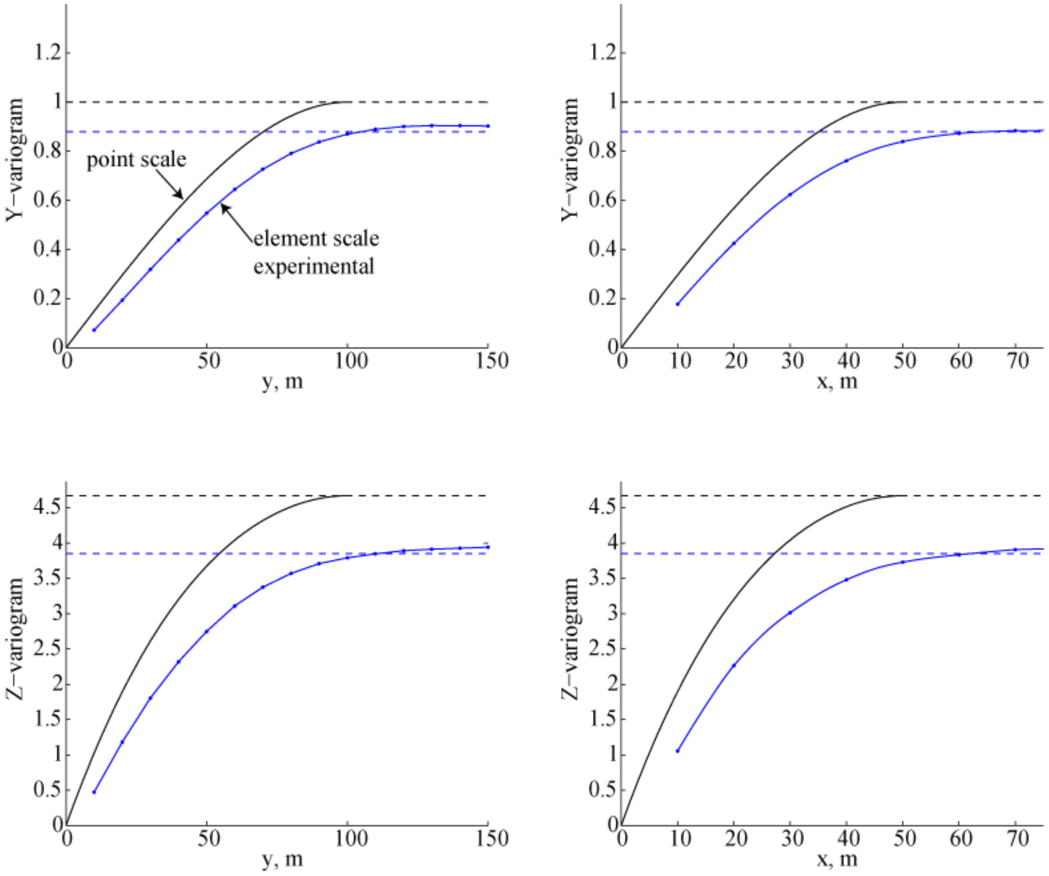


Figure 5-2: Variograms of averaged Gaussian, *Y*, and lognormal, *Z*, variables. Sills are shown with dashed lines.

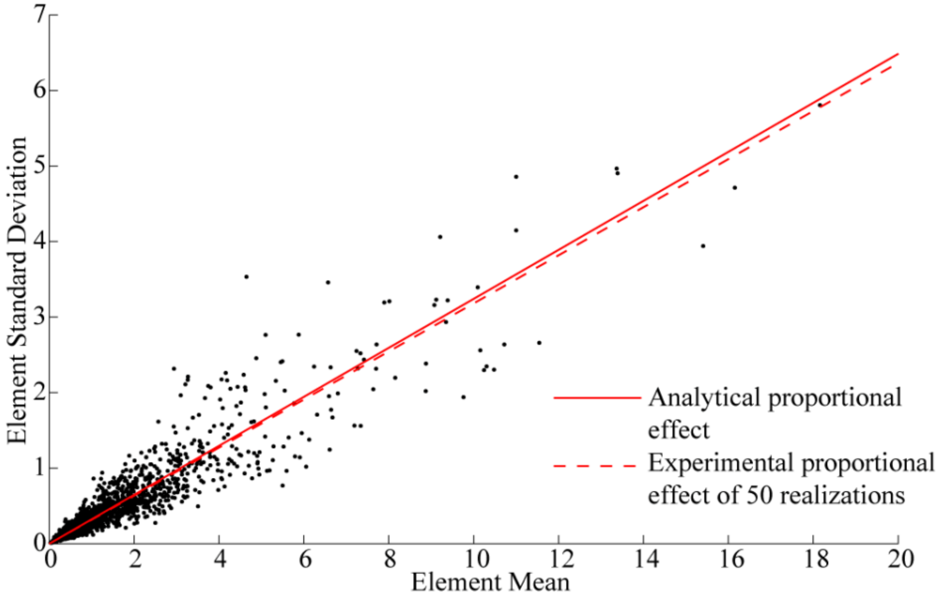


Figure 5-3: Proportional effect of an averaged lognormal realization.

Reproduction of statistics using this regular grid example indicates the methodology of refinement with simplex grids for averaging is sound and that psgsim is generating Gaussian fields with the correct spatial structure. A regular grid was used in this case to allow computation of the experimental variogram at the element scale. Equivalently, a grid of hexagons or triangles with the same volume could have been used.

5.1.3 Global Statistics

This section compares the reproduction of global statistics such as the mean and variance of a realization using current practices and regular grids with unstructured grids and the discretization approach. Since the targeted use of unstructured grids is flow simulation, it is necessary that the new approach does not introduce a bias in porosities, permeabilities, or initial fluid volumes in place. In Chapter 3, it was shown that psgsim yields Gaussian realizations with zero mean and unit variance on average. This indicates there is no bias in realizations though a transform and estimates of pore volume for example would be accurate; however, conditioning data was not involved. The impact of conditioning data is that the variation of a reservoir property within the vicinity of the data is reduced. For example, a large grid element that is consistently assigned a low value due to its proximity to a low porosity sample draws the mean of the realizations down. The discretization process mitigates this type of effect by accounting for the scale and shape of grid elements.

A simple two dimensional example is used to demonstrate the impact of conditioning data and grid element distribution on estimates of oil in place (OIP). OIP is calculated by Equation 5.17, where N is the number of grid elements, ϕ_i, V_i, and S_{wi} are the porosity, volume, and water saturation of element i respectively.

$$OIP = \sum_{i=1}^{N} V_i \phi_i \left(1 - S_{wi}\right) \qquad 5.17$$

The field is 100 by 100 meters with a thickness of 1 meter and random porosity samples are placed on a 9 by 9 regular grid within the area. An arbitrary PEBI grid with very coarse elements in the Northwest quadrant and smaller elements in the Southeast quadrant discretizes the space (Figure 5-4). The porosity value near the center of the largest grid element is consistently chosen to be high with a value of 0.7 while all others are between 0 and 0.5. This is to force a bias in OIP estimates because this value will receive a high kriging weight when estimating the largest grid element. A porosity of 0.7 is somewhat physically unrealistic, but it does make the bias more pronounced for this example.

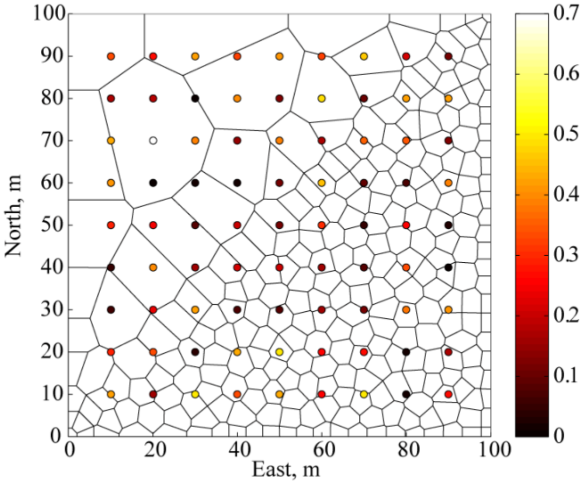


Figure 5-4: Random porosity samples (circles) and PEBI grid.

Estimates of OIP are made using ordinary kriging with the porosity samples as conditioning data and assuming a constant water saturation of 0.33. An isotropic spherical variogram with a range of 50 meters is used. The coarsest possible discretization to consider involves only the PEBI elements. A bias in the estimates of OIP in this case is expected because of the high value near the center of the largest grid element. However, if discretization is applied based on the error criteria developed in Chapter 4, no bias is expected because the scale and shape of the grid elements will be accounted for more accurately. The true OIP value is estimated using a 100 by 100 element regular grid.

Discretization area constraints are derived using the cbarerror program from Chapter 4 with the variogram used for kriging and the areas covered by the PEBI elements. The smallest element is 3.3 m^2 and the largest is 646.6 m^2. The resulting error surface and number of discretization elements curve for an error of 0.003 are shown in Figure 5-5. Converting the discretization constraints to area constraints for the program Triangle results in 3,019 triangular elements. A quality constraint of 20 degrees was used.

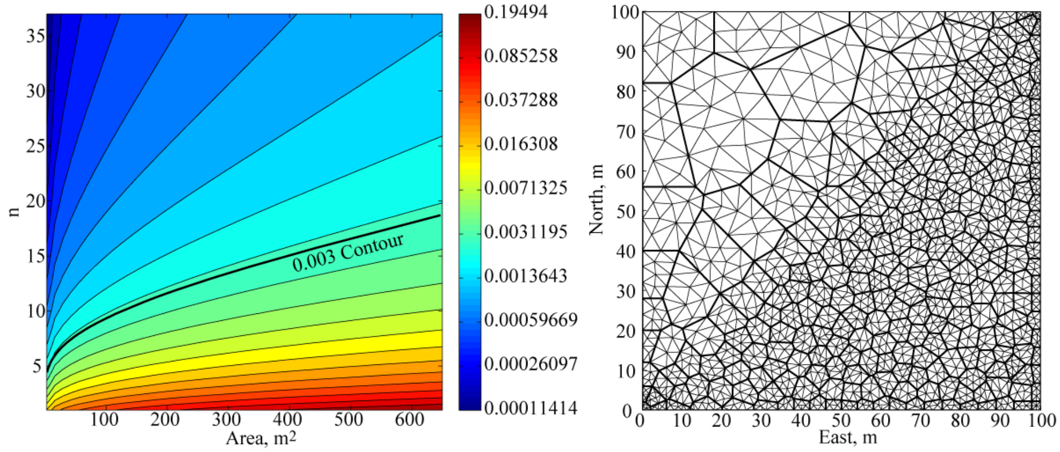


Figure 5-5: Error surface (left) and resulting discretization using the 0.003 error contour.

Fifty different random porosity data sets were used to assess the distribution of OIP estimate errors. Several different discretization constraint contours were used as well. The error is the difference between the value obtained using the regular grid and that obtained for the unstructured grid (Table 5-1).

Table 5-1: Bias due to grid coarseness.

Constraint Contour	Grid elements	Mean Error (OIP)	Mean Error (%)	Standard Deviation
None	266	107.0	5.92	36.2
0.050	1,378	7.8	0.43	9.6
0.007	1,901	-0.3	-0.02	6.3
0.003	3,019	0.8	0.05	5.5

Discretization with low enough error criteria (in this case 0.007) reduces the bias in global statistics that may occur due to a grid design. In this example, one very large grid element located close to a high porosity value gave a 6% overestimation of OIP if using no discretization. Since grid designs in this dissertation are intended primarily for flow simulation, element volume is not necessarily related to the magnitude of a reservoir property and biases in statistics may be either above or below the mean. Sufficiently fine discretization reduces the bias in any case.

An advantage related to the flexibility of unstructured grids is also present for this example. Consider that the portion of the reservoir below a northeast line is more important, i.e. it has smaller elements. The average area of triangular elements for the 0.007 case in Table 1 in this region is 3.4 m^2. A regular grid with equivalent areas would be roughly 54 by 54 (elements with side-length of 1.85 m), having a total of 2,916 elements, which is 53 % higher than the triangular grid.

5.1.4 Near-well Heterogeneity

A noted advantage of unstructured grids is their flexibility allowing smaller elements to be placed in areas where higher accuracy of the flow field is required. This is common near wells. The correctness of the flow field is also important and it depends on the underlying permeability model. Therefore, it is desirable to generate reservoir property models that represent the underlying heterogeneity. Properties should reflect the distribution and covariance structure that have been interpreted for them. This is accomplished using discretization.

It was shown in Section 5.1.1 that the covariance is reproduced using the proposed discretization approach; however, the covariance is incorrect using the common approach with regular grids. In the common approach, regular grid elements are partitioned by the unstructured grid and all components are assigned the value of the element (Figure 5-6). Of interest is when the volumes of unstructured grid elements are smaller than the regular grid elements. This is possible near wells where radial grids or other forms of local grid refinement are used.

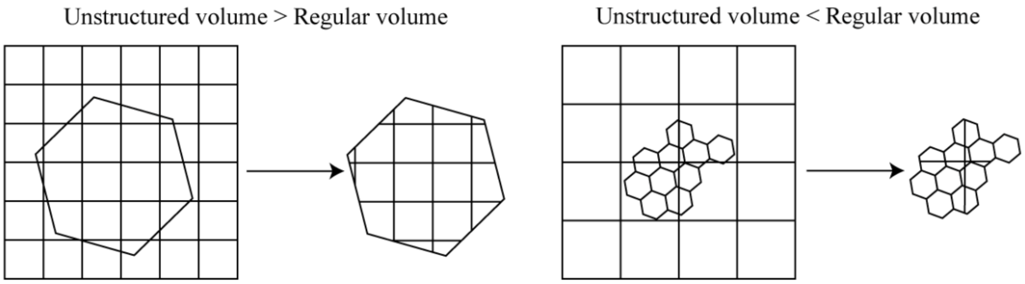


Figure 5-6: Intersection of a regular grid with an unstructured grid.

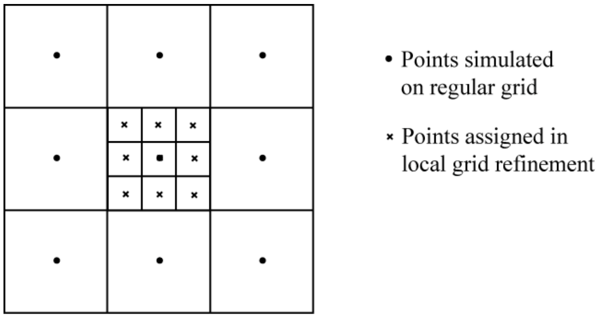


Figure 5-7: Regular grid and local grid refinement.

To show that the covariance between unstructured grid elements is incorrect the LU formalism is used. Consider a case where the unstructured grid aligns exactly with the regular grid, except where there is a well some local grid refinement is used (Figure 5-7). Values simulated on the regular grid are defined by $\mathbf{y}_s$ and values assigned to the local grid refinement as $\mathbf{y}'_s$. Denote y_{si} the simulated value of the regular element that contains the refinement and y_{sj} the value for any other element. All $y'_{sk} = y_{si}$, where y'_{sk} is the value assigned to an element of the refinement.

Using point scale LU simulation the $\mathbf{y}_s$ values are defined by Equation 5.4, reiterated here as Equation 5.18.

$$\mathbf{y}_s = \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y} + \mathbf{L}_{22}\mathbf{r} \qquad 5.18$$

The covariance between a refined point and regular grid point are derived in Equation 5.19 and between two refined points in Equation 5.20. The vectors $\mathbf{p}_i$ and $\mathbf{p}_j$ are used to extract the i and j elements of $\mathbf{y}_s$: $\mathbf{p}_i = [0 \cdots 0 \; 1_i \; 0 \cdots 0]^T$. Note that $E\{\mathbf{r}\mathbf{r}^T\}$ is the identity matrix since $\mathbf{r}$ is a vector with independent random values with zero mean and unit variance. Using point scale Gaussian simulation, the only case where Equation 5.19 is correct is when the refined point and regular grid element center are coincident. Equation 5.20 is only correct when there is only one refined point, in other words when there is no refinement. These results apply to any case where unstructured elements are embedded within a regular grid element. In all cases, the random field is too smooth in these areas and the resulting pressure gradient and flux will not be representative of the actual heterogeneity in those areas.

$$\begin{aligned}
Cov\left\{y_{sj}, y'_{si}\right\} &= E\left\{y_{sj} \cdot y'_{si}\right\} - E\left\{y_{sj}\right\} E\left\{y'_{si}\right\}, \text{ where } E\left\{y_{sj}\right\} = E\left\{y'_{si}\right\} = 0 \\
&= E\left\{\left(\mathbf{p}_j^T \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y} + \mathbf{p}_j^T \mathbf{L}_{22}\mathbf{r}\right)\left(\mathbf{p}_i^T \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y} + \mathbf{p}_i^T \mathbf{L}_{22}\mathbf{r}\right)^T\right\} \\
&= E\left\{\left(\mathbf{p}_j^T \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y} + \mathbf{p}_j^T \mathbf{L}_{22}\mathbf{r}\right)\left(\mathbf{r}^T \mathbf{L}_{22}^T \mathbf{p}_i + \mathbf{y}^T \mathbf{C}_{11}^{-1}\mathbf{C}_{21}^T \mathbf{p}_i\right)\right\} \\
&= \mathbf{p}_j^T \mathbf{C}_{21}\mathbf{C}_{11}^{-1} E\left\{\mathbf{y}\mathbf{y}^T\right\} \mathbf{C}_{11}^{-1}\mathbf{C}_{21}^T \mathbf{p}_i + \mathbf{p}_j^T \mathbf{L}_{22} E\left\{\mathbf{r}\mathbf{r}^T\right\} \mathbf{L}_{22}^T \mathbf{p}_i \\
&= \mathbf{p}_j^T \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{C}_{21}^T \mathbf{p}_i + \mathbf{p}_j^T \mathbf{L}_{22}\mathbf{L}_{22}^T \mathbf{p}_i \text{ , let } \mathbf{K}_{22} = \mathbf{L}_{22}\mathbf{L}_{22}^T \\
&= \mathbf{p}_j^T \left(\mathbf{C}_{22} - \mathbf{K}_{22}\right)\mathbf{p}_i + \mathbf{p}_j^T \mathbf{K}_{22}\mathbf{p}_i \\
&= Cov\left\{y_{sj}, y_{si}\right\} \neq Cov\left\{y_{sj}, y'_{si}\right\} \qquad \square
\end{aligned} \tag{5.19}$$

$$\begin{aligned}
Cov\left\{y'_{sk}, y'_{si}\right\} &= E\left\{y'_{sk} \cdot y'_{si}\right\} - E\left\{y'_{sk}\right\} E\left\{y'_{si}\right\}, \text{ where } E\left\{y'_{sk}\right\} = E\left\{y'_{si}\right\} = 0 \\
&= E\left\{\left(\mathbf{p}_i^T \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y} + \mathbf{p}_i^T \mathbf{L}_{22}\mathbf{r}\right)\left(\mathbf{p}_i^T \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y} + \mathbf{p}_i^T \mathbf{L}_{22}\mathbf{r}\right)^T\right\} \\
&\vdots \\
&= \mathbf{p}_i^T \left(\mathbf{C}_{22} - \mathbf{K}_{22}\right)\mathbf{p}_i + \mathbf{p}_i^T \mathbf{K}_{22}\mathbf{p}_i \\
&= Cov\left\{y_{si}, y_{si}\right\} \neq Cov\left\{y'_{sk}, y'_{si}\right\} \qquad \square
\end{aligned} \tag{5.20}$$

5.2 Flow-based Upscaling

In Chapter 4, it was shown that transmissibility upscaling using MPFA and triangular interaction regions yields the correct transmissibility for homogeneous media. This was done for one interaction region. Validation for problems involving larger grids is done in this section. Flow simulation of water flooding is done on the models. Results are summarized using pore volume injected, PVI, and water fraction produced, F_w, plots. It is assumed that the correct solution is that obtained using a high resolution regular grid and the two-point flux approximation. Examples include homogeneous media, layered media, and general heterogeneous media. These are shown in two dimensions. A similar set of examples is used by Prevost (2003) for validating permeability upscaling and streamline simulation methods. Curves are checked visually for two characteristics: agreement of breakthrough time, that is, the instant when the fraction of water produced becomes positive; and overall closeness of the curves for the duration of the flow simulation.

5.2.1 Homogeneous Media

A 100 m by 100 m domain is considered with a constant permeability of 1 Darcy. Two scenarios are considered: 1 – boundary conditions to produce flow in the x direction; 2 – quarter five-spot boundary conditions. Reference F_w curves are based on a 100 by 100 regular grid of 1 m elements. Grid element centers for the unstructured grid were generated using the program Triangle (Shewchuk, 1996) and an area constraint of 33 m^2 which results in 480 interaction regions and 272 polygonal grid elements (Figure 5-8). Average volume of the polygonal elements is 36.8 m^2. The grid was discretized using

Triangle as well with an area constraint of 1 m^2 and quality constraint of 27 degrees. A quality of 27 degrees was used to give good quality elements for flow based upscaling, but not cause excessive discretization near vertices of the coarse grid that was occurring when using an angle of 30 degrees. Each interaction region is discretized into approximately 37 triangles (Figure 5-9).

Transmissibilities for the coarse scale interaction regions were computed using the local upscaling technique developed in Chapter 4. Linear pressure boundary conditions were used for each local problem. Resulting transmissibilities are used to build the system of equations for flow simulation. Good agreement is observed between the regular grid and coarse grid curves in the PVI plots (Figure 5-10). Some deviation in breakthrough time is expected because the coarsening factor (36.8) is high and this is observed for the flow in x scenario. The curves are very similar in the quarter five-spot scenario.

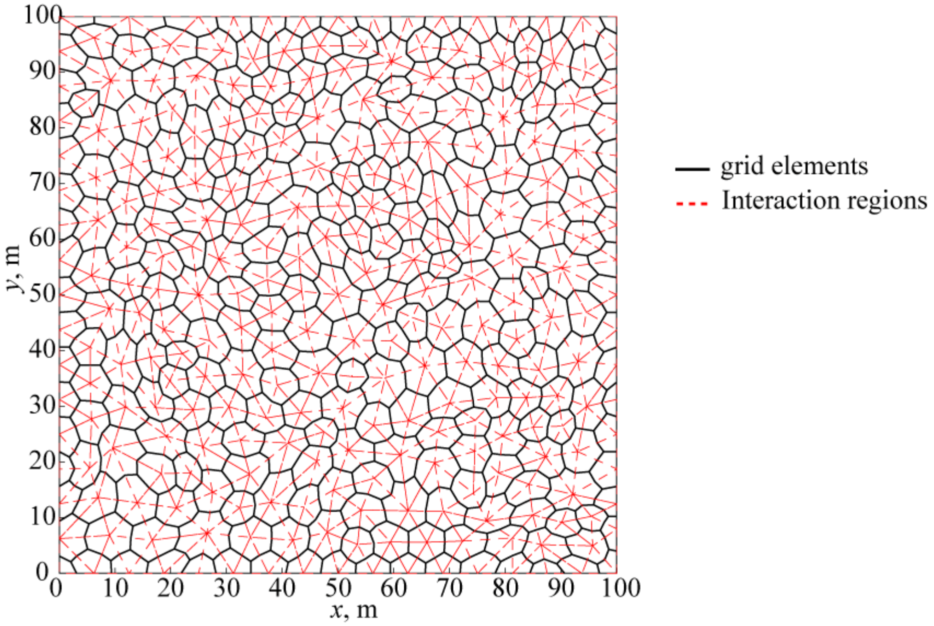


Figure 5-8: Unstructured grid for homogeneous flow validation.

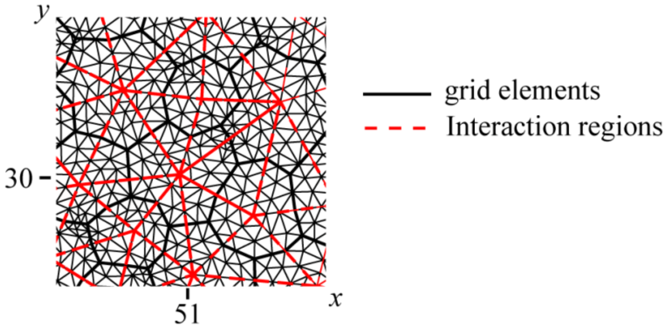


Figure 5-9: A portion of the unstructured grid at $x = 51$ m and $y = 30$ m showing discretization.

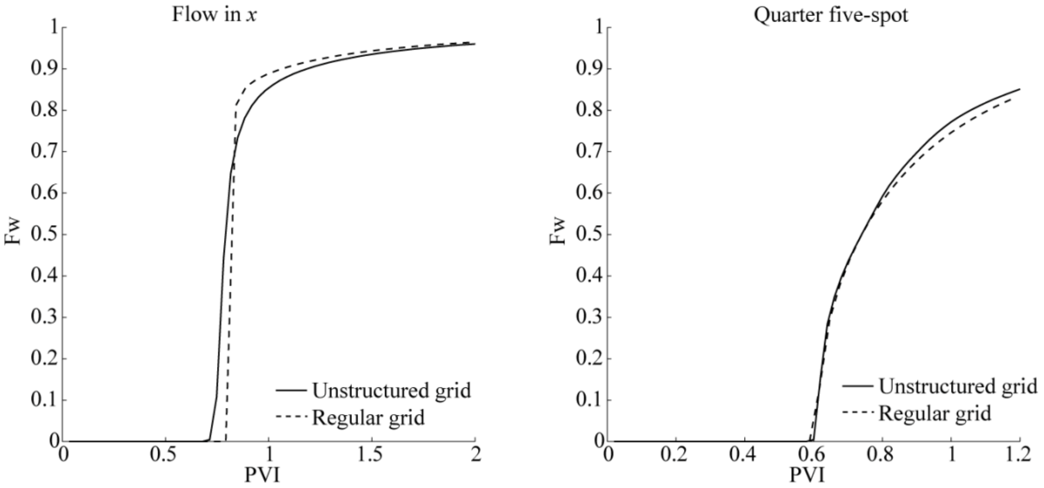


Figure 5-10: PVI-F_w curves for the homogeneous flow example.

5.2.2 Layered Media

The same domain and scenarios that were used for the homogeneous example are used here, with the addition of flow in y. The domain is divided into four equal horizontal layers with the same permeability in y, but a different permeability in x for each layer (Figure 5-11). Generating a grid that does not take into account the permeability anisotropy may result in a non-monotone system of equations and unrealistic fluctuations in the pressure. This is a documented problem with MPFA (Eigstad et al, 2002; Prevost, 2003; Mlacnik and Durlofsky, 2006) and to account for it, the space is transformed to remove the anisotropy. In this case, anisotropy is removed by scaling the space in y. Adequate scaling to achieve a stable system of equations for flow simulation was accomplished by scaling layer 1 by 2, layer 2 by 3, layer 3 by 4.5, and layer 4 by 6. Although layer 1 did not require scaling, this was done to inflate the number of elements generated for that layer to smooth the transition in element volume from layer 1 to layer 2.

Layers are only approximately matched by the polygonal grid, which consists of 1,075 elements (Figure 5-12). Reference PVI-F_w curves were generated using a 100 by 100 element regular grid. Good agreement is observed for flow parallel and perpendicular to the layers (Figure 5-13) and for the quarter five-spot problem (Figure 5-14). For flow parallel to the layers, the breakthrough time of each layer indicated by a vertical jump in F_w on the reference curve is smoothed out by the unstructured grid. The initial breakthrough time is similar in all three cases.

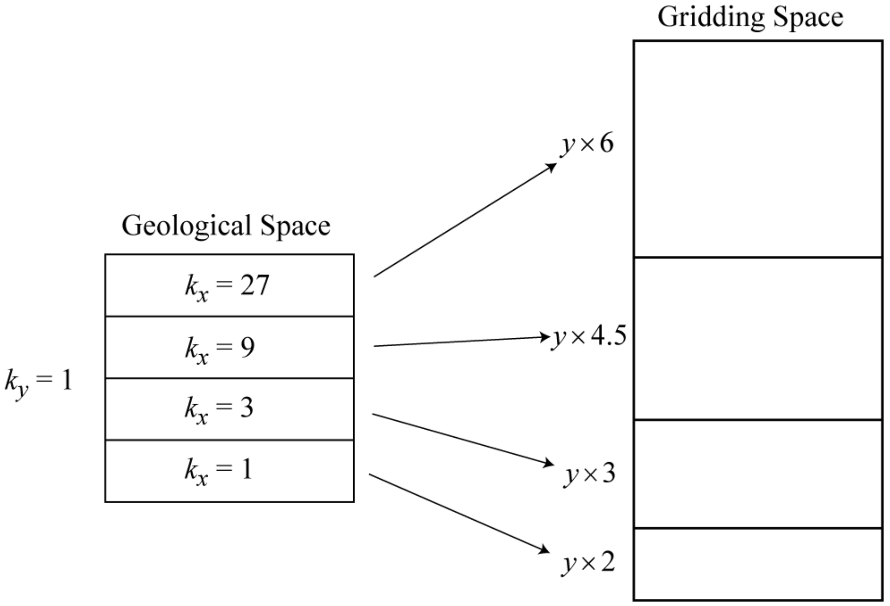


Figure 5-11: Layered media in geological and gridding space.

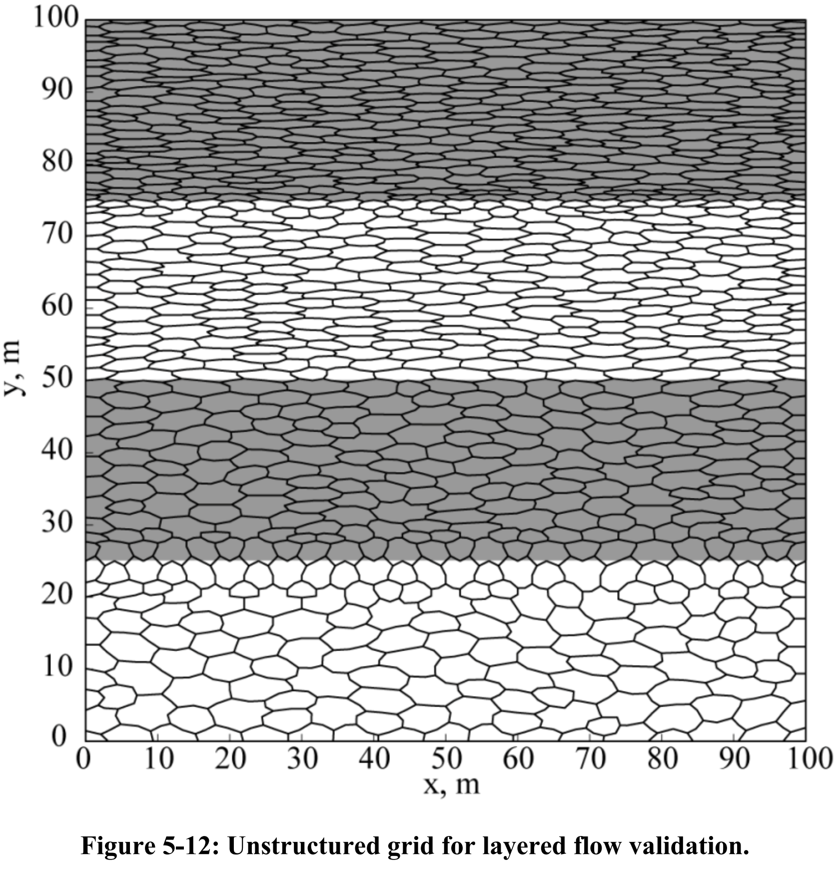


Figure 5-12: Unstructured grid for layered flow validation.

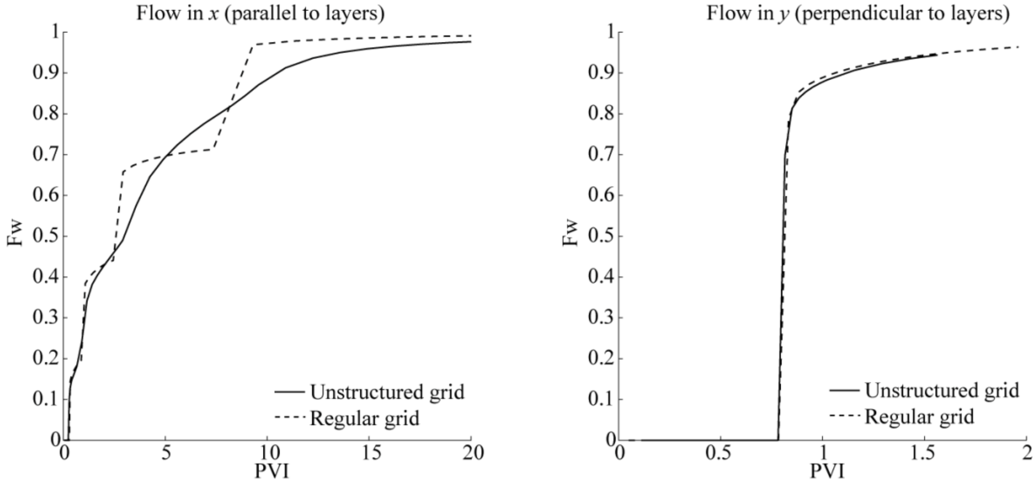


Figure 5-13: PVI-F_w curves for flow parallel and perpendicular to the layers.

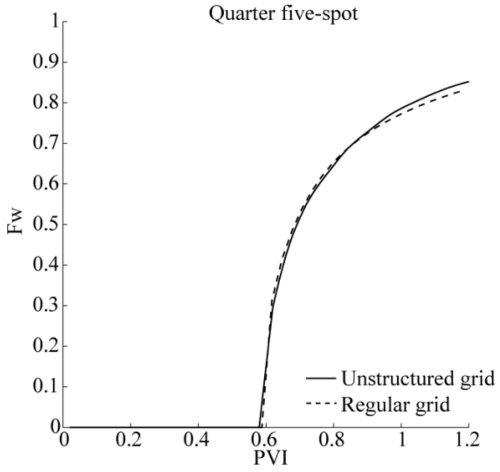


Figure 5-14: PVI-F_w curves for quarter five-spot flow in layered media.

5.2.3 Heterogeneous Media

Using the same domain and quarter five-spot boundary conditions, flow simulation on general heterogeneous media between a regular grid and unstructured grid are compared. The unstructured grid consists of 853 elements. Discretization resulted in 27,003 triangular elements that were used to generate a permeability model. Unconditional simulation on the triangle centers was carried out to give a model with three facies and permeability (Figure 5-15). Permeability was resampled using nearest neighbour interpolation onto a 100 by 100 element regular grid for the reference flow curve.

Two comparisons are done for this example: 1 – the water front near breakthrough time and 2 – the PVI-F_w curves. Water fronts (Figure 5-16) are similar; however, two issues caused by upscaling are observed. At $(x = 60, y = 10)$, water is flowing into a region in the unstructured grid where it is not in the regular grid. The permeability in this region is high and a connected segment of high permeability can be traced along the lower right corner of the model; however, this segment is not connected to the producer. Upscaling causes an increased connectivity at $(x = 90, y = 50)$ that draws water through the conduit that does not appear in the regular grid. The opposite issue occurs at $(x = 50, y = 95)$, where flow that appears on the regular grid does not on the unstructured grid. The thin segment of high permeability on the top edge of the model in this area has been mitigated by upscaling, leading to reduced flow. Despite these differences, the PVI-F_w curves are close (Figure 5-17), with the unstructured grid having a lower breakthrough time and lower F_w for the second half of the flow simulation. Lower F_w is likely caused by the excess water that is flowing through the $(x = 60, y = 10)$ region.

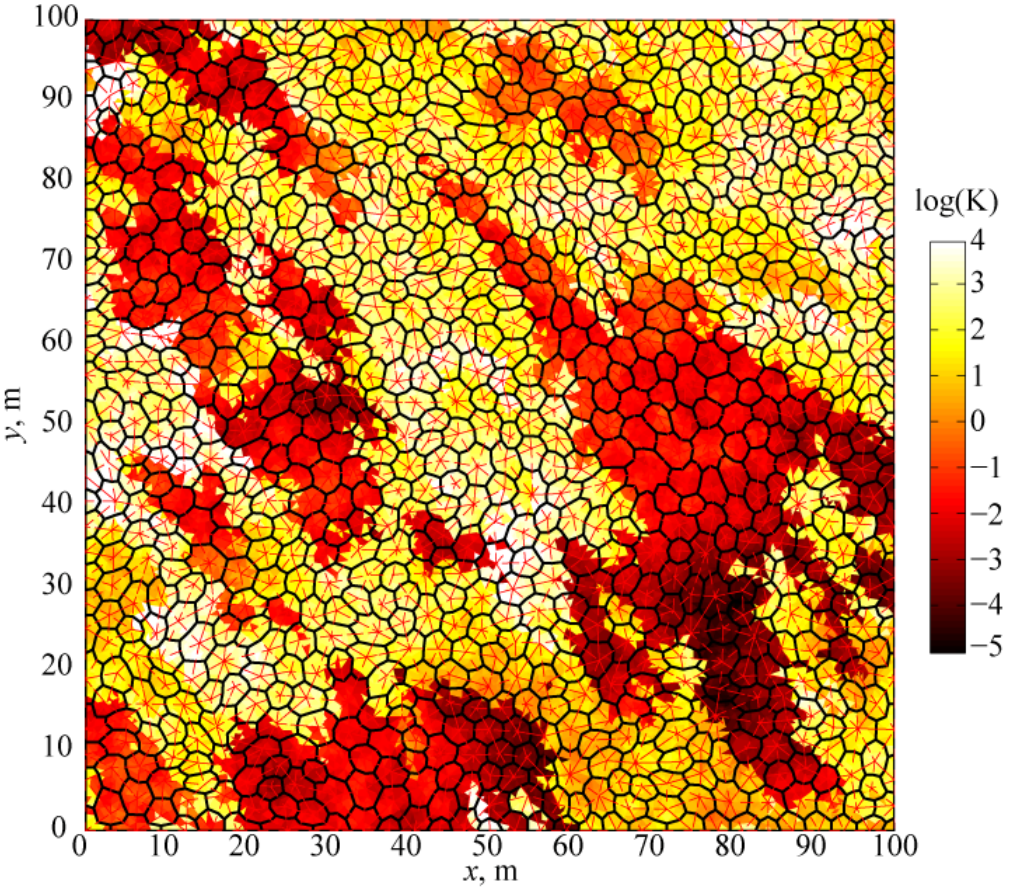


Figure 5-15: Unstructured grid and heterogeneous permeability model.

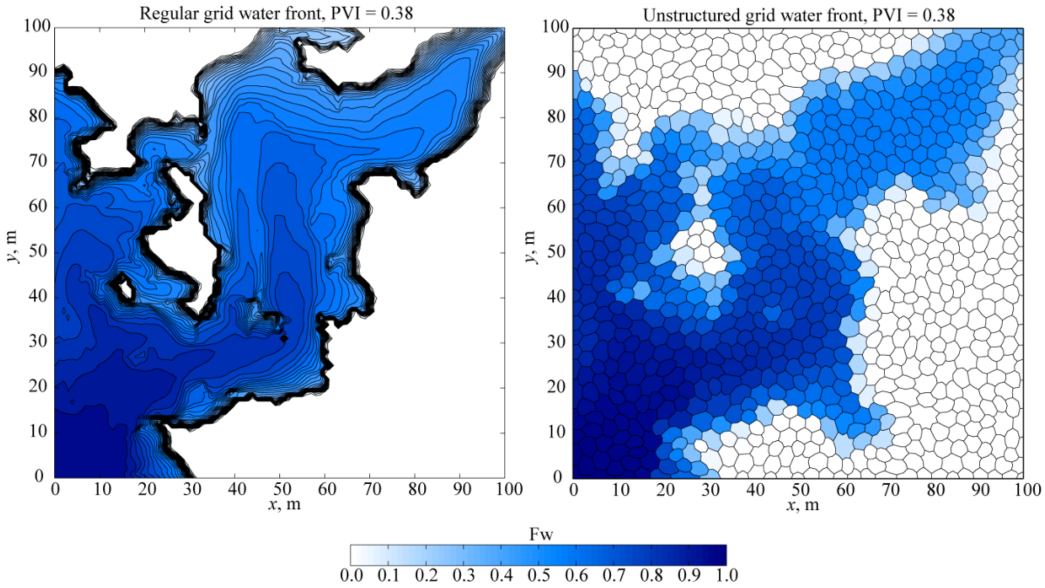


Figure 5-16: Water front comparison near breakthrough time.

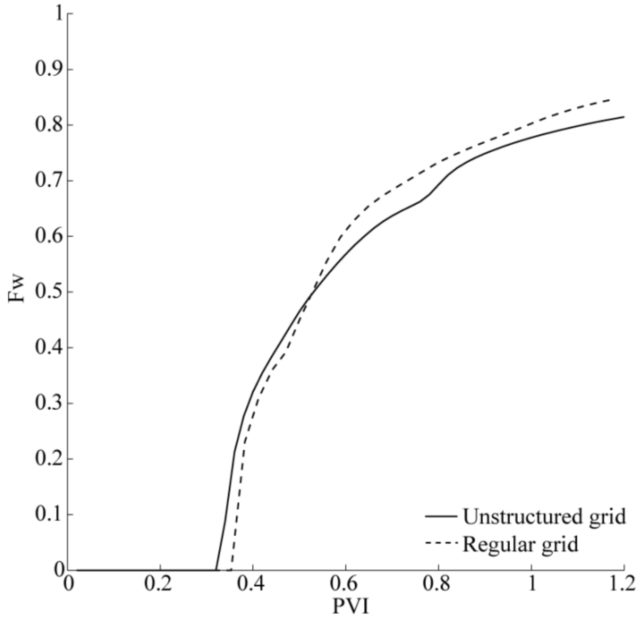


Figure 5-17: PVI-F_w curves for quarter five-spot flow in heterogeneous media.

5.3 Case Study

The focus of this study is to demonstrate the geostatistical components of the proposed reservoir modeling workflow for unstructured grids. Components of the workflow that are covered include:

1. Data assembly
2. Structural modeling of horizons and faults
3. Space transformation: geological space ↔ depositional space
4. Upscaling from log scale to modeling scale
5. Statistical analysis of model scale, including:
    a. Declustering
    b. Correlations
    c. Variography
    d. Discretization error
6. Preliminary mapping using a regular grid
7. Unstructured grid design
8. Discretization
9. Property modeling and validation
10. Upscaling arithmetic properties

Data made available by the Rocky Mountain Oilfield Testing Centre (RMOTC) and U.S. Department of Energy are used for this study. These data have been used in other research and studies as well (Olsen et al, 1993; Garcia, 2005; Friedmann and Stamp, 2006 to list a few) and it is from the federally owned Naval Petroleum Reserve No. 3, also called the Teapot Dome field. It consists of several geological units. The Tensleep formation is the focus of this case study.

Three types of data are available: core, geophysical logs, and 3D seismic. Variables derived from the core that are used include horizontal permeability, k_h, vertical permeability, k_v, porosity, ϕ, water saturation, S_w, and oil saturation, S_o. Of the available logs, those used include gamma ray, GR, bulk density, RHOB, neutron porosity, NPHI, deep resistivity, RD, and shallow resistivity, RS. There are 37 wells that intersect the Tensleep formation, 31 of them with logs and 18 with core. The boundary of the Tensleep Dome formation, topography, and wells are shown in Figure 5-18. Seismic covers the whole area and is used for structural modeling. Horizons for the Tensleep top and base are available in the time domain. Available well markers are used to convert the horizons from time to depth.

Log scale data are used for modeling. The logs are processed into four reservoir properties including lithology, porosity, water saturation, and oil saturation. All properties are modeled using geostatistical methods. Models of permeability are generated based on a relationship with porosity in the preliminary mapping stage to be used in flow based grid design. Lithology consists of sand and dolomite and is assigned based on well markers. There are four markers, which ordered by increasing depth are ASand, BDolo, BSand, and C1Dolo (Figure 5-19). Lithology is assigned as upper sand in the ASand – BDolo and lower sand in BSand – C1Dolo intervals. BDolo – Bsand is assigned as dolomite. Well markers are typically used to model surfaces that separate different stratigraphic units; however, for this study the BDolo – Bsand layer is assumed to be discontinuous so indicator simulation is used to model it. Other logs are processed using standard procedures and calibrated using the core data (Ellis and Singer, 2007).

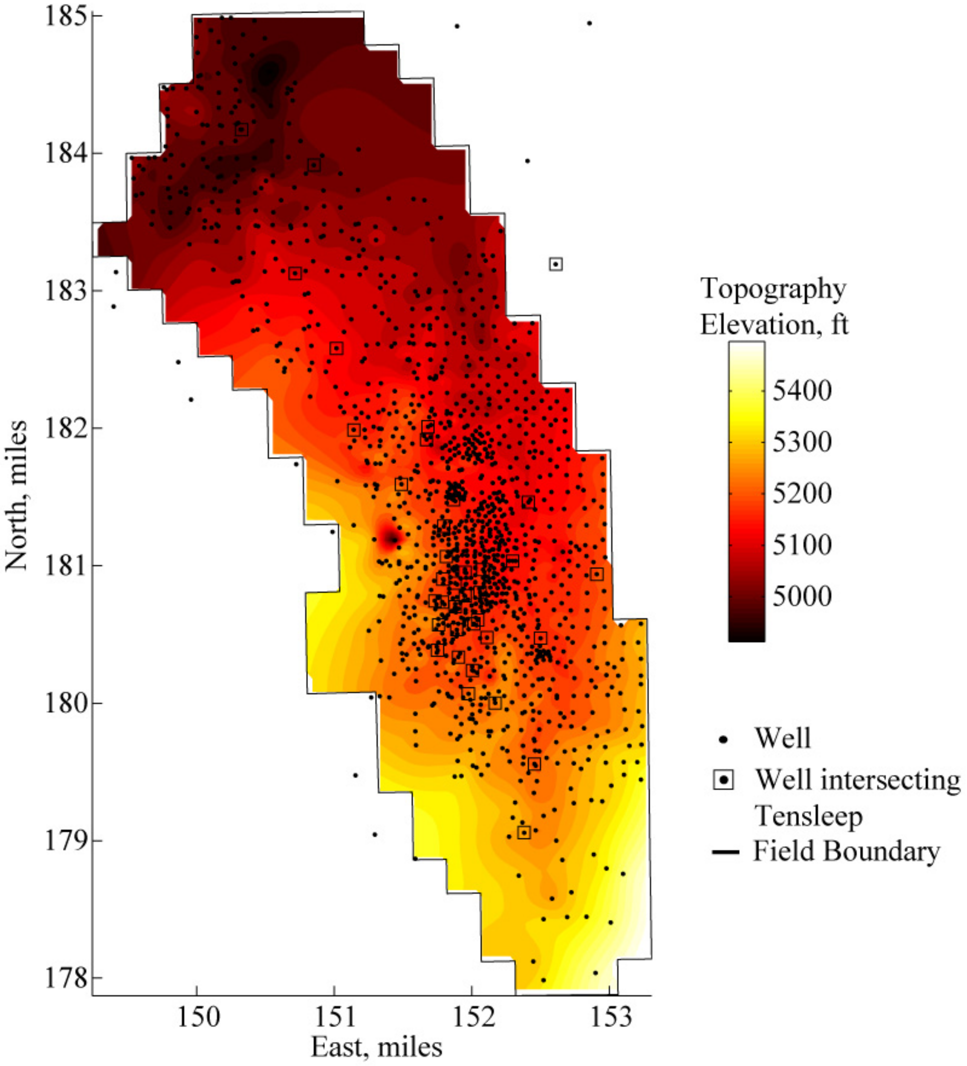


Figure 5-18: Teapot Dome field boundary and topography.

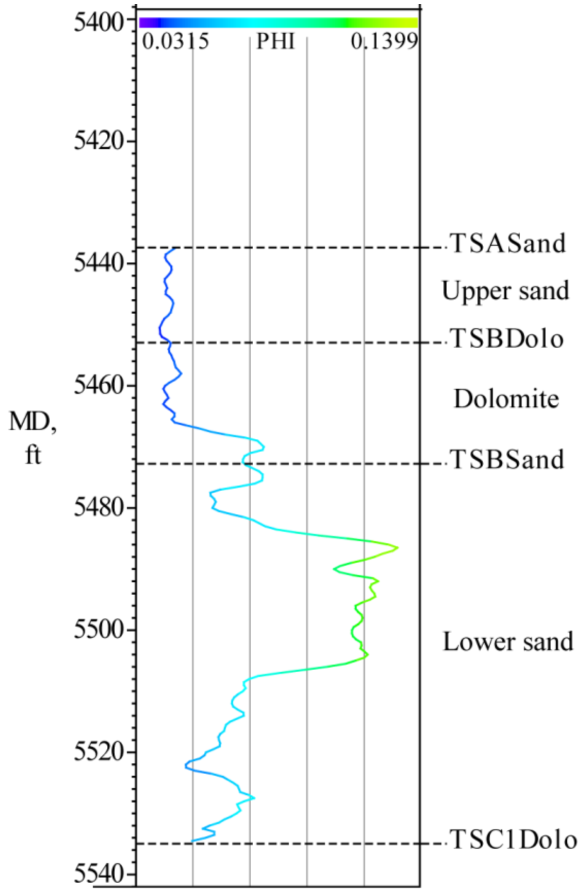


Figure 5-19: Well tops in measured depth for the Tensleep formation shown with a porosity log curve.

5.3.1 Structural Modeling

Horizons of the Tensleep top and base, respectively the ASand and C1Dolo markers, interpreted from the seismic data are available in the RMOTC data set in the time domain. These horizons were converted to the depth domain using a velocity model generated from the well markers. Resulting surfaces match the well markers. Based on the seismic, there are also several faults in the formation; however, the data set did not include any fault interpretations. Six faults were interpreted from the seismic by hand. More information on structural modeling can be found in Mallet (2002) or Groshong (2006).

Contour maps of the top and base structure are shown in Figure 5-20. Contours terminate at the limit of the horizon interpretations and the locations of faults are indicated with trace lines. The structure is an anticline that trends North North-West. Its trend can be traced along the shallowest points on the contour maps. Faults in the modeling area identified in Figure 5-20 have the largest vertical offset, up to 100 feet near mile 151.8 in East and 180.5 in North.

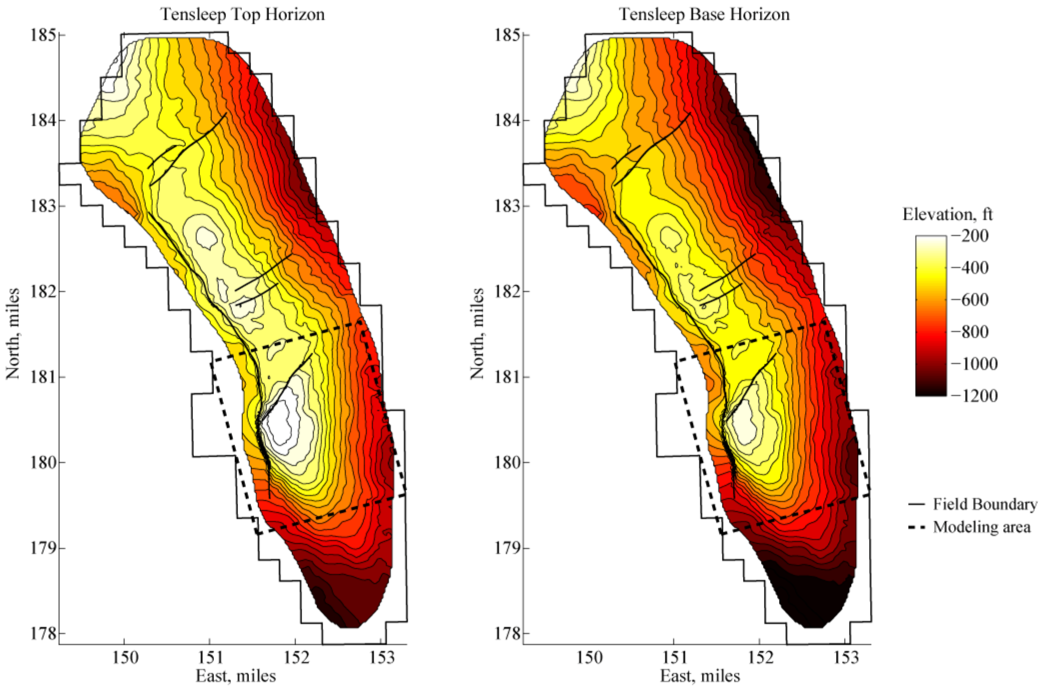


Figure 5-20: Tensleep horizons, fault traces, and modeling area.

5.3.2 Space Transformation

Transforming from geological space to depositional space involves undoing the faulting and the fold of the anticline. Stratigraphic coordinates are used, but to account for the faulting and deformation of the horizons, they are applied with a background tetrahedral grid. The stratigraphic coordinate system axes are labelled u, v, t. Within the modeling area, the following surfaces were triangulated using the program Triangle: top, base, two faults, and model boundary. These surfaces are stitched together to form a sealed model or piecewise linear complex (PLC), which is used as input for the program TetGen to produce the background tetrahedral grid (Figure 5-21).

Nodes of the tetrahedral grid are assigned depositional coordinates. All nodes on the top receive a value of zero while all those on the base receive a value of -110 feet, which is approximately the average thickness of the formation multiplied by negative one. Stratigraphic coordinates of interior tetrahedron nodes are interpolated using kriging and an isotropic Gaussian variogram with a range of 316 ft that is equivalent to a variance of 100,000 ft^2. The variance is chosen large to give a smooth interpolation. Deviations to flatten the model in z are computed as $dz = t - z$. There is some deviation in the horizontal plane due to faulting that must also be accounted for. Deviations in x and y, respectively dx and dy, are computed for any node associated with a fault by projecting it onto the fault center line in the xy-plane. Resulting deviations are smoothed away from the faults so that when mapping between coordinate systems, tetrahedra do not fold in on themselves. Stratigraphic coordinates in the uv-plane are then $u = x + dx$ and $v = y + dy$. Horizontal deviations due to faults do not exceed 63 ft and are not shown.

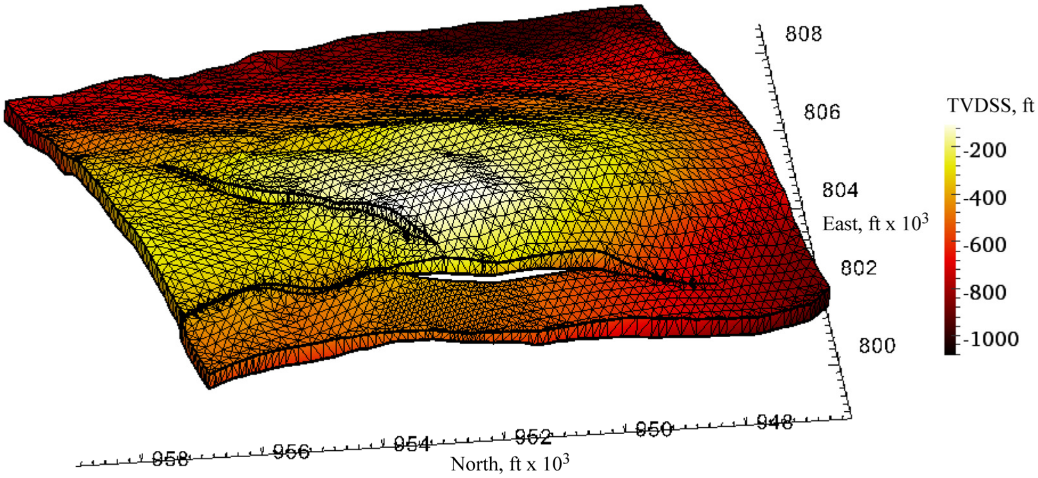


Figure 5-21: Isometric view of the tetrahedral grid generated for mapping from geological space to depositional space. A vertical exaggeration of three is used.

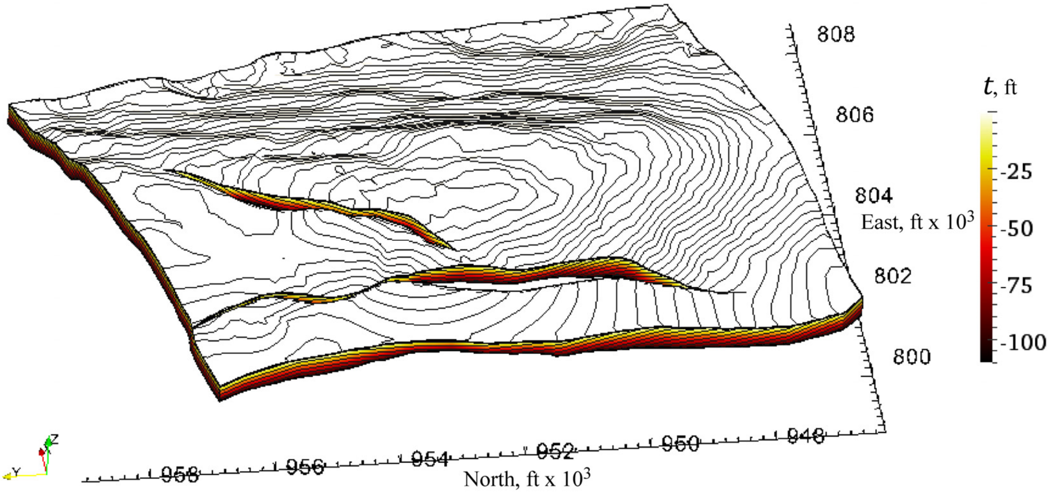


Figure 5-22: Depositional coordinates interpolated throughout the reservoir. A vertical exaggeration of three is used. Contours shown to aid in display of relief.

Deviations are used to flatten the model into depositional space, whereas the stratigraphic coordinate values and actual coordinates of the tetrahedral nodes are used to map points between the spaces. Mapping a point from geological space to depositional space involves the following steps:

1. Find the tetrahedron containing the point, $\mathbf{u}$.
2. Compute the barycentric coordinates of the point, $\boldsymbol{L} = \{L_1, L_2, L_3, L_4\}$.
3. Interpolate the depositional coordinates from the tetrahedron nodes, $\mathbf{t} = \{t_1, t_2, t_3, t_4\}$ to the point: $t_p = \boldsymbol{L}^T\mathbf{t}$.

The same process is used to shade the model by depositional coordinates (Figure 5-22). Statistics used for reservoir modeling including the distribution and variogram are derived in depositional space. To accomplish this all well data is mapped to u, v, t space.

Wells that are inside the modeling area are transformed using the tetrahedral grid method, but for purposes of variogram modeling, it is desirable to retain all well data when the whole domain is stationary. Because the deviations in x and y are small with the maximum observed being 63 ft and with the wells being nearly vertical, the wells outside the modeling area do not necessarily have to be mapped using the tetrahedral grid approach. Standard stratigraphic coordinate mapping discussed in Chapter 2 is applied to these wells (Equation 5.21).

$$t = -\frac{z - z_{ASand}}{z_{ASand} - z_{C1Dolo}} \cdot 110 \qquad 5.21$$

5.3.3 Upscaling: Log Scale to Modeling Scale

Chapter 2 proposed that well log data should be upscaled into equal intervals along the well bore instead of into grid elements, which are not available at this point in the workflow. The interval is chosen with two goals: to unify hard data to a single scale and to limit the amount of variability that is averaged out or smoothed in upscaling process. The highest level of detail possible is based on the log sample spacing, which is sampled with 6 inch spacing along the well bores. Considering all wells, there are 6,579 samples at this resolution.

A variety of approaches can be used to determine a scale for geological modeling. One approach is to attempt to determine an REV (Bear, 1972). Another possibility is to increase the scale until a specified amount of variability is removed from the data. In applications with regular grids, the scale may simply be chosen arbitrarily to limit the size of model that results or it may be based on the resolution to be used in a flow simulator. In typical applications, regular grid cells are chosen between 10 and 50 meters horizontally and 0.1 to 1 meter vertically (Aarnes et al, 2007).

In this study, the change in standard deviation is observed for porosity and water saturation as the interval size along the wells is increased from 6 inches up to 7 ft. They are upscaled using an arithmetic average. Figure 5-23 indicates that a consistent change in variance is observed for porosity up to a length of 5 ft, whereas the change varies significantly for water saturation above a length of 2.5 ft. Vertical resolution for upscaling the logs is chosen at 1.64 ft or 0.5 m. Since the lithology is defined by well markers, no upscaling intervals are chosen to contain both indicators. Upscaling is done in the ASand-BDolo, BDolo-BSand, and BSand-C1Dolo intervals separately for each well.

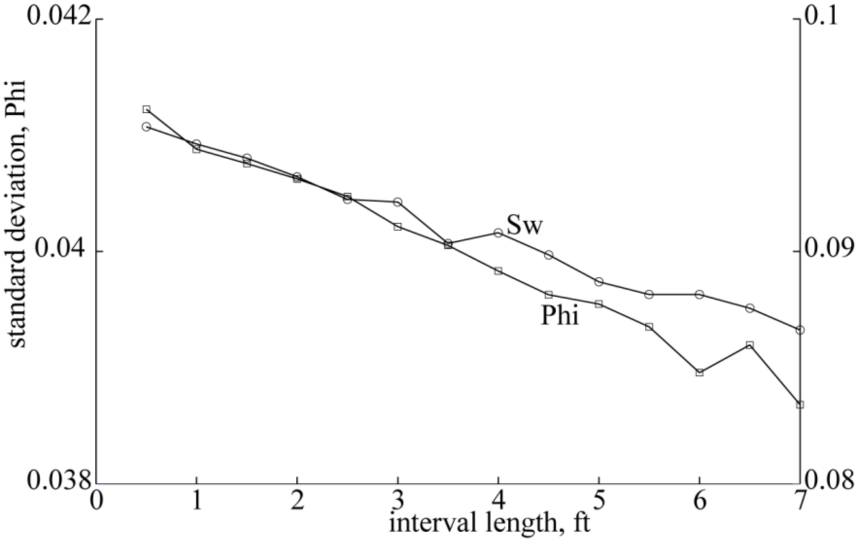


Figure 5-23: Standard deviation as a function of upscaling length.

5.3.4 Statistical Analysis

All statistical analyses are done in depositional space. Upscaled data are first mapped to depositional space using the tetrahedral grid for wells inside the modeling area and Equation 5.21 for all other wells. Processes that are covered in this section include declustering, data transformation, correlation analysis, and variogram modeling. A different data transformation process is used for oil and water saturation since these variables form a composition.

5.3.4.1 Declustering

Declustering is a technique to account for the irregular spacing of wells that intersect the Tensleep formation. Wells are clustered around the peak of the anticline structure and fairly sparse elsewhere. Cell declustering (Deutsch and Journel, 1998) is applied in two dimensions since wells are nearly vertical and the formation thickness is relatively constant. Cell size choice is based on maximizing the mean of Vsh, which is the only continuous property that is defined at all well locations. Some wells were missing logs necessary for deriving ϕ, S_w, and S_o so these are not used for declustering. A declustering cell size and mean plot (Figure 5-24) indicates a maximum plateau is reached for a cell size of approximatey 7,000 ft. This spacing coincides with the average well spacing of 7,124 ft. For this cell size the mean of Vsh is 0.243, an increase of 15% from the non-declustered mean.

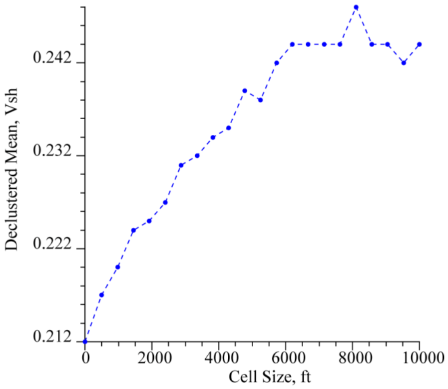


Figure 5-24: Declustering results.

5.3.4.2 Data Transformation

Continuous reservoir properties are often modeled using Gaussian simulation and the variogram of the properties in Gaussian space are required. The normal score transform is most common. Categorical properties are often transformed to indicators. A common practice in geostatistical modeling is to partition the data by lithology and model each independently. Tools such as cross plots and histograms can be used to determine if the data come from different populations, indicating a need to partition the data (Figure 5-25). The data are partitioned for this case study. Models of ϕ, S_w, and S_o are constructed within each lithology independently.

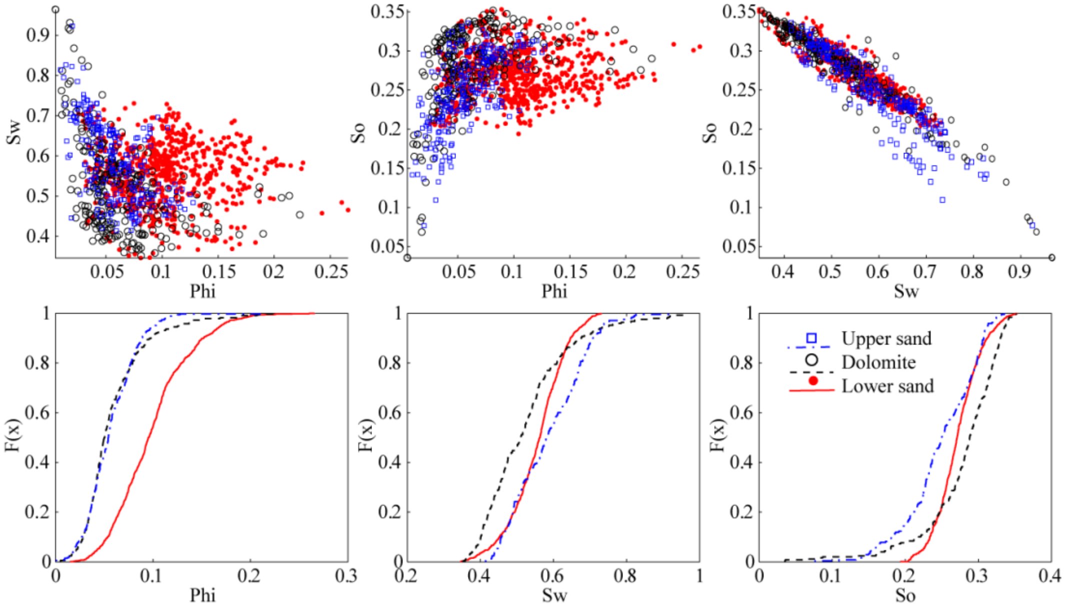


Figure 5-25: Cumulative distributions and cross plots of ϕ, S_w and S_o by lithology.

Water saturation and oil saturation are not directly transformed to Gaussian space. In cases where the reservoir rock is fully saturated with two phases such as oil and water, it is only necessary to model one and the normal score transform can be used directly. In this case, there is water, oil and an additional phase because S_w and S_o do not sum to one.

The unknown phase is labelled S_f and its value is $S_f = 1 - S_w - S_o$. These three phases form a composition and modeling the saturations as a typical continuous variable provides no control on their sum. Several techniques exist to deal with compositional data (Pawlowsky-Glahn and Olea, 2004; Pawlowsky-Glahn and Egozcue, 2006). The additive logratio (ALR) transformation is used prior to the normal score transform to remove the sum constraint. Equation 5.22 defines the forward ALR transform and Equation 5.23 the inverse ALR transformation.

$$R_x = ALR(S_x) = \log\left(\frac{S_x}{S_f}\right), x = oil, water \tag{5.22}$$

$$S_x = ALR^{-1}(R_x) = \frac{\exp\left(R_x\right)}{\exp(R_w) + \exp(R_o) + 1}, x = oil, water \tag{5.23}$$

Prior to the ALR transform, saturations are properly visualized in a simplex or ternary diagram with its vertices defining each component (S_w, S_o or S_f) at 100% saturation and the opposite sides defining 0% saturation. Post ALR transform, R_w and R_o, are plotted on Cartesian axes (Figure 5-26). Partitioning has no effect on the ALR transformation. For modeling variograms and computing correlations between variables, the normal score transformation is implemented on the partitioned data without using declustering weights.

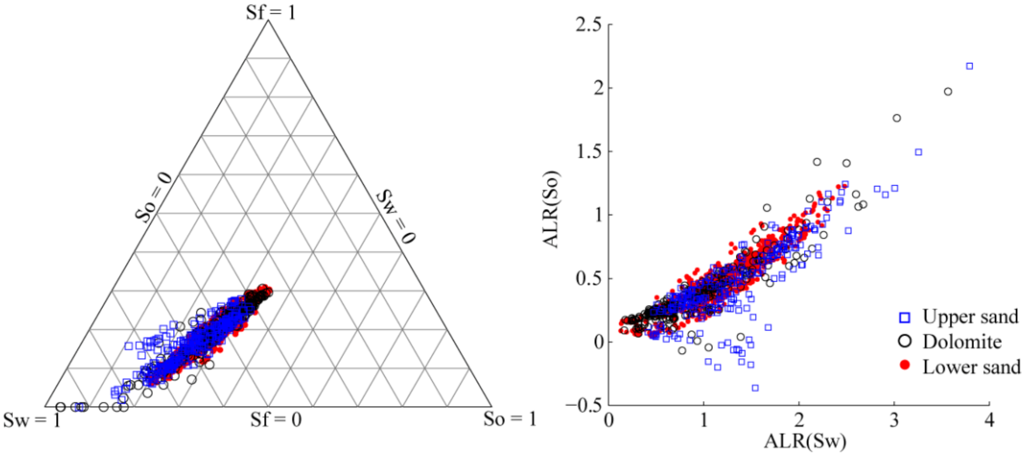


Figure 5-26: Ternary plot of saturations and ALR transformed cross plot.

5.3.4.3 Correlations

Correlation coefficients between ϕ, ALR(S_w), and ALR(S_o) are computed for each lithology using the normal scores (Figure 5-27). Permeability models will be constructed based on a regression equation between these properties rather than through Gaussian simulation because permeability at the log scale was not derived and permeability samples from core are sparse. Using linear regression, a quadratic relationship between ϕ and k_h is derived (Equation 5.24) and a linear relationship between k_h and k_v (Equation 5.25; Figure 5-28). Although this relationship was derived from core scale data, it is assumed to be valid for log scale data as well.

$$\log(k_h) = -164.07\phi^{3/2} + 135.79\phi - 7.23 \qquad 5.24$$

$$\log(k_v) = 0.79\log(k_h) - 1.06 \qquad 5.25$$

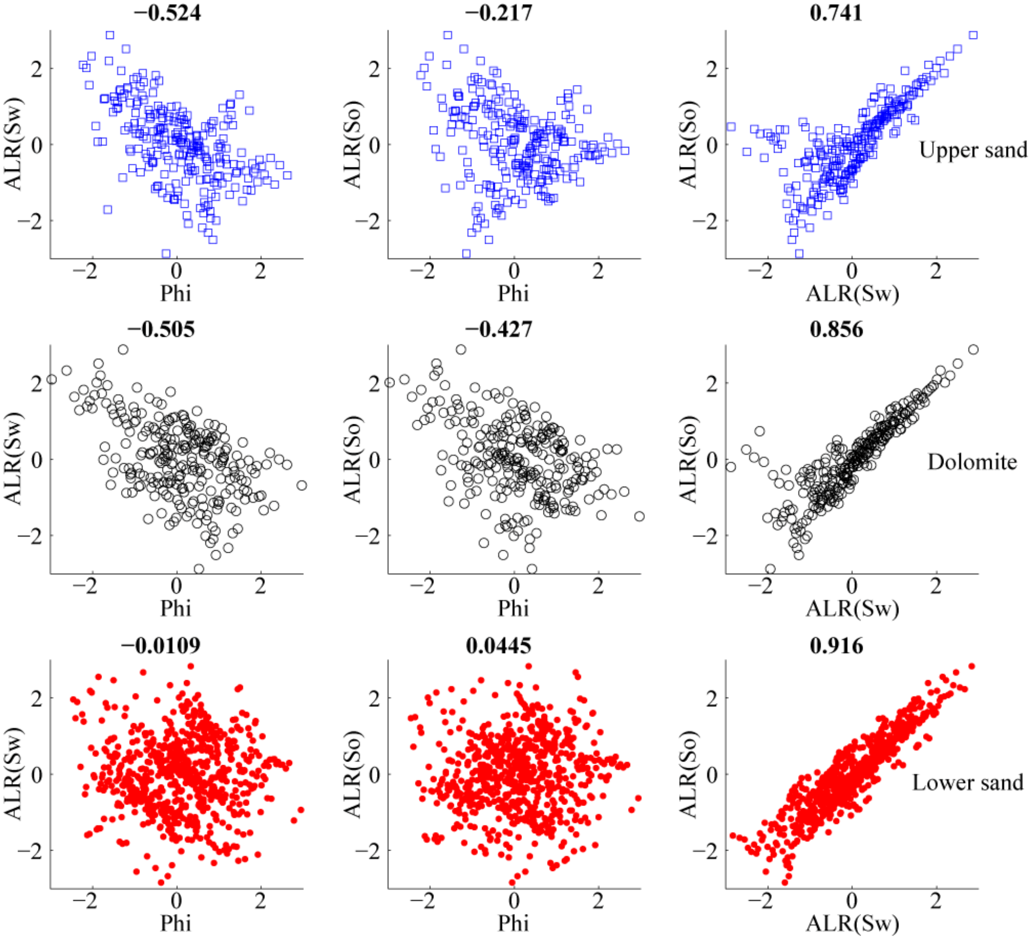


Figure 5-27: Cross plots of normal score properties. Correlation coefficients are indicated above each plot.

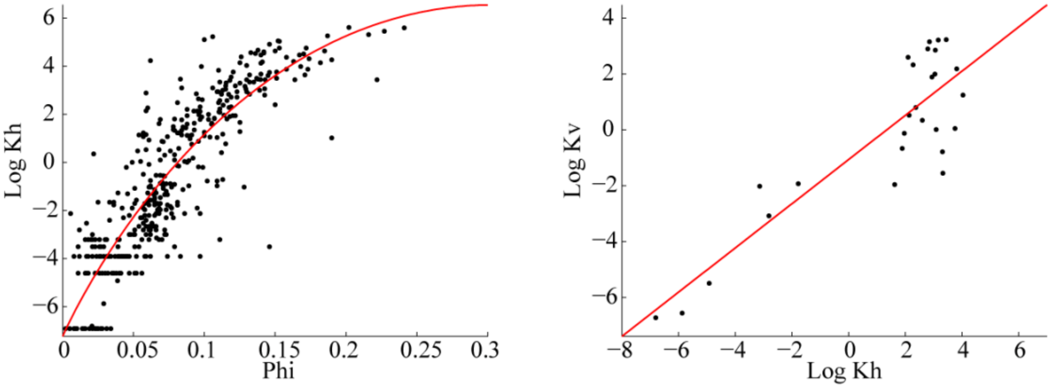


Figure 5-28: ϕ-k_h and k_h-k_v relationships.

5.3.4.4 Variograms

Variogram modeling of lithology and of the normal scored variables results in the models defined in Table 5-2 and Table 5-3 respectively. Indicator variograms of lithology depend on the proportion of each lithology, which are provided in Table 5-2. All lithology variogram models have a nugget effect of 0.002. For continuous variables, all models have a nugget effect of zero. All variograms are oriented with an azimuth of 160 degrees, corresponding to Range 1 in the table. Range 2 corresponds to an azimuth of 70 degrees and Range 3 is vertical.

Table 5-2: Indicator variograms for lithology.

Lithology Proportion, Sill	Structure	Variance	Range 1, ft	Range 2, ft	Range 3, ft
Upper Sand	Spherical	0.049	100,000	25,000	40
0.219, 0.171	Spherical	0.120	100,000	25,000	40
Dolomite	Spherical	0.050	1,500	4,000	25
0.204, 0.162	Spherical	0.110	100,000	8,000	25
Lower Sand	Spherical	0.154	100,000	25,000	45
0.577, 0.244	Spherical	0.090	100,000	25,000	100

Table 5-3: Continuous property variogram models.

Lithology	Property	Structure	Variance	Range 1, ft	Range 2, ft	Range 3, ft
Upper Sand	ϕ	Exponential	0.2	500	1,000	5
		Spherical	0.8	5,000	2,500	40
	ALR(S_w)	Exponential	0.5	2,000	3,000	50
		Spherical	0.5	8,000	3,000	50
	ALR(S_o)	Exponential	0.4	2,000	3,000	50
		Spherical	0.6	8,000	3,000	50
Dolomite	ϕ	Exponential	0.6	3,000	1,500	20
		Spherical	0.4	5,000	3,000	400
	ALR(S_w)	Exponential	0.5	2,000	3,000	18
		Spherical	0.5	8,000	3,000	20
	ALR(S_o)	Exponential	0.4	2,000	3,000	25
		Spherical	0.6	8,000	3,000	30
Lower Sand	ϕ	Exponential	0.2	500	1,000	15
		Spherical	0.8	5,000	2,500	80
	ALR(S_w)	Exponential	0.5	2,000	3,000	18
		Spherical	0.5	8,000	3,000	80
	ALR(S_o)	Exponential	0.4	2,000	3,000	20
		Spherical	0.6	8,000	3,000	100

5.3.4.5 Discretization Error Function

The discretization error function can be assessed at this point because it only depends on the variogram models and can be used to determine grid resolution for preliminary mapping and for typical geological modeling workflows involving structured grids. It is useful to assess the expected error when using one point to populate a grid cell as well as for choosing discretization criteria. In typical modeling workflows, the upscaled data is used to populate a structured grid with one point per element; therefore, the discretization error function can be used to check the error involved for the range of usual geological modeling grid element scales identified in Section 5.3.3.

To ensure the error specified by the function is achieved for all properties, a variogram that is above all other variogram models is defined. Variograms for lithology, ϕ, ALR(S_w), and ALR(S_o) are plotted on the same axes in Figure 5-29 along with the enclosing model that is defined in Table 5-4.

Table 5-4: Variogram model for discretization error.

Nugget	Structure	Variance	Range 1, ft	Range 2, ft	Range 3, ft
0.002	Exponential	0.1	400	1,700	18
	Exponential	0.5	3,300	1,700	18
	Spherical	0.4	5,000	2,500	20

Error curves that quantify the upscaling errors involved with existing modeling workflows when only one point is used are generated for hexahedral elements. Volumes are explored with a fixed vertical extent and varying horizontal extent. The smallest vertical extent is equal to the upscaling length of 1.64 ft and several other multiples of it are explored. Horizontally, the extent does not exceed 164 ft or 50 m. For this element size, the error is not expected to be large because the variogram range is relatively large. Error curves for $n = 1$ and vertical extents ranging from 1.64 ft to 6.56 ft (0.5 to 2 m) are shown in Figure 5-30.

Error curves are converted to an approximate number of hexahedral grid elements required to fill the study domain based on the domain extents (Figure 5-31). Extents are oriented at an angle of 15 degrees from the x axis and are equal to 9,060 ft in x', 11,023 ft in y', with x' and y' the rotated coordinates. For this case study, the maximum number of grid elements that is considered for preliminary mapping is 10^6. For example, if a vertical resolution of 3.28 ft is chosen, it is not possible to do better than an expected error of roughly 0.029, which corresponds to elements having a horizontal extent of 74 ft in x' and y'. These results are used in the next section on preliminary mapping. The error surface for a variety of n is revisited after an unstructured grid is designed.

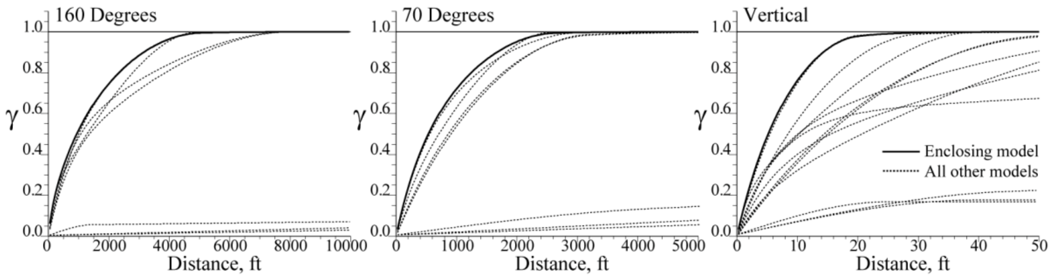


Figure 5-29: Enclosing variogram model for discretization error.

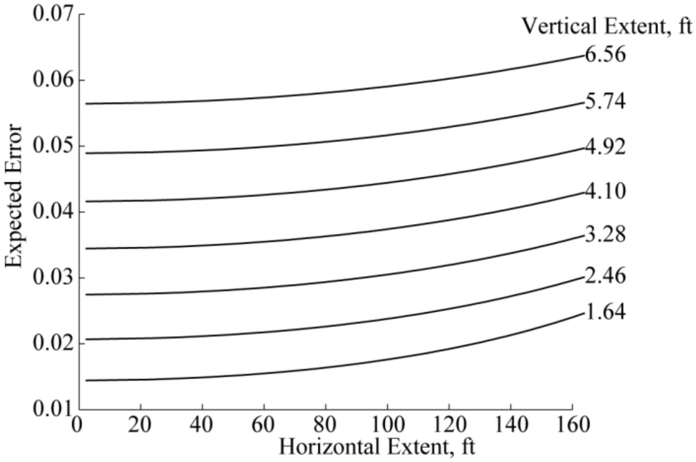


Figure 5-30: Error curves for hexahedral elements represented with one discretization point.

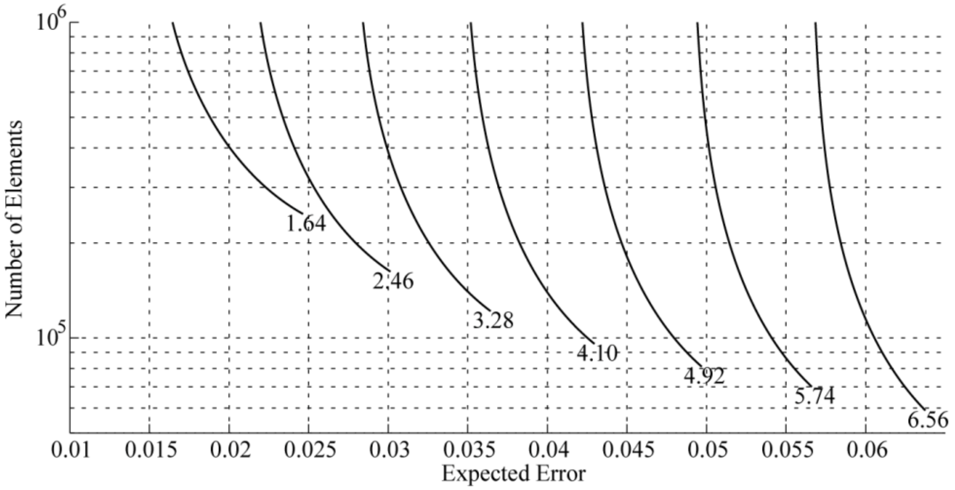


Figure 5-31: Number of elements as a function of expected error.

5.3.5 Preliminary Mapping

All of the necessary statistics are available to generate models of the reservoir properties; however, a grid is required to generate models. To provide insight into the spatial distribution of properties throughout the reservoir and how the reservoir might perform, preliminary mapping on a regular grid is used. Models from preliminary mapping are used for unstructured grid design. The parameters including the conditioning data, histograms, and variograms that are used to generate final geostatistical models are used at this stage for consistency. Using a regular grid in geological space for preliminary mapping has several advantages: very little effort is required to generate the grid; visualization is straightforward; and basic flow equations can be solved to give an idea of pressure and velocity distributions for unstructured grid design. The last point only applies to regular grids generated in geological space. A regular grid generated in depositional space, which is usually the case for existing reservoir modeling workflows, is not regular in geological space and can lead to grid orientation effects for flow simulation.

A regular grid for preliminary mapping is generated in this case study for solving pressures and flow velocities so that a flow based unstructured grid can be designed. Flow based grids are often designed using a single realization of reservoir properties; however, since only 1 grid design is considered here, kriging results are used. Kriging results should not be used to evaluate the flow response of a reservoir because the variance of the model is too low, that is, the heterogeneity of the reservoir properties is too smooth; however, kriging results can be used to generate a grid design that reflects how reservoir properties are spatially distributed in expected value. The grid is held static for all realizations.

In cases where only one pressure solve is done, a higher resolution grid can be used than when a flow simulation study requiring several solves over time is required. For this study, a grid with no more than 150,000 elements is deemed adequate for preliminary mapping. This keeps the pressure solve time and discretization error using 1 point per grid element low. Using a vertical element size of 3.28 ft, a regular grid in geological space with 140,490 elements is achieved by using horizontal extents of 151 ft in x' and y' (Figure 5-32). The grid of elements is 60 by 73 in x' and y' and nominally 32 in z. The discretization error for each element is roughly 0.035 for this grid design.

To obtain a general idea of the spatial distribution of permeability throughout the reservoir for flow based grid design, the following workflow is used:

1. Apply indicator kriging to generate a probability model of lithology.
2. Convert the probabilities to lithology by selecting the maximum. This can cause a significant bias for lithologies with a small proportion, but for this problem the lowest proportion is 20%, so this is not a concern.
3. Apply ordinary kriging to generate estimates of ϕ in each lithology.
4. Use the regression equations to compute k_h and k_v from ϕ.

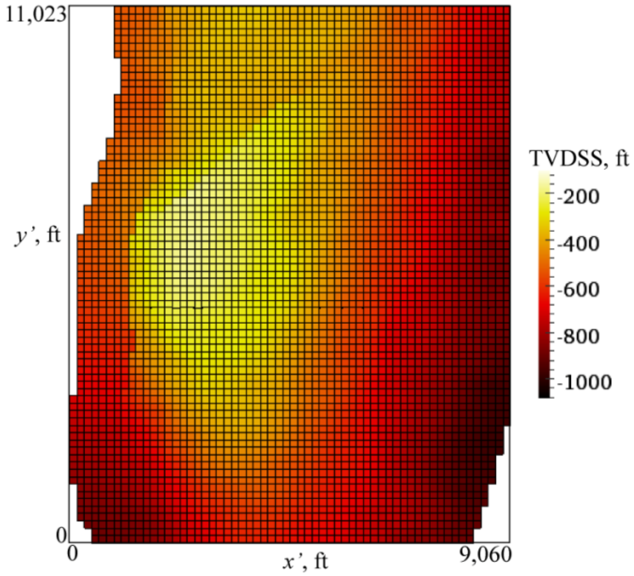


Figure 5-32: Plan view of grid for preliminary mapping in rotated coordinates.

A plan view of the resulting k_h through a depositional coordinate of $t = -65$ ft is shown in Figure 5-33. This is primarily in the lower sand, which involves the highest porosity and permeability. Computing pressures and velocities based on this permeability model also requires boundary conditions; however, choosing them depends on the context of the problem. For example, if one assumes the final grid design will get used for a variety of flow simulation studies, general boundary conditions are appropriate; in other cases where a specific well configuration and injection process are to be evaluated, boundary conditions for the specific process are appropriate.

Existing studies involving the Tensleep formation such as Friedmann and Stamp (2006) have evaluated it for enhanced recovery using CO_2 injection and CO_2 storage. The targeted area is in the shallowest portion of the anticline structure, which, along with the faults, acts as a trap for injected CO_2. Assuming this is the context for grid design, boundary conditions that highlight the flow paths of fluid injected in the anticline are considered. All exposed faces except those along the model extremities in y' are assumed sealed and are assigned no flow boundary conditions (Figure 5-33). Resulting pressure and velocity magnitude fields along the plane with deposition coordinate of $t = -65$ ft are shown in Figure 5-34. These were determined using the TPFA method. Since areas and lengths are in ft, pressures were assigned as pounds per square foot (psf) and millidarcies were converted to ft^2.

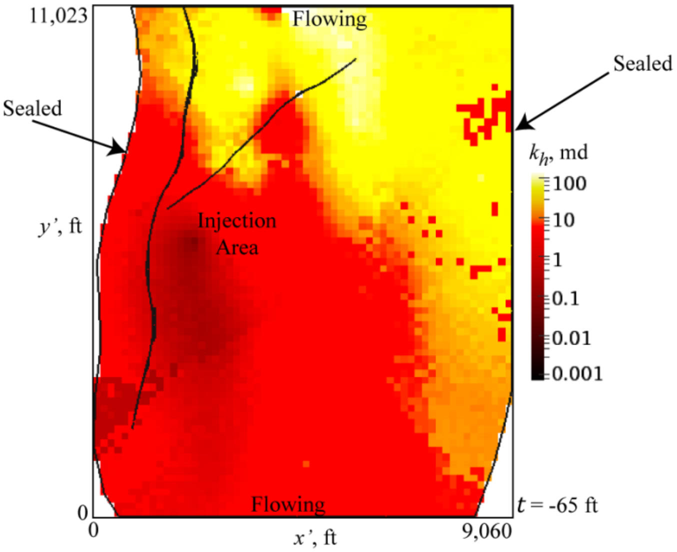


Figure 5-33: Permeability model created during preliminary mapping and boundary conditions used to solve for the pressure.

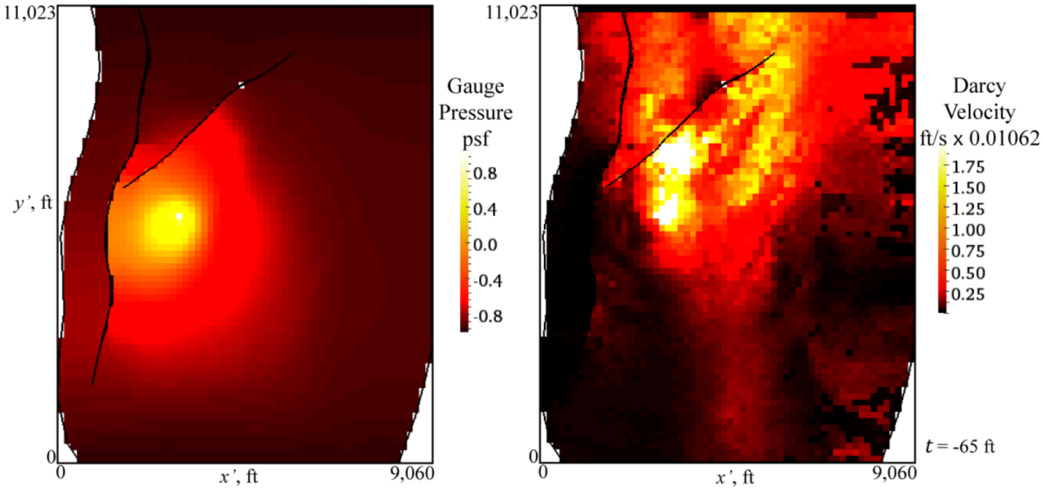


Figure 5-34: Pressure and velocity fields from preliminary mapping. Velocities shown as magnitudes without direction.

5.3.6 Unstructured Grid Design

A basic procedure for unstructured grid design is implemented for demonstration purposes. A tetrahedral grid is built using a point set generated with a density that depends on the magnitude of velocity determined from preliminary mapping. Points are inserted into the model based on the curve in Figure 5-35 that is a geometric function with a factor of 0.998: $q(i) = q(i-1) \cdot 0.998$, where q is the Darcy velocity. This was chosen so that the final grid design results in a manageable number of grid elements. 3,109 points were added based on the curve. Applying the program TetGen (Si, 2006) using the bounding surfaces, additional points, and a quality constraint of 1.5 resulted in a tetrahedral grid with 24,247 points and 121,994 elements (Figure 5-36). The dual polygonal grid is shown in Figure 5-37. Elements with smaller volume are concentrated in areas of expected high velocity and around the faults (Figure 5-38). It is important to consider areas like these for flow simulation.

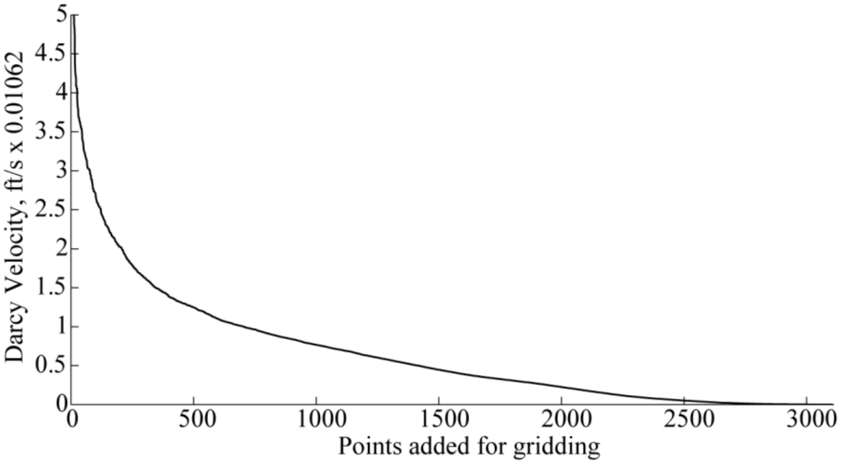


Figure 5-35: Curve relating velocity to number of points added for tetrahedral grid generation.

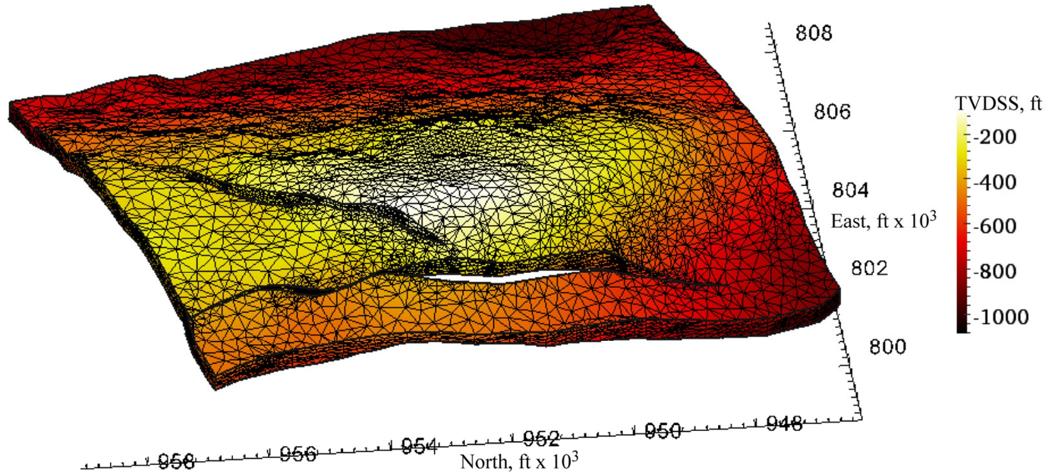


Figure 5-36: Isometric view of tetrahedral grid design.

If flow simulation were to be applied, either the tetrahedral grid or dual polygonal grid could be used. Moreover, discretization leads to the same high resolution grid for both coarse scale grids; therefore the geostatistical models will also be identical. The advantage of the tetrahedral grid is that the faults are honoured by the element interfaces. Polygonal elements are bisected by the faults and incorporating information such as the flow character of the faults is difficult. On the other hand, the advantage of the polygonal grid is substantially fewer control volumes. Flow simulation would be done using 24,247 elements as opposed to 121,994.

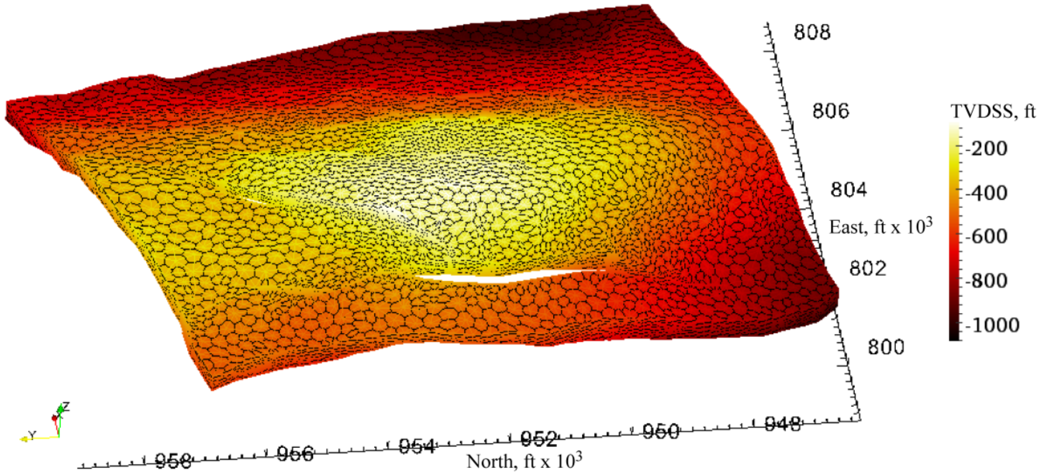


Figure 5-37: Isometric view of the polygonal grid.

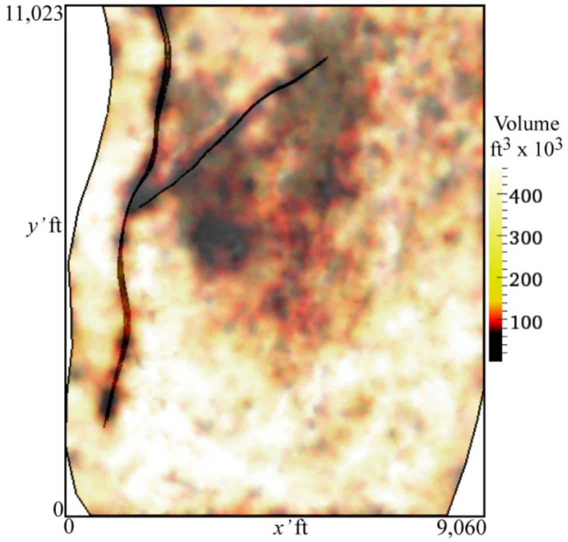


Figure 5-38: Rendering of the tetrahedron element volumes. Volumes larger than 200,000 ft^3 are transparent.

5.3.7 Discretization

Up to this point, the significant differences between existing workflows and the proposed one is that the unstructured grid was designed based on a preliminary mapping stage using a kriged model and there was no structured grid design for geostatistical modeling. Discretization is another significant difference between the workflows. Existing techniques generate reservoir models on a structured grid design, and then upscale; whereas this workflow generates reservoir models on a discretized unstructured grid prior to upscaling.

The discretization error is evaluated using the same variogram defined in Section 5.3.4.5 of this chapter. The tetrahedral grid was generated with the *z* axis scaled by 10, so most tetrahedral elements have an aspect ratio of 10 to 1. Element volumes range from less than 10 ft^3 to 900,000 ft^3. However, the distribution of element volumes is highly skewed and 80% of the volumes are less than 150,000 ft^3 (Figure 5-39). Discretization error is evaluated up to this volume. Elements with volumes above this exist in areas of the reservoir that involve low flow velocities as determined by preliminary mapping and are therefore deemed unimportant for the case study. The number of discretization elements required for 150,000 ft^3 elements is applied to all larger volumes. A close analysis of the error function shows this is necessary because there is a vertical asymptote for the volume defined by Equation 5.26.

Regression resulted in Equation 5.27 with a mean squared fit error of 6.08×10^{-5}. An error of 0.005 is chosen to control the discretization process. This choice was made to keep the number of discretization elements close to 1 million. The associated contour is show in Figure 5-40. Volume constraints are setup for the program TetGen by dividing the element volume by the number of discretization points computed from the error contour.

$$V = \left(\frac{\log \varepsilon_Z^2 - A_0}{A^1} \right)^{\frac{1}{\beta}} = 446,260\,\text{ft}^3 \tag{5.26}$$

$$\begin{aligned} \log \varepsilon_Z^2(n,V) &= A_0 + A_1 V^{\beta} + A_2 \left(n+\tau\right)^{-1} \\ &= -8.64 + 0.044 V^{0.333} + 29.34\left(n+5.667\right)^{-1} \end{aligned} \tag{5.27}$$

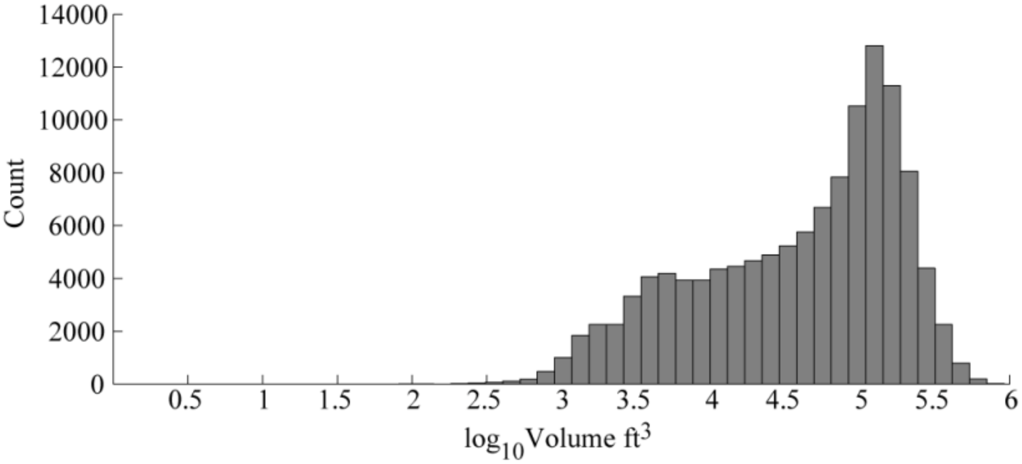


Figure 5-39: Histogram of tetrahedral grid volumes.

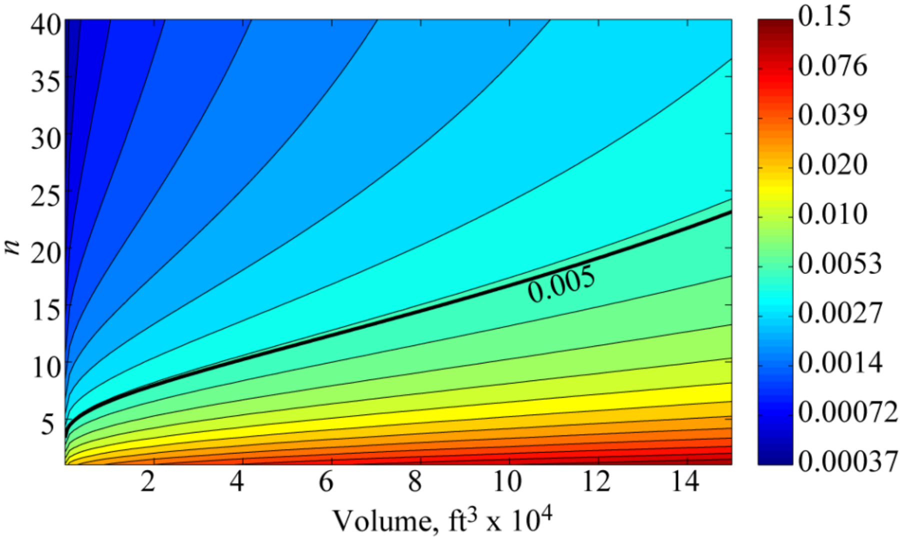


Figure 5-40: Tetrahedral grid discretization error surface.

Tetrahedron elements are discretized by the subdivision process shown for a triangle in Figure 5-41. The program TetGen did not run to completion for this problem and requires further study. Information that is not retained by the subdivision approach used is the tetrahedral topology of the discretized grid. Flow based upscaling using MPFA or MFEM requires this information; however, flow based upscaling is not done in this case study. Flow based upscaling was developed and demonstrated in two dimensions for triangular grids. Computing the arithmetic average does not require the topology since discretization points in each tetrahedron represent the same volume and all weights are equal using the subdivision approach.

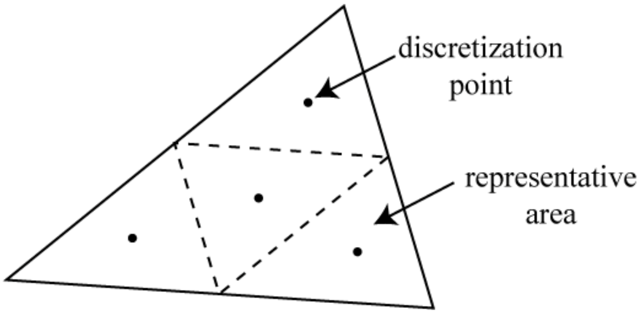


Figure 5-41: Analytic discretization of a simplex.

A total of 1,241,152 points resulted from the subdivision discretization process with an average of 10 points per volume. No volumes received more than 20 points to keep the total number reasonable. The volume represented by a point is equal to the coarse scale tetrahedral element volume divided by the number of points used to discretize it. In areas where smaller elements are located (Figure 5-38), the average tetrahedron volume is 4,389 ft^3 and they are discretized by 4 points on average. Each point represents a volume of 1,097 ft^3. To obtain such a resolution with a regular or structured grid would require a high resolution everywhere in the modeling domain. The whole reservoir has a volume of 10,495×10^6 ft^3, which would require a structured grid having upwards of 9.5 million elements. Based on the smallest tetrahedron volume that is less than 10 ft^3, a regular grid achieving this resolution would require more than one billion elements.

To compare existing workflows that would discretize the reservoir using a structured grid, it is fair to consider a grid with 1.2 million elements. Based on the extents of the modeling domain and using a vertical element size that is equal to the upscaled log data of 1.64 ft and horizontal extents of roughly 75 ft, a regular grid with 1,217,744 elements is achieved in depositional space. Utilizing this grid for upscaling to the tetrahedral grid offers little control in terms of discretization error. At one end of the spectrum, tetrahedral elements are discretized by many regular grid elements, while at the other they are discretized by less than 1. The largest element is 920,433 ft^3 and 103 elements of the regular grid intersect it based on center points only (Figure 5-42). In the middle range with elements of 150,000 ft^3, roughly 20 regular grid elements intersect the tetrahedral elements (Figure 5-43). This is similar to the proposed approach in terms of the number of discretization points. For volumes of 1,000 ft^3, less than 1 regular grid element intersects the tetrahedral element (Figure 5-44). In this case, the discretization error is higher than achieved with the proposed approach. A higher discretization error implies two results: the error in the average value of the tetrahedral element is higher; the error in the approximation of the average covariance of the element and between the element and its neighbours is higher. The later further implies the heterogeneity at the scale of the small tetrahedral elements is inaccurate.

Another issue that is indicated by these examples of discretization with a regular grid is the elements do not align with the unstructured grid element interfaces. Computing the arithmetic average or a property involves computing all intersections of regular grid elements with unstructured grid elements. This is more complex than computing the weighted average using the proposed discretization approach. Moreover, when flow-based upscaling is used, determining the velocity across the tetrahedral interfaces is not exact, especially when the tetrahedra are smaller than the regular grid elements. Velocities obtained are known across the regular grid interfaces and must be resolved to the interfaces of the tetrahedra. Unstructured grids often have smaller elements in areas of significant flow so the effort should be made to accurately determine the velocities through these elements.

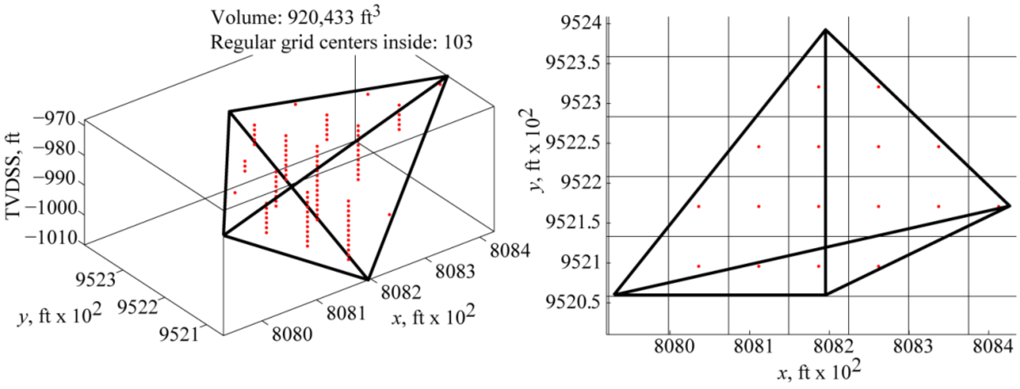


Figure 5-42: Largest tetrahedron discretized by a regular grid.

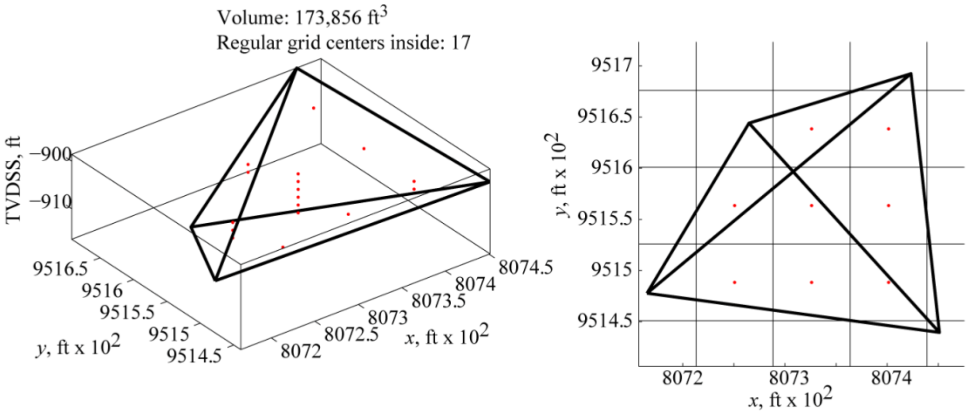


Figure 5-43: Intermediate tetrahedron discretized by a regular grid.

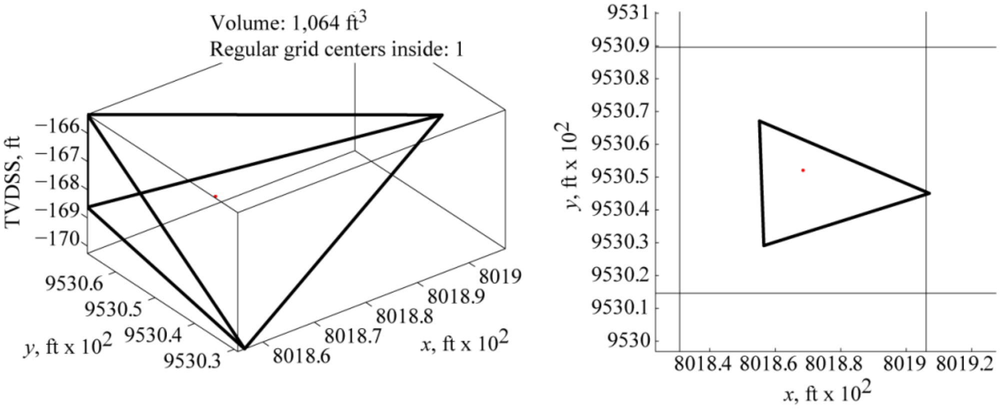


Figure 5-44: Small tetrahedron discretized by a regular grid.

5.3.8 Property Modeling

For geostatistical modeling, discretization points are transformed to depositional space where variogram modeling is done. Lithology is modeled first, then ϕ, ALR(S_w), and ALR(S_o) within each lithology. The program psgsim (see Chapter 3 and the Appendix) does this using indicator kriging for lithology modeling and collocated cokriging for multiple continuous variables. Since indicator simulation provides the lithology model, an additional feature was added to psgsim so that a lithology realization is cleaned prior to continuous variable simulation. For example, no upper sand should be simulated in the lower sand interval and cleaning is used to remove such occurrences. Lithology cleaning was done with a moving average filter. Within each lithology, ϕ is modeled using simple kriging, followed by ALR(S_w) using collocated cokriging conditional to the simulated ϕ value, and finally ALR(S_o) conditional to ALR(S_w). Correlation coefficients shown in Figure 5-27 were used.

Twenty realizations were generated to show reproduction of distributions, bivariate relationships, and variograms. 18 data were used to construct systems of equations for kriging. Each realization took approximately 7.15 minutes on an Intel 2.13 GHz processor, which is acceptable given that for each of the 1.2 million points, 6 linear systems of equations are constructed and solved: 3 for lithology proportions, and 1 for each continuous property. Each set of 3 systems requires a search for the nearest neighbours as well.

Distributions of each property are checked for shape and summary statistics including the mean and standard deviation. For lithology, the proportions are checked (Table 5-5). Proportions are reproduced well, with larger errors occurring for dolomite and upper sand, which is expected since they have smaller proportions than lower sand. Reproduction of the mean and standard deviation of continuous properties within each lithology shows low factional difference for the mean, but S_w and S_o show large differences in standard deviation (

Table 5-6). The fraction difference between true and simulated fluid saturations is high for Dolomite and upper sand, but this is due to the ALR transformation process and is caused by the occurrence of zeros in the data. When S_w and S_o sum to 1, which occurred primarily when one of S_w and S_o was equal to 1, S_f is zero causing issues with the ALR transform. Resulting distribution tails are observed in the distribution reproduction plots (Figure 5-45). Long tails on the saturation distributions account for a very small proportion of the realizations and only cause inflation in the variance; reproduction of the mean remains accurate. Otherwise, all realization distributions reproduce the input distributions well.

Variogram reproduction is not perfect in this case study: some properties and directions match very well, while others have a mismatch (Figure 5-46 and Figure 5-47). The types of mismatch observed include: variograms having the correct shape, but not reaching the sill; variograms matching short range structure well, but departing from the input model with increasing range; trouble with reproducing anisotropy; and variograms going over the sill, indicating variance inflation, which is a known issue with collocated cokriging. Causes of these are using a limited number of data for kriging, 18 in this case, possibly random path related, and attempting to enforce variogram reproduction for a multivariate study without using direct variograms from a licit linear model of coregionalization. This is an area requiring further research. Another contributor of mismatch for lithology variograms is the cleaning step discussed previously.

A final check is to ensure that the relationships between variables were preserved, which includes correlation coefficients and the shape of scatter plots. Unlike variograms and histograms, it is not feasible to create a scatterplot for all realizations so the tenth realization is shown. It is also difficult to see relationships when over 1.2 million points are plotted on a scatterplot. Instead, the realization results are converted to a density and plotted as contours. True data are plotted as points over the contours for a visual check (Figure 5-48). Contours appear to represent the density distribution from which the samples originated in all cases. Correlation coefficients are checked for all realizations (

Table 5-7). Most coefficients are reproduced well, although the coefficient for ϕ-S_o in lower sand is somewhat low.

Table 5-5: Reproduction of lithology proportions.

Lithology	True Proportion	Mean Simulated Proportion	Difference
Upper sand	0.577	0.584	0.007
Dolomite	0.204	0.219	0.015
Lower Sand	0.219	0.197	-0.021

Table 5-6: Reproduction of mean and standard deviation.

Property and Lithology		True Mean	Mean of Models	Fraction Diff	True σ	σ of Models	Fraction Diff
ϕ	Upper S.	0.061	0.061	0.000	0.025	0.025	0.000
	Dolo	0.072	0.064	-0.111	0.040	0.035	-0.125
	Lower S.	0.106	0.105	-0.009	0.037	0.037	0.000
S_w	Upper S.	0.581	0.575	-0.010	0.088	0.106	0.205
	Dolo	0.505	0.515	0.020	0.106	0.144	0.358
	Lower S.	0.574	0.570	-0.007	0.078	0.072	-0.077
S_o	Upper S.	0.249	0.257	0.032	0.045	0.081	0.800
	Dolo	0.286	0.294	0.028	0.049	0.122	1.490
	Lower S.	0.262	0.264	0.008	0.033	0.029	-0.121

Table 5-7: Reproduction of correlation coefficients.

Pair	Lithology	True Correlation	Mean Simulated Correlation	Difference
$\phi - S_w$	Upper sand	-0.037	0.005	0.042
	Dolomite	-0.442	-0.537	-0.095
	Lower Sand	-0.591	-0.643	-0.052
$\phi - S_o$	Upper sand	0.089	-0.001	-0.090
	Dolomite	0.432	0.432	0.000
	Lower Sand	0.627	0.444	-0.183
$S_w - S_o$	Upper sand	-0.914	-0.946	-0.032
	Dolomite	-0.959	-0.928	0.031
	Lower Sand	-0.912	-0.857	0.055

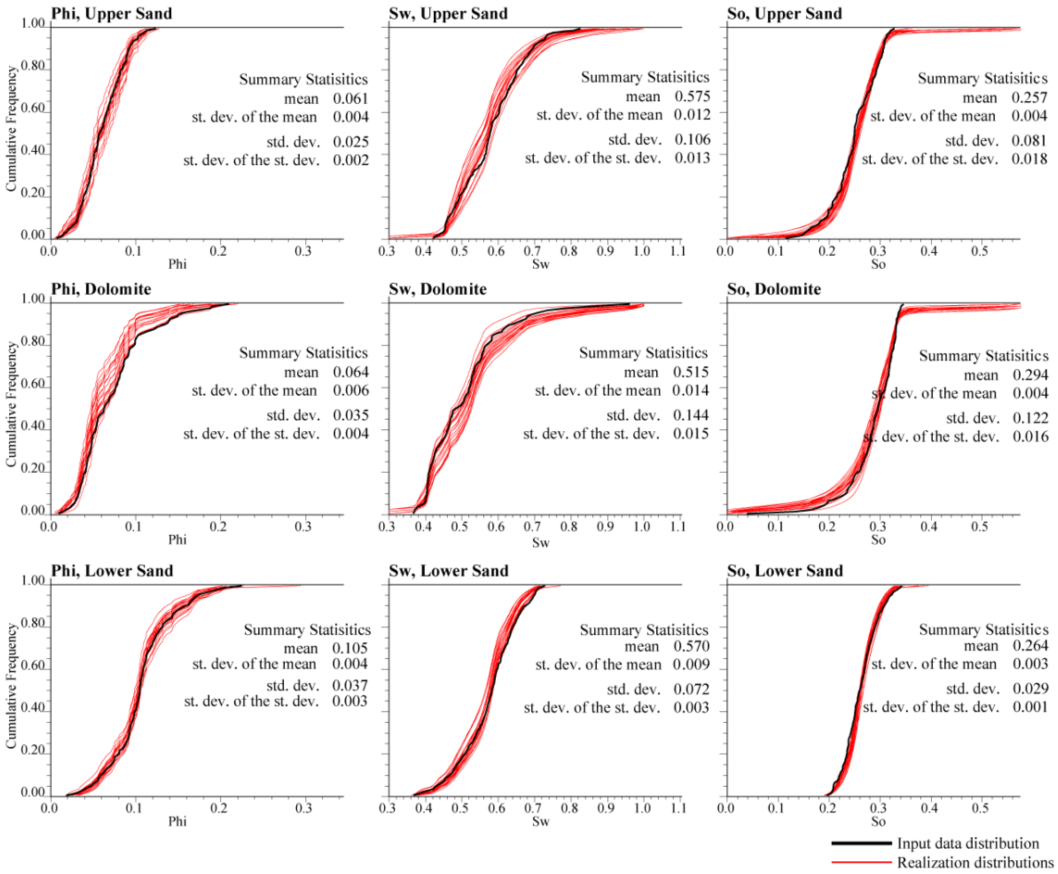


Figure 5-45: Reproduction of distributions from psgsim realizations.

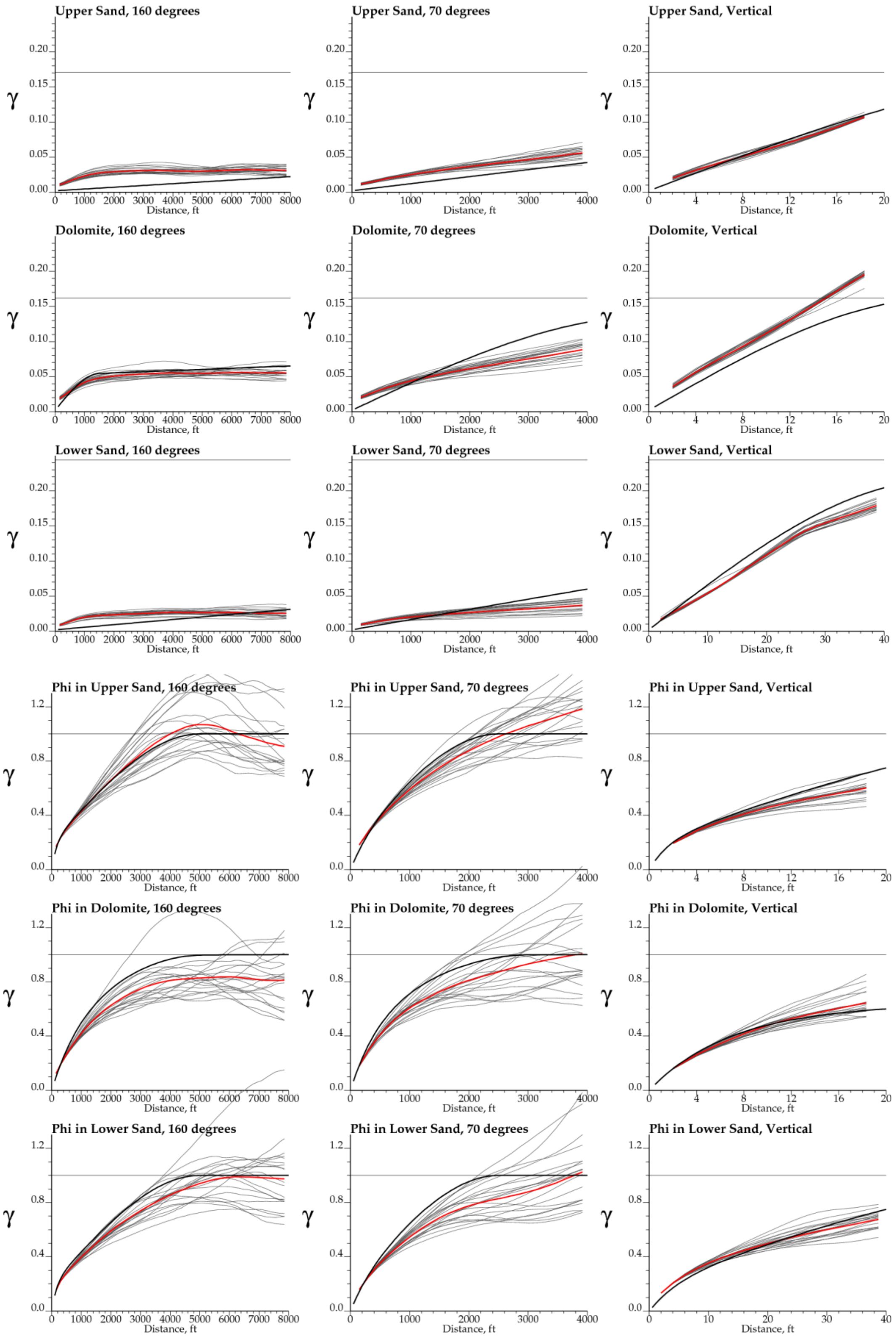


Figure 5-46: Lithology and Phi variogram reproduction.

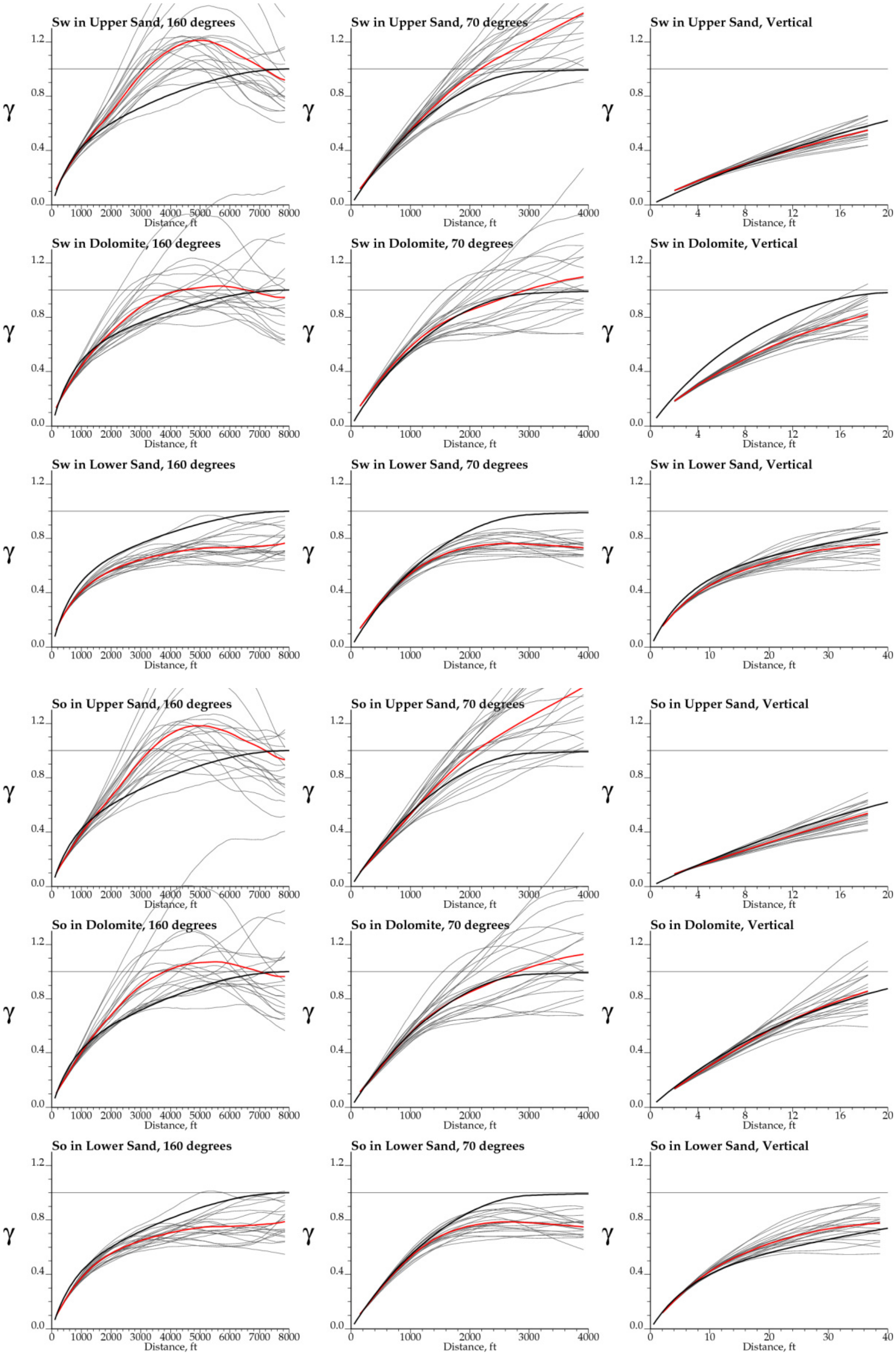


Figure 5-47: S_w and S_o variogram reproduction.

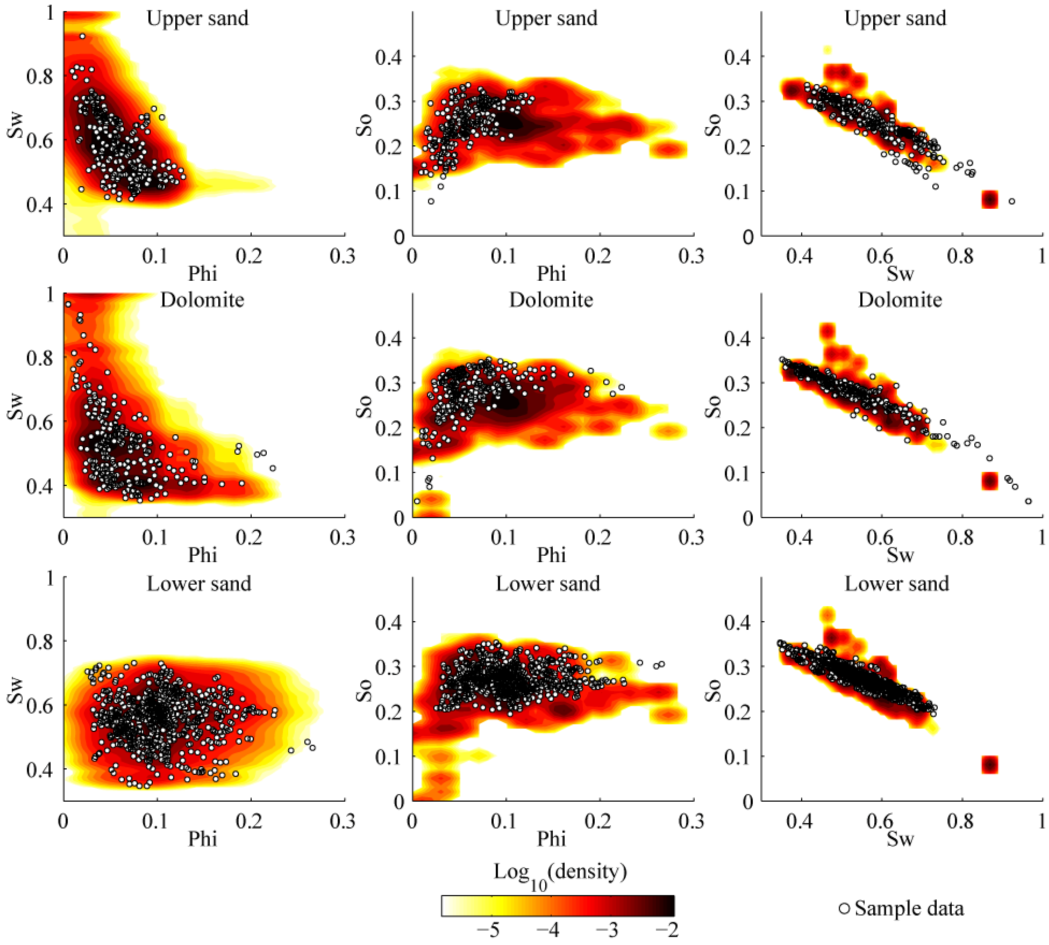


Figure 5-48: Reproduction of bivariate relationships.

5.3.9 Upscaling

An advantage of the proposed workflow is simplicity of upscaling. There are no complex element intersections to be concerned with and covariance across all scales is preserved. Upscaling a structured grid to an unstructured grid requires many volume-volume intersection operations that are time consuming. In the case of flow based upscaling, velocities do not have to be interpolated to unstructured grid element interfaces because the fine scale velocities align exactly with the coarse element interfaces.

Realizations can be upscaled to either the tetrahedral grid or the dual polygonal grid. An advantage of upscaling to the polygonal grid is ease of visualization. Upscaled properties can be represented by the polygonal element centroids that coincide with the tetrahedral grid vertices. Rendering models, generating cross sections, and performing other visualization operations are straightforward and computationally efficient in this case.

Lithology is upscaled to proportions, and ϕ, S_w and S_o are upscaled arithmetically. A realization of porosity is shown before and after upscaling (Figure 5-49). To move on with flow simulation, permeability would be computed from porosity and upscaled. After upscaling, realizations would be used in flow simulation studies to assess the uncertainty in a CO_2 injection process. This was not done since flow-based upscaling was only developed in two dimensions.

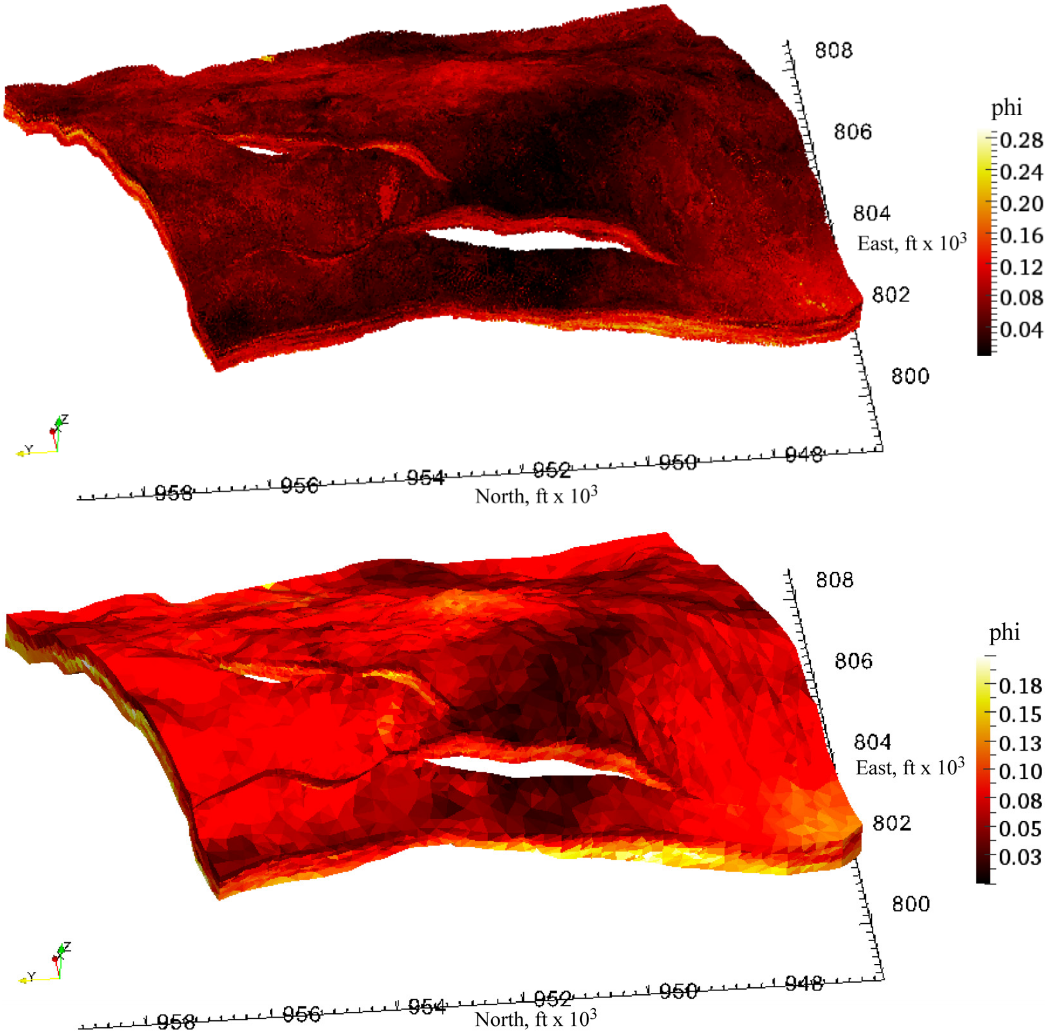


Figure 5-49: Point scale realization of porosity (top) and upscaled equivalent. A vertical exaggeration of five is used.

It is possible to use geostatistical models constructed on unstructured grids in many of the same ways as on regular grids, for example, assessing expected values and global uncertainty. Results of this case study are used to compute the expected value of the hydrocarbon volume ($HCV = V \cdot \phi \cdot S_o$) in the highest part of the anticline. It is arbitrarily assumed that injection cannot recover reservoir fluids below a depth of -300 ft subsea. The volume of reservoir above this depth is 1.3 billion ft^3. The mean porosity is 6.3 percent and mean pore volume is 82.9 million ft^3. A histogram of HCV is shown in Figure 5-50. The mean and standard deviation are 4.017 million barrels and 83,477 barrels respectively. Data density is high in this area (Figure 5-18) relative to other parts of the reservoir and parameter uncertainty was not considered, leading to the low variance in HCV estimates. Using the expected HCV model, computed from expected ϕ and S_o models, pockets where HCV is expected to be highest are identified (Figure 5-51). A $\phi \cdot S_o$ cutoff was chosen arbitrarily.

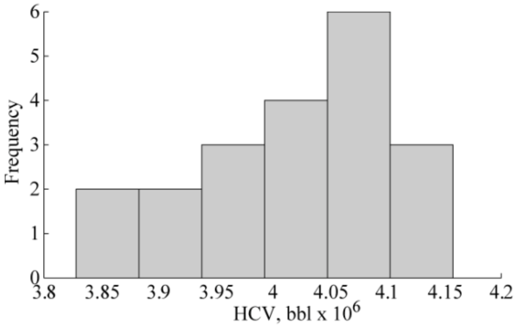


Figure 5-50: Histogram of HCV.

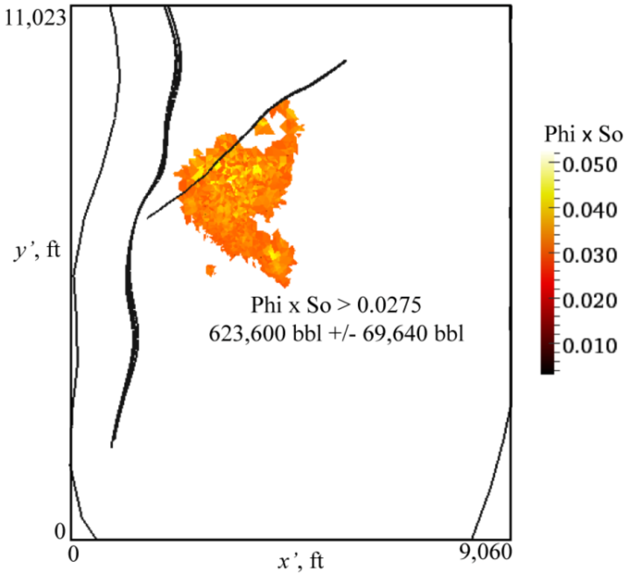


Figure 5-51: Tetrahedral elements with high expected $\phi \cdot S_o$.

5.4 Summary

This chapter showed that the discretization approach to geostatistical modeling of unstructured grids reproduces the estimate, estimation variance, and covariance between coarse grid elements and conditioning data and between multiple coarse scale elements. Results apply to Gaussian random variables; however, the advantage of the discretization approach is it applies even when non-linear transformations such as the normal score transform are implemented on non-Gaussian reservoir properties. Average covariance approaches that apply to Gaussian variables do not apply through such transformations. It was also demonstrated that the discretization approach reproduces the variogram at the fine scale and coarse scale in both Gaussian and original space. In the example, the variable in original space was lognormal; therefore it also exhibits a proportional effect.

The method reproduces this feature. Further, the method was shown to be unbiased when a low enough error is chosen for controlling the discretization process.

The second section of this chapter validated the flow based upscaling approach developed in Chapter 4. Three types of media were used: homogeneous, layered, and general heterogeneous. Good agreement between PVI-F_w curves computed using a high resolution regular grid and coarse unstructured grids was found in all cases. These were done in two dimensions. A case study demonstrating the geostatistical components of the proposed workflow was given last. The scenario involved three lithologies, porosity, water saturation, and oil saturation variables. Two advantages over methods using structured grids were identified: fewer elements are involved for geological modeling and coarse unstructured elements are discretized more uniformly. A weakness with variogram reproduction and complications in discretization were identified also.

Chapter 6 Concluding Remarks

Unstructured grids are utilized in reservoir analysis for flow simulation. The goal is to obtain accurate flow response predictions with reasonable computational effort. Unstructured grids are typically designed based on geostatistical property models that are available on structured grids. This dissertation proposed that structured grids are not the optimal spatial decomposition structure to use when properties are to be upscaled to an unstructured grid for flow simulation. Problems arise when the structured grid is coarser than the unstructured grid, which is possible in the important areas adjacent to wells and along other significant geological features and boundaries. Moreover, upscaling is complicated because the two grids do not align.

This dissertation proposed a solution to the problem of modeling reservoir properties on unstructured grids using geostatistical techniques. One concern is the range of scales introduced by the grid elements. Different element scales were accounted for by discretizing the coarse unstructured grid with a higher resolution unstructured grid. Triangular and tetrahedral grids were used. A workflow was proposed that uses geostatistics to populate the high resolution unstructured grid and upscale to the coarse grid.

6.1 Summary of Contributions

There are three principle contributions of this thesis: an approach to discretize an unstructured grid into a higher resolution triangular or tetrahedral grid; an algorithm for geostatistical simulation of categorical and continuous properties on high resolution unstructured grids; and a flow based upscaling technique based on the multipoint flux approximation.

6.1.1 Unstructured Grid Discretization

Discretizing a coarse unstructured grid for flow simulation using a higher resolution unstructured grid was chosen because of the following advantages: the grids are nested and aligned, agreeing with dispersion variance concepts of geostatistics; static upscaling including arithmetic, harmonic, geometric, and power averaging and dynamic or flow-based upscaling approaches are straightforward as compared to grids that do not align; the covariance between all coarse scale grid elements is correct up to the accuracy obtainable by discretization; and the accuracy of properties upscaled to the coarse grid can be controlled by the resolution of the fine grid. Discretization was done using triangular and tetrahedral grids that can be constrained to the coarse grid element geometry. The discretization can also be constrained to other features in the reservoir that are important to flow simulation and have not been incorporated into the coarse grid design, such as fractures or cross-bedding. These features are accounted for in the upscaling process.

The scale of triangular and tetrahedral grids used to discretize coarse unstructured grids was determined to ensure upscaled values retain a user-specified level of accuracy. It was shown that the accuracy of arithmetic averaging properties is dependent on the variogram. Expressed as the mean squared error between the true value for a coarse unstructured grid element and the value obtained using the discretization, the error expression given in Chapter 4 is equivalent to the estimation variance. Evaluating the error is computationally expensive because it involves average covariance calculations so an approximating function was developed. A program was written to generate data so that the best fit approximating function can be recovered using linear regression. The error approximating function was found to be accurate for three popular variogram structures: the spherical, exponential, and Gaussian models. Geostatistical modeling is often a multivariate problem involving facies and several continuous properties and by using one variogram model that is greater than all other variograms involved, the accuracy can be guaranteed for all properties simultaneously with one discretization. The user-specified accuracy is obtained for all coarse grid elements regardless of their volume. Situations where the fine grid is coarser than the coarse grid, such as near wells, do not occur.

For Gaussian random variables, it was shown that the conditional distributions of uncertainty from kriging are the same when average covariance is used and when point scale simulations are averaged. Both approaches account for the coarse grid element geometry and scale. Using the average covariance approach, it is possible to draw a Gaussian value from the conditional distribution; however, the value cannot be back-transformed to the space of a reservoir property that is non-Gaussian or that does not average arithmetically. The explanation is the normal score transform is a non-linear transformation when the reservoir property is non-Gaussian. Using the discretization approach, Gaussian values are drawn for the set of fine scale grid elements within the coarse scale element. These values can be back-transformed, since there has been no change in scale, and then upscaled.

6.1.2 Geostatistical Simulation

The second contribution is adaptations of geostatistical simulation algorithms for unstructured grids. Sequential indicator and sequential Gaussian simulation were adapted to work with irregular sets of points that are the barycenters of triangular or tetrahedral grids. Kriging is the underlying generator of conditional probabilities for both sequential algorithms and it is applicable to any configuration of points. Of particular concern for sequential simulation is random path generation, the search for data to use in kriging, and covariance computation.

Random path selection is typically done on regular grids using the multigrid approach to improve variogram reproduction. On an irregular set of points, a similar type of random path is generated using a quadtree in two dimensions or an octree in three dimensions. This approach gave good results for several test cases; however, variogram reproduction in the case study was not as good as anticipated. Two search structures were tested: the superblock search and the *k*d-tree. The *k*d-tree was found to be faster and it requires fewer parameters. Covariance computation was the principal loss of efficiency. Unlike a regular grid where a covariance lookup table can be constructed and referenced during simulation, a lookup table cannot be constructed for irregular point sets. The covariance between points is always calculated from the specified equation. The final modeling algorithm was found to run in a reasonable time despite the losses in efficiency compared to regular grid algorithms. A realization of 1.2 million points with 3 lithologies and 6 variables from the case study took just over 7 minutes.

Geostatistical models must be checked for consistency with the input data and parameters. Two common checks are histogram and variogram reproduction. The simulation algorithm was shown to produce random fields with a Gaussian distribution for unconditional simulation. This is important since we assume the data are multivariate Gaussian after a normal score transformation for use in sequential Gaussian simulation and guarantees the field will have the same distribution as the input data after a back-transformation to the original variable. Computing the histogram of a random field on an unstructured grid is accomplished by weighting the property values by the grid element volumes. This approach is similar to polygonal declustering. An existing program for computing the experimental variogram (Deutsch and Journel, 1998) was adapted to work on very large point sets so that the variogram can be checked. The new algorithm was found to be effective for checking the variogram on sets of more than 1 million irregularly spaced points.

6.1.3 MPFA Upscaling

A local flow-based upscaling approach based on the MPFA method was the third contribution. The MPFA method is applicable to triangular and tetrahedral grids. The upscaling approach involved solving the flow equations on the discretization for a series of boundary conditions, then using the results to compute coarse scale MPFA transmissibility matrices for flow simulation on the coarse unstructured grid. Developments were made in two dimensions. The extension to three dimensions is possible, but there are challenges with grid design and quality to ensure the flow equations are well conditioned.

Using the upscaling approach developed requires that the unstructured grid is discretized to honour the control volumes and interaction regions. Substantially more edges and interfaces are involved as constraints for the discretization. In two dimensions, this is not a concern; however, in three dimensions it is anticipated that the number of elements in the discretization could grow large to meet constraints and quality criteria.

The convergence of upscaled transmissibilities with different permeability scenarios was explored. This included homogeneous permeability, permeability that was different in each coarse scale control volume, and general heterogeneous permeability with different variograms. All cases indicated convergence with relatively few fine scale elements. 12 to 18 elements seem appropriate in two dimensions. It was shown that the upscaling approach was correct for homogeneous permeability. Four flow simulation examples were used to validate the proposed upscaling approach including: flow through homogenous media; flow through layered media parallel to the layers; through layered media perpendicular to the layers; and through general heterogeneous media with three facies. Resulting flow curves closely matched the reference curves that were generated using a high resolution regular grid and the TPFA approach.

6.2 Limitations

Several limitations exist for implementing the proposed workflow and algorithms. The workflow involves two modeling steps: a preliminary mapping step that is used for grid design and a final modeling step that populates the discretized grid. This is somewhat redundant, although in cases where the unstructured grid design will be influenced by reservoir properties, some degree of modeling is required to provide a property model. In existing approaches, unstructured grid design is often based on a single realization or estimated model of properties that was generated on a structured grid. Generating a design based on a single realization in the proposed workflow is not straightforward because the realization generated during preliminary mapping will be different from the one generated on the discretization. Features of the unstructured grid that are based on the preliminary mapping realization may no longer apply to the discretized realization. The proposed approach is restricted to more general grid designs, those that are not based on a single realization of reservoir properties.

Using a single grid design rather than a different grid for each realization implies the approach may be limited when considering large scale heterogeneity, that is, uncertainty in faults, horizons, channels, and other geological features that are incorporated into an unstructured grid. In this case, each realization has different large scale features. If the grid is to conform to such features, then a different grid is necessary for each realization.

Another limitation is the type of discretization that can be used to refine an unstructured grid for geostatistical modeling. Triangular and tetrahedral grids were used in this dissertation. Other discretizations can be used, but they are limited by the type of flow simulation algorithm that is available if flow based upscaling will be used. The type of discretization chosen can also restrict the types of flow simulation algorithms that are applicable. In choosing triangular grids, the TPFA cannot be used for flow based upscaling. The discretization approach used in the case study excludes all simulators that require a grid.

The MPFA method allows flow simulation on several grid types including triangular grids; however, the stability of the system of equations to solve for the pressures is a limitation. Stability is sensitive to the quality of the discretization and it depends on the geometry of the grid elements and the permeability tensors within them. Discretization in this thesis was done based on geometric quality alone and although stability problems were not encountered in two dimensions with flow based upscaling studies, they may be a concern in three dimensions. Also, shifting to three dimensions using tetrahedral discretization causes the number of grid elements to increase dramatically from two dimensions and triangulations. Depending on the size of the reservoir modeling problem, the number of elements in the discretized grid may grow too large and property modeling will take an excessive length of time.

Another grid-related limitation is the flow simulation grid must be static. Discretization depends on the grid and so do the resulting reservoir property models. Methods involving adaptive grids are not applicable. Each step in an adaptive process that changes the grid requires re-computing the discretization, generating another property model, and upscaling. Although this could be automated, it does not seem feasible for two reasons: extensive amounts of computer time are involved; and the generated property model is not guaranteed to preserve information from the initial realization.

Using unstructured discretizations also limits the type of geostatistical modeling algorithms that are applicable. Sequential indicator and sequential Gaussian simulation using kriging as the underlying estimator are applicable to any spatial configuration of points; however, two geostatistical algorithms are restricted to regular grids: multipoint statistics (MPS) and spectral simulation based on the Fast Fourier transform (FFT). Both MPS and the FFT cannot be applied directly; however, they may be applied in an approximative manner using an underlying regular grid. Other facies modeling techniques such as object based modeling and process mimicking techniques introduce a challenge because the unstructured grid design should conform to the resulting geometry of the objects or facies. Objects are placed into a model with uncertainty and processes proceed with uncertainty. A different unstructured grid design would be required for each realization in these cases.

6.3 Future Work

The presented approach to reservoir modeling using unstructured grids could benefit from further research. Four issues were identified through the case study: 1 – variogram reproduction was not as good as anticipated; 2 – a vertical asymptote exists in the discretization error function; 3 – discretizing the coarse tetrahedral grid with a finer scale tetrahedral grid was unsuccessful; 4 – the permeability upscaling technique is not developed in three dimensions. Components of sequential simulation that impact variogram reproduction are the random path, anisotropy of the search; and the number of data used for kriging. Research into these areas may identify the source of variogram reproduction issues identified in the case study. For the second issue, different functions to approximate the error could be looked into, or more efficient methods to numerically evaluate the average covariance integrals of the exact equation. Future work related to the third issue is to explore existing software to solve the problem or develop new methods that are specialized for dealing with the geometries involved in reservoir engineering and geological modeling.

Not all geostatistical modeling algorithms were adapted for unstructured grids. Other approaches for modeling categorical and continuous variables could be adapted, with MPS and spectral simulation as exceptions. Methods such as truncated Gaussian simulation and truncated pluri-Gaussian simulation for categorical variables are possible on any spatial configuration of points. Methods that handle non-stationary behaviour of reservoir properties would be beneficial as well including simulation with a locally varying mean and external drift. Multivariate techniques such as LU simulation and full cokriging are also possible.

Work could be done to handle the scale of discretized grid elements with more accuracy. In this thesis, the focus was on limiting the error involved in the upscaled values on the coarse unstructured grid. Improving accuracy at the finer scale will improve accuracy on the coarse scale. This can be taken a step further and methods to account for the scale of coarse unstructured grid elements directly without the need for a discretization step could be considered.

Flow-based upscaling using the MPFA was developed as a local upscaling method. A few variants of boundary conditions were covered, but not thoroughly tested. Boundary conditions that would make the approach useful in a coupled local-global upscaling regime are an area for further work. The method also requires extension to three dimensions. This will require further work on generating three dimensional grids and discretizations that result in a stable system of equations for the MPFA.

The discretization scheme developed in this thesis is ideal for multiscale methods. The grids align exactly. Multiscale methods using the MPFA could be developed for coarse polygonal grids discretized by triangles or tetrahedra. Instead of upscaling, the fine scale information is retained and used to simulate fluid flow on the fine scale grid, while the coarse scale grid is used to solve for the pressure. The last future work item is geostatistical modeling and flow based upscaling using meshfree methods (Liu and Gu, 2005). Discretization in this case does not require a grid of elements and the approach used in the case study could be used in meshfree flow-based upscaling.

6.4 Final Remarks

Discretizing unstructured grids using the presented approach allows the accuracy of upscaled properties to be controlled. For flow based upscaling, computing pressures and resolving flux across coarse element interfaces is straightforward and introduces no inaccuracies due to interpolation. As with any flow based upscaling method, the goal is to minimize the loss in accuracy that is a result of using a coarse grid for flow simulation. Coarse grids are used to reduce the amount of computational effort that is required for flow simulation. This is a necessity when the uncertainty in a flow response is assessed, since quantifying uncertainty requires several flow simulation runs with different realizations of reservoir properties.

Minimizing the reduction in accuracy of the average covariance is also important. Using the discretization approach, the covariance between coarse grid elements is correct to a known accuracy and realizations on the coarse scale have the correct spatial structure. The spatial structure of reservoir properties influences the path fluids take in flow simulation and this affects the uncertainty in a flow response. Making decisions such as injection well placement depends on the path injected fluids will take from the injector to a producing well. By having the correct spatial structure, the uncertainty in production information such as breakthrough time and water cut relates back to the variograms used for modeling the reservoir properties.

The variogram is just one of many parameters that effects flow simulation results. The histogram of reservoir properties, characteristics of the fluids involved, surface conditions of the pore spaces, and boundary conditions are a few others. Flow simulation sensitivity studies are used to understand how these parameters effect flow performance. By ensuring that the accuracy of other processes including reservoir modeling and upscaling is high, they will not bias the effects other parameters have on flow response, and these parameters can be better understood.

A challenge for any research is transferring findings to industry. This research builds on unstructured grids that have been undergoing research for some time in reservoir flow simulation; however, unstructured grids are still uncommon in industrial reservoir analysis applications. Research related to modeling unstructured grids with geostatistics will take time to be applied. There are a few findings from this research that would be useful in existing workflows with regular grids and can be incorporated sooner.

The approach for upscaling well data to a scale for geological modeling could be used instead of upscaling to regular grid elements. Some existing techniques for upscaling well data to regular grid elements do not account for the scale in the direction perpendicular to wells. The approach developed for assessing the error in upscaled values could be used to determine a discretization to use to upscale well data to the regular grid elements so that the scale of the elements is accounted for more accurately. In this case, the discretization would take on the form of a micro-model or mini-model (Deutsch, 2009; Hosseini et al, 2008). Errors in values upscaled to the flow simulation scale could also be assessed to aid in selecting an appropriate grid resolution for geological modeling.

Flow simulation using the MPFA is applicable to structured grids used for geological modeling. The flow based upscaling method developed in this dissertation could be adapted to upscale from a high resolution geostatistical model on a structured grid to a coarser structured or unstructured grid. Both grids could be designed to align and the advantages identified for aligned unstructured grids would apply.

Bibliography

Aarnes JE, Kippe V, Lie K-A, Rustad AB (2007) Modelling of multiscale structures in flow simulations for petroleum reservoirs. Hasle G, Lie K-A, Quak E (eds), Geometric Modelling, Numerical Simulation, and Optimization, Springer, 307-360

Aavatsmark I (2002) An introduction to multipoint flux approximations for quadrilateral grids. Journal of Computational Geosciences, 6, 405-432

Aavatsmark I, Barkve T, Bøe Ø, Mannseth T (1998a) Discretization on unstructured grids for inhomogeneous, anisotropic media. Part I: derivation of the methods. SIAM Journal of Scientific Computing, 19(5), 1700-1716

Aavatsmark I, Barkve T, Bøe Ø, Mannseth T (1998b) Discretization on unstructured grids for inhomogeneous, anisotropic media. Part II: discussion and numerical results. SIAM Journal of Scientific Computing, 19(5), 1717-1736

Aavatsmark I, Barkve T, Bøe Ø, Mannseth T (1996) Discretization on non-orthogonal, quadrilateral grids for inhomogeneous, anisotropic media. Journal of Computational Physics, 127, 2-14

Arbogast, T (2002) Implementation of a locally conservative numerical subgrid upscaling scheme for two-phase Darcy flow. Computational Geosciences, 6(3), 453-481

Asmussen S, Glynn PW (2007) Stochastic simulation: algorithms and analysis. Springer Science+Business Media, 482

Aziz K (1993) Reservoir simulation grids: opportunities and problems. Journal of Petroleum Technology, 45(7), 658-663

Bear J (1972) Dynamics of fluids in porous media. American Elsevier Publishing Company, 765

Berg M de, Kreveld M van, Overmars M, Schwarzkopf O (2000) Computational geometry, algorithms and applications. Springer-Verlag Berlin Heidelberg, 375

Cardwell WT, Parsons RL (1945) Averaging permeability of heterogeneous oil sands. Petroleum Transactions, AIME, 160, 34-42

Caumon G, Grosse O, Mallet J-L (2005) High resolution geostatistics on coarse unstructured flow grids. O. Leuangthong, C. Deutsch (eds), Geostatistics Banff 2004, vol. 1, Springer Science+Business Media, 703-712

Chen Y, Li Y (2009) Local-global two-phase upscaling of flow and transport in heterogeneous formations. Multiscale Modeling & Simulation, 8(1), 125-153

Chen Y, Wan J, Yang Y, Mifflin RT (2008) Enriched multi-point flux approximation for general grids. Journal of Computational Physics, 277, 1701-1721

Chen Y, Mallison BT, Durlofsky LJ (2007) Nonlinear two-point flux approximation for modeling full-tensor effects in subsurface flow simulations. Computational Geosciences, 12(3), 317-335

Chen Y, Durlofsky LJ (2006) Adaptive local-global upscaling for general flow scenarios in heterogeneous formations. Transport in Porous Media, 62(2), 157-185

Chen Y, Durlofsky LJ, Gerritsen M, Wen XH (2003) A coupled local-global upscaling approach for simulating flow in highly heterogeneous formations. Advances in Water Resources, 26, 1041-1060

Chiles JP, Delfiner P (1999) Geostatistics: modeling spatial uncertainty. John Wiley and Sons, 720

Christensen JR, Dechelette DB, Ma H, Sammon PH (2004) Applications of dynamic gridding to thermal simulations. SPE 86969, 10

Christie MA (2001) Flow in porous media – scale up of multiphase flow. Current Opinion in Colloid & Interface Science, 6, 236-241

Christie MA (1996) Upscaling for reservoir simulation. SPE 37324, 1004-1010

Cormen TH, Leiserson CE, Rivest RL, Stein C (2002) Introduction to algorithms. Massachusetts Institute of Technology, 1056

Cressie NAC (1993) Statistics for spatial data. John Wiley & Sons, 928

David M (1977) Geostatistical ore reserve estimation. Elsevier, Amsterdam, 364

Davis MW (1987) Production of conditional simulations via the LU triangular decomposition of the covariance matrix. Mathematical Geology 19(2), 91-98

Dubrule O, Damsleth E (2001) Achievements and challenges in petroleum geostatistics. Petroleum Geosciences, 7, S1-S7

Deutsch CV (2009) Estimation of vertical permeability in the McMurray formation. Proceedings of the Canadian International Petroleum Conference, Calgary, Alberta, June 16-18, 10

Deutsch CV, Lan Z (2008) The beta distribution for categorical variables at different support. GF Bonham-Carter, Q Cheng (eds), Progress in Geomathematics, Springer-Verlag Berlin Heidelberg, 445-456

Deutsch CV, Tran TT, Pyrcz MJ (2002) Geostatistical assignment of reservoir properties on unstructured grids. SPE Annual Technical Conference and Exhibition, San Antonio, 29 September – 2 October, 10

Deutsch CV (2002) Geostatistical reservoir modeling. Oxford University Press, 384

Deutsch CV, Journel AG (1998) Geostatistical software library and user's guide. Oxford University Press, 384

Deutsch CV, Srinivasan S, Mo Y (1996) Geostatistical reservoir modeling accounting for precision and scale of seismic data. SPE 36497, 9-19

Deutsch CV (1989) Declus: a FORTRAN 77 program for determining optimum spatial declustering weights. Computers & Geosciences, 15(3), 325-332

Deutsch CV (1989) Calculating effective absolute permeability in sandstone/shale sequences: SPE Formation Evaluation, September, 343-348

Deutsch CV (1987) Estimating block effective permeability with geostatistics and power averaging. SPE, 15991, 32

Donaldson EC, Chilingarian GV, Yen TF (1989) Enhanced oil recovery, II: processes and operations. Elsevier, 618

Du Q, Wang D (2006) Recent progress in robust and quality Delaunay mesh generation. Journal of Computational and Applied Mathematics, 195, 8-23

Du Q, Faber V, Gunzberger M (1999) Centroidal voronoi tessellations: applications and algorithms. SIAM Review, 41(4), 637-676

Durlofsky LJ (2003) Upscaling of geocellular models for reservoir flow simulation: a review of recent progress. 7th International Forum on Reservoir Simulation, Buhl/Baden-Baden, Germany, June 23-27, 58

Edelsbrunner H (2001) Geometry and topology for mesh generation. Cambridge University Press, 190

Edwards MG (2002) Unstructured, control-volume distributed, full-tensor finite-volume schemes with flow based grids. Computational Geosciences, 6, 433-452

Edwards MG, Rogers CF (1998) Finite volume discretization with imposed flux continuity for the general tensor pressure equation. Computational Geosciences, 2, 259-290

Eigestad GT, Klausen R (2005) On the convergence of the multipoint flux approximation O-method: numerical experiments for discontinuous permeability. Numerical Methods for Partial Differential Equations, 21, 1079–1098

Eigestad GT, Aavatsmark I, Espedal M (2002) Symmetry and M-matrix issues for the O-method on an unstructured grid. Computational Geosciences, 6, 381-404

Ellis DV, Singer JM (2007) Well logging for earth scientists. Springer, 692

Farmer CL (2005) Geological modelling and reservoir simulation. Iske A, Randen T (eds), Mathematical Methods and Modelling in Hydrocarbon Exploration and Production, Springer, 451

Farmer CL (2002) Upscaling: a review. International Journal for Numerical Methods in Fluids, 63-78

Flandrin N, Borouchaki H, Bennis C (2006) 3D hybrid mesh generation for reservoir simulation. Int. Journal for Numerical Methods in Engineering, 65, 1639-1672

Frey PJ, George P-L (2000) Mesh generation: application to finite elements. HERMES Science Europe Ltd., 814

Friedmann SJ, Stamp VM (2006) Teapot Dome: characterization of a CO_2-enhanced oil recovery and storage site in Eastern Wyoming. The American Association of Petroleum Geologists, 181-199

Frykman P, Vejbæk OV, Rasmussen R (2005) Scaling relations and sampling volume for seismic data. O. Leuangthong and C.V. Deutsch (eds), Geostatistics Banff 2004, 731-736

Frykman P, Deutsch CV (2002) Practical application of geostatistical scaling laws for data integration. Petrophysics, 43(3), 153-171

Frykman P, Deutsch CV (1999) Geostatistical scaling laws applied to core and log data. SPE Annual Technical Conference and Exhibition, Houston, Texas, 3–6 October, 12

Fujimoto T, Ranade RR (2004) Two characterizations of inverse-positive matrices: the Haukins-Simon condition and the Le Chatelier-Braun principle. Journal of Linear Algebra, 11, 59-65

Fung LS-K, Buchanan L, Sharma R (1994) Hybrid-CVFE method for flexible-grid reservoir simulation. SPE Reservoir Engineering, August, 188-194

Garcia RG (2005) Reservoir simulation of CO_2 sequestration and enhanced oil recovery in the Tensleep formation, Teapot Dome field. Texas A&M University, 97

Geiger S, Roberts S, Matthäi SK, Zoppou C, Burri A (2004) Combining finite element and finite volume methods for efficient multiphase flow simulations in highly heterogeneous and structurally complex geologic media. Geofluids, 4, 284-299

Gomez-Hernandez JJ, Cassiraga EF (1994) Theory and practice of sequential simulation. Geostatistical Simulations, M. Armstrong and P.A. Dowd (eds), Kluwer, Dordrecht, Netherlands, 111-124

Gomez-Hernandez JJ, Journel AG (1990) Stochastic characterization of grid-block permeabilities: from point values to block tensors. D. Guerillot and O. Guillon (eds), 2nd European Conference on the Mathematics of Oil Recovery, Paris, 83-90

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, 496

Goovaerts P (1994) Comparative performance of indicator algorithms for modeling conditional probability distribution functions. Math Geol 26(3), 389–411

Gringarten E, Arpat B, Haouesse A, Dutranois A, Deny L, Jayr S, Tertois AL, Mallet JL, Bernal A, Nghiem L (2008) New grids for robust reservoir modeling. Society of Petroleum Engineers 116649, 11

Groshong RH (2006) 3-D structural geology: a practical guide to quantitative surface and subsurface map interpretation. Springer, 400

Haas A, Dubrule O (1994) Geostatistical inversion – a sequential method of stochastic reservoir modelling constrained by seismic data. First Break, 12, 561-569.

Hale D (2002) Atomic meshes: from seismic imaging to reservoir simulation. Proceedings of the 8th European Conference on the Mathematics of Oil Recovery, Freiberg, Germany, ISBN 90-73781-24-8

He C, Edwards MG, Durlofsky LJ (2002) Numerical calculation of equivalent cell permeability tensors for general quadrilateral control volumes. Computational Geosciences, 6, 29-47

Heinemann ZE, Brand CW, Munka M, Chen YM (1991) Modeling reservoir geometry with irregular grids: SPE Reservoir Engineering, 225-232

Hosseini AH, Leuangthong O, Deutsch CV (2008) An integrated approach to permeability modeling using micro-models. International Thermal Operations and Heavy Oil Symposium, Calgary, Alberta, October 20-23, 10

Hrenikoff A (1941) Solution of problems in elasticity by the framework method: Journal of Applied Mechanics, ASME, A8, 169-175

Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, 592

Jensen JL (1991) Use of the geometric average for effective permeability estimation. Mathematical Geology, 23(6), 833-840

Jones TA, Hamilton DE, Johnson CR (1986) Contouring geologic surfaces with the computer. Springer, 336

Journel AG, Kyriakidis PC (2004) Evaluation of mineral reserves. Oxford University Press, 232

Journel AG, Deutsch CV, Desbarats AJ (1986) Power averaging for block effective permeability: 56th SPE California Regional Meeting, Oakland, April 2-4, 6

Journel AG (1983) Nonparametric estimation of spatial distributions. Math. Geology, 14(3), 445-468

Journel AG, Huijbregts C (1978) Mining geostatistics. Academic Press, New York City, 600

Karimi-Fard M, Durlofsky LJ (2009) Detailed near-well Darcy-Forchheimer flow modeling and upscaling on unstructured 3D grids. SPE 118999, 15

Katzmayr M, Ganzer L (2009) An iterative algorithm for generating constrained voronoi grids. SPE Reservoir Simulation Symposium, Woodlands, Texax, 2-4 February, 11

Kennel M (2004) KDTREE 2: Fortran 95 and C++ software to efficiently search for near neighbors in a multi-dimensional Euclidean space. Kennel, Matthew B. eprint arXiv:physics/0408067

Kitware Inc. (2006) VTK user’s guide. Kitware, Inc., 382

Klinger A (1972) Patterns and search statistics. Optimization Methods in Statistics, Rustagi, J.D. (ed), New York, Academic Press, 303-339

Knabner P, Angermann L (2003) Numerical methods for elliptic and parabolic partial differential equations. Springer-Verlag New York, 415

Kocberber S (1997) An automatic, unstructured control volume generation system for geologically complex reservoirs. SPE Reservoir Simulation Symposium, 241-252

Kupfersberger H, Deutsch CV, Journel AG (1998) Deriving constraints on small-scale variograms due to variograms of large-scale data. Mathematical Geology, 30(7), 837-852

Kythe PK, Schaferkotter MR (2004) Handbook of computational methods for integration. Chapman & Hall, 624

Lee SH, Wolfsteiner C, Durlofsky LJ, Jenny P, Tchelepi HA (2003) New developments in multiblock reservoir simulation: black oil modeling, nonmatching subdomains and near-well upscaling. SPE Reservoir simulation symposium, February, 11

Leuangthong O (2004) The promises and pitfalls of direct simulation. O. Leuangthong and C. Deutsch (eds), Geostatistics Banff 2004, vol. 1, Springer Science+Business Media, 305-314

Levy B, Caumon G, Conreaux S, Cavin X (2001) Circular incident edge lists: a data structure for rendering complex unstructured grids. Presented at IEEE Visualization, San Diego, October 21 – 26, 191-198

Liner C (2004) Elements of 3-D seismology. Pennwell Publishing, 450

Liu GR, Gu YT (2005) An introduction to meshfree methods and their programming. Springer 500

Maliska CR, Cordazzo J, Silva AFC (2007) Petroleum reservoir simulation using EbFVM: the negative transmissibility issue. Proc. 4th Int. Conf. on Computational Methods in Multiphase Flow, WIT Press, 133-142

Mallet J-L (2004) Space-time mathematical framework for sedimentary geology. Mathematical Geology, 36(1), 32

Mallet J-L (2002) Geomodeling. Oxford University Press, 624

Manchuk JG, Leuangthong O, Deutsch CV (2009) The proportional effect. Mathematical Geosciences, 41(7), 799-816

Manchuk JG, Leuangthong O, Deutsch CV (2004) Direct geostatistical simulation of unstructured grids. O. Leuangthong and C. Deutsch (eds), Geostatistics Banff 2004, vol. 1, Springer Science+Business Media, 85-94

Matsumoto M, Nishimura T (1998) Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Transactions on Modeling and Computer Simulation, 8(1), 3-30

McLennen JA (2007) The decision of stationarity. University of Alberta, 191

Miller GL, Talmor D, Teng S-H, Walkington N, Wang H (1996) Control volume meshes using sphere packing: generation, refinement and coarsening. Proceedings of 5th International Meshing Roundtable, 47-61

Miller GL, Talmor D, Teng S-H, Walkington N (1995) A Delaunay based numerical method for three dimensions: generation, formulation, and partition. Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing, Las Vegas, Nevada, 683-692

Mlacnik MJ, Durlofsky LJ (2006) Unstructured grid optimization for improved monotonicity of discrete solutions of elliptic equations with highly anisotropic coefficients. Journal of Computational Physics, 216, 337-361

Mlacnik MJ, Durlofsky LJ, Heinemann ZE (2006) Sequentially adapted flow-based PEBI grids for reservoir simulation. SPE 90009, SPE Journal, September, 317-327

Mlacnik MJ, Harrer AW, Heinemann ZE (2003) Locally streamline-pressure-potential-based PEBI grids. SPE Reservoir Simulation Symposium, Texas, 3-5 February, 12

Olsen DK, Sarathi PS, Hendricks ML, Schulte RK, Giangiacomo LA (1993) Case history of steam injection operations at Naval Petroleum Reserve No. 3, Teapot Dome Field, Wyoming: a shallow heterogeneous light-oil reservoir. SPE 25786, 93-110

Oz B, Deutsch CV, Tran TT, Xie Y (2003) DSSIM-HR: a FORTRAN 90 program for direct sequential simulation with histogram reproduction, Computers and Geosciences, 29, 39-51

Palagi CL, Aziz K (1994) Use of voronoi grid in reservoir simulation. SPE Advanced Technology Series, 2(2), 69-77

Paluszny A, Matthäi SK, Hohmeyer M (2007) Hybrid finite element-finite volume discretization of complex geologic structures and a new simulation workflow demonstrated on fractured rocks. Geofluids, 7, 186-208

Pathak P, Winterfeld PH, Davis HT, Scriven LE (1980) Rock structure and transport therein: unifying with voronoi models and percolation concepts: First Joint SPE/DOE Symposium on Enhanced Oil Recovery, SPE 8846, 381-395

Pawlowsky-Glahn V, Egozcue JJ (2006) Compositional data and their analysis: an introduction. Buccianti, A., Mateu-Figueras, G., and Pawlowsky-Glahn, V. (eds), Compositional data analysis in the geosciences: from theory to practice: Geological Society, London, Special Publications, 264, 1-10

Pawlowsky-Glahn V, Olea RA (2004) Geostatistical analysis of compositional data: Oxford University Press, 304

Pedrosa OA, Aziz K (1986) Use of a hybrid grid in reservoir simulation. SPE Reservoir Engineering, November, 611-621

Pirzadeh S (1993) Structured background grids for generation of unstructured grids by advancing front method. AIAA Journal, 31(2), 257-265

Ponting DK (1992) Corner point geometry in reservoir simulation. King, P.R. (ed) The Mathematics of Oil Recovery, Institute of Mathematics and its Applications, 45-65

Potella RCM, Hewett TA (2000) Upscaling, gridding, and simulation using streamtubes. SPE 49071, 361-371

Prevost M, Lepage F, Durlofsky LJ, Mallet J-L (2005) Unstructured 3D gridding and upscaling for coarse modeling of geometrically complex reservoirs. Petroleum Geoscience, 11(4), 8

Prevost M (2004) Accurate coarse reservoir modeling using unstructured grids, flow-based upscaling and streamline simulation. Stanford University, 248

Renard Ph, de Marsily G (1997) Calculating equivalent permeability: a review. Advances in Water Resources, 20(5), 253-278

Samet H (1990a) Applications of spatial data structures: Computer graphics, image processing, and GIS. Addison-Wesley, 507

Samet H (1990b) The design and analysis of spatial data structures. Addison-Wesley, 510

Sammon PH (2003) Dynamic grid refinement and amalgamation for compositional simulation. SPE 79683, 11

Schroeder W, Martin K, Lorensen B (2002) The visualization toolkit: an object-oriented approach to 3D graphics. Kitware, Inc., 521

Shewchuk JR (1996) Triangle: engineering a 2D quality mesh generator and Delaunay triangulator. Applied Computational Geometry: Towards Geometric Engineering (Ming C. Lin and Dinesh Manocha, eds.), vol. 1148 of Lecture Notes in Computer Science, Springer-Verlag, Berlin, 203-222

Si H (2006) TetGen: A quality tetrahedral mesh generator and three-dimensional Delaunay triangulator. Numerical Mathematics and Scientific Computing Research Group, Weierstrass Institute for Applied Analysis and Stochastics

Skiena SS (2008) The algorithm design manual. Springer-Verlag London, 736

Skoreyko F, Sammon PH, Melichar H (2003) Use of PEBI grids for complex advanced process simulators. SPE 79685, 10

Squillacote AH (2008) The ParaView guide. Kitware, Inc., 366

Tran TT, Deutsch CV, Xie Y (2001) Direct geostatistical simulation with multiscale well, seismic, and production data. SPE Annual Technical Conference and Exhibition, New Orleans, 30 September – 3 October, 8

Verma SK (1998) Flexible grids for reservoir simulation. Stanford University, 280

Verma S, Aziz K (1997) A control volume scheme for flexible grids in reservoir simulation. SPE 37999, Proc. 14th SPE Symposium on Reservoir Simulation, Dallas, Texas, 215-227

Verma S, Aziz K (1996) Two- and three-dimensional flexible grids for reservoir simulation: 5th European Conference on the Mathematics of Oil Recovery, Leoben, Austria, 3-6 September

Wackernagel H (2003) Multivariate geostatistics: an introduction with applications. Springer-Verlag Berlin Heidelberg, 403

Warren JE, Price HS (1961) Flow in heterogeneous porous media. Society of Petroleum Engineers Journal, September, 153-169

Webster R, Oliver MA (2007) Geostatistics for environmental scientists. John Wiley & Sons, 330

Wen X-H, Gomez-Hernandez JJ (1996) Upscaling hydraulic conductivities: an overview. Journal of Hydrology, 183, 9-32

White CD, Horne RN (1987) Computing absolute transmissibility in the presence of fine scale heterogeneity. SPE 16011, 9th SPE Symposium on Reservoir Simulation, San Antonio, Texas, pp. 209-220

Xu W, Journel AG (1994) DSSIM: a general sequential simulation algorithm. Report 7, Standford Center for Reservoir Forcasting, May

Xu W, Tran TT, Srivastava RM (1992) Integrating seismic data in reservoir modeling: the collocated cokriging alternative. SPE 24742, 833-842

Yadecuri ME, Mahani H (2009) Unstructured coarse grid generation for reservoir flow simulation using background grid approach. SPE 120170, 13

Yamamoto JK (2000) An alternative measure of the reliability of ordinary kriging estimates. Mathematical Geology, 32(4), 489-509

Zezula P, Amato G, Dohnal V, Batko M (2006) Similarity search: the metric space approach. Springer Science+Business Media, 220

Zienkiewicz OC, Taylor RL, Zhu JZ (2005) The finite element method: its basis & fundamentals. Butterworth-Heinemann, 752

Appendix

A FORTRAN program called psgsim was developed to perform sequential simulation of categorical and continuous properties on irregular sets of points. Sequential indicator simulation and sequential Gaussian simulation are possible. Program flow is as follows:

1. Load conditioning data if it exists into main memory.
2. Load point locations for simulation into main memory.
3. Partition the data into categories if categorical data is used.
4. Data transformation.
   a. Categorical variables transformed to indicators.
   b. Continuous variables normal score transformed.
5. For each realization:
6. Generate a random path through the point set.
7. For each point:
   a. Search for the set of nearest neighbours including conditioning data and previously simulated points.
   b. Kriging. For indicators, the proportion of each category is calculated and for continuous variables to conditional mean and variance.
   c. Draw from the conditional distribution and add to the pool of previously simulated points.
8. Back transformation.
   a. Indicators are converted to categories.
   b. Gaussian values are converted to the original continuous variables.
9. Output results to disk.
10. Go to step 5 to generate another realization.

Input parameters for psgsim are defined in a keyword based parameter file that is parsed upon program execution. Not all keywords are required and their order in the parameter file is irrelevant. Table 1 describes keywords and associated parameters that must be space delimited. The table is followed by more detailed explanations where more information may be required. Comments can be added to the parameter file by preceding them with pound sign (#). An entire block of parameters can be commented out by preceding the keyword with # as well. Several execution modes are also possible depending on which keywords are specified in the parameter file (Table 2).

Table 1: Parameters for psgsim.

Keyword	Purpose	Line	Parameters
DATA	Specify that conditioning data exists.	1	`DATA`
		2	`File with conditioning data`
		3	`Columns for x, y, z, weight`
		4	`Number of variables` ($nvar$)
		5	`Columns for variables` ($1,\ldots,nvar$)
		6	`Variable type`
		7	`Transformation flag`
UNDEF	Specify undefined values.	8	`UNDEF value`
PROPS	Specify global proportions for categorical variables. As many proportions as categories must be specified.	9	`PROPS`
		10	prop1 prop2 … prop$k$
TAILS	Specify minimum and maximum values for continuous variables[3].	11	`TAILS` nt
		12	vnum1 `minimum maximum`
		13	…
		14	vnumnt `minimum maximum`
SEARCH	Specify search type and ellipsoid.	15	`SEARCH`
		16	`type`
		17	`ang1 ang2 ang3`
		18	`rad1 rad2 rad3`
SIMP	Specify locations for simulation and related parameters.	19	`SIMP`
		20	`File with point locations`
		21	`Columns for x, y, z`
		22	`Maximum conditioning data`
		23	`Number of realizations`
		24	`Random number seed`
		25	`File with specified path`
		26	`File for output`
KRIG	Specify the kriging type.	27	`KRIG`
		28	$ktype$
IVARGS	Specify variograms for categorical variables	29	`IVARGS`
		30	nst `nugget` hf tf `sill`
		31	$type_k$ cc_k $ang1_k$ $ang2_k$ $ang3_k$
		32	$range1_k$ $range2_k$ $range3_k$
CVARGS	Specify variograms for continuous variables	33	`CVARGS`
		34	nst `nugget category` $h\,t\,corr$
		35	$type_k$ cc_k $ang1_k$ $ang2_k$ $ang3_k$
		36	$range1_k$ $range2_k$ $range3_k$

Line 2 and 20 – Input data files have the following format:

Line	**Content**	**Example**
1	Title	`Data for psgsim`
2	Number of columns, n	`3`
3	Column 1 title	`East, m`
4	Column 2 title	`North, m`
5,...,$n+2$	Column n title	`porosity`
$n+3$	Start of data records	`10.0   5.0   0.22`

Line 6 – Specify the type as continuous with a 0 or categorical with a 1. Only 1 categorical variable is supported, but there is no limit to the number of categories in that variable.

Line 7 – Normal score transform continuous variables from Line 6. For each variable, the transform is applied if a 1 is input and not if a 0 is input.

Line 9 – If the PROPS keyword is not included or is commented out, then proportions are calculated based on a categorical variable from input conditioning data.

Line 11 – The number of distribution tails, nt, does not have to match the number of continuous variables. The variable number coincides with the input order used to specify columns for variables in the DATA block. If tails are not specified, the range of the conditioning data is used.

Line 12 – Minimum and maximum tail values are used to extrapolate the tails of continuous variable distributions.

Line 15 – Two search methods are available: the superblock search specified with a 0 or a kd-tree with a 1. For the superblock search, the angles and radii define an ellipsoid which is used to determine the superblock grid specification and search template. For the kd-tree, ellipsoid parameters are used only to define the anisotropy for distance calculations. Radii do not constrain search results when using the kd-tree.

Line 22 – The maximum conditioning data define the number of conditioning data and previously simulated points to use in kriging.

Line 24 – Uniform real random numbers in [0,1] are generated using the Mersenne Twister (Matsumoto and Nishimura, 1998). This requires a non-zero integer for a random number seed.

Line 25 – Random paths are generated automatically, but other random paths may be specified in a file and passed to the program. The file must only contain a set of space delimited integers that index into the file with point locations. Integers coincide with the order the point locations are stored in the point file.

Line 28 – Three kriging types are supported: simple kriging specified with a 0; ordinary kriging with a 1; and collocated cokriging with a 2.

Line 29 – As many variogram specifications as categories are required. The block from nst to $range3$ is repeated for each variogram, where nst is the number of nested structures, hf and tf are the head category and tail category, $type$ defines the model to use from a set of available models (0 – spherical, 1 – exponential, 2 – Gaussian, 3 – power, 4 – hole effect), cc is the variance contribution, and the angles and ranges define the anisotropy of the model. Parameters from $type$ to cc are repeated for $k = 1, \ldots, nst$.

Line 33 – Specified similarly to indicator variograms, except the facies or category that the variogram applies to is specified and h and t refer to the head and tail continuous variable number, and $corr$ is the correlation between them. Full linear models of coregionalization and cokriging are not supported; however, for collocated cokriging ($ktype = 2$), any arbitrary cross variogram can be input with the correct correlation that is used by the program. The cross variogram is not used.

Table 2: Execution modes of psgsim.

Keywords present	Execution mode
DATA with continuous and categorical variables, CVARGS, IVARGS	Conditional simulation of categorical and continuous variables.
DATA with continuous variables only	Conditional simulation of continuous variables.
DATA with continuous variables only, PROPS, IVARGS	Unconditional simulation of categorical variables and conditional simulation of continuous variables.
DATA with categorical variables only	Conditional simulation of categorical variables.
DATA with categorical variables only, CVARGS	Conditional simulation of categorical variables and unconditional simulation of continuous variables.
CVARGS	Unconditional simulation of continuous variables.
PROPS, IVARGS	Unconditional simulation of categorical variables.
CVARGS, PROPS, IVARGS	Unconditional simulation of categorical and continuous variables.