DISSERTATION

EFFECTIVE TRANSMISSIVITY IN TRANSIENT STREAM DEPLETION

Submitted by

Calvin D. Miller, Jr.

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2015

Doctoral Committee:

Advisor: Deanna S. Durnford

Luis A. Garcia
William E. Sanford
John D. Stednick

ABSTRACT

EFFECTIVE TRANSMISSIVITY IN TRANSIENT STREAM DEPLETION

Quantifying the timing of streamflow depletion caused by groundwater pumping wells is a central issue in the conjunctive management of groundwater and surface-water resources. It is an important consideration in regions where water supplies and demands are offset annually by season and interannually through variable wet and dry years. Increased water demands, regulatory policy shifts, and aquifer changes have brought scrutiny to this type of stream-aquifer interaction. From this, analytical models of stream depletion have received renewed attention and numerous refinements which have focused primarily on a variety of complex boundary conditions. The question of representing aquifer heterogeneity through simplified input parameters in analytical models has not been directly examined for transient stream depletion.

The objective of this research is to identify upscaled transmissivities that effectively model transient stream depletion rates caused by pumping groundwater wells in heterogeneous aquifers. Two-dimensional heterogeneity is considered, with horizontal anisotropy in spatial correlation ranges as a primary independent variable. The subject aquifer is relatively narrow, with an impermeable boundary parallel to a fully connected river boundary.

Using numerical flow simulation and the Monte Carlo approach, stream depletion rate curves were computed for transmissivity fields constructed under various geostatistical models of heterogeneity. Effective (expected) and equivalent transmissivities—referring to stochastic ensemble-mean behavior and to individual realizations, respectively—were interpreted from the depletion curves using the Glover analytical solution for stream depletion in a homogeneous,

bounded aquifer. The interpreted effective and equivalent transmissivities were related to statistical moments of the heterogeneous fields through power averaging.

Effective transmissivity results ranged between the bounding arithmetic and harmonic means, varying with the spatial correlation structure of the transmissivity field, partly as a function of geometric statistical anisotropy. Notably, the shape of that function was similar to what has been derived analytically for steady-state, mean-parallel flow conditions in unbounded domains. Additionally, there was no apparent difference in effective transmissivity results between transient stream depletion conditions and steady-state, mean-parallel flow conditions simulated in the same test domain. Also, unlike some studies on effective permeability under various transient conditions, no time dependency was observed in effective transmissivity for the transient stream depletion case.

Results were sensitive to including a nugget effect in the spatial correlation model and to non-stationarity of the transmissivity field. Results were only mildly sensitive to field variance. Ensemble-mean behavior was mildly sensitive to correlation scale, but ensemble variance was strongly sensitive to correlation scale. The latter is to be expected, but is notable for stream depletion considering that transmissivity correlation is often regional in scale and thus often large relative to the scale of the pumping well depletion problem. In such cases, the equivalent transmissivity for a given field and well location was often case-specific and not well-predicted by the expected transmissivity.

ACKNOWLEDGEMENTS

First, I offer sincere thanks to my advisor, Dr. Deanna Durnford. Through two graduate degrees, now spread over two decades, I have benefited from her instruction and guidance. This included high standards for the knowledge one in our profession should have, but tempered with sharp cautions about thinking one can ever have *the* answers in groundwater. Her expectations were sometimes intimidating, but rightfully motivating. I have deeply appreciated her extended patience with me as I took detours along the way.

I credit Dr. Durnford, along with the overall CSU water programs (both times!), for my interest in the groundwater field and the professional career I enjoy from it. I thank my graduate committee: Drs. Luis Garcia, Bill Sanford, and John Stednick, for their positive support, interest, and patience. I also would like to acknowledge the late Dr. James Warner for his kind personal support, his instruction on numerical modeling, and for telling some good stories along the way.

I thank Jon Altenhofen, Mary Halstead, and Bruce Kroeker, among others, for their openness to work with me on the Tamarack Project and extended endeavors. They helped introduce me to interesting questions and issues in conjunctive water management.

I offer particular acknowledgment and gratitude to Dr. Ayman Alzraiee. His optimistic spirit, intelligence, and diligence have often inspired me over the years. On a more tangible note, Ayman provided great help with scripting my simulation approach in MATLAB. That was enormously valuable. Without that help and the extra kick it provided, this work might not have been completed.

DEDICATION

To my wife Katrina

and my daughter Madeleine

without whom this work might have been completed years earlier!

Or, perhaps never completed at all.

Either way, their encouragement and patience have been appreciated

and they have made life rich along the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS and ABBREVIATIONS

a	distance between pumping well and river
K	hydraulic conductivity
K_{eff}	effective hydraulic conductivity
K_A	arithmetic mean K
K_G	geometric mean K
K_H	harmonic mean K
λ_i	correlation scale in the i-direction
p	exponent of the power-average function
p_{eq}, p_{eff}	equivalent p and effective p
S	storage coefficient
SSMP	steady-state, mean-parallel
t'	non-dimensional characteristic response time, a^2S/T
T	transmissivity
T_{eff}	effective transmissivity
T_{eq}	equivalent transmissivity
T_A	arithmetic mean T
T_G	geometric mean T
T_H	harmonic mean T
TSD	transient stream depletion
Y	natural-log-transformed T or K
σ^2_Y	variance of Y
q	depletion rate from stream (volume/time)
Q	well pumping rate (volume/time)
W	aquifer width (on one side of river)

CHAPTER 1 INTRODUCTION

1.1 Research Motivation

Quantifying the timing of streamflow depletion caused by pumping wells is a central issue in the conjunctive management of groundwater and surface-water resources. This delay between pumping operations and consequent stream depletion is an important consideration where natural resource management policies may limit depletions seasonally and where water-rights systems may restrict out-of-priority depletion. It can be a crucial factor in regions where water supplies and demands are offset annually by season and interannually through variable wet and dry years. Likewise, this timing between aquifer changes and stream responses is a design and management tool in aquifer recharge projects created for streamflow re-timing and augmentation.

Increasing water demands, regulatory policy shifts, and aquifer water-level changes have brought renewed scrutiny to stream depletion mechanics and depletion estimation methods. This has led to many refinements in a wide array of stream depletion models—in both analytical solutions and numerical methods—but the refinements have primarily focused on alternative boundary conditions and more-detailed treatments of complex boundary conditions (see Huang et al. 2014). The issue of representing heterogeneous aquifer characteristics as simplified inputs to such models has not received as much attention. Upscaling local permeability estimates in heterogeneous aquifers for use in transient stream depletion models has not been directly examined in the literature.

1.2 Problem Basics

Pumping a groundwater well initially removes water only from aquifer storage, lowering water levels to create a so-called *cone of depression* in water levels around the well (Freeze and

Cherry 1979). Over time, the depression expands radially from the well and may interact with hydraulic and hydrogeologic boundaries in the area. If the aquifer is hydraulically connected to a surface-water boundary such as a river, then the rate of water removed from aquifer storage (volume/time) will decrease over time as water is increasingly derived from the river. If the river is the only head-dependent aquifer boundary present in the area, then the stream depletion rate (q) will eventually approach the pumping rate (Q). Stream depletion will also continue for a period after the cessation of pumping while the cone of depression is filled by water derived from the river (e.g., Jenkins 1968).

In the Glover analytical solution for stream depletion by wells, the depletion rate q (volume/time) as a function of time depends on the aquifer transmissivity (T), the storage coefficient or specific yield (S), the distance between the well and the river, and the location and shape of other flow boundaries (Glover and Balmer 1954; Glover 1977). For higher transmissivities, the time lag between pumping and depletions is shorter, with $q/Q = 1$ being approached more quickly, including post-pumping depletion rates that decrease more quickly.

1.3 Research Context

The Glover solution assumes homogenous aquifer transmissivity. A large body of literature has been devoted to the estimation of single-valued permeabilities that can effectively represent heterogeneous porous media in analytical models that assume homogeneity or that otherwise must simplify or upscale heterogeneity. Wen and Gomez-Hernandez (1996), Renard and Marsily (1997), and Sanchez-Vila et al. (2006) provide extensive reviews of that literature.

In the effective transmissivity concept, the goal is to replace a medium that has spatial variability in transmissivity by a medium with a single value of effective transmissivity (T_{eff}) while preserving selected hydraulic behavior(s) of the original medium (Smith and Freeze 1979;

El-Kadi and Brutsaert 1985, and others). A complete definition of T_{eff} requires specification of which hydraulic behavior is intended to be preserved (El-Kadi and Brutsaert 1985; Sanchez-Vila et al. 2006). For example, targeted behavior could include average head over the flow domain, drawdown near pumping wells, flux at hydrogeologic boundaries, etc. Specification of the targeted behavior is necessary since T_{eff} is not necessarily an intrinsic property of the medium; rather, it may depend on the type of targeted behavior plus flow and boundary conditions for a given application. In this research, the targeted behavior is the rate of stream depletion (q), i.e., the flux rate at the river boundary.

The body of available research has shown that effective permeability behavior is complex. In addition to it being a function of certain geostatistical characteristics of the heterogeneous medium, it is also a function of flow-domain dimensionality (i.e., one-, two-, or three-dimensional), of whether the flow regime is generally mean-parallel or convergent, and whether the flow is transient or steady-state. Effective permeability can also be a function of boundary conditions and location within the domain with respect to boundaries (Sanchez-Vila et al. 2006; Gomez-Hernandez and Gorelick 1989; Sanchez-Vila 1997; Riva et al. 2001). For example, for a pumping well in a bounded domain and steady-state flow, Riva et al. (2001) found effective transmissivity to be a non-monotonic function dependant on location between the well and the boundary. In summary, estimating an effective transmissivity that represents heterogeneous aquifer conditions is not simply a function of the heterogeneous medium characteristics; it can be specific to the flow scenario being considered.

Effective permeability estimates for various flow regimes and boundary scenarios are discussed in the literature review (Chapter 2). Despite some cases with general similarity—such as three studies of steady-state flow and pumping wells in bounded domains—to this author's

knowledge, the effective permeability literature does not include a direct analogy for the case of transient stream depletion from pumping wells.

1.4 Objective

The primary objective of this research is to identify effective transmissivities for the case of transient stream depletion caused by pumping a groundwater well in a heterogeneous aquifer that is hydraulically connected to a river. A secondary objective is to evaluate the sensitivity of that effective transmissivity behavior over a range of geostatistical model assumptions and conditions.

1.5 Scope

As noted previously, transient stream depletion refers to the change in the depletion rate over time. This refers to the *total* rate of river depletion at a given time. The depletion rate is also variable spatially and is distributed over a river reach that is several times longer than the distance between the well and the river (Glover 1977). Stream depletion is, therefore, a relatively large-scale phenomenon.

Spatial variations in the depletion rate along a river can be relevant in certain applications (e.g., Fleckenstein et al. 2006), and such local-scale spatial variability may be quite high due to heterogeneities along the river. Considering spatial variations is not within the scope of this work. The spatially variable local depletion rates are aggregated into a total depletion rate that is the subject of this research and the interest of many water management applications. The goal is to identify an appropriate averaging process for heterogeneous aquifer transmissivity such that the upscaled (averaged) transmissivity effectively predicts the timing of that total depletion rate.

The flow regime and boundary-condition scenario considered in this research is a relatively narrow, two-dimensional aquifer bounded on one side by a straight constant-head

boundary representing a river, and on the other side by a zero-flux boundary, oriented parallel to the river, representing an impermeable aquifer boundary such as the edge of a valley-fill aquifer. Changes in saturated thickness due to pumping are assumed to be small relative to the initial saturated thickness. The problem domain scale considered, and more importantly the correlation scale of the heterogeneous field relative to the domain scale, is representative of the typically large correlation scales reported for transmissivity (Rubin 2003; Anderson 1997; Dagan et al. 2009).

Simulations are limited to the case where the pumped groundwater volume is equal to the sum of changes in aquifer storage plus induced leakage from a single straight river boundary, i.e., there are no other head-dependant flux boundaries present. The river is treated as fully penetrating the aquifer depth and with a permeable riverbed, or otherwise sufficiently connected to the aquifer to function as a constant-head boundary.

This research applies equally to stream accretions (i.e., streamflow gains) caused by localized aquifer recharge, such as managed aquifer recharge operations. To be concise, the term *stream depletion* or the more general term *response curve* is used herein to refer to both stream depletion and accretion.

1.6 Methods Synopsis

Geostatistics offers tools to characterize, quantify, and model subsurface heterogeneity. It does this by accounting for the spatial continuity that is an identifiable feature of many natural phenomena (Isaaks and Srivastava 1989; Anderson 1997). Geostatistical modeling supports the stochastic approach to characterizing groundwater behavior, in which parameter variability and uncertainty are acknowledged and accounted for by specifying an assumed or estimated mean and variance of aquifer parameters (Anderson 1997). A goal of the stochastic approach is to

identify effective behavior—meaning the outcome expected on average—and to consider uncertainty around that expected behavior.

An *effective transmissivity* models expected behavior, and as such is an average over probability space, referring to the stochastic view of aquifer characterization. *Equivalent transmissivity* refers to a transmissivity that models response behavior for a given aquifer realization. Invoking ergodicity, at large enough domain scales the equivalent and effective transmissivities are expected to be similar (Sanchez-Vila et al. 2006).

Effective transmissivity for transient stream depletion is examined in this research through a Monte Carlo numerical simulation approach. Aquifer transmissivity is treated as a statistically anisotropic, spatially correlated, random variable. Approximately 11,000 realizations of synthetic heterogeneous transmissivity fields were generated using unconditional sequential Gaussian simulation (Deutsch and Journel 1998). The primary cases considered are stationary fields with different degrees of heterogeneity and geometric anisotropy in the spatial correlation patterns. Fields with non-stationarity, zonal anisotropy, and other geostatistical model variations are also evaluated.

Equivalent transmissivity, again referring to a single realization, is an average over physical space, such as averaging local transmissivity values over the domain of a given aquifer. In this research, however, that equivalent transmissivity is first identified through interpreting observed (simulated) response behavior for each realization. Indeed, the motivation for this research is the absence of literature on estimating equivalent or effective transmissivity from the local field properties for the specific case of transient stream depletion. The goal herein is to address that limitation empirically by simulating and interpreting the depletion rate response and relating those results to the spatial statistics of the transmissivity field.

A transient finite-difference numerical model simulation of each heterogeneous aquifer realization was used to determine depletion rates from the river boundary occurring in response to a specified-flux boundary condition (well pumping) started in the aquifer at time zero. The equivalent transmissivity for each field realization was identified from the simulated response as that transmissivity that results in a best fit between the heterogeneous-aquifer's depletion rate curve and the Glover analytical solution for stream depletion in a bounded homogenous aquifer (Glover and Balmer 1954; Glover 1977; Knight et al. 2005). Statistical moments of the ensemble simulation results then identify effective (expected) transmissivity (Sanchez-Vila et al 2006).

In this research, effective and equivalent transmissivities are related to the first and second statistical moments of the transmissivity fields through the power-average or power-mean function (Gomez-Hernandez and Gorelick 1989; Desbarats 1992b; Wen and Gomez-Hernandez 1996; Ronayne and Gorelick 2006; de Dreuzy 2010). Ensemble and individual realization results are presented in the form of power-mean exponents which yield the effective and equivalent transmissivities when applying the power-mean function to the heterogeneous transmissivity fields. In essence, the ensembles of power-mean exponents produced in this research indicate, as a function of certain geostatistical properties of heterogeneous transmissivity fields, what form of average to use for an effective transmissivity in models of transient stream depletion.

1.7 Organization

Chapter 2 reviews the literature on effective permeability estimation for different groundwater flow regimes. Recent publications on stream depletion modeling are also noted. Chapter 3 describes the methodology used in performing the Monte Carlo simulations and in

interpreting and quantifying simulation output. Chapter 4 presents the primary results of this research. Those results are relationships observed between effective transmissivity (expressed as effective power-mean exponents) and various geostatistical models of aquifer heterogeneity. Chapter 5 provides a sensitivity analysis of some of the geostatistical conditions assumed for the simulated heterogeneous fields. Chapter 6 presents field case examples that were analyzed for comparison to the synthetic aquifer results. Last, Chapter 7 presents discussion and conclusions.

CHAPTER 2 LITERATURE REVIEW

A few recent publications on stream depletion by pumping wells are highlighted in Section 2.1. Earlier stream depletion publications and a general history of analytical stream depletion models are noted in Miller et al. (2007).

A review of effective permeability literature is presented in Section 2.2, organized by type of groundwater flow regime. Presentation of the power average, a key part of the methods presented in Chapter 3, is also included (Section 2.2.4). The literature available for effective permeability in statistically anisotropic, heterogeneous porous media under steady-state mean-parallel flow is discussed in Section 2.2.5. That flow regime provides an import point of reference for the results presented in Chapter 4 for transient stream depletion.

2.1 Recent Publications on Stream Depletion by Pumping Wells

2.1.1 General Reviews and Analytical Solution Developments

A general overview and background on natural resource management issues and groundwater concepts involved in the topic of stream depletion by pumping groundwater wells can be found in Barlow and Leake (2012). Kendy and Bredehoeft (2006), Bredehoeft and Kendy (2008), and Bredehoeft (2011) discussed the transient nature of stream depletion by pumping wells and associated implications on water management.

Hunt (2014) reviewed the capabilities of most of the recent advances in analytical stream depletion models. Butler et al. (2007) and Zlotnik and Tartakovksy (2008) present analytical modeling that includes induced leakage from an underlying aquitard. They note such leakage can significantly impact stream depletion rates at certain scales. Huang et al. (2014) have also presented new analytical solutions generalized for a few different aquifer and boundary

configurations. It is noteworthy, considering part of the motivation for this dissertation (Section 1.1), that Huang et al. compiled tables listing 25 variants of analytical models of stream depletion, 22 of which were published within the last 16 years.

2.1.2 Stream-Aquifer Interactions with Heterogeneity

Fleckenstein et al. (2006) showed that intermediate-scale heterogeneity—meaning in their case the arrangement of hydrofacies underneath and near the river—can create significant spatial differences in the stream seepage (loss) rate on the Cosumnes River in California. Such spatial variability was significant for low-flow management by leading to variability in which sections of the river were hydraulically connected to water table changes, a complexity that a homogenous model failed to capture. They found total-river annual net depletions to be generally similar in their six different hydrogeologic model realizations, but variability in the location and rates of seepage loss led to different estimates about how many days the river would be open for salmon migration runs. Their work has implications for parameter upscaling in such applications (see also Fleckenstein and Fogg 2008) and for questions about data sufficiency in characterizing stream-aquifer interactions at their scale of focus.

Engdahl et al. (2010) constructed numerical groundwater flow models of the Rio Grande alluvial aquifer in New Mexico, with one homogeneous model and six heterogeneous conditional realizations. They simulated steady-state river losses under different river stages. Among other things, they found a decrease in stream loss with increasing heterogeneity and that a homogenous model of that system may over-predict stream loss.

2.2 Effective Permeability Literature

The estimation of single-valued effective or equivalent parameters that can be representative of heterogeneous porous media has been considered repeatedly in the literature.

Research has included specific yield and unsaturated-flow parameters, among other things, but the largest body of research has focused on hydraulic conductivity (*K*). Wen and Gomez-Hernandez (1996) and Renard and Marsily (1997) provided general reviews of the topic. A variety of specific aspects are found in numerous works (e.g., Vermeulen et al. 2006; Neuman and Federico 2003; Gelhar 1993). Sanchez-Vila et al. (2006) provided an extensive review on representative hydraulic conductivities, including discussion of the concepts and methods commonly used in that body of research. At 46 pages and containing around 270 citations, their review is a testament to the breadth and depth of the subject.

2.2.1 Steady-State Uniform Flow

Cardwell and Parsons (1945) showed, analytically and with electric-analog experiments, that equivalent permeability, conceptualized for a heterogeneous oil sand in their case, was bounded by the arithmetic and harmonic means of the volume-weighted permeabilities in the formation. Actual values were noted to lie between the bounds. These are sometimes referred to as the Wiener bounds in hydrogeology (Renard and Marsily 1997) and in fields such as composite materials engineering.

Many groundwater textbooks demonstrate where the Weiner bounds are precisely correct: the arithmetic mean for steady, uniform, one-dimensional (1D) flow oriented parallel to lithologic layering, and the harmonic mean for such flow oriented perpendicular to the layers. Those bounds are strictly correct only if the layers are laterally extensive or flow is constrained from lateral flow around low-*K* layers by boundaries (e.g., permeameter tube walls). For other flow regimes and heterogeneity types, such as a series of lenses instead of extensive layers, the averaged permeability falls between the bounds (Gelhar and Axness 1983; Sanchez-Vila et al 2006).

Warren and Price (1961) showed through statistical analysis of digital modeling experiments that "the most probable behavior of a heterogeneous system with single-phase flow approaches that of a homogeneous system having a permeability equal to the geometric mean of the individual permeabilities." Matheron (1967) is frequently cited (e.g., Lachassagne et al. 1989; Wen and Gomez-Hernandez 1996) for demonstrating mathematically that the effective permeability lies between the harmonic and arithmetic means and that the average permeability in two dimensions (2D) is equal to the geometric mean under certain restrictions.

Equivalent permeability has been shown to vary below and above the geometric mean for one and three dimensions, respectively, but in 2D the geometric mean has been frequently observed to reasonably represent average behavior in heterogeneous systems (Dagan 1979; Gomez-Hernandez and Gorelick 1989; Lachassagne et al. 1989; Wen and Gomez-Hernandez 1996; Sanchez-Vila 1997). For effective hydraulic conductivity (K_{eff}) in unbounded domains under *n*-dimensional flow, the following relationships were developed by Gutjahr (1978) and Dagan (1979):

$$K_{\text{eff}} = [\exp(\mu_Y)]\left(1 - \frac{\sigma_Y^2}{2}\right) \qquad n = 1 \qquad \text{(2-1a)}$$

$$K_{\text{eff}} = \exp(\mu_Y) \qquad n = 2 \qquad \text{(2-1b)}$$

$$K_{\text{eff}} = [\exp(\mu_Y)]\left(1 + \frac{\sigma_Y^2}{6}\right) \qquad n = 3 \qquad \text{(2-1c)}$$

Or, equivalently:

$$K_{\text{eff}} = [\exp(\mu_Y)]\left(1 + \left(\frac{1}{2} - \frac{1}{n}\right)\sigma_Y^2\right) \qquad \text{(2-2)}$$

where μ_Y and σ^2_Y are the mean and variance, respectively, of the natural-log-transform of K (Y = ln(K)) (Gomez-Hernandez and Gorelick 1989; Sanchez-Vila et al. 2006). The relationship given for 2D yields the geometric mean. These relationships were developed for low variance. Gelhar and Axness (1983) conjectured the following as an extrapolation valid for large σ^2_Y, as written by Sanchez-Vila et al. (2006):

$$K_{eff} = [\exp(\mu_Y)]\exp\left(\left(\frac{1}{2}-\frac{1}{n}\right)\sigma_Y^2\right) \tag{2-3}$$

2.2.2 Steady-State Radial Flow

Cardwell and Parsons (1945) reported that the arithmetic and harmonic bounds also applied to the radial case with the addition of weighting permeability values by their inverse square distance from the well. For drawdown at a well, Desbarats (1992a, 1993) suggested a geometric mean of point (local) T values weighted by the inverse square distance from the well and found agreement with a numerical model. According to Renard and Marsily (1997), Matheron (1967) noted that his proof for using the geometric mean in heterogeneous media does not necessarily hold for radial flow.

Sanchez-Vila (1997), using an analytical approach, found for steady-state radial flow in a statistically isotropic field that apparent effective transmissivity (T_{eff}) was a monotonically increasing function of distance from the well. It was approximately the harmonic mean near the pumping well and asymptotically approached the geometric mean with distance from the well. It was stressed that T_{eff} is not a value solely defined by the statistical properties of the heterogeneous field, but "rather it depends on the choice of boundary conditions".

Riva et al. (2001) used analytical methods, supported by numerical Monte Carlo simulations, to consider apparent transmissivity under steady-state radial flow to a well in a radially bounded heterogeneous aquifer. Their apparent transmissivity, which was defined as the ratio of expected flux to head gradient as a function of radial distance from the well, was a non-monotonic function of distance found to be the harmonic mean at the well, to trend toward the geometric mean (T_G) at intermediate radial distances, and then to increase to above T_G near the boundary.

Not all studies have found consistent conclusions for T_{eff} for flow toward wells (Sanchez-Vila et al. 2006). The results of Sanchez-Vila (1997) noted above, and the spatial variability of apparent transmissivity found by Riva et al. (2001) highlight that effective permeability in porous media may be specific to the particular location and boundary conditions considered.

Desbarats (1992a) found "that the expected block transmissivity decreased from the ensemble arithmetic mean to the ensemble geometric mean as the field size became large compared to the range of transmissivity spatial correlation." Desbarats (1993) considered steady-state interwell transmissivity between an injection well and a pumping well in a heterogeneous aquifer and found good agreement with a harmonic average of the effective conductivity around each well via the averaging scheme of Desbarats (1992a), the latter being a radial-distance-weighted geometric average around each well.

On the surface, Desbarats (1993) looks to be a perfect analogy for stream depletion since it uses the same dipole image well configuration as, for example, the Theis and Glover solutions for stream depletion (Theis 1941; Glover and Balmer 1954). However, Desbarats (1993) was relating steady-state flow between the wells to drawdown at the wells, while the stream depletion

case is interested in the transient flux evaluated at the stream boundary which would be mid-way between the two wells of the Desbarats analysis.

2.2.3 Transient Flow and Other Conditions

Numerous researchers have reported effective hydraulic conductivity values that vary with time, distance from the well, or other scale effects (Freeze 1975; Vandenberg 1977; El-Kadi and Brutsaert 1985; Indelman and Dagan 2004; Sanchez-Vila et al. 1996; Sanchez-Vila 1997; Desbarats 1992a; Ababou and Wood 1990; Naff 1991; Dagan 1982; Riva et al. 2001). Freeze (1975) was an early and influential paper using stochastic modeling to demonstrate the significant impacts of heterogeneity (Anderson 1997). Among other things, Freeze reported that effective hydraulic conductivity for transient flow changed with time. Vandenberg (1977) found through a numerical model that effective transmissivity varied with time and distance from the well and that, for the normally distributed transmissivity field used, an arithmetic average closely matched the Theis curve for transient drawdown in an equivalent homogeneous aquifer. Desbarats (1992a) noted that it isn't clear if Vandenberg's arithmetic average accounted for his progressively finer model grid spacing near the pumping well. The result, therefore, was possibly an arithmetic average effectively weighted by distance from the well.

Meier et al. (1998) found T_{eff} to be close to T_G for several numerically modeled cases of transient radial flow, although cases with well-connected high transmissivity zones yielded T_{eff} greater than T_G. Meier et al. also observed that T_{eff} estimated from late-time drawdown data at monitoring wells in simulated pumping tests was generally close to the T_{eff} expected for parallel flow. This was viewed as an encouraging result due to it partly validating the common use of T derived from radial-flow pumping tests to uniform flow applications. Similarly, Indelman (2003), using analytical methods and considering three-dimensional heterogeneity, reportedly

showed transmissivity from late-time pumping test drawdown to be precisely the effective conductivity for uniform horizontal flow.

In a transient analysis, Dagan (1982) found that T_{eff} was time dependent, dropping from the arithmetic mean initially to an expected steady-state value (e.g., Equations 2-1 to 2-3) during a relaxation time. Dagan found this relaxation time to be small in 3D under certain realistic aquifer parameters, but cautioned that for 2D it could be large due to the typically large correlation scales of transmissivity.

El-Kadi and Brutsaert (1985) also observed K_{eff} to vary with time in transient numerical simulations, with this temporal variation increasing with increased heterogeneity. They evaluated K_{eff} for transient discharge to a river boundary under gravity drainage of an adjacent, connected, and bounded aquifer (i.e., no pumping, just natural drainage to the river under a falling water table condition, such as after a sudden river stage change). They found K_{eff} to be greater than K_G and close to the arithmetic mean for early time, and approximately equal to K_G for late time. El-Kadi and Brutsaert noted that in their studies this relaxation time was shorter for smaller variance in K.

It appears to this author that the El-Kadi and Brutsaert (1985) observation of a high (arithmetic) early-time T_{eq} may have been a consequence of their flow domain and relatively coarse model discretization. At early time, most flow was taking place from drainage of, and flow across, the first one or two columns of their model, the columns along the river boundary. That flow would have been similar in nature to 1D flow across a parallel set of resistances and not through a series of resistances, thus behaving as an arithmetic mean of the K block values in the first one or two columns.

2.2.4 The Power Average

The *power mean* or *power average* has been used in many works to upscale or otherwise characterize effective permeability (e.g., Gomez-Hernandez and Gorelick 1989; Desbarats 1992b; Ronayne and Gorelick 2006; de Dreuzy 2010; and other examples listed in Wen and Gomez-Hernandez 1996 and de Dreuzy 2010). The power average can be written as:

$$K(p) \quad = \left[\frac{1}{N}\sum_i^N (K_i)^p\right]^{1/p} \tag{2-4}$$

in which $K(p)$ is the power-averaged K and the exponent p is used in the range $-1 < p < 1$.

The power average includes the following classical means: the arithmetic mean for $p = 1$, the harmonic mean for $p = -1$, and the geometric mean in the limit of $p \rightarrow$ zero (Gomez-Hernandez and Gorelick 1989, and others). But, it is not limited to those means; it also provides a continuum of intermediate values. The following inequality holds between the three means: harmonic < geometric < arithmetic. Note that power averaging is done with the K or T data values in real space, even though some of the mathematical relationships of the power average are based on normality of the log-transformed data.

The power average is sometimes referred to as a generalized mean, and also as the power norm yielding the p-norm of the input data (e.g., Gomez-Hernandez and Gorelick 1989). However, since the p-norm term appears to have a specific meaning in vector mathematics, the function is discussed herein as a power average or power mean and noted with $K(p)$ and $T(p)$. The power-average exponent (p) yielding the appropriate T_{eff} or T_{eq} from the T-field data is referred to as an effective p (p_{eff}) or an equivalent p (p_{eq}), respectively.

Ababou and Wood (1990) and Sanchez-Vila et al. (2006) note the following relationships apply when the distribution of K (or T) is log-normal: $K(p) = K_G(K_G/K_H)^p$, $K(p) = K_G(K_A/K_G)^p$ and:

$$K(p) = K_G \exp(p\,\sigma^2_Y/2). \tag{2-5}$$

A power mean can therefore be computed simply from knowing the statistical moments of the data (σ^2_Y and K_G), rather than empirically computing it from the field data of a given realization (Ababou and Wood 1990). The former approach can provide a lower computational burden in cases such as Monte Carlo simulations.

It is useful to characterize T_{eq} and T_{eff} by the power-mean function. First, it allows for intermediate values between the discreet classical Pythagorean means (arithmetic, geometric, and harmonic), a capability which is needed for media that is correlated but not perfectly stratified (Gelhar and Axness 1983; Sanchez-Vila et al. 2006). Another advantage is that it scales for variance, σ^2_Y. In characterizing effective permeability relationships, T_{eff} and K_{eff} are commonly normalized by the geometric mean of the heterogeneous field, but this ratio (T_{eff}/T_G) increases with increasing variance σ^2_Y even with all other things equal. (For example, $T_A/T_G = 1.649$ for $\sigma^2_Y = 1.0$, and $T_A/T_G = 2.718$ for $\sigma^2_Y = 2.0$. This T_{eff}/T_G ratio is illustrated in Figure 2 of Sanchez-Vila et al. (2006) as a function of σ^2_Y and as a function of p and σ^2_Y in Chapter 3 herein.) In contrast, the arithmetic mean is given by $p = 1$ no matter the variance level, allowing for a more direct comparison of averaging processes for fields of different variance. On the other hand, use of the power average could lead one to overlook the fact that for low heterogeneity (low σ^2_Y) there may be little practical difference between T_A, T_G, and T_H.

Gomez-Hernandez and Gorelick (1989) determined best-fit p exponents from a numerical model that generally corresponded with a heterogeneous field site. Their model of an

unconfined aquifer included 10 pumping wells, zoned heterogeneity potentially representing different geologic zones, and a riverbed with spatially variable leakance. They found *p* exponents of -0.4 and -0.2—values between the harmonic and geometric means—to best match modeled heads on the average. However, this was noted as being influenced by the behavior near the wells: K_G was reported as appropriate away from the wells and K_{eff} for their overall system reverted (increased) to K_G when the wells were turned off. Ababou and Wood (1990) commented that Gomez-Hernandez and Gorelick's model may have also behaved partly as a 1D system, thus lowering T_{eff} below the geometric mean expected for a 2D system (Equation 2-1).

Ronayne and Gorelick (2006) used power averaging and numerical simulation to characterize effective permeability in two-dimensional, binary-media systems of branching channel networks. They examined dependence of the effective power-average exponent on certain metrics that characterize the structure of the channel network. The purpose was to facilitate predictive modeling that didn't require numerical modeling.

The focus of de Dreuzy et al. (2010) was to evaluate the validity of power averaging for permeability upscaling. They considered a variety of lattice percolation networks. Their work did not include anisotropy or large-range correlation. A conclusion was that power averaging was strictly valid only for a limited case, but that it gives an approximation of upscaling at a precision within a few percent of the possible exponent range.

2.2.5 Steady-State, Mean-Parallel Flow in Statistically Anisotropic Media

Gelhar and Axness (1983) developed an analytical estimate for directional K_{eff} in a statistically anisotropic, 2D, infinite domain under mean-parallel flow conditions. Mean-parallel or mean-uniform flow refers to flow lines that are parallel on average at the marcro scale, i.e., no source/sink areas with convergent flow, but flow lines are not necessarily parallel at the local

scale in heterogeneous media. Sanchez-Vila et al. (2006) state that Gelhar and Axness may have been the first to address that statistically anisotropic problem, and they wrote Gelhar and Axness's 2D result more explicitly as:

$$K_{eff,i} = K_G \left[1 + \sigma_Y^2 \left(\frac{1}{2} - \frac{\lambda_j}{\lambda_1+\lambda_2}\right)\right] \quad (2\text{-}6)$$

where $i,j = 1,2$; $i \neq j$; and λ_1 and λ_2 are the correlation scales in the principal directions. The following generalization was proposed to extend Equation 2-6 to the domain of $\sigma^2_Y > 1$ (Gelhar and Axness 1983; Paleologos 1994; Sanchez-Vila et al. 2006):

$$K_{eff,i} = K_G \, exp\left[\sigma_Y^2 \left(\frac{1}{2} - \frac{\lambda_j}{\lambda_1+\lambda_2}\right)\right] \quad (2\text{-}7)$$

Ababou (1991) conjectured the following estimate for the power-average exponent (Equation 2-4) for the case of steady-state mean-parallel (SSMP) flow in statistically anisotropic aquifers (Ababou 1991; Sanchez-Vila et al. 2006):

$$p_i = 1 - \frac{2}{n}\frac{\lambda_H}{\lambda_i}; \qquad \lambda_H = \left(\frac{1}{n}\sum_{i=1}^{n} {\lambda_i}^{-1}\right)^{-1} \quad (2\text{-}8)$$

where n is the dimensionality of the problem, and λ_H is the harmonic average of the n directional correlation scales, λ_i. For the two-dimensional case this can be written as:

$$p_i = 1 - \frac{2\lambda_j}{\lambda_1+\lambda_2} \quad (2\text{-}9)$$

If Equation 2-5 is rewritten as:

$$p_{eff} = 2 \ln(K_{eff}/K_G)/ \sigma^2_Y \quad (2\text{-}10)$$

and Equation 2-7 is substituted for K_{eff}/K_G in Equation 2-10, then it is seen that Equations 2-9 and 2-7 are equivalent. Desbarats (1992b) noted they are equivalent for the three dimensional case as well. These functions are plotted in Chapter 4.

2.2.6 Existence of a Comprehensive Effective Permeability

In a summary paper, Anderson (1997) noted that many researchers have questioned the existence of a unique value for effective hydraulic conductivity, including Smith and Freeze (1979), El-Kadi and Brutsaert (1985), Gomez-Hernandez and Gorelick (1989), Neuman and Orr (1993), and Williams (1988). Williams (1988) was a qualitative critique, including questioning the validity of treating heterogeneities such as fractured rock zones as random processes. Others based their conclusions on the results of Monte Carlo simulations of hypothetical aquifers.

In light of the sensitivity of K_{eff} to pumping well operations, boundary conditions, and other aspects specific to their simulations, Gomez-Hernandez and Gorelick (1989) concluded there was not a unique best-fit effective conductivity for their modeled scenario. Ababou and Wood (1990) concurred and added that different possible fit-assessment criteria (e.g., overall head distribution, drawdown at wells, or flux) also complicate the concept of a single best-fit effective conductivity.

Sanchez-Vila et al. (2006), Fogg (1986), Meier et al. (1998), and others have noted that the degree of connectivity of the transmissivity field strongly influences effective parameters and that stochastic theory has not been able to characterize this connectivity outside of using numerical experiments. Wen and Gomez-Hernandez (1996) cite several studies which conclude the effective p exponent to be case specific (e.g., Desbarats 1992b).

A summary point is that one can find, with theoretical and physical basis, various estimates of the expected relationships between an upscaled or averaged K and a stochastic representation of heterogeneity, but caution is in order when using those relationships. The relationships, including those presented in Chapter 4, are not solely a characteristic of the media

but, rather, are also a function of the flow and boundary conditions. The degree of applicability will also depend on how well a given heterogeneous field can be characterized stochastically.

2.2.7 Evaluation Methods and Conceptual Assumptions

Analytical approaches on estimating effective parameters are often restricted to low degrees of heterogeneity (low variance in the log-transformed K field) and domains that are unbounded or with idealized boundary conditions. For less restrictive approaches, or to test the analytical derivations, numerical simulations and Monte Carlo methods are commonly used. Anderson (1997) provides this discussion:

> *"…a number of researchers have used numerical models to calculate an effective hydraulic conductivity (or transmissivity) and then to compare the result with the value expected from theory. The approach is to generate a hydraulic conductivity field and then calculate flow rates through the system under an imposed head gradient. Most of these experiments use a hydraulic conductivity field generated by a random-field generator using a Gaussian model. Multiple versions of the hydraulic conductivity field may be produced using Monte Carlo simulations, or a single realization may be used."*

At least one researcher used a physical experiment, a heterogeneously packed sand tank with one realization, for testing equivalent hydraulic conductivity (K_{eq}) and compared the result to theoretically expected values (Danquigny et al 2004).

The result of a pumping test is an upscaled T_{eq} that is in reality comprised of smaller-scale vertical and horizontal heterogeneity. Given that the pumping test is fundamental in hydrogeology, the meaning of T_{eq} obtained from pumping tests has been explored in the literature for heterogeneous conditions (e.g., Meier et al. 1998, and others). It is noted that the validity of treating T_{eff} simply as K_{eff} multiplied by saturated thickness appears to be an open question in some cases (Tartakovsky et al. 2000; Dagan et al. 2009), but for this work it is assumed the point-scale or model grid-scale T values have been determined (upscaled) by

pumping tests and therefore are valid local T_{eq} estimates. Numerous investigations in effective permeability have worked on the 2D flow case and have treated T as a large-scale, spatially correlated random variable as is done in this work (Dagan et al. 2009; Sanchez-Vila et al. 2006).

2.2.8 Summary Points from Effective Permeability Literature Review

The following summary is made from the preceding literature review:

- The geometric mean of local-scale transmissivity has been frequently shown to approximate larger-scale system behavior in heterogeneous isotropic fields, away from flow boundaries, in steady-state mean-parallel flow (e.g., Gutjahr 1978; Dagan 1979; Wen and Gomez-Hernandez 1996; Sanchez-Vila et al. 2006).
- Drawdown near wells may be best approximated by mean-T values lower than the geometric mean, but, with increasing distance from the well in isotropic unbounded conditions, the apparent transmissivity has been reported to approach the geometric mean, and to show non-monotonic changes with increasing distance from the well in a bounded domain (Sanchez-Vila 1997; Riva et al. 2001)
- Based on early stochastic research, it had been proposed that there may be no comprehensively applicable T_{eff} in transient cases since effective parameters were reported to vary with time. In several cases, T_{eff} was reported to start near the arithmetic mean and then approach a lower T_{eff} value expected for steady-state (e.g., Dagan 1982; El-Kadi and Brutsaert 1985).
- Estimates of K_{eff} and T_{eff} are specific to given boundary conditions, flow regime (i.e., convergent or mean-parallel), and dimensionality of the domain (Sanchez-Vila et al. 2006). Even within a given flow regime and boundary scenario, apparent transmissivity has been reported to vary with location between a pumping well and a domain boundary (Sanchez-Vila 1997; Riva et al. 2001).

Estimating stream depletion rates during the transient phase of the stream-aquifer pumping response is the interest of this dissertation since, at later time, the depletion rate approaches the pumping rate. Spatially, it is a relatively large portion of the aquifer that

influences the transient response. The stream depletion flow regime is a mix of radial and near-uniform flow (mean-parallel) conditions, and a stream boundary and pumping well are inherent problem conditions. The literature suggests effective transmissivity is potentially a function of all these conditions.

As stated by Sanchez-Vila et al. (2006), the concept of representative hydraulic conductivities, including the existence and actual values of these parameters, "depend on the flow regime, boundary conditions, and assumptions regarding the underlying structure of the spatial random function hydraulic. This has led to a large variety of methods and results."

Of the published theoretical and experimental evaluations of effective hydraulic conductivity and transmissivity, the case of transient stream depletion induced by groundwater pumping in an adjacent aquifer, or conditions directly analogous to that case, have not been addressed.

CHAPTER 3 METHODS

3.1 Scale and Behavior of Interest

This research considers the case of transient stream depletion caused by initiating a point stress (e.g., pumping well or recharge facility) in a bounded, heterogeneous aquifer that is hydraulically connected to a river. Transient stream depletion refers to the change in the total rate of depletion to a river over time. The depletion rate also varies spatially along the river, but in the realm of conjunctive use and stream-flow management, one impact of interest is the total reduction in stream flow, i.e., integration of the spatially variable depletion rate over the length of the river.

The length of the impacted river reach is proportional to, and several times longer than, the distance between the well and the river (Glover 1977). Stream depletion is, therefore, a relatively large-scale phenomenon. For wells far from the river, it may be regional in scale.

Transmissivity (T) typically exhibits spatial correlation at scales of hundreds to thousands of feet (Rubin 2003; Dagan et al. 2009). The correlation scale relative to the scale of the flow problem is relevant in stochastic work as it affects ergodic behavior, i.e., how well deterministic spatial averaging approaches probabilistic averages as the flow area or averaging area increases (Deutsch and Journel 1998). For this reason, absolute problem dimensions are relevant in this work and are described in the following section.

3.2 Key Concepts in Stream Depletion Mechanics

A contour map of simulated groundwater drawdown around a pumping well, in an aquifer bounded by a river and an impermeable boundary, is shown in **Figure 3-1**. This illustrates a pumping well's cone of depression reaching and interacting with the river and

aquifer boundaries. This could equally be contours of groundwater mounding around a recharge facility. Examples of drawdown patterns in a homogeneous aquifer (top) and a heterogeneous aquifer (bottom) are included.

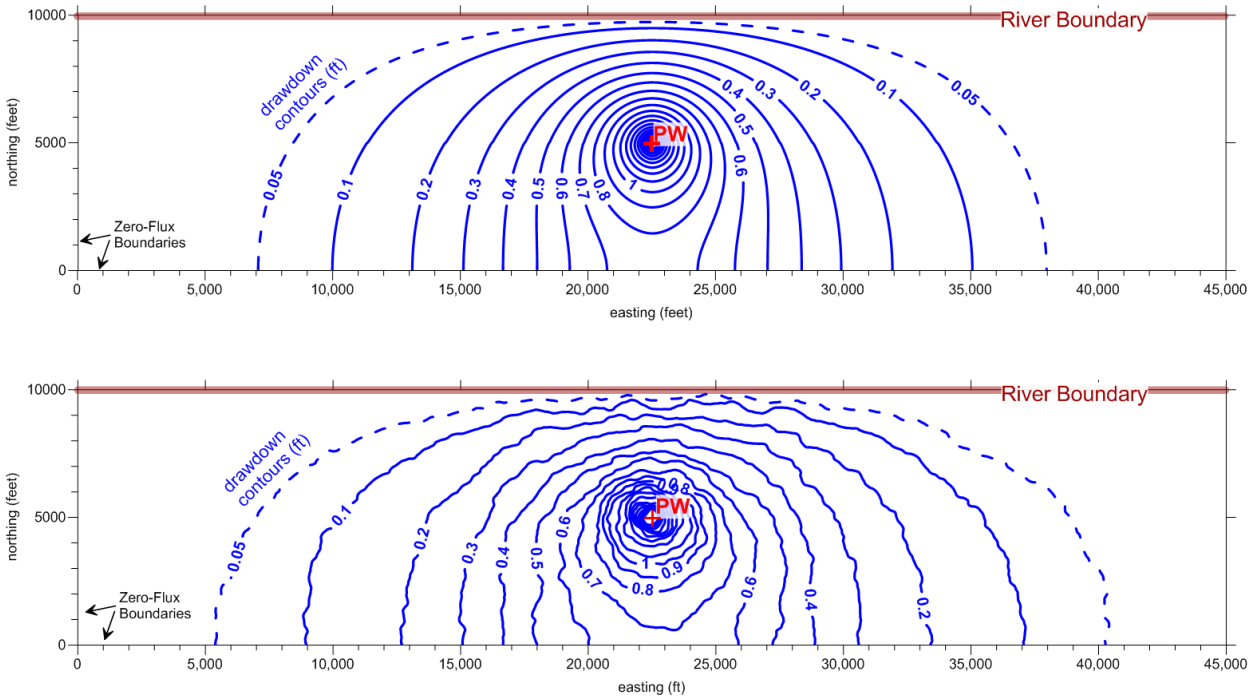

Figure 3-1. Contour map of drawdown around a pumping well illustrating interaction with a river boundary. Included are examples of drawdown patterns in homogenous (top) and heterogeneous (bottom) aquifers.

Pumping creates a perturbation in the aquifer (the cone of depression) which expands radially from the well with time. Interaction between the perturbation and the river creates a change in the rate of water exchange between the river and the aquifer. The total stream depletion rate increases, and a longer river reach is affected, as time increases (Glover 1977). The perturbation expands until a new equilibrium with the river is reached.

The change in the preexisting river-aquifer exchange rate occurs due to the groundwater level and gradient *changes* created along the river by the well. Stream depletion is a reduction in

the streamflow from that which would have been present absent the groundwater pumping. For example, streamflow depletion occurs if there is a reduction in the rate of groundwater discharge to the river, even if the river is still gaining groundwater discharge. Depletion can also be an increased loss rate in a river that was already losing water to the aquifer.

Since the behavior of interest is the change in exchange rate and not the absolute exchange rate, and since the governing differential equations are linear, this permits superposition of depletion patterns in time and space, such as for multiple wells and intermittent pumping schedules. It permits assessing depletion location and rates generally without regard for the prior regional gradient and the prior gain/loss patterns. Gradient changes caused by pumping can be superimposed on a prior regional gradient, or, as in the simulations herein, they can be imposed on an initially flat water table that has zero river-aquifer exchange as the pre-pumping baseline condition.

3.3 Heterogeneous Behavior of Interest

As a cone of depression expands around a pumping well and reaches a river (Figure 3-1), the portion of the river closest to the well is first affected. Over time, the loss rate at that point increases and an increasingly wider reach of the river is also affected. This is illustrated in **Figure 3-2** and **Figure 3-3** which plot local (per unit length) stream depletion rates along the river for the model simulations depicted in Figure 3-1.

In these examples, the river boundary is 45,000 ft long, the aquifer width between the river boundary and a parallel impermeable aquifer boundary is 10,000 ft, and the pumping well is located in the middle at 5,000 ft from the river (Figure 3-1). Model output is plotted at two different simulation times. Time is normalized by the characteristic response time, $t' = a^2S/T_G$, where a is the distance between the well and the river, T_G is the geometric mean of the aquifer

transmissivity, and S is the storage coefficient. For this example, T_G =30,000 ft^2/day, S = 0.2, and a = 5000 ft, for a^2S/T = 167 days.

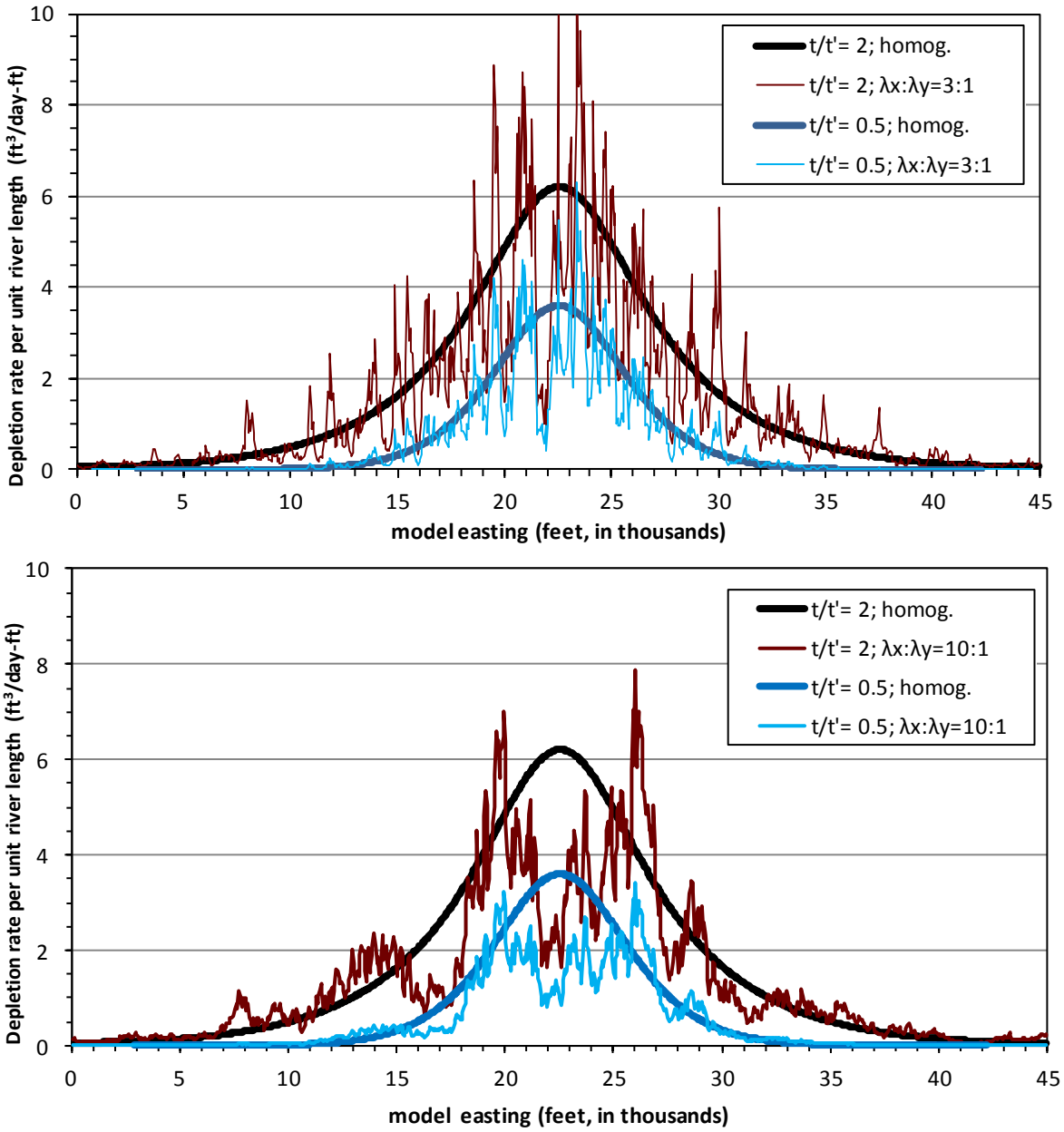

Figure 3-2. Spatial distribution of stream depletion along river boundary. Curves are shown for two normalized points in time (t/t' = 0.5 and 2.0) for both homogeneous and heterogeneous conditions. The heterogeneous cases are from single realizations with geometric anisotropy (λ_x:λ_y) equal to 3:1 (top set) and 10:1 (bottom set).

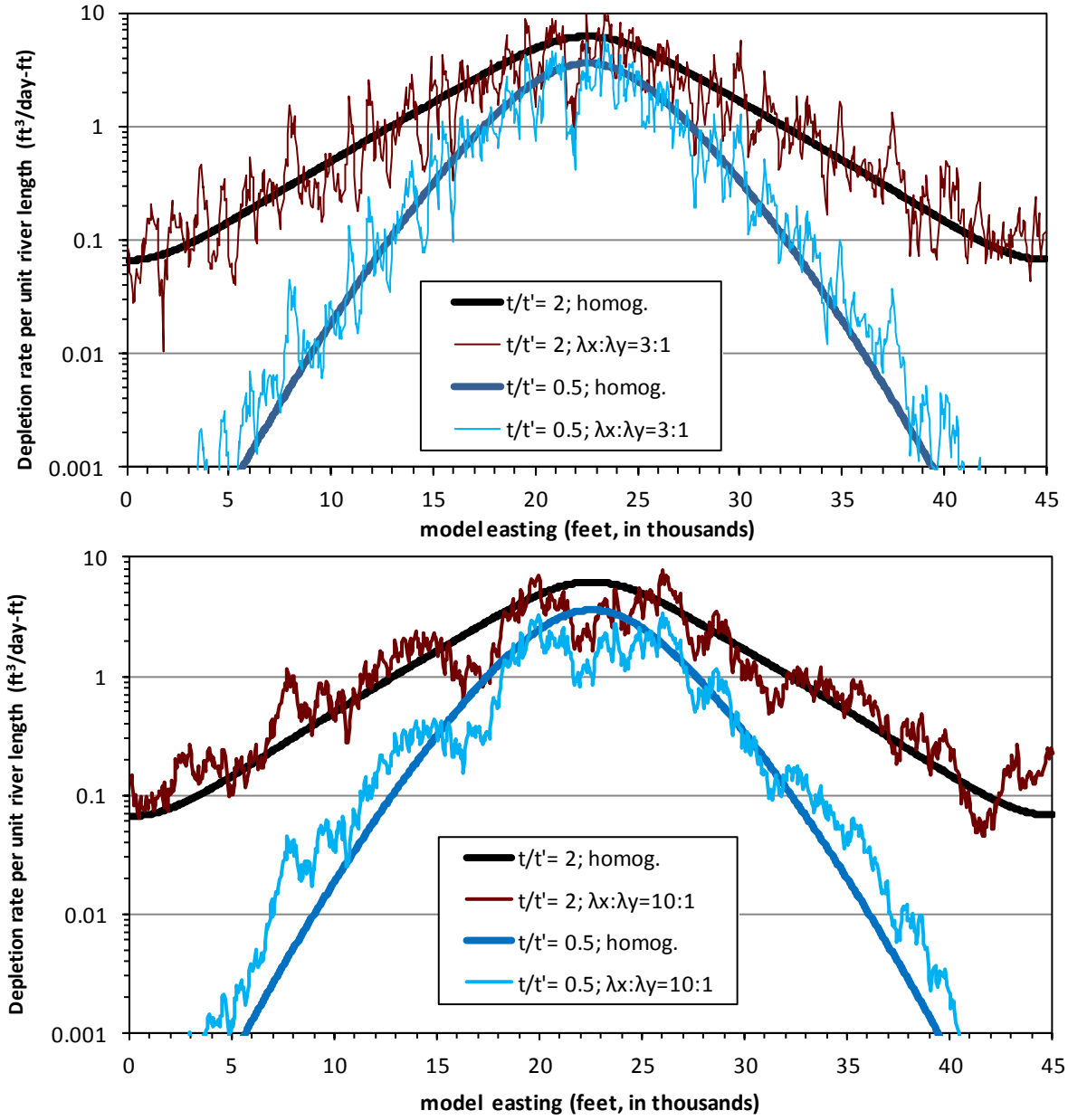

Figure 3-3. Same as previous figure, semi-log scale.

The smooth curves in Figures 3-2 and 3-3 are for a homogenous aquifer. Similar plots for homogeneous conditions were constructed by Glover (1977). Curves from two heterogeneous cases are also shown for comparison: one with statistical geometric anisotropy equal to 3:1 (top plot) and the other at 10:1 (bottom plot). Geometric anisotropy is the ratio of

transmissivity spatial correlation scales in the two principal axis directions. This is defined further in the following section. The principal direction of transmissivity correlation in these simulations was aligned with the x-coordinate direction, thus parallel to the river. Transmissivity of the homogenous case was set equal to T_G of the heterogeneous case for these simulations.

The area under a given curve in these figures, i.e., the integration of the spatially variable depletion rate along the length of the river, gives the total river depletion rate at that time. That total rate is a single point in time on the depletion rate response curve. Several characteristic depletion rate response curves are presented by Jenkins (1968) and Miller et al. (2007).

Note that the well was located 5,000 ft from the river in the simulations used to produce Figures 3-2 and 3-3 and that depletions occur over a reach with a practical length up to about 35,000 ft. The term practical length refers to depletion being non-zero mathematically even at great distances (Figure 3-3), but being small as a practical matter at a finite distance. This is analogous to the effective radius of a pumping well being finite when defined by a small drawdown threshold (McWhorter and Sunada 1977).

The small-scale spatial variability of depletion rates is clearly high for the heterogeneous case (Figures 3-2 and 3-3). These local variations are, however, aggregated into the total depletion rate. The spatial variability is not as important in this work as the question of how heterogeneity influences the growth of the total depletion rate curve over time. It is, therefore, meaningful to work with spatially averaged aquifer parameters even for a spatially heterogeneous depletion rate if those averaged parameters effectively predict timing of total depletions. To identify upscaled (spatially averaged) transmissivities that effectively

approximate the timing of total stream depletion rates in heterogeneous aquifers is the goal of this work.

A few interesting behaviors can be seen in these examples (Figure 3-2 and Figure 3-3). The case with higher spatial correlation (10:1) in aquifer properties has higher spatial correlation in the local depletion rates. One can also see in the semi-log plots (Figure 3-3) that the 10:1 heterogeneous curve affects a longer river reach than the homogeneous case and the 3:1 case. It is also the case, although not clear visually in these examples, that the area under each heterogeneous curve is less than the corresponding homogeneous case and this difference is greater for the 10:1 case than the 3:1 case. As will be shown in Chapter 4, when the direction of maximum spatial correlation of a heterogeneous transmissivity field is oriented parallel to the river, the depletion rate is delayed relative to the homogeneous case even with equal mean transmissivities of the two cases.

Note in Figure 3-1 that the drawdown pattern shown for the heterogeneous case is different from the homogeneous case in local detail, but the macroscale drawdown pattern is roughly similar. In contrast, the spatial distribution of groundwater flux at the river is very different between the homogeneous and heterogeneous cases (Figure 3-2). Similarly, it can be shown that groundwater flux in the aquifer varies smoothly in the homogeneous case and varies sharply in space in the heterogeneous case. That spatial variability in flux can have important consequences in water quality transport (i.e., preferential flowpaths) and other applications were averaging may not be valid. Still, the fact that the drawdown pattern in the heterogeneous case is distorted locally more than globally supports the validity of using spatially averaged parameters in the homogenous Glover equation (Glover and Balmer 1954; Glover 1977) to approximate

macroscale drawdown patterns across a heterogeneous aquifer. The degree of validity depends on the scale of heterogeneity (Section 3.5.2).

3.4 Experimental Design

3.4.1 Generation of Heterogeneous Fields

Isaaks and Srivastava (1989) described geostatistics as "a way of describing the spatial continuity that is an essential feature of many natural phenomena." That description implicitly acknowledges heterogeneity as a ubiquitous condition, and emphasizes the spatial continuity that is often identifiable in such systems. Geostatistics provides a way to characterize, quantify, and model that heterogeneity (Anderson 1997). Geostatistical modeling supports the stochastic approach to understanding groundwater behavior, in which variability and uncertainty are accounted for by specifying an inferred or assumed mean and variance of parameters (Anderson 1997). Geostatistical simulation was used in this work to create a variety of synthetic heterogeneous fields in which stream depletion behavior was examined.

The sequential Gaussian simulator SGSIM (version 3.001) of the geostatistical library GSLIB (Deutsch and Journal 1998) was used to generate multiple realizations of random, unconditional, spatially correlated, stationary, log-normal hydraulic conductivity (K) fields. SGSIM generates the field values from a standard normal distribution (unit variance and zero mean) which are then scaled to have the desired mean and variance of log-transformed K data. The log K field (natural log) is then back transformed for use in groundwater flow simulations.

One hundred K-field realizations, each generated from the same specified geotstatistical model, were created for most of the ensembles presented herein. The realizations in a given ensemble are randomly different locally but are equally probable and exhibit approximately the same geostatistical characteristics globally (mean, variance, and spatial correlation structure).

The number of realizations created per ensemble was occasionally adjusted in the range of 30 to 200 depending on the effect of the given geostatistical parameters on variance of the ensemble results. For example, simulations of fields having larger spatial correlation ranges lead to greater variance in ensemble results, which was handled by increasing the number of realizations.

The SGSIM program includes the standard variogram models: Gaussian, spherical, and exponential (Deutsch and Journal 1998). All three are common, with the spherical model thought to be the most commonly used (Isaaks and Srivastava 1989). These three were tested for potential influence on the methods and simulation results (Section 4.1). Except where noted differently, the spherical variogram model with variance of the log-transformed field equal to 1.0 ($\sigma^2_Y = 1.0$) and zero nugget effect is used herein as the default and reference case.

An example variogram, constructed from one *K*-field realization, is provided in **Figure 3-4** with a comparison to the specified spherical model. A variogram is defined as one-half the average squared difference between paired data values (Isaaks and Srivastava 1989), and is a measure and model of spatial variability as a function of distance (the lag distance) between the paired data points.

Note: The term *semivariogram* is frequently viewed as more precise language for the quantity that was defined above and referred to by Isaaks and Srivastava as a variogram, but this distinction is inconsistent in the literature. There has been some debate about the actual mathematical meaning and best semantics in practice, with Bachmaier and Backes (2008, 2011) arguing the term semivariogram should not be used. In general, the terms are used synonymously, with variogram commonly chosen for general usage and for conciseness (e.g., Gringarten and Deutsch 2001; Remy et al. 2009). That generalized usage is adopted herein as well.

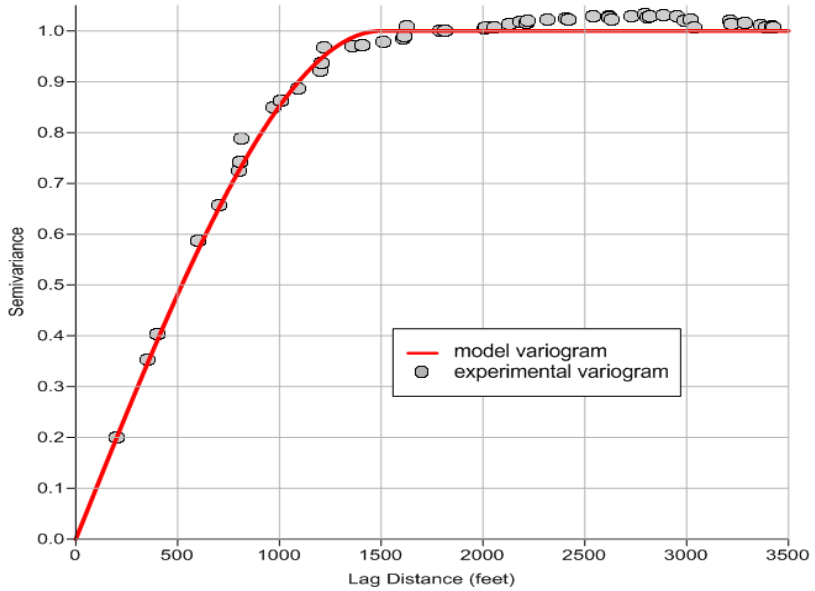

Figure 3-4 Example variogram constructed from one SGSIM-generated *K*-field realization.

At large enough separation distance, the paired data become uncorrelated and the variogram reaches a plateau called the *sill*. The sill is equal to the population variance (Barnes 1991). The distance at which the sill is reached is called the *range*, beyond which there is no spatial correlation (Isaaks and Srivastava 1989; Gringarten and Deutsch 2001). For the exponential and Gaussian models which approach the sill asymptotically, an effective range is defined as the distance at which the variogram reaches 95% of the sill (Deutsch and Journal 1998). The spherical variogram in Figure 3-4 has a range of 1,500 ft. The range and effective range are often referred to informally as the correlation range (e.g., Gringarten and Deutsch 2001). That language is also used herein as it adds clarity in certain contexts.

A *nugget effect* refers to having a non-zero variogram value at zero or near-zero lag distance. This discontinuity arises from measurement error and/or small-scale variability below

the resolution of the sampling density (Isaaks and Srivastava 1989; Gringarten and Deutsch 2001). This is modeled with a non-zero intercept at zero distance in a model variogram.

The following parameters used in the geostatistical modeling of heterogeneity were examined as variables affecting stream depletion timing: (i) geometric anisotropy, (ii) the absolute correlation range relative to the problem scale, (iii) degree of heterogeneity (as measured by σ^2_Y), (iv) orientation of principal directions with respect to the river, and (v) nugget effects in the variogram model. Additional structural forms of heterogeneity are discussed in Chapter 5.

In the course of methods development, geometric anisotropy was identified as a primary variable influencing stream depletion timing. A heterogeneous field exhibits geometric anisotropy when the correlation range changes with direction while the sill remains constant (Isaaks and Svrivastava 1989). The degree of geometric anisotropy is commonly expressed as a ratio of the correlation scale (λ) in the principal directions (λ_{max}:λ_{min}, or λ_x:λ_y when the principal directions are aligned with the coordinate system).

The SGSIM domain was defined to be 10,000 x 45,000 ft and discretized into 200 rows and 900 columns of block-centered grid nodes, thus a uniform node spacing of 50 x 50 ft. This grid was defined to match the flow model grid presented in the next section. Three example heterogeneous fields generated with SGSIM are shown in **Figure 3-5**. Other example fields, for various spatial correlation models, are included in Appendix A. The three fields shown in Figure 3-5 used a spherical variogram model with $\sigma^2_Y = 1.0$ and with the direction of maximum correlation oriented with the grid rows. The minimum correlation range was 500 ft (10 rows) in all cases. The maximum correlation ranges were set at 500 ft, 1500 ft, and 5000 ft (10, 30, and

100 columns, respectively), for geometric anisotropy ratios (λ_x:λ_y) equal to 1:1, 3:1, and 10:1 (top, middle, and bottom images, respectively, in Figure 3-5).

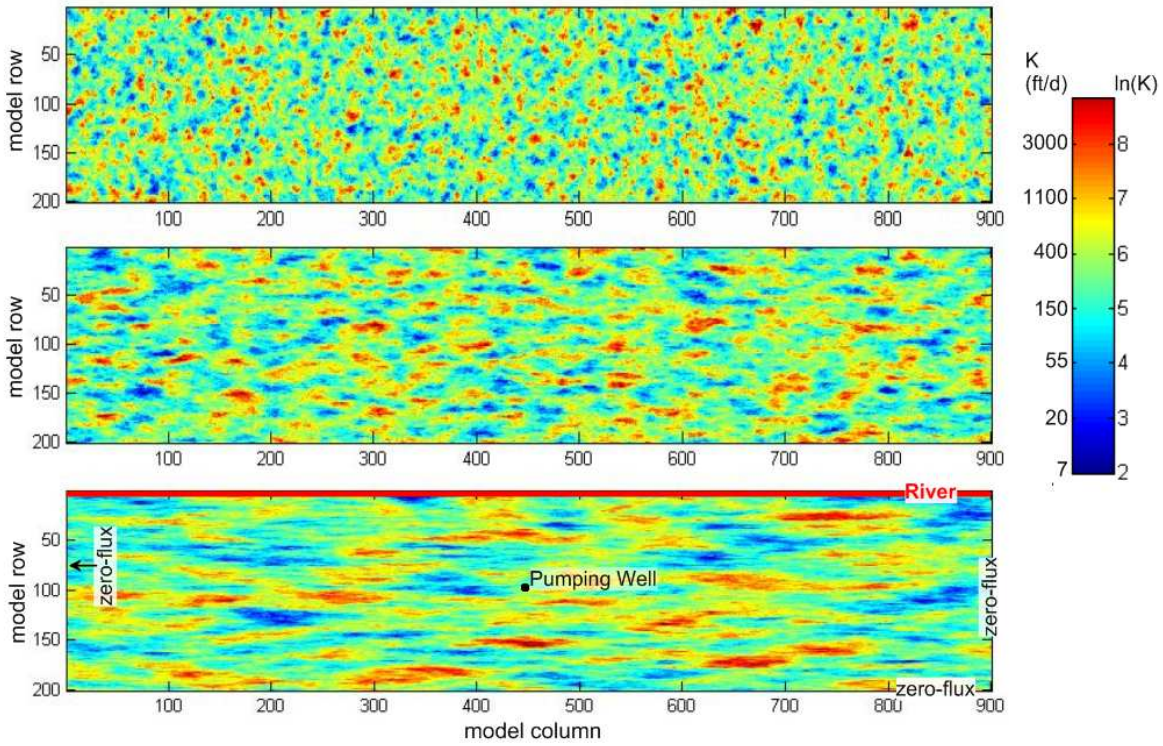

Figure 3-5. Example realizations generated with SGSIM and used in the groundwater flow simulations. Shown are *K* fields having geometric anisotropy ratios of 1:1 (top), 3:1 (middle), and 10:1 (bottom). Bottom image notes location of the flow model boundaries.

3.4.2 Flow Model Design

Each SGSIM-generated *K*-field realization was used to populate a transient groundwater flow simulation performed with MODFLOW-2005 (Harbaugh 2005). The flow model grid was constructed with 200 rows and 900 columns with uniform nodal spacing of 50 length units. In this case, the nodal spacing was 50 feet giving a 45,000 x 10,000 ft model domain. A single model layer (thus 2D) was used with a uniform aquifer saturated thickness set at 100 ft and specified to be constant. The storage coefficient (*S*) was also uniform and set equal to 0.2. The

orientation of the model domain with respect to the river and aquifer boundaries is shown in Figure 3-1 and in the bottom image of Figure 3-5.

Model Row 1 was set as a constant-head boundary representing an idealized river boundary. A zero-flux boundary condition was used for the other three model domain boundaries. The zero-flux boundary condition in Row 200 represents an impermeable aquifer boundary oriented parallel to the river, such as the edge of a valley-fill alluvial aquifer.

A pumping well was placed in the middle of the model domain at 5,000 ft from the river in most simulations. The well was represented by a specified-flux boundary condition in one model cell, Row 101 and Column 450. At this position relative to the two hydrogeologic boundaries, approximately the lower half ($q/Q < 0.5$) of the stream depletion response curve is the same as an unbounded aquifer response (Miller et al. 2007). The second half is influenced by both boundaries. Other well positions were tested (e.g., 2,500 and 7,500 ft from the river) which resulted in earlier and later onset of the bounded-aquifer type response curve. In those sensitivity tests, the well position and consequent variations in boundary influence did not have an apparent effect on the methods and primary results. The primary results are the ensemble-mean T_{eff} values and trends that are presented in Chapter 4.

Simulating a wider aquifer—width referring to the distance, W, between the river and the parallel impermeable aquifer boundary—at the same grid resolution would increase simulation CPU time which was a constraint for running several thousand simulations. Simulating a narrower aquifer was not optimal since it would compress the subject transient period and would decrease the model domain size relative to the range of correlation scales to be tested, which is a constraint on minimum domain size. In either case, the absolute model dimensions and aquifer

parameters are arbitrary. Stream depletion response curves are scalable by an aquifer's characteristic transient response time (a^2S/T) and the relative well position (a/W).

The zero-flux boundary condition in Columns 1 and 900 (oriented perpendicular to the river) were placed by design to be far enough away from the pumping effects to not significantly influence results in most simulations. These lateral model distances were based on the length of the river reach expected to be affected during the transient phase. The affected river length for homogeneous isotropic conditions is approximately 10 times the distance between the well and the river at near-steady conditions, and less during the transient phase (Glover 1977), with most river depletions occurring over a shorter reach nearest the well (Figure 3-2).

The initial condition in all groundwater flow simulations was static, meaning the head was constant throughout the domain and equal to the river boundary head, thus a baseline condition of no groundwater flow and no river-aquifer exchange (Anderson and Woessner 2002). The well began pumping at time zero and pumping remained constant. The model dimensions and aquifer properties were selected so that in most cases the river depletion rate (q) was greater than 90% of the pumping rate (Q) by the end of the simulation. Note that all realizations and all ensemble cases would asymptotically approach $q/Q = 1.0$ (the steady-state) in longer simulations, so it is the earlier, more-transient portions of the response curve that is of interest and in which the most variability between different realizations is seen.

The saturated thickness in all simulations was uniform in space and specified to be constant in time. Holding saturated thickness constant is a common modeling technique used to improve model stability and reduce simulation time (Faunt et al. 2011; Sheets et al. 2015). It was advisable in this stochastic work since K in the pumped cell and surrounding cells would vary widely, leading to dry up of the pumped cell and consequent model simulation failure in

certain realizations. Reducing the computation time was also a consideration for those cases with high variance which required more time for the solver to converge on a head solution. With the constant saturated thickness simplification, the simulation results are applicable to confined aquifers and to the unconfined case when drawdown is small relative to the initial saturated thickness (Faunt et al. 2011; Sheets et al. 2015).

The geostatistical parameters used to model and characterize heterogeneity were identical whether considering *T* or *K* since saturated thickness was uniform and constant. The population variance, spatial correlation functions, and ratios between the harmonic, geometric, and arithmetic means for a given *T* or *K* field were the same. One can therefore refer to *K* or *T* interchangeably in this work when discussing geostatistical parameters or effective permeability results (i.e., effective power averages). From this point forward, *T* is discussed rather than *K* since the grid values used in the flow model are assumed conceptually to be local *T* estimates at the scale of a pumping test (Section 2.2.7). Discussing *T* is also appropriate given that the flow simulations are two-dimensional and with correlation scales representative of *T*. Correlation scales simulated in this work ranged from 500 ft to 40,000 ft. Rubin (1993) and Anderson (1997) include Gelhar's (1993) table of example correlation scales for *T* and *K*.

For context, it is noted that the geometric means of the simulated *K* fields used in this work ranged from 100 to 1,000 ft/day, with most cases having K_G = 300 ft/day, thus T_G = 30,000 ft^2/day and *t*' = 167 days. Other parameters were defined previously. The flow simulation period was 1250 days, for a simulation time of *t*/*t*' = 7.5 in most cases.

Though arbitrary, parameter units and model dimensions must be assigned. They were chosen to represent an aquifer with transmissivity suitable for supplying high-capacity wells and with transient depletion timing that would be of the greatest interest from a technical and

conjunctive management perspective. For example, much faster depletion timing, such as for wells much closer to the river and in narrower aquifers, could be nearly trivial due to that case having a short and steep response curve that tracks pumping schedules more directly. And for much slower depletions, such as for wells far from the river in a wide aquifer, intra-annual depletion-rate fluctuations from seasonal irrigation are smoothed out as depletions are delayed and spread over multiple years. In that latter case, the slow build up in the depletion rate can be of interest, but it eventually reaches a relatively steady depletion rate equal to the the long-term *annualized* extraction rate, even for seasonal pumping (e.g., Bredehoeft 2011; Kendy and Bredehoeft 2006). Having a characteristic response time on the order of 167 days, as used in most simulations in this work, represents a case potentially with both interannual and intra-annual fluctuations in depletion rates due to seasonal irrigation pumping.

3.4.3 Example Flow Model Output

Simulating stream depletion response curves under a variety of heterogeneity types was a fundamental step. Response curve characterization (Section 3.5) was necessarily automated due to the thousands of simulations involved, but the shape of individual realizations and the shape and variability of curve ensembles were directly evaluated as part of methods development. The curves were periodically inspected again during ensemble production runs. Since the curves are fundamental, yet ultimately an intermediate step in the analysis, this section is provided to document the appearance of that model output. Differences between the example response curve ensembles are highlighted.

Example flow model output is shown in **Figure 3-6** where stream depletion response (q/Q) is plotted as a function of non-dimensional time ($t/t' = tT/a^2S$). Output is plotted for three ensembles each comprising 100 realizations. These simulated T fields had equal geometric

mean transmissivity (T_G) as well as equal variance, which together means they also have equal arithmetic and harmonic means, T_A and T_H (Section 2.2.4). The direction of maximum spatial correlation was oriented parallel to the river with correlation lengths specified at 500, 1,500, and 5,000 feet. The minor-direction correlation length was set at 500 feet (10 model nodes) in all three cases, giving geometric anisotropy ratios (λ_x:λ_y) of 1:1, 3:1 and 10:1 (top, middle, and bottom plots, respectively). Fields with these correlation scales and anisotropy ratios were illustrated in Figure 3-5.

Figure 3-7 compares the timing and distribution of each ensemble shown in Figure 3-6 through histograms of the q/Q depletion rates constructed for the four time points, $t/t' = 1, 2, 3$ and 5. It is apparent in these figures that the depletion timing was delayed, and the ensemble response variability increased, as geometric anisotropy and correlation scale increased. Note again that the arithmetic, geometric, and harmonic means for these fields were equal, yet the depletion timing was different. Characterizing these differences in ensemble behavior and identifying potential relationships between them and heterogeneity structure is a goal of this work.

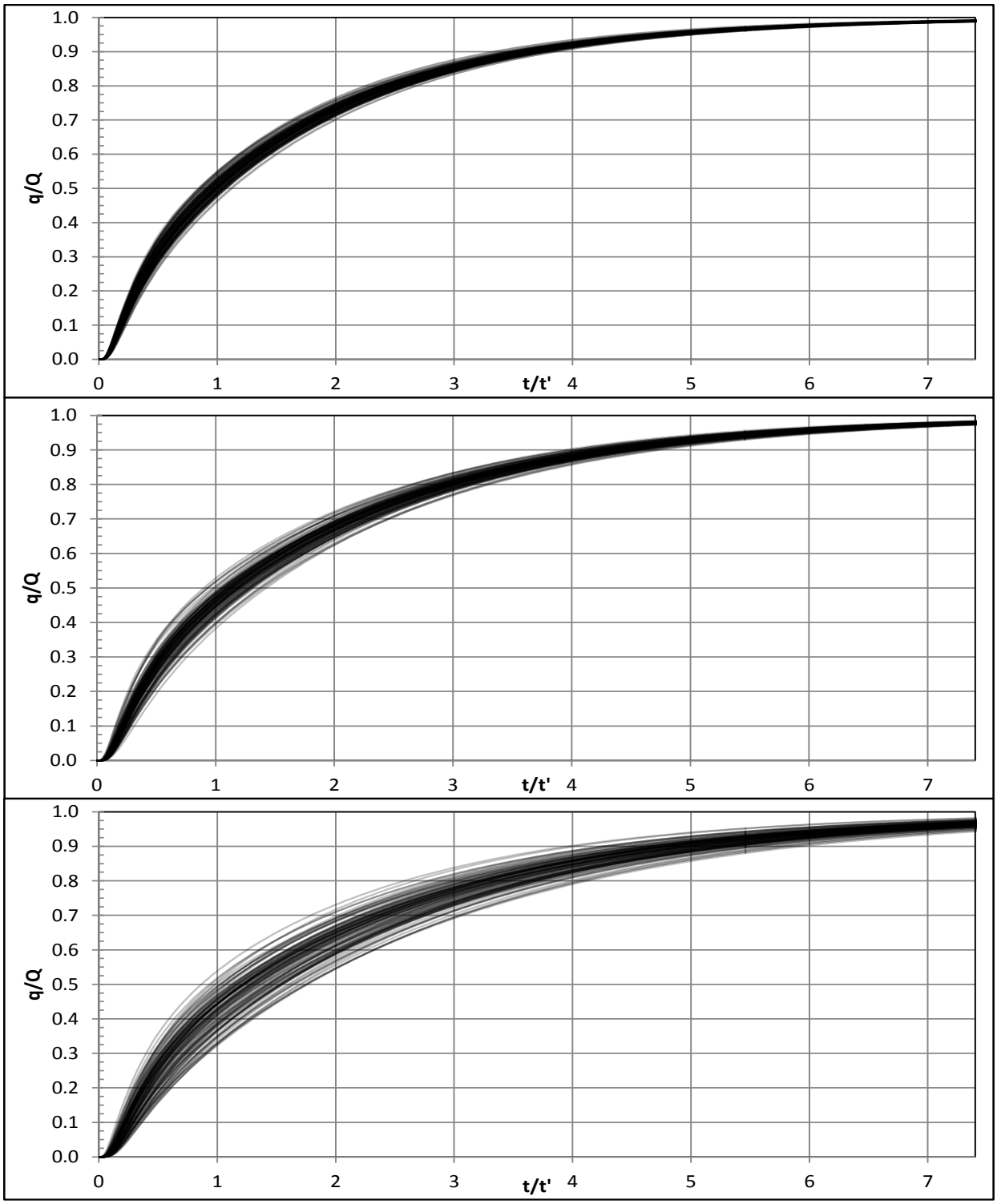

Figure 3-6. Stream depletion rate (q/Q) as a function of time (t/t') for three ensemble simulations of T fields that have the same T_G, T_A, and T_H but with different geometric anisotropy ratios equal to 1:1 (top), 3:1 (middle) and 10:1 (bottom).

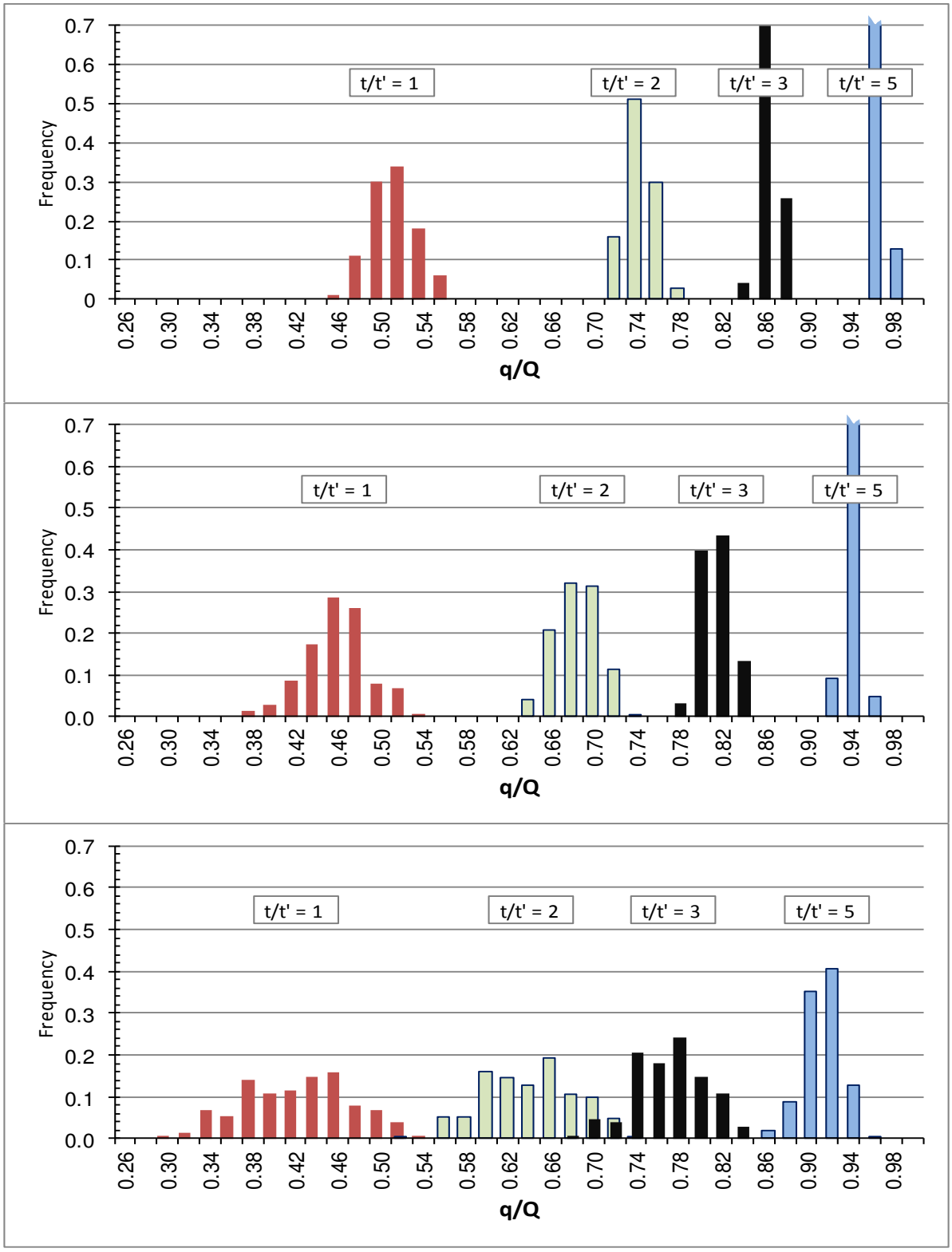

Figure 3-7. Stream depletion rate histograms at time t/t' = 1, 2, 3, and 5 for the three ensembles shown in Figure 3-6, comparing differences in depletion response for T fields with geometric anisotropy ratios equal to 1:1 (top), 3:1 (middle) and 10:1 (bottom).

3.5 Characterizing the Simulated Response Curves

3.5.1 Analytical Solution Fit to Simulated Depletion Response

The simulated stream depletion rate response curve for each heterogeneous realization was fit to the Glover and Balmer (1954) analytical solution for stream depletion from wells as it is applied for the bounded aquifer case (Glover 1977; Knight et al. 2005; Miller et al. 2007). The unbounded Glover solution for depletion rates can be written as the following:

$$q/Q = erfc\left(\frac{a}{\sqrt{4tT/S}}\right) = erfc\left(\sqrt{\frac{a^2 S}{4tT}}\right) \tag{3-1}$$

where, as previously defined, q is the stream depletion rate (volume/time), Q is the pumping rate, a is the distance between the well and the river, T is aquifer transmissivity, S is aquifer storage coefficient or specific yield, and t is time since pumping began.

The simulated response curves in this research were fit with the bounded form of the Glover solution as it was written by Knight et al. (2005):

$$\frac{q}{Q} = erfc\left(\frac{a}{\sqrt{4tT/S}}\right) + \sum_{n=1}^{i}\left[(-1)^{n+1}\cdot\left(erfc\left(\frac{2nW-a}{\sqrt{4tT/S}}\right) - erfc\left(\frac{2nW+a}{\sqrt{4tT/S}}\right)\right)\right] \tag{3-2}$$

where W is the distance between the river and the impermeable boundary and all other variables are as defined previously. Knight et al. (2005) present the derivation of Equation 3-2, while Glover (1977) and Miller et al. (2007) illustrate the image well pattern it represents.

The unknown variable in Equation 3-2 was T; all other parameters were specified in the numerical model construction. The Glover T that yielded a match to the simulated depletion curve identifies the equivalent transmissivity (T_{eq}) for that realization. The *equivalent* label

indicates T_{eq} it is obtained for a given realization and is not necessarily an expected (effective) value stochastically.

The curve fitting was necessarily automated, but the quality of the fit was periodically verified. For a homogeneous case, changes in T shift the Glover stream depletion response curve forward or backward in time, but keep the same curve shape, while changes in aquifer width (W) reshape the curve (Miller et al. 2007). A decrease in W steepens the curve, mostly in the late-time response. The characteristic curves are a function of a^2S/T and well position relative to the aquifer boundaries (a/W). In developing and testing the methods, and in checks of the ensemble production simulations, the characteristic curves for stream depletion in two-dimensional, homogeneous, bounded aquifers were observed to match, in most cases, the curve shapes from the simulated heterogeneous fields.

Note that T_{eq} is defined by Sanchez-Vila et al. (2006) and others as being determined by spatially averaging over a realization, such as the T_{eq} obtained from the ratio of averaged Darcy velocity to averaged gradient. It is also used to describe an upscaled T obtained from spatially averaging local T values directly. Sanchez-Vila et al. noted the additional category of *interpreted T*, which is T deduced from analyzing an observed response such as drawdown from a pumping test. In the pumping test case, one fits the observed drawdown response to the characteristic curve for drawdown in a homogeneous 2D aquifer. The pumping test, in effect, physically averages smaller-scale T variability. In the work herein, T_{eq} is identified from fitting the simulated observations (the depletion response curve) to the appropriate Glover curve for stream depletion in a homogeneous 2D aquifer. In that respect, the T_{eq} identified herein is comparable to an interpreted transmissivity (Sanchez-Vila et al 2006).

3.5.2 Ergodicity and Statistical Homogeneity

Response curves for realizations occasionally deviated from the characteristic shapes discussed in the previous section. This occurred more often when simulating large correlation ranges, a scenario in which the condition of statistical homogeneity is not well met. Although these shape deviations were generally small, they were investigated during methods development. The purpose of this section is to discuss the nature of those deviations. Understanding their nature was also important in making conclusions about the time-invariant nature of T_{eff} in stream depletion observed in this work.

Statistical homogeneity is related in this case to the concept of ergodicity. When a property averaged over space in one realization coincides with that property averaged over multiple realizations, the process is said to be ergodic (Sanchez-Vila et al. 2006), i.e., the spatial average matches the ensemble average. As described more specifically by Deutsch and Journal (1998), a spatial random function model of T is ergodic for T if the realization statistics tend toward the mean T of the function *as the size of the field increases*. If the T field is large enough, the statistics of the field should match the specified model. The larger the correlation range, the larger the field must be to achieve ergodicity (Deutcsh and Journal 1998).

Ergodic fluctuations refer to discrepancies between realization statistics and specified-model statistics (Goovaerts 1997). These are expected. The only match between the model statistics and realization statistics guaranteed by stochastic simulation theory is an average (expected value) over a large number of realizations (Deutsch and Journal 1998). The discrepancies fluctuate around the specified geostatistical model. The longer the correlation range, the less ergodic are the realizations, and the more realizations are needed to approach the expected value (Deutsch and Journal 1998). These ergodic fluctuations have been viewed as a

desirable reminder of uncertainty, such as when the model statistics themselves are uncertain due to being inferred from sample data (Goovaerts 1997; Deutsch and Journel 1998).

As noted, an SGSIM simulation only guarantees that the assigned correlation model, including mean and variance, is reproduced over the average of multiple realizations and not necessarily in each realization. This can be extended to subareas of the realization domain. Consider a field with dimensions many times larger than the correlation range. This field could be sub-divided into areas that are still ergodic, meaning their statistics match the specified model and match each other. In fact, if the initial field were large enough, the sub-divided areas could be treated as independent realizations themselves. As that field is further divided into smaller areas, however, ergodicity is lost and the statistics of the subareas will differ from the model and from each other. This becomes more pronounced when the subarea of interest is small relative to the correlation range.

Differences between the statistics of small subareas of a given realization are referred to herein as statistical inhomogeneity. These differences are a form of ergodic fluctuations as argued in the preceding paragraph, but the different term is used herein since ergodic fluctuations are typically discussed in the context of differences between realizations (Goovaerts 1997; Deutsch and Journel 1998) and not within a given realization.

Departures from the characteristic response curve shapes were observed for some realizations, particularly when simulating fields having larger correlation ranges. Based on inspection of the *T* field in many of those cases, this was attributed to sub-domain-scale spatial fluctuations in field properties, i.e., statistical inhomogeneity. In those cases, the process of producing spatially correlated but random fields sometimes resulted in mean *T* being distinctly different for a subarea on one side of the well than the other side, for example. That affects the

response curve shape. As the cone of depression expands over time, it is affected by different subareas of the aquifer in shifting proportions. If the subareas have different mean *T* and are large relative to the cone of depression, then the concept of averaging heterogeneity does not hold up well. A mean *T* that is applicable to the entire cone of depression may not exist in that case. When that scenario arose in random simulations, it resulted in the response curve not being well fit by a single T_{eq} value.

The ensemble-mean response curves were still well matched even for those ensembles containing realizations that deviated from the characteristic curve. That was additional evidence for attributing the shape-deviation of certain realizations to statistical inhomogeneity caused by local ergodic fluctuations: Even for conditions that are poorly ergodic such as simulations with large correlation ranges, the ensemble average of those subareas is expected to match the specified-model statistics, given a sufficient number of realizations. The fluctuations are averaged out over the ensemble and the ensemble-mean curve is thus well-described by single-valued average *T*.

3.5.3 Potential Transient Effects in T_{eq}

Some studies have observed effective permeability to vary with time under transient flow conditions, such as decreasing from arithmetic means to geometric means as time increased (Section 2.2.3). Dagan (1982) referred to this as a relaxation time, and noted it was short in 3D cases, but could be long for the 2D case given the large correlation ranges for *T*. Other studies did not report their T_{eq} to be time-dependant (Section 2.2.3). With those mixed reports in the literature in mind, the fit between the characteristic depletion curves for the homogeneous case and the modeled curves for the heterogeneous case was examined closely during methods development.

In short, the simulated depletion curves were well-fit by the characteristic analytical curves using a time-invariant T_{eq}. The T_{eff} relationships as identified in this study do not appear to be time dependant. This conclusion could be made since the curve shape deviations discussed in the previous section were occasional, were not observed in the ensemble response, and, as noted, were attributed to statistical fluctuations in subareas of the random field.

3.5.4 Computing Power-Average Exponents

Power averaging was used to relate each *T*-field realization to the T_{eq} identified from its simulated depletion response curve. The power-mean is an upscaling function producing a range of power-averaged transmissivities, $T(p)$, from a field of local-scale *T* estimates. The task described in this section was to identify the equivalent exponent, p_{eq} , such that $T(p_{eq}) = T_{eq}$. In more qualitative terms, the resulting p_{eq} is an index indicating where T_{eq} lies between the bounding arithmetic and harmonic means.

The power-mean function used herein is:

$$T(p) \quad = \left[\frac{1}{N}\sum_i^N (T_i)^p\right]^{1/p} \tag{3-3}$$

where T_i are the local, point-scale transmissivity values (grid-cell values in the model), and p is the power-mean exponent. The function yields an arithmetic mean for $p = 1$, a harmonic mean for $p = -1$, a geometric mean for p near zero, and yields intermediate averages for non-integer p exponents between +1 and -1 (Gomez-Hernandez and Gorelick 1989; Desbarats 1992b; Wen and Gomez-Hernandez 1996; Ronayne and Gorelick 2006; de Dreuzy 2010).

Example $T(p)$ curves for four different variance levels ($\sigma^2_Y = 0.25$, 0.5, 1.0, and 2.0) are illustrated in **Figure 3-8**. The $T(p)$ values are normalized by T_G, with both $T(p)$ and T_G in real space. For example, for $\sigma^2_Y = 1.0$ and $p = -0.5$, $T(p)/T_G = 0.78$.

Logistically, the $T(p)$ curve and p_{eq} could have been estimated from the statistical moments of the specified geostatistical model of the heterogeneous field, using Equation 2-5 (Ababou and Wood 1990). That approach would have been essentially exact for the ensemble and reasonably accurate for most realizations. For this work, however, a $T(p)$ table was computed directly from each field realization, constructed in p increments of 0.1, between the theoretical -1 and +1 bounds. Cubic-spline interpolation was used to read p_{eq} from the tables (Mathworks 2013).

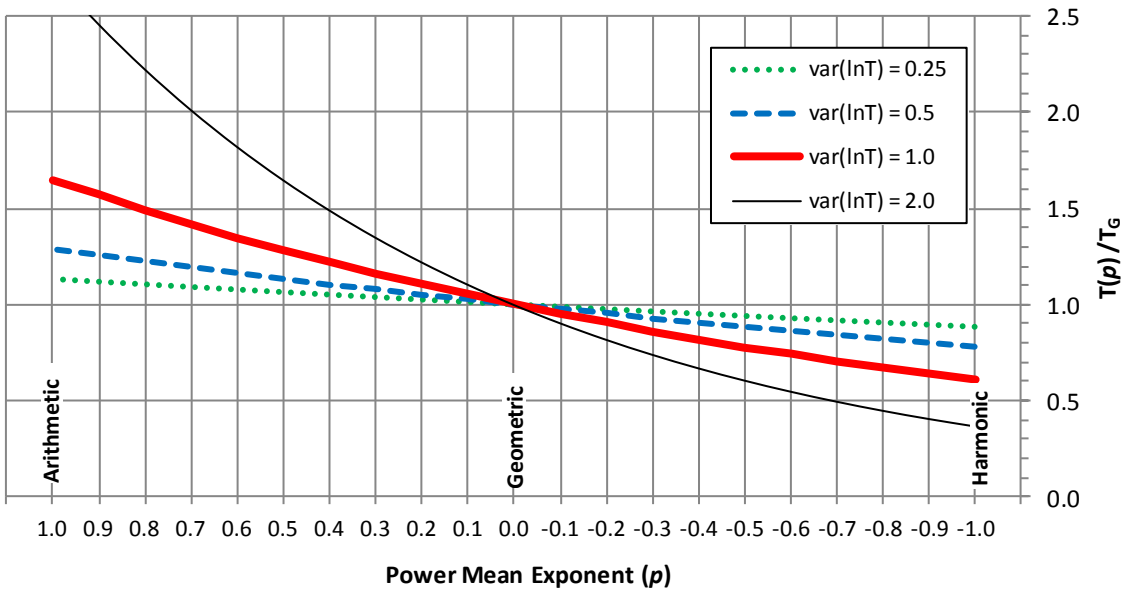

Figure 3-8. Shape of power-average curves.

3.5.5 Ensemble Behavior

An effective parameter, like T_{eff}, is by definition based on averaging behavior over an ensemble of realizations (Sanchez-Vila et al. 2006). Two approaches were tested for producing T_{eff}. In one, the deterministic T_{eq} was identified for each realization's simulated depletion response curve through numerical modeling, as described in Section 3.5.1, and each T_{eq} was

characterized through a power-averaging exponent, p_{eq}, as described in Section 3.5.4. The statistical moments of the p_{eq} ensemble therefore provided T_{eff} through the power-averaging exponent p_{eff}. In the other approach, T_{eff} was determined directly from fitting the ensemble depletion response to the Glover solution and its p_{eff} was determined from the moments of the specified statistical model of T (Equation 2-5, Page 18), thus skipping the quantification of each p_{eq}.

These two approaches were found to produce essentially equal results. First, it was observed that p_{eq} for individual realizations compared favorably to p_{eff} of either method when spatial correlation scales were small compared to domain scales, as is the expectation (Section 3.5.2), i.e., there was low p_{eq} variability when ergodic conditions were well met. More importantly, it was confirmed that mean and variance of the p_{eq} ensembles approached the mean and variance of p_{eff} determined from the ensemble even for the larger correlation scales that yielded wide variability of p_{eq}.

Although the two methods produced the same results, computing p_{eff} only from ensemble response curve moments lost certain information about p_{eq} variability among the realizations. No information was lost statistically—the ensemble moments were determined either way—yet it was informative to record p_{eq} for individual realizations, such as to inspect properties of fields that produced outlier p_{eq} values. For most simulations, therefore, the first approach was used in which a p_{eq} was determined and plotted for reach realization, and the statistical moments of that p_{eq} ensemble provided p_{eff}.

Terminology note: T_{eq} and p_{eq} refer to outcomes for individual realizations and T_{eff} and p_{eff} to the ensemble behavior. The given T and power-average exponent p could be discussed interchangeably at this point. The general convention hereinafter will be to refer to the power-

average exponents when discussing specific quantitative results and plots, and T_{eff} and T_{eq} for more the general or qualitative concepts of effective and equivalent transmissivity.

Example distributions of ensemble p_{eq} results are illustrated in **Figure 3-9**. These correspond to the ensembles shown in Figure 3-7. The curves in Figure 3-6 represent the entire depletion response over time, and the histograms in Figure 3-7 are for responses at four different points in time. Time is not a factor in Figure 3-9, however, since it was observed that each response curve could be well-fit with a single time-invariant p_{eq}.

The p_{eq} values in Figure 3-9 indicate, essentially, how the heterogeneous T field in each realization would be averaged for use in computing its response curve. For example, for the isotropic case (λ_x:λ_y = 1:1), p_{eq} ranged from 0.15 to -0.15 with an ensemble mean near $p = 0$, the geometric mean. The p_{eff} was lower for the two anisotropic cases, with those indicating an expected power-average between the geometric and harmonic means ($-1 < p_{eff} < 0$).

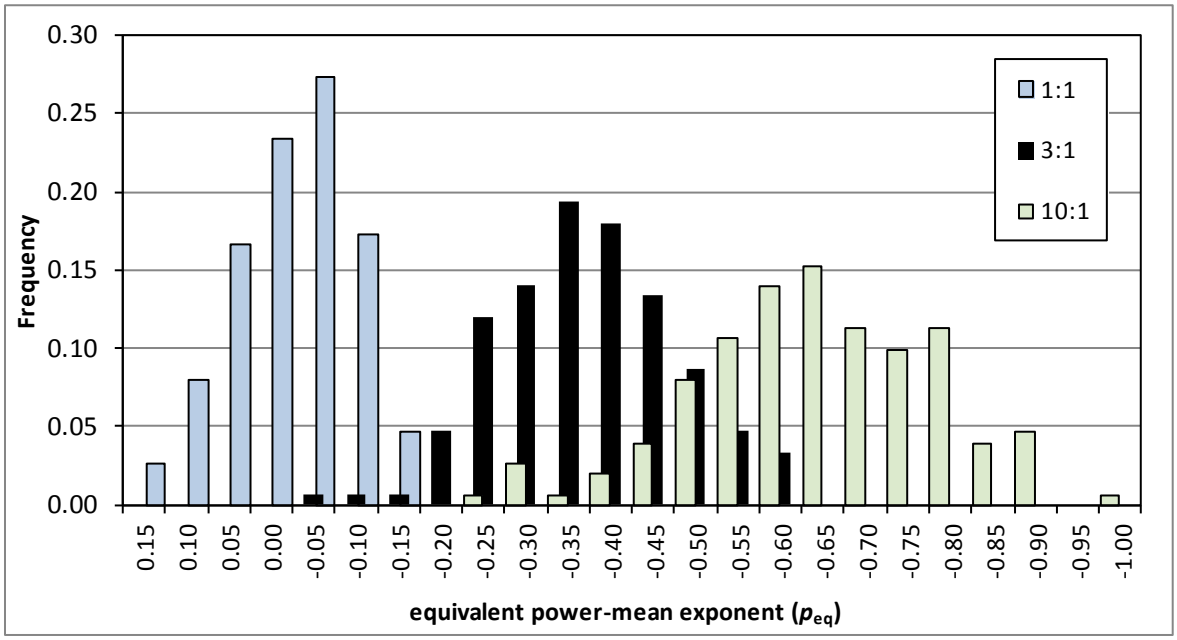

Figure 3-9. Equivalent transmissivity (as power mean exponent, p) for T fields with three different geometric anisotropy ratios (1:1, 3:1, and 10:1)..

3.6 Simulation Logistics

3.6.1 Spatial Weighting of the Area of Influence

In the geostatistical simulation of heterogeneous fields, the statistical moments of domain subareas may not match the specified model exactly; the moments fluctuate around the specified model (Section 3.5.2). In groundwater hydraulics, domain subareas will be influenced unequally by a pumping well. If area of influence is defined by magnitude of drawdown, as in the case of the effective radius of a pumping well, then areas far away from the well have little influence (Figure 3-1). In the case of stream depletion, if area of influence is defined by groundwater flux, then a wide area along the river has greater influence than other portions of the aquifer, and this influence increases with proximity to the well (Figure 3-2).

Different spatial weighting methods were tested in the power-averaging process. This was done out of consideration that certain subareas of the aquifer have greater influence on the stream-aquifer response and that local statistical moments in those areas could fluctuate. The intention was to compute aquifer statistics from those areas most influencing the response. This spatial testing was possible due to the decision to determine a T_{eq} for each realization and to relate T_{eq} to p_{eq} based on power-averaging each field realization directly rather than using the moments of the specified geostatistical model (Section 3.5.4).

For the theoretical case of pure statistical homogeneity and ergodic realizations, the averaging area would not matter since the statistical moments and correlation model would be the same regardless of the size and location of the averaging area. In practice, however, whether in the field or in stochastic simulation, there are spatial differences unless the correlation range is small relative to the area of concern (Deutsch and Journel 1998; Sanchez-Vila et al. 2006).

It was found in methods-testing simulations that the variance of a p_{eq} ensemble was reduced when the $T(p)$ curve used to determine p_{eq} (Section 3.5.4) was computed from a limited region along the river reach where most of the depletion occurs, instead of computing $T(p)$ from the entire domain. This observation supported the reasoning that selective spatial weighting better captured local fluctuations in mean transmissivity in the area of influence.

Ensemble variance was also reduced when the pumping well was placed farther from the river while keeping correlation range and other conditions equal. The area of influence was then larger relative to the correlation range, which results in a smaller degree of fluctuation in that area around the specified statistical model. Similarly, and underscoring the point, ensemble variability was most reduced for steady-state mean-parallel flow simulations (Section 4.2) because the area of influence was then the entire domain with all areas contributing equally to T_{eq} and p_{eq}. The only relevant statistical fluctuations in that case would be from the overall ergodic fluctuations between realizations and not spatial fluctuations.

Energy dissipation analogies were initially tested by assigning weights to the spatial averaging process in proportion to an expected flow velocity at each location (e.g., Molz et al. 2005). Velocity-weighting schemes, however, put nearly all the weight in the high-velocity, convergent-flow zone near the well. This introduced extreme variability in the weighted T-field statistics since the moments in such small subareas do not reflect the statistics of the wider area. Transmissivity is essentially random at that highly local scale. That method introduced apparently random variability in p_{eq}, since the averaging area used to compute p_{eq} was a small part of the aquifer zone influencing the depletion response. This remained the case when testing various upper-limit velocity cut-off thresholds to exclude the high-velocity areas near the well. Proportional velocity weighting was therefore not used.

Positive results, indicated by reduced ensemble variance, were found with a simpler averaging area having uniform weights within it and zero weight outside it. The averaging area is shown in **Figure 3-10**. It was based on covering that portion of the aquifer along the river reach that contributes more than 90% of depletion during the transient period from time zero through $t/t' = 6$. For $t/t' < 1$, nearly 100% of the depletion occurs within the highlighted reach. The zone covers a river reach length equal to five times the well-to-river distance (thus 5*a*). This depletion-reach length criterion was assessed using the Glover (1977) analytical solution for the spatial distribution of depletion along a river for pumping in an isotropic homogeneous aquifer.

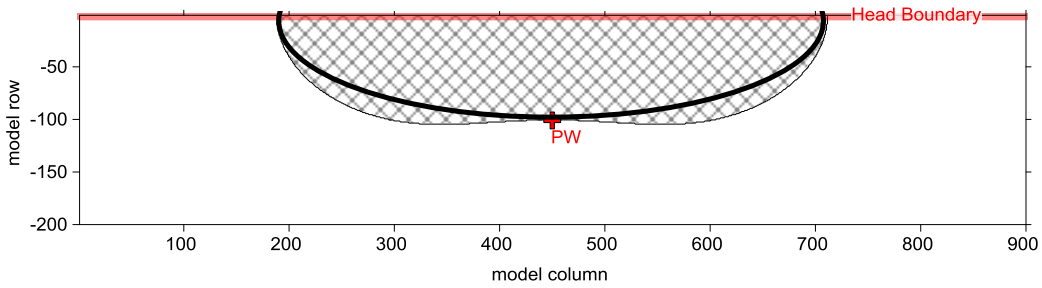

Figure 3-10. Spatial averaging area with its river reach length based on Glover (1977), and its width and shape as suggested by groundwater flux maps (hatched) approximated with a semi-ellipse (heavy line).

As a strict matter, this averaging area would not be unchanging. It would be rescaled for aquifer anisotropy and it would also expand along the river with increasing time (see Glover 1977 or Figure 3-2). However, even at later time, the majority of depletion still occurs within the reach length equal to 5*a*. For isotropic conditions and $t/t' = 10$, more than 85% of the depletions occur within this zone. For a few test realizations of fields with a statistical geometric anisotropy of 10:1 with the principal direction originated parallel to the river—a condition which spreads the depletions over a wider reach and corresponds approximately to a hydraulic

anisotropy of 2:1 (Section 5.5)—approximately 80% of depletion occurred within the 5*a* reach during the transient period up through about $t/t' = 3$.

The 5*a* length of the selected averaging zone along the river has a clear physical and mathematical basis, and is reasonably robust in time and to aquifer anisotropy. A clear basis for the width (perpendicular to the river) and shape of the zone, however, was not identified. An indication of the width of an area of influence was taken from groundwater seepage velocity maps produced during the initial velocity-weighting tests. Those maps indicated an elliptical area of influence, especially in statistically anisotropic fields, that covered the affected river reach length and the pumping well. An example of a velocity-based area is shown as the hatched area in Figure 3-10. A similar shape is obtained by drawing a flow net on the drawdown contour maps provided in Figure 3-1, with flow lines originating within the affected reach length *5a.*

Considering that the affected reach length as identified by Glover (1977) is a function of the well-to-river distance, an ellipse provides a simple definition, even if somewhat arbitrary, for an averaging area. The averaging area in Figure 3-10 is a semi-ellipse with its major axis positioned along the river boundary, its semi-major axis positioned and defined by the distance between the well and the river (a), and its major axis length being five times that of semi-major axis ($5a$).

The precise shape of the averaging zone was not significant in this work. Ensemble variance outcomes were insensitive to changes in the averaging area as long as it was near this scale or slightly larger (Figure 3-10). Additionally, it is stressed that the primary results of this research were not affected by the decision to use a weighting area or by the shape of the weighting area. The ensemble means, the p_{eff} results, were consistent with or without spatial weighting. Indeed, ensemble-mean results shouldn't be affected since the subarea statistical

fluctuations addressed by this spatial weighting are expected to average-out over the ensemble to match the specified geostatistical model, given sufficient realizations.

A caveat to the above is that when the weighting area was very small, such as immediately around the well or weighted highly toward the center of the river reach, then it failed to capture the relevant aquifer area and thereby increased p_{eq} variability. In those cases, it was not clear if running additional realizations would converge on an expected value.

Although it did not affect ensemble-mean behavior, using the selective averaging area did appear to provide a more accurate link between the modeled depletion curve and the statistics of the portion of the aquifer most influencing that curve. It reduced ensemble variance and narrowed the confidence interval around the ensemble means. Again, this was done by taking actual *T* statistics in the area of influence into account, rather than assuming they were consistent with the overall aquifer or the specified geostatistical model.

3.6.2 Interblock Transmissivity in the Finite-Difference Formulation

In the finite-difference method, as used in MODFLOW, the flow between model nodes in response to a head difference is proportional to a computed interblock transmissivity (IBT) between the nodes. A commonly used IBT scheme is the block-size-weighted harmonic mean of the transmissivity at each node (McDonald and Harbaugh 1988; Goode and Appel 1992; Romeu and Noetinger 1995). This is theoretically appropriate for steady-state one-dimensional flow and when it is assumed that *T* is constant within the given cell block, changing abruptly at the interface between cell blocks. Goode and Appel (1992) proposed what they called a *logarithmic mean* to be theoretically correct for IBT for the case where *T* varies linearly between model nodes. The logarithmic mean and other IBT options have been added to post-1992 versions of MODFLOW.

Goode and Appel (1992), Romeu and Noetinger (1995), and others have shown how the IBT scheme choice is consequential when the model discretization is coarse relative to the scale of heterogeneity. Both found the harmonic IBT to be biased low in that case. Romeu and Noetinger (1995) reported that: (i) the geometric, logarithmic, and arithmetic IBT schemes were each biased high; (ii) all schemes converged on the theoretically correct result as the grid resolution was refined; and (iii) the harmonic scheme may converge slightly faster with increased grid refinement.

These same results were found when testing simulation design herein. IBT bias was tested in the flow model domain described previously, primarily by using an isotropic *T* field under steady-state mean-parallel (SSMP) flow conditions. SSMP flow was used since the expected value, T_{eff}, is well established to be close to T_{G} for that case (Section 2.2).

The bias-testing results are illustrated in **Figure 3-11**. The simulated effective power-average exponent, p_{eff}, is plotted as function of the correlation range in the minor direction. The correlation range is normalized by the node spacing. The constraint identified was the minimum correlation range that could be simulated in the minor principal direction, relative to the fixed nodal spacing, without introducing bias from the choice of the IBT scheme.

The harmonic and logarithmic IBT schemes were tested (labeled LogIBT and HarIBT in Figure 3-11) since they are standard options in MODFLOW-2005 (Harbaugh 2005) and since they are reported to be biased low and high, respectively. Three different variance levels were tested (labeled var0.5, var1, and var2, for $\sigma^2_{\text{Y}} = 0.5$, 1.0, and 2.0, respectively), along with three different geometric anisotropy ratios (1:1, 3:1, and 10:1) each with the major principal axis oriented parallel to the river.

Bias in T_{eff} was significant with either of the IBT choices for isotropic fields with nodal spacing set at 1/5th the correlation range (Figure 3-11). Nodal spacing refined to at least 1/10th the correlation range was needed to reduce bias in that case, and the harmonic IBT converged faster. For the two higher anisotropy ratios (3:1 and 10:1), the harmonic mean was not strongly sensitive to grid refinement, and was less sensitive to logarithmic IBT overall. Based on these results, all simulations used a minimum correlation range equal to 10 grid nodes or higher, and all used the harmonic IBT scheme, except where noted differently.

Note: For the isotropic case, T_{eff} converged slightly above the geometric mean, near p = +0.05. This deviation from theory is possibly due to having a bounded domain (e.g., Paleologos et al. 1996), though sensitivity tests were inconclusive. This deviation is small relative to the range of p_{eff} trends of interest in this research.

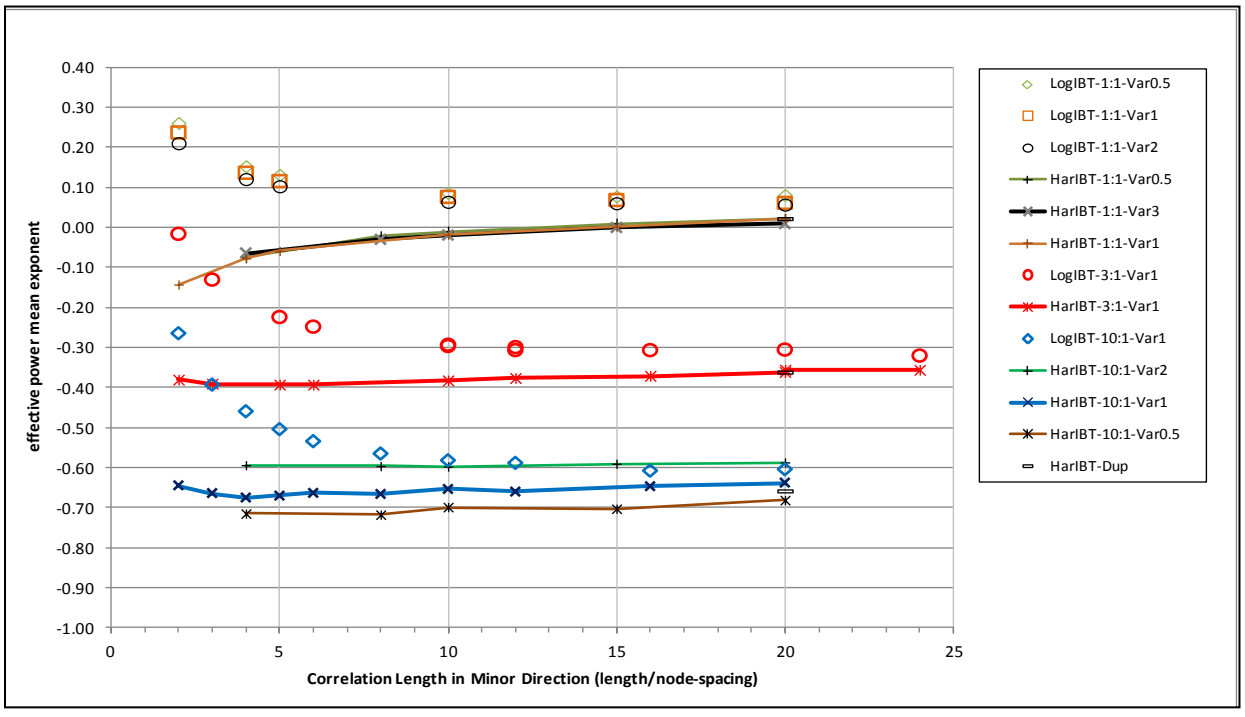

Figure 3-11. Plot illustrating sensitivity of effective power mean exponent to grid refinement and interblock transmissivity scheme.

3.6.3 Correlation Range Constraints

A minimum correlation range constraint, as established in the previous section, leads to other simulation constraints, which is why the minimum was not set higher. For example, to reduce influence from the arbitrary lateral boundaries (Columns 1 and 900), it is suggested that the correlation scale not be much more than one-third the domain size to reduce edge effects in the SGSIM simulations (Xu and Dowd 2012), and perhaps smaller to reduce flow-boundary effects (Rubin and Dagan 1989; Paleologos et al. 1996). In this model, that limitation would be about 300 nodes when the principal correlation direction is parallel to the river. Combining that with the minimum-range constraint of 10 nodes limits maximum geometric anisotropy to 30:1.

Fortunately, the constraint on maximum correlation range noted above is mostly a concern for SSMP simulations, because that flow field covers the entire domain. For the transient stream depletion case, the active flow area is smaller than the domain and focused away from the lateral boundaries, so lateral boundary effects are less significant. Additionally, having a *T* correlation scale larger than the transient stream depletion flow area is a realistic scenario. Rubin (1993) and Anderson (1997) include Gelhar's (1993) table of example correlation scales for *T* and *K*, which notes large transmissivity correlation scales.

The constraints suggested by Xu and Dowd (2012), Rubin and Dagan (1989) and Paleologos et al. (1996) would be relatively severe when simulating a principal correlation direction oriented perpendicular to the river, since that model dimension is 200 nodes. The one-third guidance in that orientation would put the maximum range at 66 nodes, constraining anisotropy to under 1:7. An SGSIM edge-effect was not apparent on T_{eq} results, however, based on a few test simulations performed with a 900 x 800 node domain. Furthermore, it is proposed that flow-boundary effects under that perpendicular orientation are acceptable since the

boundaries in Row 1 and Row 200 are not arbitrary domain boundaries. Those boundaries have physical meaning and are fundamental to the flow regime being tested.

In summary, the lateral model-domain boundaries may introduce bias to results for anisotropy ranges greater than about 30:1 in the SSMP flow simulations. That potential bias was not quantified since SSMP flow beyond that anisotropic limit was not a focus of this research.

CHAPTER 4 PRIMARY RESULTS

4.1 Dependence on Statistical Anisotropy

Range of Results

Figure 4-1 is a plot of effective power-mean exponents (p_{eff}) determined for transmissivity fields constructed with various geometric anisotropy ratios (λ_x/λ_y) ranging from 0.01 to 100. The plot includes ensemble sets with variance of ln(T), σ^2_Y,equal to 0.25, 0.5, 1.0, and 2.75. Anisotropy ratios less than one ($\lambda_x/\lambda_y < 1$) indicate the principal direction of correlation was oriented perpendicular to the river, while ratios greater than one ($\lambda_x/\lambda_y > 1$) indicate the orientation was parallel to the river. The 95% confidence intervals of the ensemble means are shown.

The heavy black curve in Figure 4-1 is Equation 2-9 (page 20), the estimate proposed by Ababou (1991) for p_{eff} under steady-state mean-parallel flow conditions. It is included for reference. As discussed in Section 2.2.5, it coincides with Equation 2-7, the derivation by Gelhar and Axness (1983) also for steady-state mean-parallel conditions (Sanchez-Vila et al. 2006).

For all cases, p_{eff} was near the geometric mean ($p = 0$) for fields with statistically isotropic correlation structure ($\lambda_x/\lambda_y = 1$) and it asymptotically approached the bounding arithmetic ($p = 1$) and harmonic ($p = -1$) means as geometric anisotropy increased. For the four heterogeneity levels tested ($\sigma^2_Y = 0.25$, 0.5, 1.0, and 2.75), p_{eff} exhibited a shift toward the geometric mean with increased heterogeneity. That sensitivity to field variance is examined more closely in Section 4.3.

Spatial Weighting

For one set of ensembles plotted in Figure 4-1, the statistical moments of the realized transmissivity fields, and thus also the idenitfied p_{eff} results, were computed using the entire

realized fields. That set is labeled *unweighted*. Analysis of all other ensemble runs was based on transmissivity statistics limited to the main area of influence (see Section 3.6.1). A direct comparison between unweighted and weighted ensemble runs can be made using the weighted case that had field variance at $\sigma^2_Y = 1$, which was the same as the unweighted set. The 95% confidence interval of the mean was generally smaller for the weighted case, but no statistical difference in the ensemble means was found between the comparable weighted and unweighted cases.

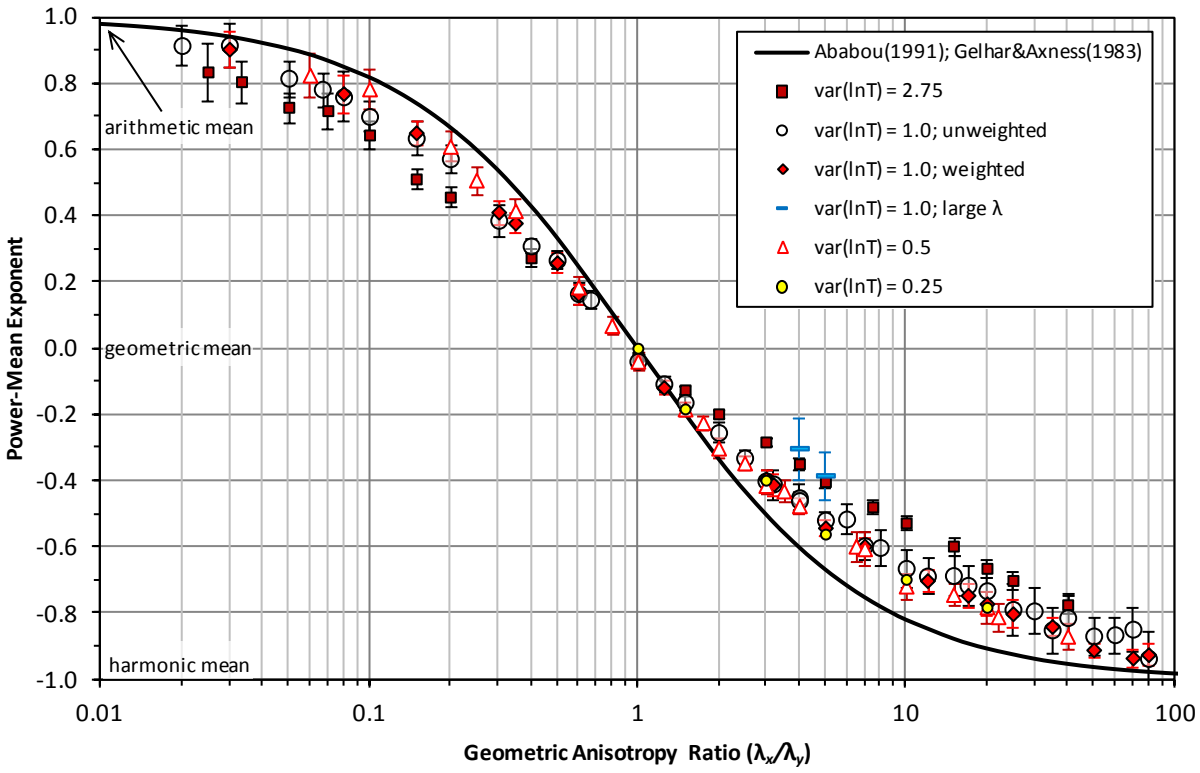

Figure 4-1. Empirical relationship between effective transmissivity (as power mean) and statistical geometric anisotropy for approximately 110 ensemble simulations of transient stream depletion in spatially correlated heterogeneous transmissivity fields.

Correlation Scale Effects

Two p_{eff} results shown in Figure 4-1, both with $\lambda_x/\lambda_y = 4$ and 5, were based on transmissivity fields that had large-scale correlation ranges relative to the other simulations. For example, one $\lambda_x/\lambda_y = 5$ ensemble had correlation ranges set at $\lambda_x = 15{,}000$ ft and $\lambda_y = 3{,}000$ ft, whereas the other $\lambda_x/\lambda_y = 5$ ensembles had $\lambda_x = 2{,}500$ ft and $\lambda_y = 500$ ft. The two large-scale correlation cases exhibited greater ensemble variance, as well as p_{eff} results shifted toward the geometric mean, relative to the other cases. This scale effect is discussed further in Chapter 6.

Insensitive Simulation Parameters

The following simulation variables were also examined for their potential influence on the effective power-mean exponents: (i) using different spatial correlation models in generating the transmissivity fields (i.e., spherical, Gaussian, and exponential model variograms); (ii) placing the pumping well at different distances from the river ($a/W = 0.5$ and $a/W = 0.75$); and (iii) using 15 nodes in the minimum correlation direction instead of using the 10-node minimum discussed in Section 3.6.2. Those simulations are not plotted since the results were the same as the other cases, i.e., no significant influence on the ensemble means was observed for those variables.

Placing the pumping well at greater distance from the river under the same transmissivity field correlation range resulted in lower ensemble variance, but no effect on ensemble means was observed. As discussed further in Chapter 6, the change in ensemble variance but not ensemble mean with well position arises from the change in the relative scales of the area of influence and the correlation range of heterogeneity.

It has been shown analytically that T_{eff} can depend in part on the correlation function shape, referring to the spherical, Gaussian, or exponential models for example (e.g., Indelman

and Abramovich 1994). Those studies also indicate the effect is small relative to other factors such as degree of statistical anisotropy. The choice of correlation model was not observed to affect p_{eff} results in this study. Such an effect, if present, may be smaller than was distinguishable by the numerical approach used in this study.

Ensemble Variability

The ensemble means plotted in the prior figure are shown again in **Figure 4-2** along with the p_{eq} results from each ensemble. This type of density plot may overemphasize outliers visually, but it is included to illustrate the distribution of the approximately 11,000 realizations completed and to document certain outliers. The Gaussian distribution of most ensembles is seen in Figure 3-9 (Page 52) and considered further in Chapter 6.

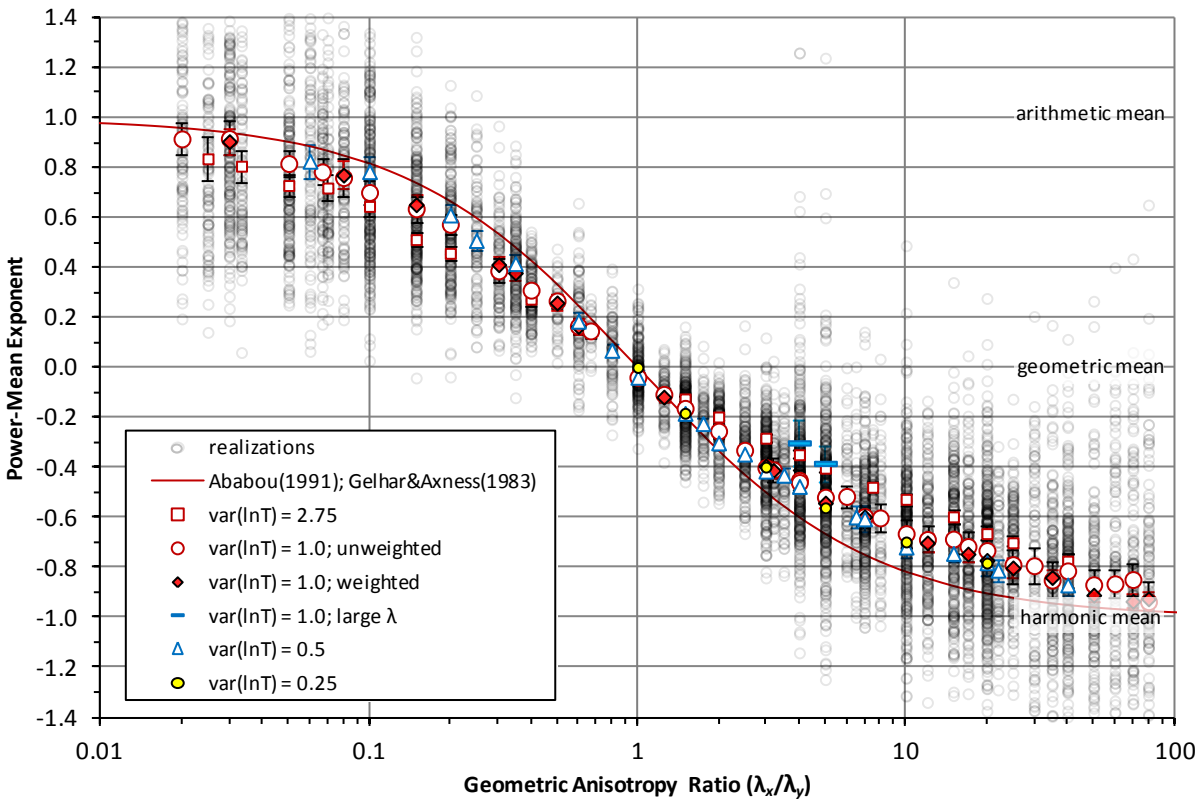

Figure 4-2. Same ensemble-mean results as previous figure, with realizations included.

Certain caveats should be noted when considering individual realizations and Figure 4-2. As discussed in Chapter 3, the SGSIM method doesn't guarantee that each realization meets the assigned geostatistical model; it guarantees only that the geostatistical model is reproduced over the ensemble (Deutsch and Journel 1998). A degree of error is therefore expected between the statistical anisotropy as plotted and what was actually realized in the SGSIM field construction.

Statistical moments of each field realization were checked against the intended statistical model and were found to be very close. But, it was computationally impractical to construct an empirical variogram for each realization to check the correlation range and anisotropy for each case. Like Sarris and Paleologos (2004) and others, the realized correlation range and anisotropy characteristics were assumed to reasonably match the conditions specified in SGSIM. This assumption applies best to fields with smaller correlation ranges. Random checks of several realizations supported that assumption. Sarris and Paleologos (2004) also provide justification.

The realized transmissivity fields for a few outlier p_{eq} results were reviewed visually and by constructing variograms. Some fields with outlier results appeared to reasonably match the assigned geostatistical correlation model, while others showed moderate deviations. No comprehensive conclusions were made about field properties of the outliers, but certain cases had a clear explanation. For example, one extreme outlier was a case with high variance (σ^2_Y = 2.75) and a large correlation scale with the principle direction oriented perpendicular to the river. A narrow zone with high transmissivity and high contrast with the surrounding media happened to be located directly at the pumping well and aligned between the pumping well and the river, (see Appendix A, Figure A-5, for illustration). This led to a fast depletion response that was not captured by a T_{eq} averaged over a wider area. The response in that case was, presumably, more like conduit flow.

Approximately 9% of the realizations had a p_{eq} outside the theoretical $-1 < p < +1$ bounds. Nearly all such instances occurred in ensembles with geometric anisotropy ratios outside the interval $\lambda_x/\lambda_y = [0.1, 10]$. Assumptions of statistical homogeneity and ergodicity are most strongly violated for fields with larger-scale correlation ranges. In those cases, it appeared that subareas had a dominant influence on the depletion timing due to having a local-average transmissivity that was significantly different than the statistics of the overall field. For example, in the outlier discussed in the previous paragraph, the transmissivity of the controlling high-permeability channel was much higher than the arithmetic average of a wider zone. That scenario produced a p_{eq} greater than +1.

4.2 Comparison to Steady-State Mean-Parallel Flow

The p_{eff} results obtained from transient stream depletion simulations were compared to results from steady-state, mean-parallel (SSMP) flow simulations. The same simulation procedures were used for the SSMP case except for the necessary configuration changes. The pumping well was removed and Model Row 200 was changed from an impermeable boundary to a specified-head boundary. The specified-head boundary applied a fixed gradient (i) across the domain in the y direction. The resulting steady-state flow (Q_{SSMP}) across the model domain was evaluated with the Darcy equation ($Q_{SSMP} = T_{eq}\, i$), where T_{eq} was the unknown.

As shown in **Figure 4-3**, p_{eff} results for SSMP simulations were not significantly different from the transient stream depletion cases. The ensemble variance was smaller for comparable SSMP simulations, seen as smaller 95% confidence intervals of the mean, but the ensemble means were not statistically different.

Ensembles for the transient stream depletion and SSMP comparisons were generated independently. They were equivalent geostatistically, but did not comprise exactly the same set

of realizations. For an additional comparison, a few individual field realizations were run under both transient stream depletion and SSMP flow. The p_{eq} difference found between transient stream depletion and SSMP flow for those field realizations was not zero, but it was small. The small differences were attributed to the SSMP flow being influenced by the entire transmissivity field while stream depletion is influenced by a smaller subarea of the field. As discussed in Chapter 3, the geostatistical properties of a field sub-area can be different than that of the wider field.

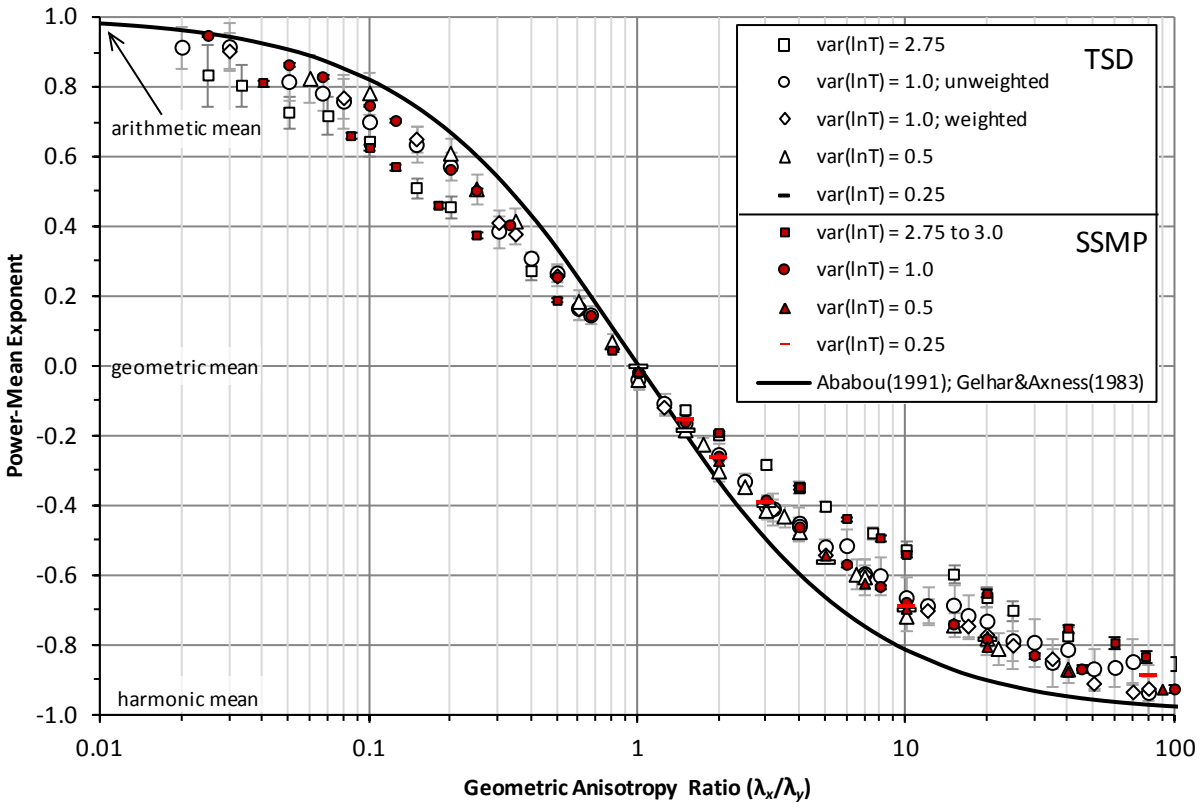

Figure 4-3. Comparison of effective transmissivity for transient stream depletion (TSD) and steady-state mean-parallel (SSMP) flow simulations.

Except for cases with mild geometric anisotropy ($0.5 < \lambda_x/\lambda_y < 2$), p_{eff} results for SSMP flow did not precisely match the Ababou (1991) and Gelhar and Axness (1983) analytical estimate (Figure 4-3). Modeled results were closer to the analytical estimate for fields with low

variance than those with high variance, but even the lowest-variance cases converged on an apparent p_{eff} curve that was shifted away from the analytical estimate toward the geometric mean.

Spector and Indelman (1998) presented a higher-order estimate for K_{eff}, as a function of geometric anisotropy, that coincides well with the p_{eff} trend modeled herein for $\sigma^2_Y = 1.0$. Additionally, their derivation departs most significantly from the Gelhar and Axness (1983) and Ababou (1991) estimate in the anisotropy range of 3 to 10, which is similar to the modeled results (Figure 4-3). It is further noted that the Spector and Indelman estimate predicts roughly similar behavior as was modeled for increasing variance—p_{eff} shifts away from the Weiner bounds toward the geometric mean with increasing variance—while the Ababou and Gelhar and Axness estimates are insensitive to variance (when T_{eff} is normalized through p_{eff}). However, the Spector and Indelman estimate does not match the modeled results for $\sigma^2_Y = 0.25$ or 0.5, with their estimate being very close to Gelhar and Axness (1983) in those cases. Nonmonotic behavior was encountered with their estimate for $\sigma^2_Y > 1.25$. Since the match at $\sigma^2_Y = 1.0$ may therefore be coincidental, it was not included in the plots herein.

4.3 Influence of Field Variance

The previous figures indicated some sensitivity of p_{eff} to the degree of heterogeneity (σ^2_Y). Modeled trends for p_{eff} as a function of σ^2_Y are plotted in **Figure 4-4** for $0.1 < \sigma^2_Y < 3.25$. Results for both transient stream depletion and SSMP flow are included for comparison. Most simulations used a correlation range in the minimum principal direction equal to 10 grid nodes. Two ensembles shown in Figure 4-4 used 15 nodes.

Low sensitivity to σ^2_Y was observed for the statistically isotropic case; its p_{eff} remained near the geometric mean ($p = 0$) for $0.5 < \sigma^2_Y < 3.25$. For the $\lambda_x/\lambda_y = 10$ case, p_{eff} increased from

about -0.7 to -0.5 over the tested variance range. For $\lambda_x/\lambda_y = 3$, p_{eff} increased from about -0.4 to -0.25.

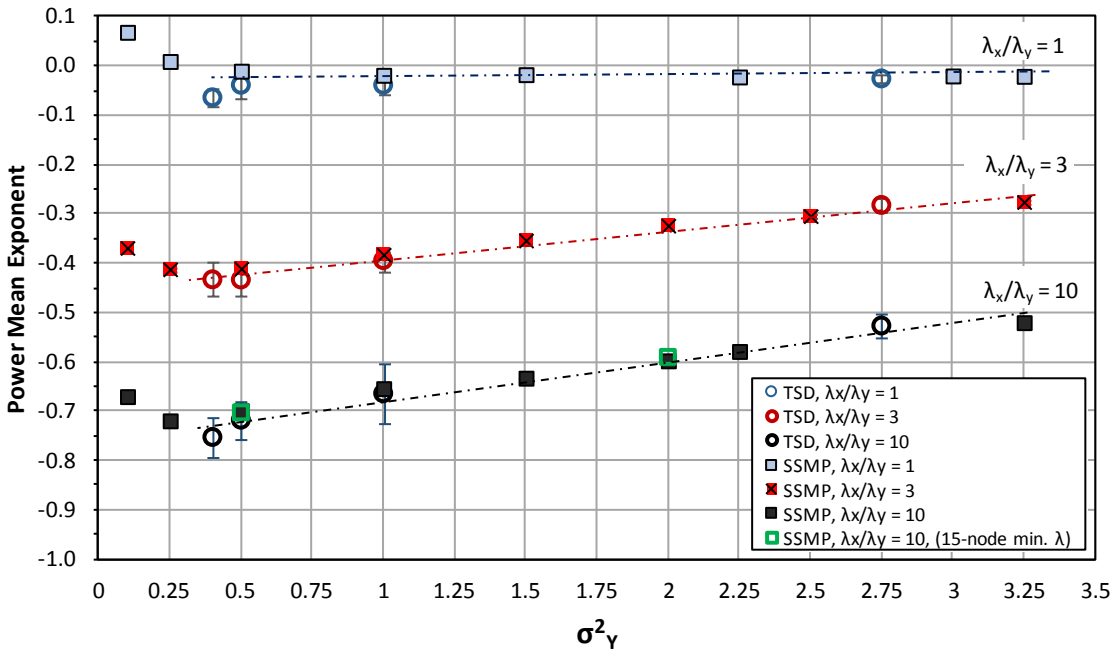

Figure 4-4. Effective transmissivity (as power mean) as a function of natural-log transmissivity variance (σ^2_Y) and geometric anisotropy ratio (λ_x/λ_y = 1, 3, and 10). Results shown for both transient stream depletion (TSD) and steady-state mean-parallel (SSMP) flow simulations.

The nonmonotonic trend starting around $\sigma^2_Y < 0.4$ is not understood (Figure 4-4). It is noted, however, that the field is relatively homogenous in that range. For example, at $\sigma^2_Y = 0.25$, T_{eq} changes by less than 1% for a p change of 0.07. The methods used to identify p_{eq} may exhibit small error at low σ^2_Y due to interpolating from a discretized $T(p)$ curve that is relatively flat for low σ^2_Y (Section 3.5.4).

Interestingly, p_{eff} estimates obtained from the Gelhar and Axness (1983) equation that was derived only for low variance ($\sigma^2_Y < 1$, Equation 2-6, page 20) did roughly match the modeled behavior shown in Figure 4-4 for $0.5 < \sigma^2_Y < 3.25$, while the Gelhar and Axness

generalization for higher variance ($\sigma^2_Y > 1$, Equation 2-7, page 20) produces p_{eff} that is invariant with changes in σ^2_Y.

Hydraulic conductivity can span several orders of magnitude in a given aquifer, but variance of ln(*T*) is generally much lower than variance of ln(*K*) since *T* is itself an up-scaled parameter. In other words, while the portion of an aquifer influencing a given pumping test may include lenses of clay, silt, sand, and gravel—thus high *K*-variance in the pumping test domain—a pumping test produces a single *T* estimate. Compilations of transmissivity data provided by Gelhar (1993), Hoeksema and Kitanidis (1985), and Rubin (2003) suggest that σ^2_Y is commonly near 1.0 and lower (with several exceptions noted). For this reason, heterogeneity was represented in most simulations at $\sigma^2_Y = 1.0$, and the range used in the sensitivity runs was limited to $0.1 < \sigma^2_Y < 3.25$.

It is noted that in aquifers with high hydraulic conductivity contrasts (thus high variance) it can be the *connectivity* of high-permeability zones, acting as channels, that influences effective macroscale permeability (e.g., Ronayne and Gorelick 2006; Fogg 1986). Channel connectivity is a condition not accounted for with the SGSIM approach used herein.

4.4 Nugget Effect

The influence on p_{eff} of a nugget effect in the spatial correlation model was examined. A nugget effect refers to having non-zero variance among points at zero or near-zero lag distance. Variance between essentially adjacent samples may be from measurement error and thus not have physical meaning, but it can also represent local-scale variability below the resolution of the sampling density (Isaaks and Srivastava 1989; Gringarten and Deutsch 2001). It is modeled with a non-zero intercept in a variogram. With an increasing nugget effect, the simulated transmissivity field looks increasingly grainy, due to having an increasing degree of

discontinuity locally, even when larger-scale organization is still evident (see Appendix A, Figure A-5).

Figure 4-5 illustrates the effect observed in p_{eff} when simulated transmissivity fields included a nugget effect in the variogram. The two different nugget cases tested here had scaled nugget contributions equal to 0.33 and 0.5, where 1.0 is the scaled total variance. The results are compared to the p_{eff} results shown previously in Figure 4-3 which had no nugget included. Relative to those prior cases, p_{eff} appears to level off further from the Weiner bounds with an asymptote shifted toward the geometric mean. In these examples, the asymptote appears to be about $p = -0.7$ and $p = -0.6$ for nuggets equal to 0.33 and 0.5, respectively.

The nugget-effect trend seen in Figure 4-5 might be explained, qualitatively, by considering that the Weiner bounds are approached with increased organization of the transmissivity field. In the stream depletion case, increased organization refers to increased stratification (in plan view) of correlated transmissivity zones. With a nugget effect, however, even with significant large-scale correlation, the continuity of the stratification is limited locally. If the nugget contribution approaches the total variance, then there would be little structural organization and the effective mean would remain close to the effective mean for uncorrelated media, which is the geometric mean.

The sensitivity of p_{eff} to a wider range of nugget contributions is illustrated in **Figure 4-6** for three cases of geometric anisotropy: $\lambda_x/\lambda_y = 1$, 3, and 10. Modeled p_{eff} for the three cases converged to a single value—ostensibly the effective mean for uncorrelated media—as the scaled nugget contribution approached 1.0.

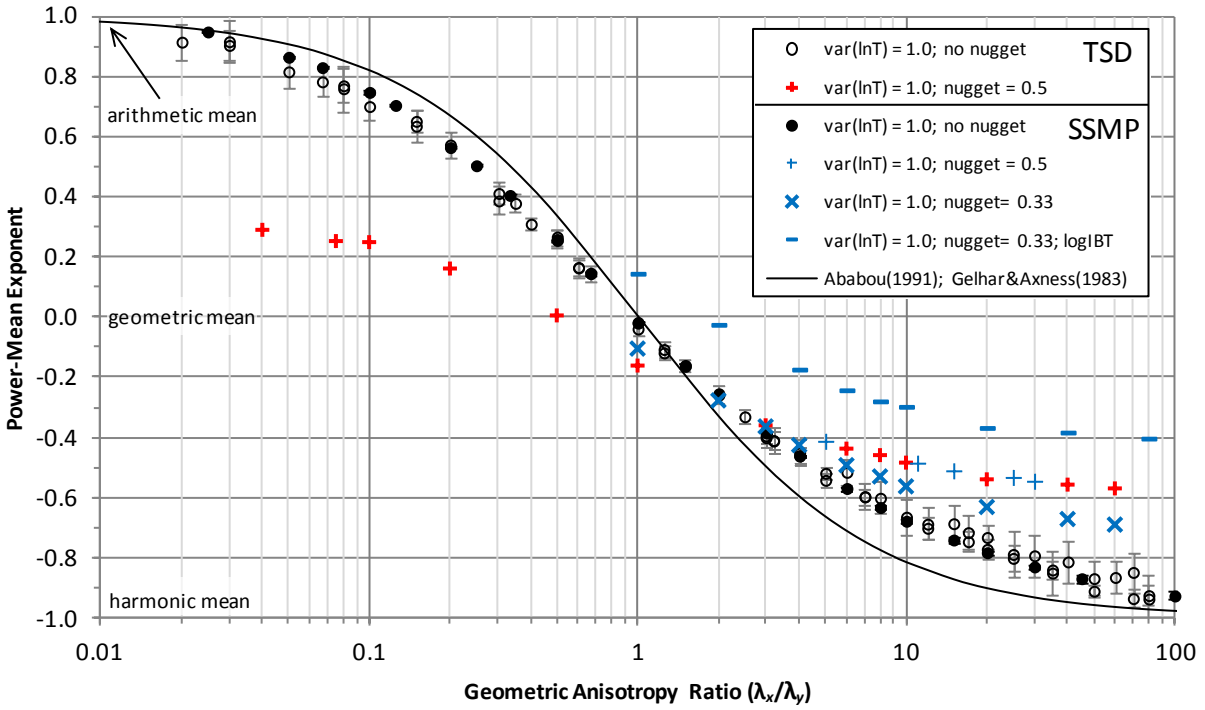

Figure 4-5. Effective transmissivity when including a nugget effect in the transmissivity spatial correlation model. Results are shown from transient stream depletion (TSD) and steady-state mean-parallel (SSMP) flow simulations. One ensemble set used the logarithmic interblock transmissivity scheme (logIBT) all others used the harmonic scheme.

The choice of interblock-transmissivity averaging scheme in the finite-difference model affected p_{eff} when a nugget effect was present in the spatial correlation function for transmissivity. As discussed previously (Section 3.6.2), the interblock-transmissivity scheme was found to bias simulation results when the transmissivity correlation scale was less than approximately 10 model grid nodes. The logarithmic scheme created a high-bias and the harmonic scheme created a low-bias in that case. Similar biases were apparent with the nugget effect (Figure 4-6). The p_{eff} converged for the different anisotropy cases as the media became uncorrelated (i.e., as the nugget contribution appraoched the total variance, 1.0) but the cases converged below the theorectical geometric mean (p = 0) for the harmonic interblock transmissivity scheme and above the geometric mean for the logirthmic scheme.

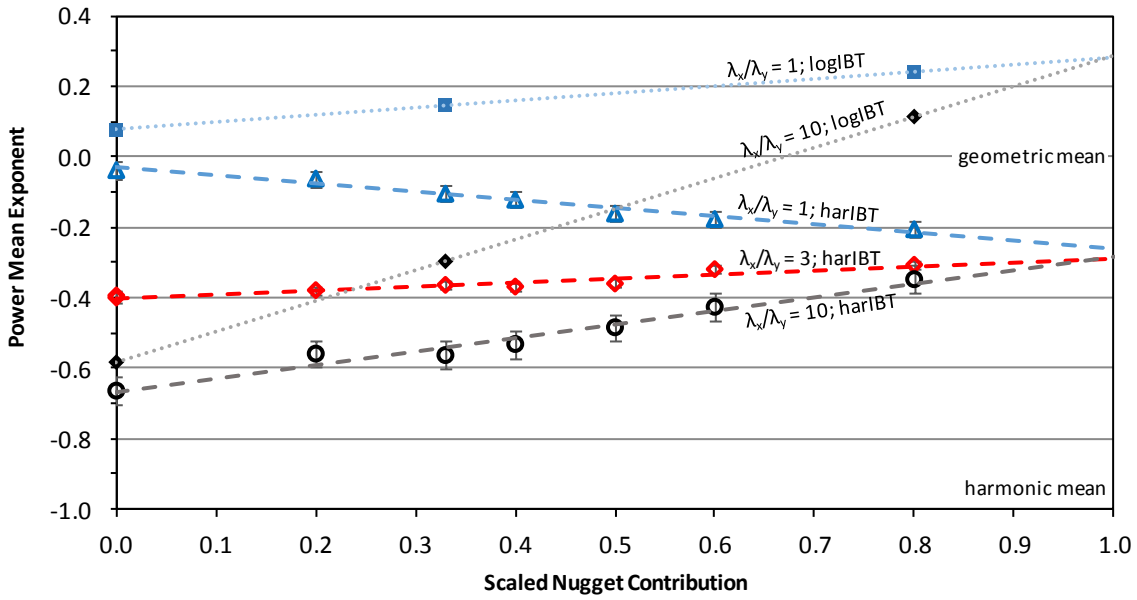

Figure 4-6. Sensitivity of effective transmissivity to a nugget effect in the transmissivity correlation model and to the logarithmic and harmonic interblock transmissivity schemes (logIBT and harIBT). Trends are shown for geometric anisotropy ratios (λ_x/λ_x) equal to 1, 3, and 10.

Figure 4-6 includes results from steady-state mean-parallel flow simulations. It is well established that p_{eff} should approach the geometric mean ($p = 0$) under that flow regime for two-dimensional, isotropic, randomly heterogeneous media (Section 2.2). Combining that fact with the established nature of the interblock-transmissivity bias, it is reasonably assumed that the trends shown in Figure 4-6 bracket the actual trend which would have p_{eff} approaching $p = 0$ with the increasing nugget contribution.

In simulations without a nugget effect in the correlation model, the interblock transmissivity bias was avoided by using an appropriate minimum correlation scale, but with a nugget present there is a degree of uncorrelated spatial variability even between adjacent grid nodes. The interblock transmissivity bias that arises from a nugget effect might be reduced by using a MODFLOW grid with substantially higher resolution than the SGSIM grid. Refining the

MODFLOW grid was not explored, however, since the general result of a nugget effect was adequately established by the simulations presented herein.

If a nugget in the transmissivity spatial correlation function is indeed just an artifact of measurement error or limited data density, rather than representing actual local-scale discontinuity, then the nugget impacts illustrated in Figure 4-5 may have limited physical meaning. In that case, model-derived power means associated with a nugget effect may not be appropriate for use in analytical models.

Aside from potential uncertainties about its physical meaning, a nugget is sometimes used in spatial correlation models. If a numerical groundwater flow model is used to simulate depletion timing, and if that flow model's transmissivity field is populated using a geostatistical model that includes a nugget effect, then the simulated depletion timing may be impacted.

The nugget-effect simulations support a general observation from the other simulated cases: The bounding arithmetic and harmonic means for T_{eff} are approached only with an increasingly strong degree of structural organization of the transmissivity field. If that degree of organization is limited, either in the large-scale correlation structure (statistical anisotropy) or by local-scale variability embedded within the overall correlation structure (nugget effect), then T_{eff} is limited from approaching the bounding arithmetic and harmonic means.

4.5 Effect of River Angle

Paleologos (1994) presented the following estimate for K_{eff} in two-dimensional, anisotropic media under steady-state mean-parallel flow, when the direction of mean flow and the principal axis of correlation are not aligned:

$$K_{eff_{ii}} = K_G \left[1 + \frac{\sigma_Y^2}{2} - \sigma_Y^2 \left(\frac{\lambda_j cos^2\theta + \lambda_i sin^2\theta}{\lambda_1 + \lambda_2} \right) \right] \quad (4\text{-}1)$$

in which θ is the angle between the direction of mean flow and the principal axis of correlation. When the flow direction and principal correlation axis are aligned, this reduces to Equation 2-6. Equation 4-1 was outlined by Gelhar and Axness (1983), but Paleologos (1994) noted that this more explicit and closed form had not been presented in the literature.

Equation 4-1 was evaluated in this work for the case of transient stream depletion when the river boundary orientation is not aligned with the principal directions of the aquifer's correlation structure. In an alluvial aquifer, the principal direction of transmissivity correlation and the river may both be aligned, generally, with the longitudinal axis of the aquifer, but there are exceptions. For example, a tributary stream may cross the alluvial valley at an oblique angle, and certain reaches of a meandering river may cut across the valley at different angles. Furthermore, stream depletion concerns are not limited to pumping from alluvial aquifers; other types of river-aquifer systems are of interest as well.

Modeled results are compared to Equation 4-1 in **Figure 4-7**. The simulations included both transient stream depletion and steady-state, mean-parallel flow conditions, under a range of river-aquifer alignments, and for cases with geometric anisotropy equal to 1, 3, and 10. Directional alignment was not expected to be relevant for the isotropic case, but it was included as a reference case.

When the aquifer's principle direction of correlation was aligned at angles near 0° and 90° with respect to the river, the simulations were the same as the cases plotted previously (e.g., Figure 4-1, Figure 4-3). At intermediate angles, p_{eff} shifted away from the bounding means toward the geometric mean. At θ = 45°, all three cases had p_{eff} equal to the geometric mean (p = 0).

A discrepancy was observed between the modeled p_{eff} results and the analytical estimate, with the modeled output exhibiting a small shift away from the estimate toward the geometric mean. This appears to be the same behavior observed in the primary simulations (Figure 4-1, Figure 4-3).

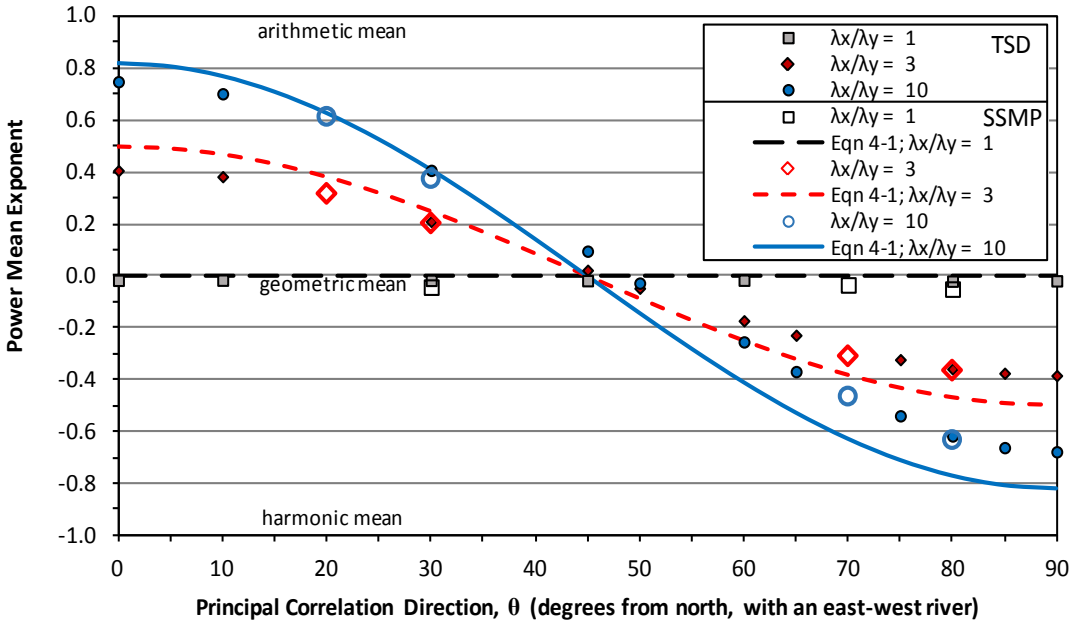

Figure 4-7. Effective transmissivity when the transmissivity field's principal direction of spatial correlation is oriented at an angle θ from the principal flow direction in steady-state, mean-parallel (SSMP) flow conditions. Included are transient stream depletion simulations where the river orientation is at θ = 90°. Trends are shown for geometric anisotropy ratios (λ_x/λ_y) equal to 1, 3, and 10.

The modeled trend for p_{eff} in the $\lambda_x/\lambda_y = 10$ case exhibited a degree of asymmetry as the alignment was rotated away from θ = 45°. The asymmetry is most easily seen in the divergence from the symmetric analytical estimate, a divergence that was observed to be larger for θ > 45° than for θ < 45° (Figure 4-7). This is possibly caused by the aspect ratio of the bounded model domain, which could lead to stronger boundary effects for θ approaching 0° than for θ approaching 90°, due to the smaller domain width in the θ = 0° direction (see Appendix A,

Figure A-3). An asymmetry caused by boundary effects would also explain why it is prominent only in the $\lambda_x/\lambda_y = 10$ case, the case with the longest correlation range.

The trends illustrated in Figure 4-7 could be relevant for the case of a pumping well located near the confluence of a river and a tributary stream. Stream depletion from the pumping well would accrue to both streams. If the aquifer exhibits geometric anisotropy, then the effective transmissivity governing the depletion timing would be different for each stream due to their different relative orientations. If the aquifer's principal correlation direction were oriented parallel to the main stem of the river, and if the tributary stream approaches at a right angle, then the expected stream depletion timing would be described by p_{eff} below a geometric mean for impacts to the main stem of the river and above a geometric mean for the tributary. This two-valued p_{eff} scenario may have applications in the analytical model presented by Yeh et al. (2008) for stream depletion from a pumping well located at the confluence of two rivers.

CHAPTER 5 - LIMITATIONS for NON-STATIONARITY and OTHER CASES

Stochastic approaches to analyzing groundwater flow in heterogeneous fields usually assume stationarity of the field (Rubin and Seong 1994; Renard and Marsily 1997; Chang and Yeh 2012). Likewise, the effective transmissivity behavior presented in Chapter 4 was simulated under the condition of stationarity. Sanchez-Vila et al. (2006) stated that field data suggest stationarity is commonly not a property of the medium but rather is an artifact of the scale of observation. Stationarity has also been characterized as a decision made in analysis, emphasizing it is a condition of the method and not necessarily a property of the data distribution (Deutsch and Journel 1998).

The effective transmissivity relationships identified in Chapter 4 are tested in this chapter for sensitivity to the stationarity condition. In particular, the case of mean transmissivity decreasing linearly with increasing distance from the river is simulated. Sensitivities to hydraulic anisotropy and zonal anisotropy, the latter being a different form of statistical anisotropy, are also evaluated.

5.1 Block-Wise Variable Mean

To support interpretation of depletion response behavior from an aquifer that has a smoothly varying trend in mean transmissivity, a simplified case was first simulated. The simpler case had two homogenous aquifer zones with dissimilar transmissivity. Transmissivity zone T_1 was located in the half of the aquifer between the river and the well (model rows 1 to 100), with zone T_2 located between the well and the impermeable boundary (model rows 101 to 200). Transmissivity of zone T_2 was varied by factors of T_2/T_1 = 0.125, 0.25, 0.5 and 2.0.

Consider that if zone T_2 has a transmissivity near zero, then it would act as an impermeable boundary that in effect narrows the aquifer. One would expect a steeper depletion

curve in that case over the homogeneous case (Miller et al. 2007). Indeed, as T_2 was decreased in these simulations, the early-time depletions were increased relative to the homogenous case in which $T_2/T_1 = 1$ (**Figure 5-1**).

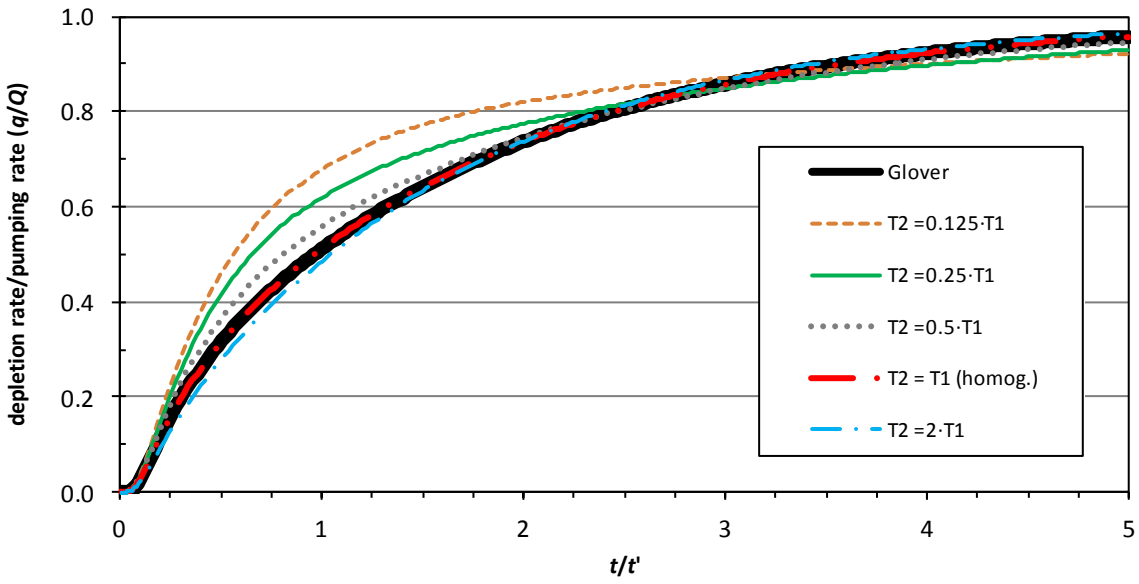

Figure 5-1. Depletion response curves for block-wise heterogeneity with transmissivity T_2 between the well and the impermeable aquifer boundary scaled relative to the transmissivity T_1 between the well and the river. Time is scaled based on T_1 ($t' = a^2S/T_1$).

In turn, late-time depletions were slightly slower than the homogenous case. Mathematically, the slower late-time depletions are required because the area under the depletion curve must approach the pumped volume over time. The pumped volume was the same for all cases shown in Figure 5-1. A physical explanation is that the T_2 zone would drain more slowly toward the steady-state condition, due to its lower transmissivity, thereby slowing the rate at which equilibrium between pumping and stream depletion is reached, i.e., slowing the rate at which $q/Q = 1$ is approached.

A pertinent observation was that the standard analytical Glover model, whether applied in bounded or unbounded form, did not match the shape of the depletion response curves from this block-wise-heterogeneous aquifer. The curve shapes could be matched if boundary distances were adjusted—invoking an effective boundary distance concept—but they could not be matched using the known geometry of the system. The shape error was modest for the milder transmissivity contrasts.

5.2 Trending Mean Transmissivity in Steady-State Mean-Parallel Flow

5.2.1 Framework for Sensitivity Test

A framework for evaluating the sensitivity of T_{eff} to a non-stationary transmissivity in the stream depletion case is presented in this section by first examining sensitivity in the steady-state, mean-parallel flow case. The basis for choosing the magnitude of the tested transmissivity trend is also described.

For valley-fill alluvial aquifers, transmissivity values often increase toward the middle of the valley, and thus often increase with proximity to the river, due simply to greater aquifer thickness toward the middle of the river valley. Such non-stationary problems are not trivial to analyze. For one issue, the block-wise heterogeneous scenario in the prior section showed that an increase in T toward the river will alter the shape of the transient stream depletion response such that current analytical models may not be strictly applicable. However, those simulations also suggested that the shape change was reasonably small for a moderate spatial change in T.

In further test simulations, response curves from a non-stationary transmissivity field were observed to have a maximum deviation from the stationary case by about 5% when T increased by a factor of three across the model domain. Those simulations used a homogenous hydraulic conductivity field with saturated thickness that increased linearly by a factor of three

from the impermeable aquifer boundary (Row 200) to the river (Row 1). Based on that result indicating modest error in the curve shape, a three-fold increase in mean *T* across the model domain was selected for testing the heterogeneous non-stationary case.

The magnitude and direction of that selected trend is realistic, yet it allowed interpretation of the depletion response with acceptable accuracy using the Glover analytical model. More to the point, even while the response still reasonably fit a Glover response curve, this trend was found to strongly affect T_{eff} results.

It is common to model a non-stationary heterogeneous field as a deterministic trend component combined with a stochastic residuals component, the latter being a stationary random function representing heterogeneity (Deutsch and Journel 1998). The deterministic treatment makes particular sense when there is an independent, even if qualitative, understanding of the nature of the trend (Isaaks and Srivistava 1989). For the simulations presented herein, a non-stationary heterogeneous field was created by first generating a stationary heterogeneous field and then adding the deterministic trend (Appendix A, Figure A-6). This can be viewed as a stationary heterogeneous *K* field with a deterministic trend in aquifer thickness, thus a non-stationary heterogeneous *T* field.

Indelman and Rubin (1996) examined effective transmissivity for steady-state mean-parallel flow in a heterogeneous field with a linear spatial trend in mean transmissivity. They showed:

> "when the trend is parallel to the gradient, the equivalent conductivity is bounded by the effective conductivity of the stationary medium on one hand and the harmonic mean conductivity on the other, with the harmonic mean pertaining to the case of a very large trend."

Similarly, when the trend is perpendicular to the gradient, effective conductivity is bounded by the stationary effective conductivity and the arithmetic mean (Indelman and Rubin 1996; Rubin and Seong 1994).

The Indelman and Rubin statement is qualitative and the trend's significance was described through its impact on the outcome. For this research, significance of the trend was gauged separately from its impact on the outcomes by considering the variance it contributed to a given heterogeneous field. By comparing the variance σ^2_Y of the trended transmissivity field to that of the de-trended field, it was found that the three-fold increase in mean transmissivity across the domain contributed approximately $\sigma^2_Y = 0.11$ to the total variance of the field, regardless of the scale of σ^2_Y of the de-trended field, i.e., regardless of the degree of stationary heterogeneity as measured by $\sigma^2_{\ln(K)}$. This held within the tested range, which was for σ^2_Y of the de-trended field ranging from 0.1 to 0.9. For example, when the stationary field variance was $\sigma^2_Y = 0.10$, the field variance with the trend included was $\sigma^2_Y = 0.21$. For stationary field variance $\sigma^2_Y = 0.89$, the field variance with the trend included was near $\sigma^2_Y = 1.0$.

5.2.2 Sensitivity for Steady-State, Mean-Parallel Flow

In steady-state mean-parallel flow simulations, the influence of a trend in transmissivity was found to be minor when the field was strongly heterogeneous relative to the variance added by the trend. For example, in that minor-impact case, the de-trended field variance was $\sigma^2_Y = 0.89$ and the total field variance was $\sigma^2_Y = 1.0$. The trend contributed about 10% of the total variance and T_{eff} remained primarily a function of the geometric anisotropy relationship shown previously in Chapter 4. However, when the field was only mildly heterogeneous, such as with a de-trended variance $\sigma^2_Y = 0.11$ and with the trend contributing about half of the total field

variance $\sigma^2_Y = 0.22$, then there was a significant shift in T_{eff} toward the harmonic mean. Results from these two cases are shown in **Figure 5-2**. This is apparently similar to Indelman and Rubin (1996) which noted a significant trend leads to T_{eff} being closer the harmonic mean than would expected from the stationary component of the field's properties.

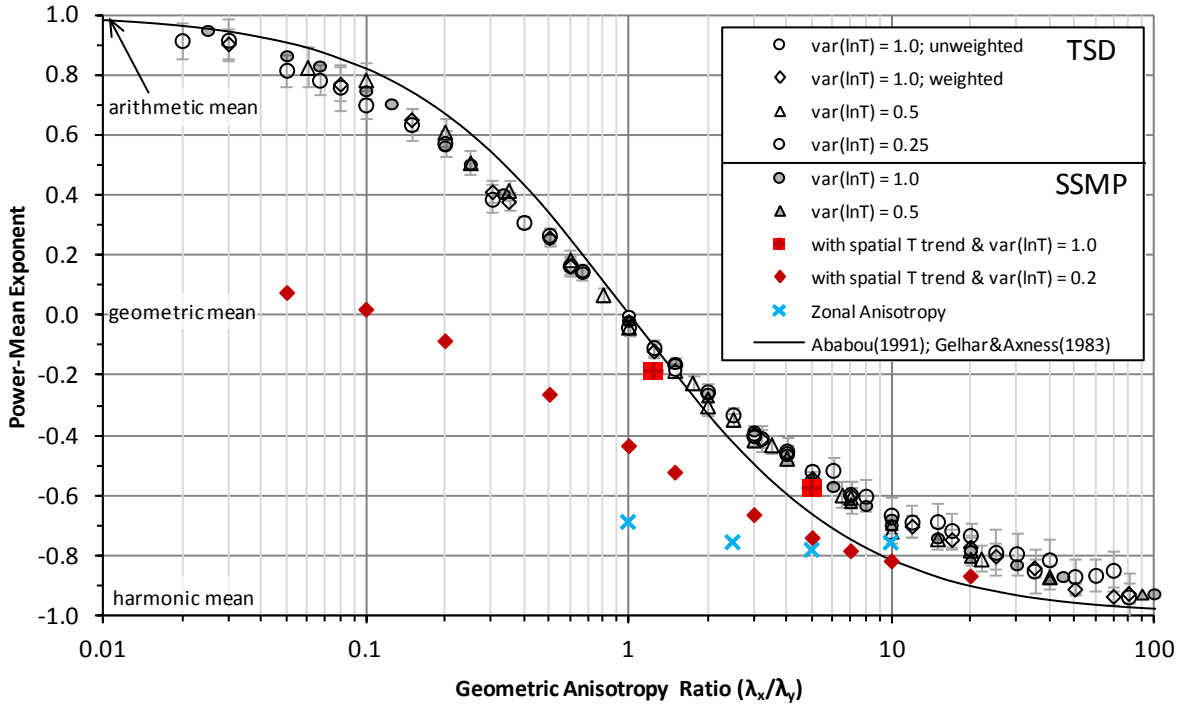

Figure 5-2. Power-mean exponents for two cases with non-stationary fields that had a linear trend in local-mean transmissivity. One case of zonal anisotropy is also included. Stationary cases shown in the prior figures are included again here for reference.

The framework described above could be used to further develop empirical relationships between stronger trends, trended and de-trended variance, and p_{eff} for SSMP flow. However, the cases shown in Figure 5-2 were sufficient for the current objective, which was to assess whether the chosen linear trend was large enough to influence T_{eff} values, and to know what that effect would be in the steady-state mean-parallel case, before advancing to transient stream depletion simulations in the same non-stationary fields.

5.3 Trending Mean Transmissivity in Transient Stream Depletion

Up to this point, no cases have been presented with significant differences in T_{eff} behavior between transient stream depletion and steady-state mean-parallel flow cases. Such comparison is problematic, however, for a field with a spatial trend in transmissivity. A strictly defined effective transmissivity value does not exist for the non-stationary case (Sanchez-Vila et al 2006). Instead, a pseudoeffective value must be defined that depends on location within the flow domain.

Defining the area of influence, and therefore the transmissivity averaging area, is straightforward in the non-stationary SSMP flow simulations—it is the entire area between the two head boundaries, and all sub-areas of the transmissivity field have equal influence on global flow behavior. In contrast, it is not straightforward for the transient stream depletion case. Hydraulic analysis and empirical tests suggested the area of primary influence in the transient stream depletion simulations is near the well along the river boundary (Section 3.6), but that area does not have sharp or specific extents. Influence decreases with increasing distance from the well and river.

The precise shape and location of the chosen weighting area (Figure 3-1) was not significant in the simulations presented so far. In principle, this is because the fields were stationary and relatively homogenous statistically. And in practice, this was confirmed through sensitivity simulations which showed the weighting area only affected the ensemble variance and not the ensemble mean, i.e., not T_{eff} itself.

In contrast, the choice of weighting area was found to dominate results for transient stream depletion when a non-stationary mean transmissivity was present. This was found through a series of non-stationary stream depletion simulations. First, if the weighting area used

previously (Figure 3-1) was expanded to include a small portion of the area between the well and the impermeable boundary, where the mean transmissivity was lower, then the resulting p_{eff} was greater than +1.0, meaning it was greater than the arithmetic average of that expanded weighting zone. Exceeding the arithmetic bound is in conflict with theory in general and specifically with analytical work on other non-stationary cases (e.g., Rubin and Seong 1994; Indelman and Rubin 1996). That suggests the expanded averaging area reached too far beyond the relevant area and/or should be weighted more toward the region between the well and the river where mean transmissivity was higher.

Next, when the weighting area was the same as used in the prior stationary cases (Figure 3-1), then p_{eff} was slightly above the trend established for those stationary cases. For example, p_{eff} was slightly but distinctly above the geometric mean under mild geometric anisotropy. Finally, if the averaging zone was narrowed slightly such that the semi-minor axis of the averaging-area ellipse included 90% of the distance between the river and the well, then p_{eff} corresponded more closely to the trend for the stationary cases shown in Figure 5-2, such as close to the geometric mean for cases with mild geometric anisotropy.

It is worth reviewing that an equivalent transmissivity, T_{eq}, was identified from the depletion response curve and related to the distribution of the realized transmissivity field through the power-mean function exponent (Section 3.5.4). That response curve, the corresponding T_{eq} value, and the transmissivity field is the same in each of these test cases; only the power-mean exponent was changing. The power-mean exponent corresponding to the interpreted T_{eq} changed as the averaging area changed because the non-stationary field's properties changed in space.

To approach this p_{eff} sensitivity to the weighting area differently, one could presume that p_{eff} for both steady-state mean-parallel (SSMP) flow and transient stream depletion might still be similar in a non-stationary field—a presumption based on the similarity in the stationary simulations—and therefore adjust the weighting area until the non-stationary p_{eff} results are similar between stream depletion and SSMP simulations. When taking that approach, p_{eff} behavior for non-stationary stream depletion simulations was close to the non-stationary SSMP cases when the weighting area's semi-minor axis was decreased to include about 80 to 85% of the distance between the river and the well.

The non-stationary stream depletion scenarios described above are not included in Figure 5-2 since they were so sensitive to arbitrary tests of the area of influence. These observations may not be directly transferrable, quantitatively, to cases with trends of different magnitudes or fields with different variance levels. However, even with these limitations, the results clearly illustrate the strong sensitivity of p_{eff} to how the pseudo-effective transmissivity area is defined when there is a realistic spatial trend in mean transmissivity. If it is accepted that the p_{eff} must lie between the arithmetic and harmonic bounds for this case, then these results bracket the outside extent of the averaging zone under the simulated trend and variance levels.

5.4 Zonal Anisotropy

With geometric anisotropy, the correlation range varies with direction, but the variogram sill, i.e., the uncorrelated variance, is isotropic. With zonal anisotropy, the sill level is anisotropic. Zonal anisotropy may represent distinct lithologic layering, which would be common in a vertical profile, such as an alternating sand, clay, and silt depositional sequence. Zonal anisotropy is conceivable in a horizontal profile where strata have been deformed and exposed. Zonal anisotropy can also be an artifact of a trending mean transmissivity, an artifact

which may fade once the data have been de-trended (Gringarten and Deutsch 2001). That latter situation was apparently the case for several locations considered in Section 6.2.

Four ensembles with zonal anisotropy were simulated. Transmissivity fields exhibiting zonal anisotropy were generated by using two nested spatial correlation functions (Deutsch and Journal 1998). The variogram sills for orientations parallel and perpendicular to the river were approximately 0.25 and 0.75, respectively. That case is shown in Appendix A.

Results from the four ensembles are plotted in Figure 5-2. As shown, the p_{eff} results were insensitive to the degree of geometric anisotropy, remaining near $p_{eff} = -0.75$ for geometric anisotropy ratios (λ_x/λ_y) ranging from 1 to 10. All four ensembles therefore behaved as strongly "stratified" cases, including those with isotropic correlation ranges.

5.5 The Homogeneous Anisotropic Case

Geometric anisotropy is a *geostatistical* anisotropy that is used to describe the structure of aquifer heterogeneity. It is based on a directionally dependant correlation of local-scale transmissivity zones, even when those zones may be isotropic locally. As suggested by Gelhar and Axness (1983), geometric anisotropy leads to hydraulic anisotropy at a larger scale, but it is not the same physical phenomena as local hydraulic anisotropy.

There are interesting differences to consider between these two forms of anisotropy. A regional-scale anisotropic transmissivity is not necessarily feasible to measure or test. At the spatial scale that is assessed by a typical aquifer test, the tested area can be isotropic, yet the larger spatial structure of the collection of isotropic zones can lead to anisotropy in regional-scale transmissivity. It is also possible in principle to have regional-scale hydraulic anisotropy that is the combined result of local-scale hydraulic anisotropy plus overall geometric anisotropy. In that case, the two behaviors would be superimposed (Sanchez-Vila et al. 2006).

If it cannot be directly tested, regional-scale anisotropic transmissivity attributable to geometric anisotropy might be estimated only by: (i) a numerical model populated with a transmissivity field interpreted from local-scale estimates or (ii) estimated from geometric anisotropy that is inferred from constructing variograms (Sanchez-Vila et al. 2006). The second case would likely be combined with geologic interpretation, such as considering depositional history.

Regional-scale hydraulic anisotropy is a nonlinear function of geometric anisotropy. Using the directional effective permeability estimate of Gelhar and Axness (1983), which was based on geometric anisotropy, Sanchez-Vila et al. (2006) noted the following relationship:

$$\frac{T_x}{T_y} = exp\left(\frac{1-e}{e+1}\,\sigma_Y^2\right) \tag{5-1}$$

in which e is a geometric anisotropy ratio defined as λ_y/λ_x (≤ 1). Note that this ratio is inverted from the convention used elsewhere herein. A plot of this relationship is shown in **Figure 5-3**.

The regional-scale hydraulic anisotropy ratio (T_x/T_y) that develops from geometric anisotropy is a function of variance, σ^2_Y (Equation 5-1). Logically, as the contrast between local-scale transmissivity zones increases, which increases the field variance, the T_x/T_y contrast that arises from the structural organization of those contrasting zones will also increase. Or, to approach the dependence on variance from another perspective: It can be shown with Equations 5-1 and 2-5 that, as geometric anisotropy increases, the T_x/T_y ratio approaches the ratio between the arithmetic and harmonic means of the local transmissivity values. The ratio between the arithmetic and harmonic means is a function of field variance (Section 2.2.4).

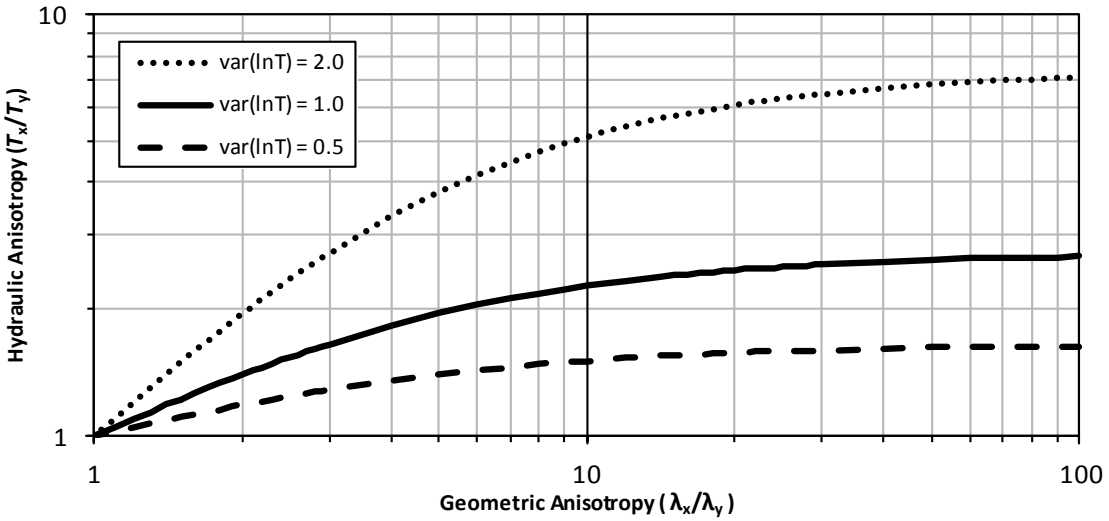

Figure 5-3. Nonlinear relationship between hydraulic anisotropy and geometric statistical anisotropy.

The power-mean function partly normalizes for field variance (Section 2.2.4). The p_{eff} is therefore a function only of the system flow behavior. Considering that p_{eff} for stream depletion timing has been shown to have only a mild dependence on field variance (Figure 4-4), while Equation 5-1 shows hydraulic anisotropy to be a function of field variance, this raises the question if hydraulic anisotropy is a factor in stream depletion timing.

The shape of the cone of depression around a pumping well is distorted by hydraulic anisotropy (T_x/T_y). It may seem intuitive that a depression distorted to be wider in the direction parallel to the river (which arises in the $T_x > T_y$ case) would interact more slowly with the river relative to the isotropic case. And, vice versa, that it would interact more quickly with the river if distorted toward the river (the $T_y > T_x$ case). This concept was tested through model runs in which only homogeneous anisotropy was simulated.

In the homogeneous anisotropic simulations, stream depletion timing was dependent only on the transmissivity in the direction perpendicular to the river (T_y). It was insensitive to

transmissivity in the parallel direction (T_x) and thus insensitive to hydraulic anisotropy ratios (T_x/T_y). It was therefore concluded that stream depletion timing is sensitive to geometric anisotropy due only to the fact that geometric anisotropy influences T_{eff} in the y-direction.

Transmissivity in the T_x direction did indeed influence the width of the cone of depression and thereby influenced the *length* of the affected river reach, even while the total-river depletion timing was not affected. That spatial distribution of stream depletion can be relevant in conjunctive water management, such as in estimating the portion of a well's depletions that impact the river above a critical point like an intervening water right, or a protected habitat zone, etc.

Figure 5-4 compares spatial depletion rates along the river for two cases of homogenous anisotropy and an isotropic reference case. In the simulations used to produce this plot, it was apparent that the shape of the cone of depression was dependent on T_x and, as shown, the length of the affected river was influenced by T_x. However, the total area under each spatial depletion curve (Figure 5-4), and thus the total depletion rate at a given time, was the same in each case. The depletion was simply concentrated nearer the well for $T_x/T_y < 1$ and was more diffuse for $T_x/T_y > 1$. It was also found that the length of the affected river reach, which can be computed analytically for the isotropic case (Glover 1977), can be successfully scaled by the square root of the hydraulic anisotropy ratio, $(T_x/T_y)^{½}$, as is done in other applications for anisotropic conditions.

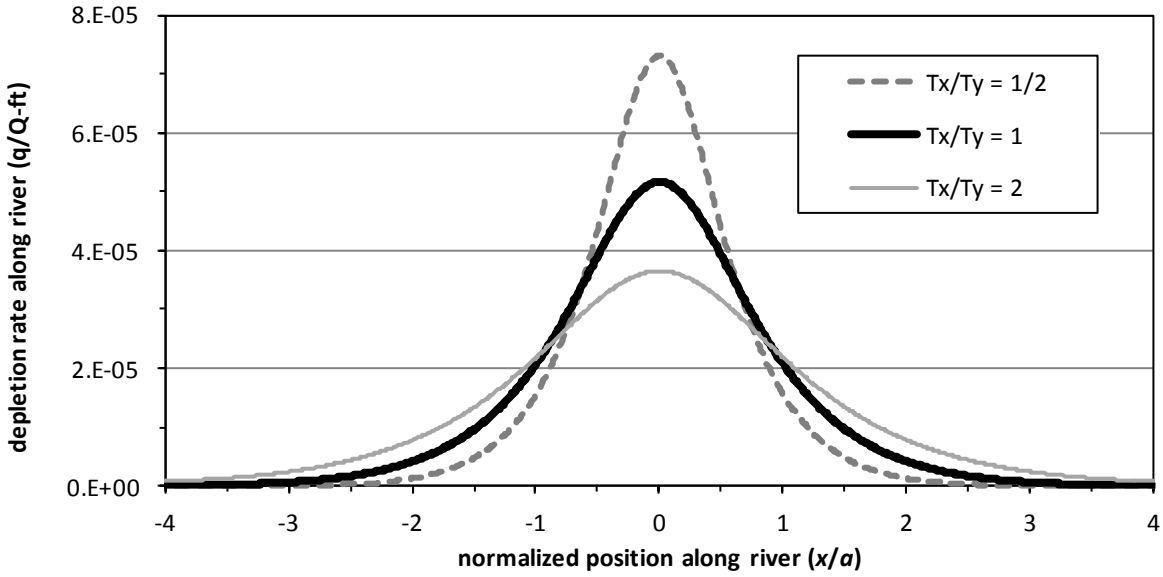

Figure 5-4. Spatial distribution of depletion along river at time t/t' = 1 for three different hydraulic anisotropy ratios (T_x/T_y). Position along river (x) scaled by well-to-river distance (a). The total river depletion rate (area under the curve) is identical for each case.

CHAPTER 6 FIELD EXAMPLE and APPLICATION LIMITATIONS

6.1 Example Geometric Anisotropy Ranges

Extensive lists of correlation ranges for various aquifers have been compiled in publications, including examples of horizontal correlation ranges at the regional scale (e.g., Gelhar 1993; Hoeksema and Kitanidis 1985). Statistical anisotropy for the vertical plane (horizontal to vertical correlation ratio, λ_h/λ_v) can be found in many reports as well. In contrast, published data about statistical anisotropy in the horizontal plane (λ_x/λ_y) appears to be uncommon.

One publication noting statistical anisotropy in the horizontal plane was a regional-scale geostatistical study by Welhan and Reed (1997) around the Idaho National Energy and Environmental Laboratory (INEEL) in the Snake River Plain aquifer. They reported a regional-scale correlation (up to 6 km) with a horizontal statistical anisotropy ratio of 4/1, oriented southeast toward the river plain. The scale of correlation and the orientation toward the river were attributed to lava flows and volcanic structural features. The INEEL site is in a complex geologic setting where groundwater flow is suggested by Welhan and Reed to be dominated by connectivity of high-contrast permeability zones, so the applicability of the Gaussian fields used in this study to the INEEL site may be limited. Still, the INEEL study had a large spatial scale, giving it some relevance as a horizontal geostatistical anisotropy example.

In further geostatistical work for INEEL, Welhan et al. (2002) commented that although the site is primarily a volcanic aquifer system, the spatial distribution and architecture of interflow zones, referring to relatively thin but laterally extensive and highly permeable zones between stacked lava flow lobes, "bear a resemblance to the distribution of coarse fluvial facies in sedimentary aquifers." Their 2002 study was focused on smaller-scale correlations (e.g., 100

m correlation lengths) and used indicator variogram analysis. At that scale, they reported horizontal statistical anisotropy at 1/1 for sediments and 7/1 for the interflow zones, with an east-west orientation.

Due to finding only one published study documenting large-scale horizontal statistical anisotropy, geostatistical correlation structures for five areas of the South Platte River alluvial aquifer in Colorado are examined in this chapter.

The headwaters of the South Platte River are located in the Rocky Mountains southwest of Denver, Colorado. The river exits the mountains near Denver and flows north and east roughly 200 miles across the plains to the northeast corner of Colorado. It meets the North Platte River in west-central Nebraska. It is generally a shallow, braided river with a sand and gravel streambed (Lindsey et al. 2005; SPDSS 2006a).

The South Platte River alluvial aquifer is a valley-fill aquifer lying in paleochannels eroded into the Pierre Shale for much of its length in Colorado, with western reaches of the alluvium in buried valleys in the Fox Hills sandstone, the Laramie Formation, and other units of the Denver basin. The aquifer is composed of unconsolidated gravels, sands, silts, and clays. Significant portions consist primarily of glaciofluvial sand and gravel deposits (Bjorklund and Brown 1957; Smith et al. 1964; Tweeto 1979; Aikin et al. 2000; Topper and Wilson 2003; Lindsey et al. 2005).

The aquifer's saturated thickness is typically in the 20 to 40 ft range near Denver, and up to 200 ft thick in northeast Colorado. The aquifer's width varies between roughly two to six miles. Typical transmissivity is around 15,000 ft^2/day in the thinner reaches near Denver, and over 100,000 ft^2/day in the thicker deposits in northeast Colorado (Hurr et al. 1972a, 1972b,

1972c). High-capacity groundwater wells are used extensively in the South Platte basin for irrigated agriculture.

Transmissivity maps for five areas selected for geostatistical evaluation are shown in Figures 6-1 through 6-5. These maps were constructed from the GIS files *tgrid1106* and *tgrid0309* which were developed by the South Platte Decision Support Study (SPDSS) using the Hurr and Schneider work (1972a, 1972b, 1972c, and others in the series), combined with other data sources, and interpolated into a 100-meter-grid data file (SPDSS 2006b, 2006c, 2009, 2013). The river configuration is from a GIS file digitized by Colorado State University from USGS maps (CSU 2006; Hurr and Schneider 1972a, 1972b, 1972c, and others in the series).

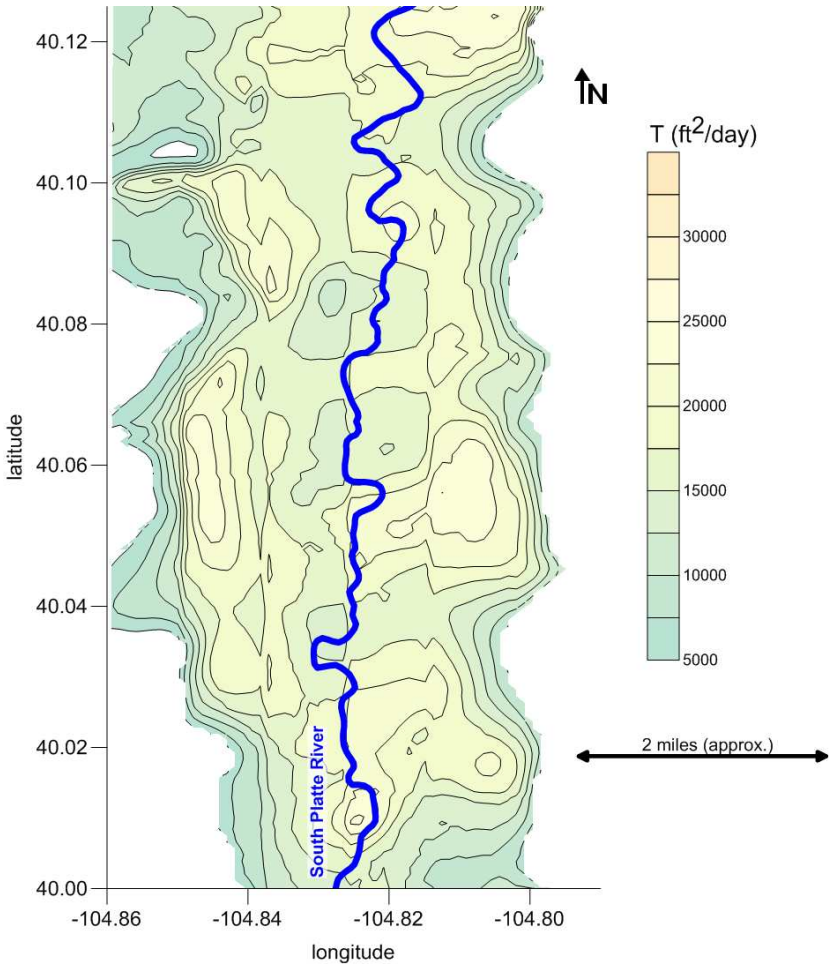

Figure 6-1. Transmissivity of the South Platte River alluvial aquifer near Fort Lupton, Colorado.

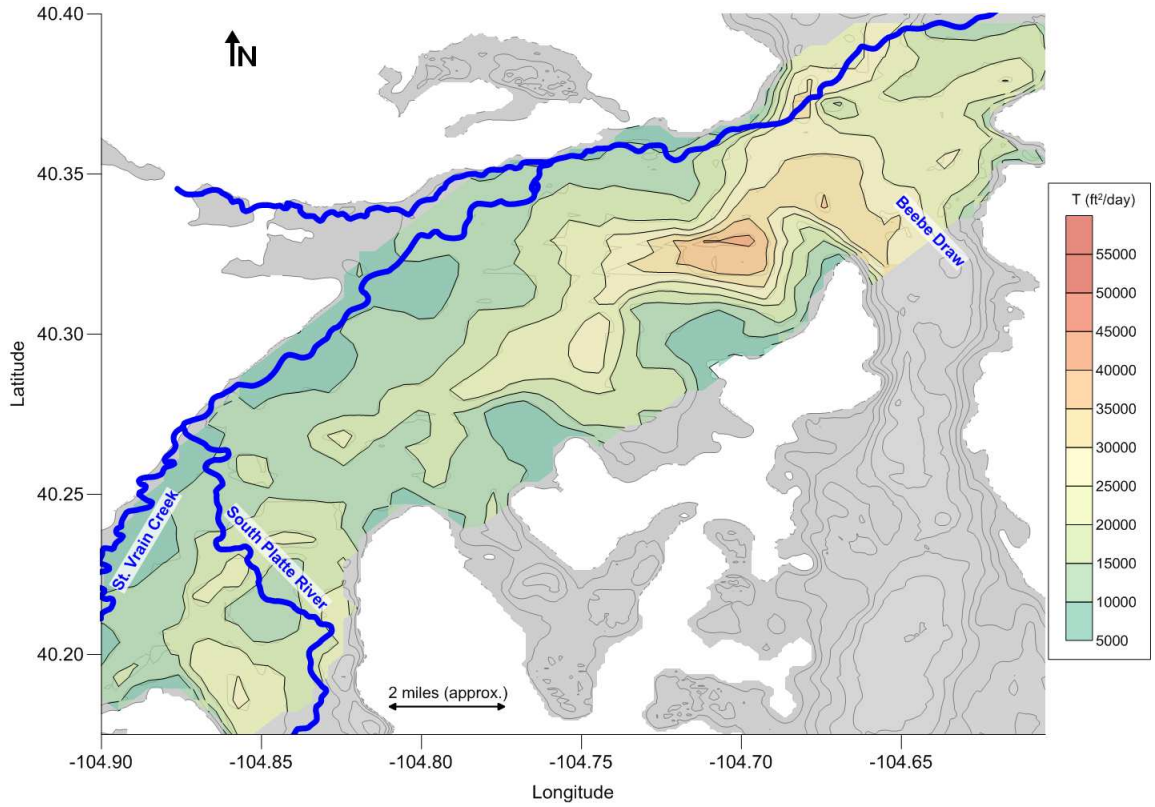

Figure 6-2. Transmissivity of the South Platte River alluvial aquifer near Greeley, Colorado.

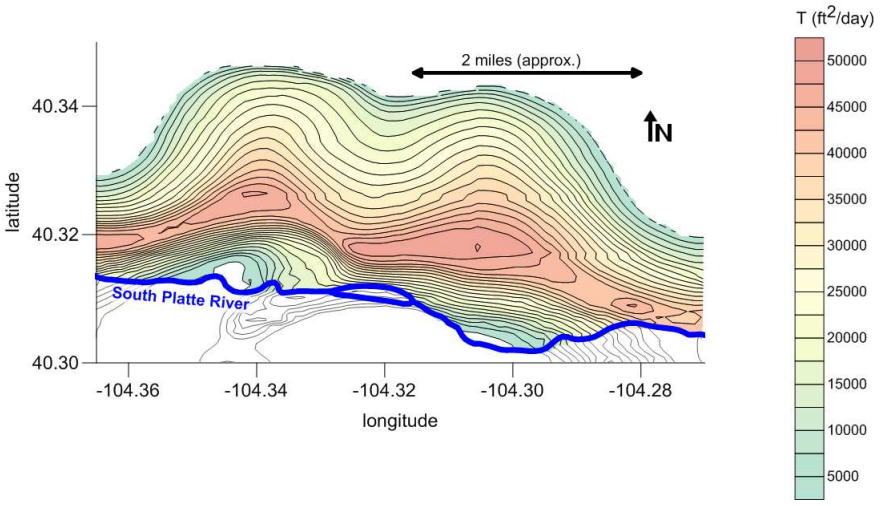

Figure 6-3. Transmissivity of the South Platte River alluvial aquifer near Riverside Reservoir, Colorado.

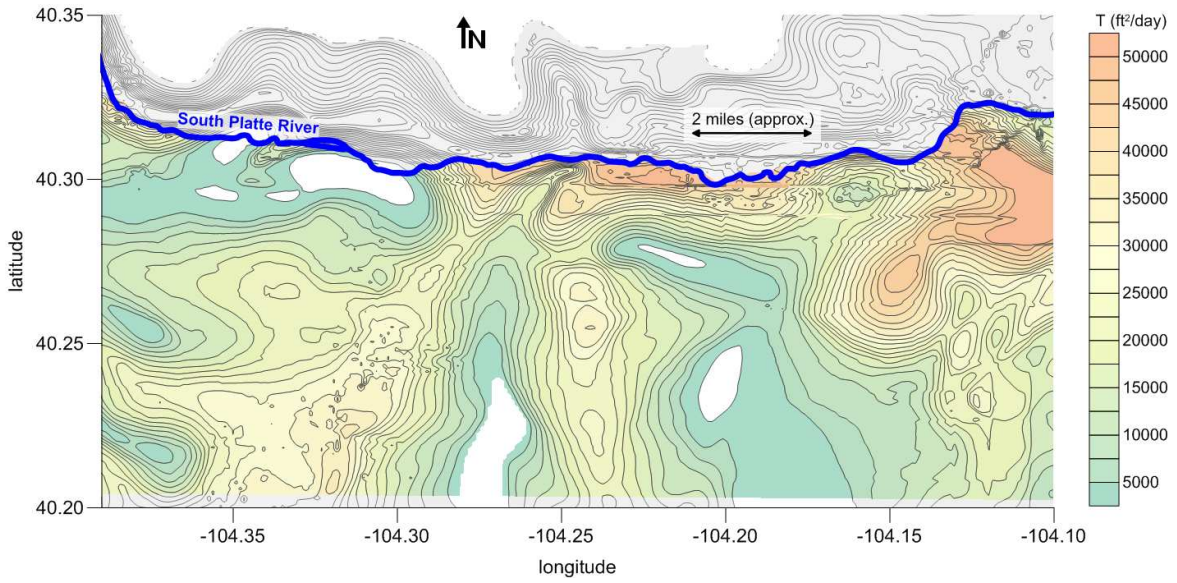

Figure 6-4. Transmissivity of the South Platte River alluvial aquifer around Lost Creek, Colorado.

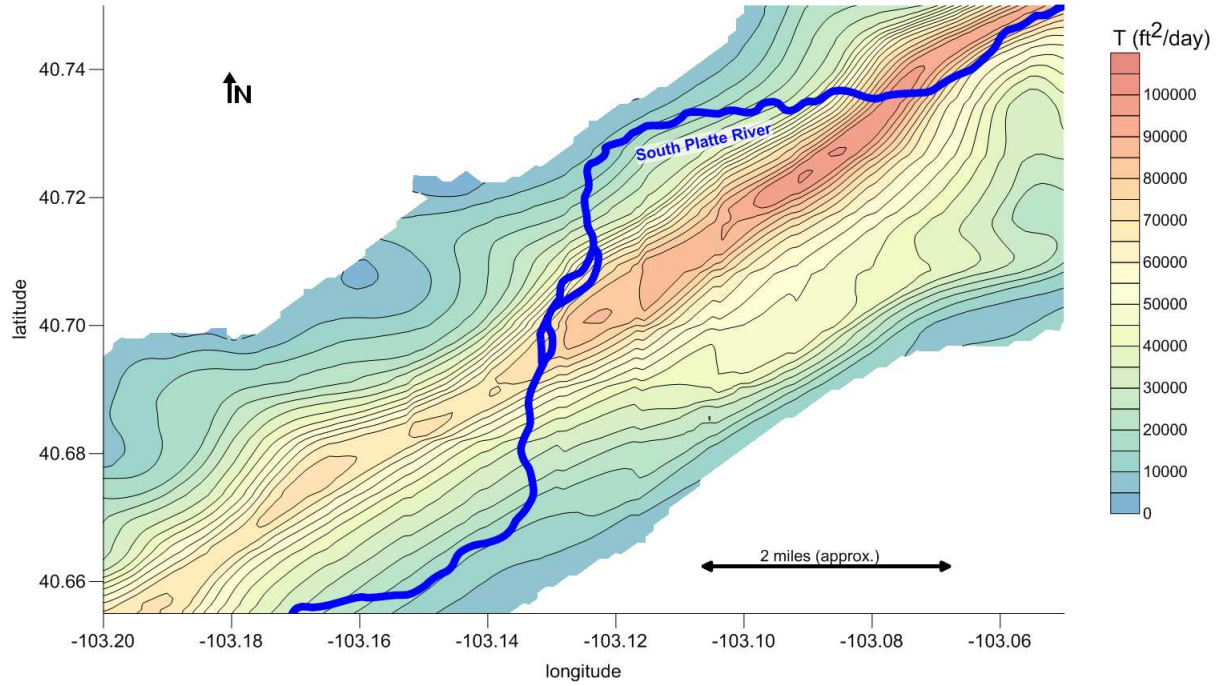

Figure 6-5. Transmissivity of the South Platte River alluvial aquifer near Iliff, Colorado.

Two programs were used to construct and analyze experimental variograms for the five South Platte areas: *Surfer* (version 11.6) by Golden Software, Golden, Colorado; and the Stanford Geostatistical Modeling Software, *SGeMS* (Remy et al. 2009). **Table 6-1** lists statistical anisotropy ratios and principal directions estimated from the experimental variograms for the five South Platte areas shown in Figures 6-1 through 6-5. **Table 6-2** lists the same information for when the fields were detrended.

Table 6-1. Estimated spatial correlation parameters and modeled power-mean exponents for five areas along the South Platte River. Power-mean exponents are shown for both transient stream depletion (TSD) and steady-state mean-parallel (SSMP) flow.

General Area	$\lambda_{max}/\lambda_{min}$	λ_{max} azimuth (from north)	λ_{max} (ft)	river azimuth (from north)	TSD p_{eq}	SSMP p_{eq}
Fort Lupton	4 to 7, greater	10°	10,000	0°	inconclusive	0.3
Greeley	2 to 4, greater	45°	13,000	45°	-0.5 to 0.5	-0.4 to 0.9
Riverside	>3	110°	>domain	110°	-0.1	0.1
Lost Creek	3 to 4	20°	15,000	90°	0.1 to 0.4	0.1
Iliff	>10	45°	>domain	12°	>1.0	0.5

Table 6-2. Estimated spatial correlation parameters for the five areas along the South Platte River after detrending.

Approximate Area	$\lambda_{max}/\lambda_{min}$	λ_{max} azimuth (from north)	λ_{max} (ft)
Fort Lupton	-	-	-
Greeley	near 1	45°	13,000
Riverside	-	-	-
Lost Creek	2 to 3	10°	15,000
Iliff	1.5 to 4	35°	5,000 to 9,000

Caveats

It is noted that experimental variograms, meaning those that are created from site data and not mathematical models of correlation, can be sensitive to construction decisions, such as choosing data-pair lag distances, data search criteria, etc. The development of representative variograms is not a straightforward process for some data sets, with solutions that may not be unique (Isaaks and Srivistava 1989). And, there are no robust or objective measures of validity for a given variogram (Goovaerts 1997). With those concepts in mind, the following observations were made during variogram construction. No obvious sensitivity to the lag search parameters was found when different tolerances were tested for both lag distance and search angles, but visual assessment of the transmissivity maps suggest there are irregular-shaped, large-scale correlation patterns at some locations (e.g., Figure 6-3) that might not be well-captured in directional variograms that have a standard type of search window. The constructed variograms also suggested short-scale correlation patterns nested within patterns at a scale larger than the domains tested. Additionally, different degrees of influence from non-stationarity in transmissivity were apparent among the five cases. For these reasons, the geostatistical correlation parameters produced for the selected South Platte areas may not be fully representative of those areas and are not the only possible interpretation.

A few other observations were noted for the results shown in Tables 6-1 and 6-2. Milder statistical anisotropy was found in some of the variograms than might be expected based on a visual assessment of the transmissivity maps. The Fort Lupton area transmissivity field was not de-trended due to the relative homogeneity of the area and since there wasn't a clear trend except near the aquifer margins. A clear variogram was not identified for the Greeley area once the field was detrended. The Riverside area appears highly correlated (Figure 6-3), but the

sinusoidal pattern is likely not captured by the typical directional variogram used. The Riverside area was not de-trended due to the change in trend direction near the river. Variograms for the Lost Creek area exhibited zonal anisotropy, for which detrending had a minor effect.

6.2 Site Application of Equivalent Transmissivity

Five numerical groundwater flow models, and several variants of those five, were built with the MODFLOW-2000 code (Harbaugh et al. 2000) for the areas shown in Figures 6-1 through 6-5. Transient stream depletion rate response curves were simulated for pumping wells placed generally near the middle of the domains shown in the figures. Steady-state mean-parallel (SSMP) flow conditions were also simulated for these domains by placing an artificial straight head boundary near the impermeable aquifer boundary to impose a fixed gradient across the aquifer toward the river. The SSMP simulations were performed since the transmissivity averaging area for that case would be less ambiguous than for the pumping well depletion case, and because the SSMP test would influence a larger portion of the aquifer, giving a more comprehensive estimate of p_{eq} for the area.

The modeled depletion response for the transient pumping simulations, and the total flow rates induced under the SSMP tests, were used with the Glover equation and a simple Darcy calculation, respectively, to compute a T_{eq} for each simulation. As was done in the synthetic-case simulations, a T_{eq} was related to the power-mean statistics from each region's transmissivity field to obtain a p_{eq}. The averaging area for the stream depletion case was based on the area expected to be most influenced by the pumping well (Section 3.6.1). For the SSMP case, it was the entire area between the artificial head boundary and the river.

The model-estimated p_{eq} results for the five South Platte areas are listed in Table 6-1. The results were highly variable, and without a clear relationship to geometric anisotropy ratios

of the fields. Factors contributing to the lack of correlation with geometric anisotropy are discussed in the next section. Observations made during the evaluation process are noted first.

For the Fort Lupton area, a clear p_{eq} was indeterminable for the transient stream depletion case due to the relative homogeneity of the region. For several tested subareas, there was only a 2% difference between harmonic and arithmetic means such that a best-fit p_{eq} between the two bounding means could not be resolved, and wouldn't have practical meaning anyway.

Qualitatively, the Greeley area (Figure 6-2) looks to be the least statistically anisotropic of the five areas examined. Quantitatively, its variogram did indeed indicate relatively mild anisotropy. The mild anisotropy may explain why the stream depletion results were generally around the geometric mean (p_{eq} = -0.5 to 0.5). However, the results were also sensitive to the location tested and the size of the averaging area, in both the TSD and SSMP tests. In particular, the higher-transmissivity zone located toward the northeast appeared to influence the depletion response for wells placed near that area, even if the wells were not directly within the zone. That high-transmissivity zone similarly influenced the SSMP case, as was noted when the size of the stressed area was varied.

That zone of higher transmissivity in the Greeley area is located where the Beebe Draw alluvium merges with the South Platte alluvium. The Beebe Draw is thought to be an ancestral route of the South Platte River (Smith et al. 1964). It can also be seen in Figure 6-2 that the principal direction of anisotropy is oriented toward the river in that area between the end of Beebe Draw and the river, whereas the areas further west have structure oriented parallel to the river. This last observation highlights the spatial variability of aquifer characteristics that influence appropriate T_{eq} averaging.

For the Riverside case, considering the irregular shape of visually apparent correlation, it is likely that the anisotropy ratio was not well-captured in the variogram modeling. Riverside is also a case with a strong trend toward the river, a trend which also reverses before the river. Such trends introduce sensitivity to the choice of spatial weighting area (Chapter 5). Effective transmissivity results were very close to the geometric mean (-0.1 to +0.1) though this was sensitive to the tested area and the averaging area, and is therefore a low-confidence estimate.

The Lost Creek region is different from the other cases because, for much of the area, the principal correlation direction is oriented toward the river. For this analysis, the upper reaches of Lost Creek itself were neglected as a relevant stream boundary. Lost Creek reaches were assumed to be ephemeral or have head-invariant leakage rates, leaving the South Platte River as the affected stream boundary. The p_{eq} results for the Lost Creek case are roughly consistent with a correlation axis oriented toward the river, as results were above the geometric mean (p_{eq} = 0.1 to 0.4).

The region southwest of Iliff was chosen because, unlike the other cases, a reach of the river there is oriented at an oblique angle to the principal axis of correlation. Based on that angle and the results of Section 4.5, p_{eq} was expected to be close to the geometric mean. That, however, was not the outcome of the Iliff flow simulations. The model-estimated p_{eq} was high, between the arithmetic and geometric means for the SSMP simulation, and above the arithmetic mean for the TSD case. The pumping well in that case was located slightly south of the high transmissivity zone in the middle of the aquifer. Further examination of both the TSD and SSMP flow simulations suggested the high transmissivity zone near the middle of the simulated area strongly influenced the stream response, a characteristics that was not well-captured by averaging transmissivity over a wider area.

6.3 Issues with Site Application for Transient Stream Depletion

There are at least two limitations in applying the results of Chapter 4 to the five South Platte areas modeled herein: (i) the large scale of correlation relative to the scale of the problem and (ii) the spatial trends in mean transmissivity. The relationships between p_{eff} and geometric anisotropy presented in Chapter 4 apply to expected (mean) behavior, but it is also clear that there is significant variability among individual field realizations (Figure 4-2). That variability increases as the correlation scale becomes large relative to scale of the flow problem.

The effect of large correlation scales is shown in **Figure 6-6** which compares p_{eq} histograms for two ensembles that have the same geometric anisotropy (5/1) but different correlation scales. Example field realizations for these two cases are illustrated in **Figure 6-7**. One case has a correlation length equal to 50 model nodes (2,500 ft) and the other equal to 200 model nodes (10,000 ft). Mean behavior from the two ensembles were roughly similar to each other (p_{eff} at -0.54 versus -0.38) and similar to the curves shown in Chapter 4. (Note the large-scale correlation cases plotted in Figure 4-1.) Yet, Figure 6-6 shows there is poor probabilistic predictability for individual realizations in the case with the large correlation scale.

The standard deviations of equivalent power-mean exponents (p_{eq}) from various ensemble simulations are plotted in **Figure 6-8** and **Figure 6-9** and as a function of correlation scale. The correlation range was scaled by the distance between the well and the river. Figure 6-8 is for ensembles with the principal correlation direction oriented parallel to the river, and Figure 6-9 is for the principal direction oriented perpendicular to the river. Considering that p_{eff} ranges only from -1.0 to +1.0 for the bounding harmonic and arithmetic means, respectively, it is clear that when the correlation range is large relative to the scale of the problem then there is significant variability relative to the bounding means for individual realizations.

A few other observations can be made from these figures. As noted in Chapter 3, applying the spatial weighting area did not affect p_{eff} results but it did reduce ensemble variability. The reduced variability is apparent in Figure 6-8 which shows the larger ensemble standard deviations in the unweighted cases compared to the weighted cases, particularly for the ensembles with $\lambda_x/a > 2$.

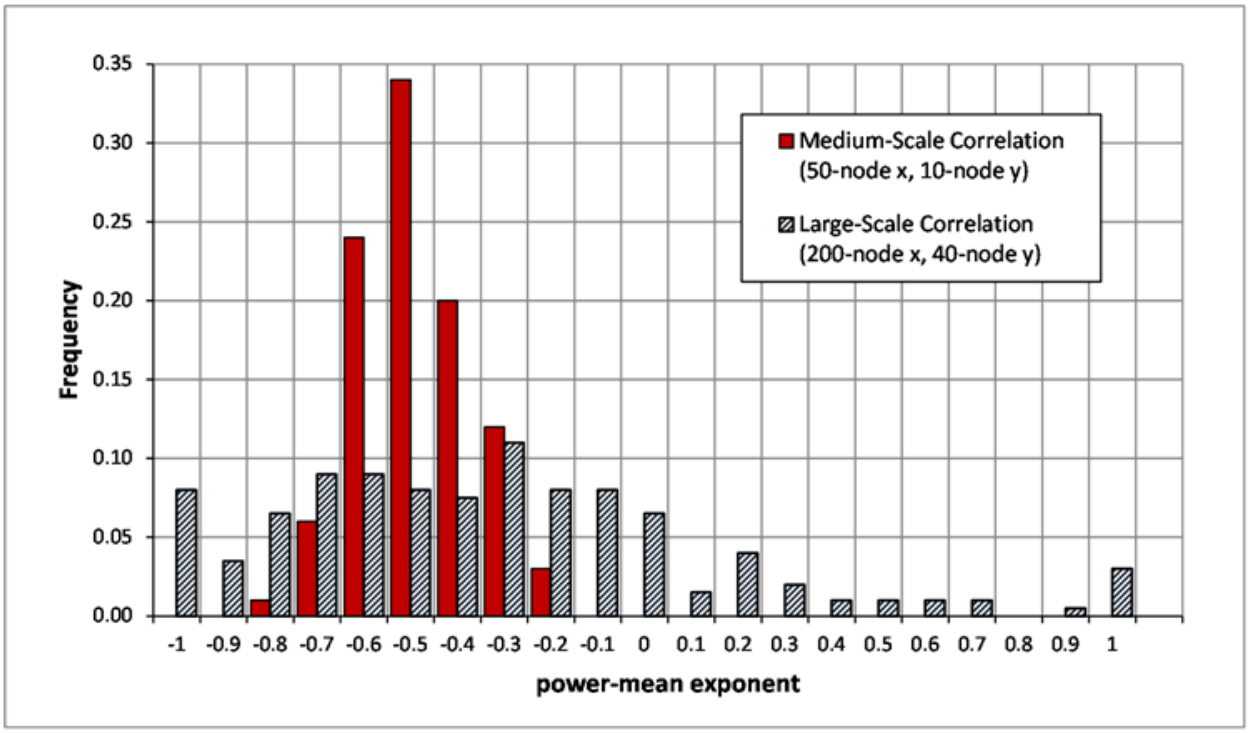

Figure 6-6. Distribution of power-mean exponents for two ensembles with the same statistical geometric anisotropy (λ_x/λ_y =5/1) but different absolute correlations scales relative to the same well-to-river distance.

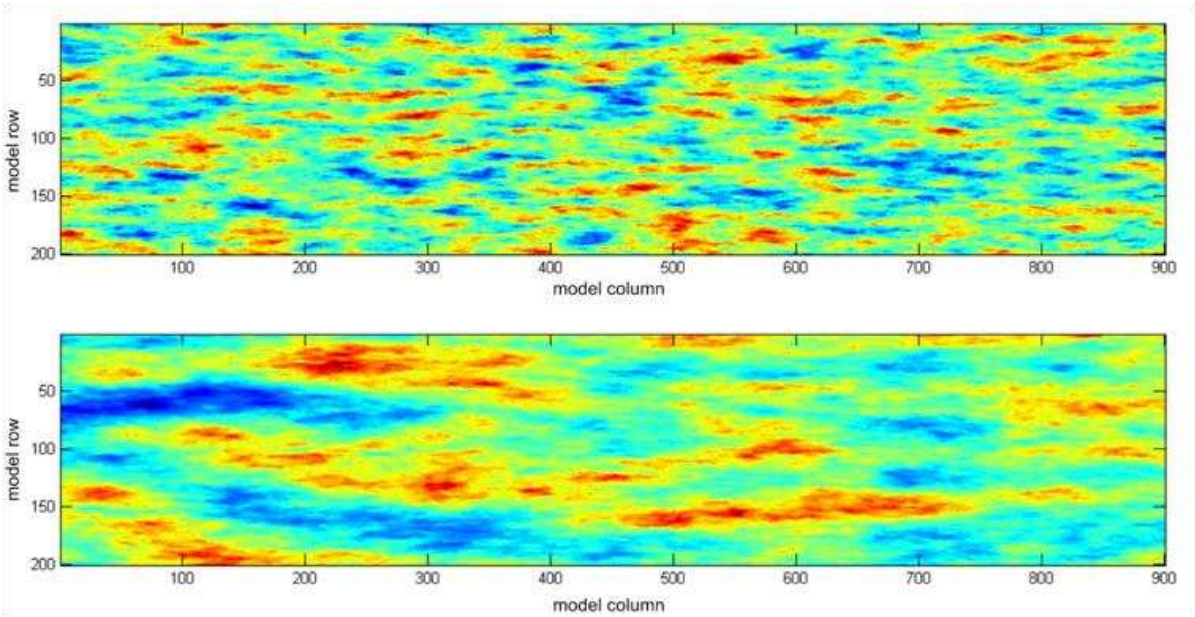

Figure 6-7. Example transmissivity field realizations for the same statistical geometric anisotropy (5/1) but different absolute correlation scales. Pumping well is at center (450, 100). Correlation scale is 50x10 nodes in the top field and 200x40 nodes in the bottom field.

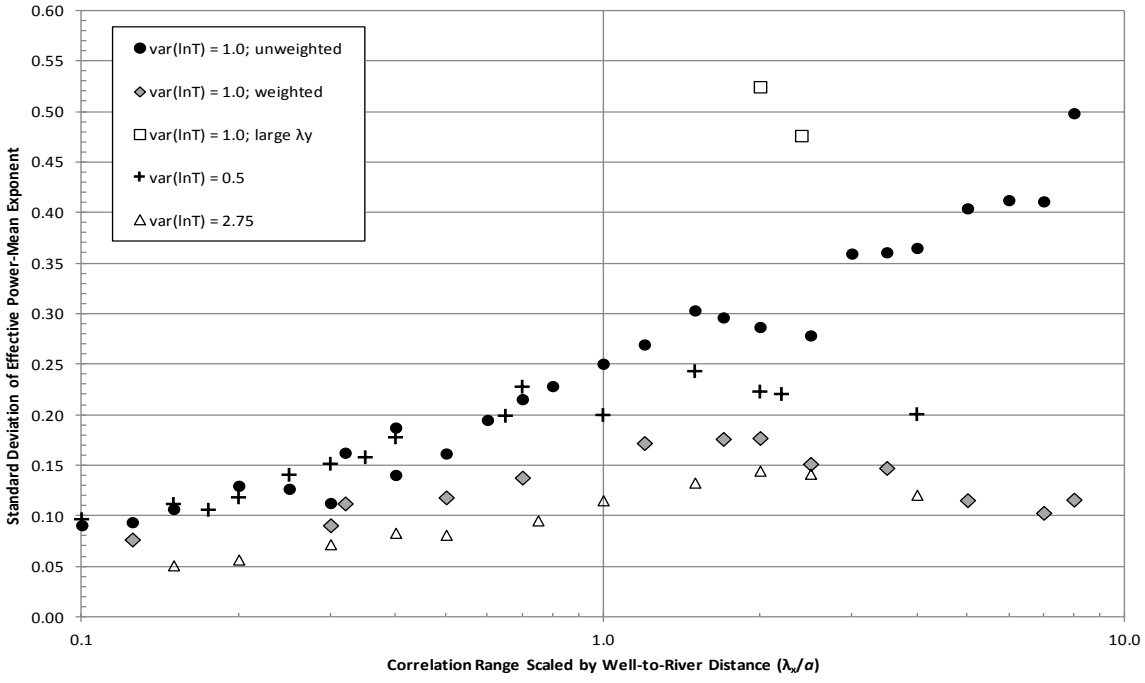

Figure 6-8. Standard deviations of power-mean exponent ensembles as a function of correlation scale. Plotted cases have principal correlation direction oriented parallel to river.

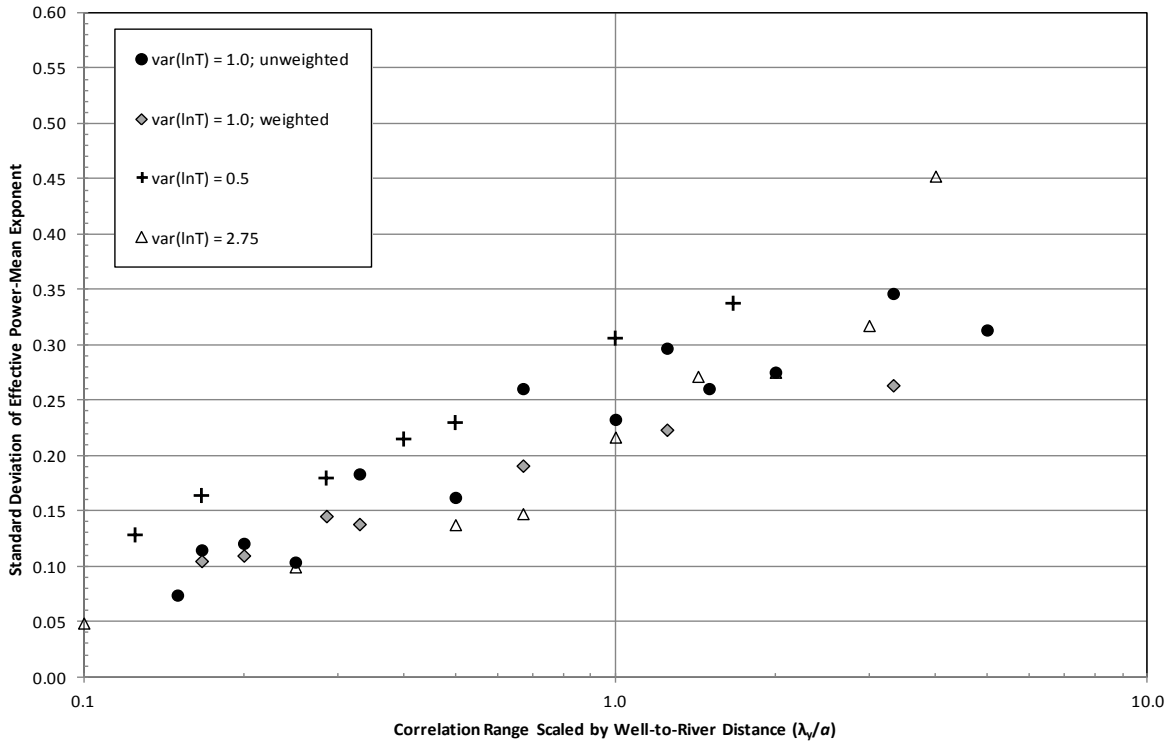

Figure 6-9. Standard deviations of power-mean exponent ensembles as a function of correlation scale. Plotted cases have principal correlation direction oriented perpendicular to river.

Two cases are plotted in Figure 6-8 for which the correlation scale in the minimum principal direction (λ_y) was increased relative to the other cases. These are the two large-scale correlation cases also plotted in Figure 4-1 and Figure 4-2, and includes the large-scale case shown in Figure 6-6. The other simulations in Figure 6-8 had equal correlation ranges in the minimum principal direction, which was the 10-node minimum range allowable to avoid grid discretization bias (Section 3.6.2). For the two large-scale cases, the minor direction correlation range was increased to 40 nodes, thus from 10% to 40% of the distance between the well and the river. As seen in Figure 4-1, Figure 4-2, Figure 6-6, and Figure 6-8, the increase in correlation scale in the minor direction significantly increased ensemble variability.

An increase in ensemble variability with increased correlation scale is not novel (Deutsch and Journel 1998). What is notable is that transmissivity correlation scales are commonly quite large (Rubin 2003; Anderson 1997; Dagan et al. 2009), which leads to correlation commonly being large relative to the scale of the problem in transient stream depletion.

This issue of predictive uncertainty being a function of problem scale and correlation scale was a clear difference in the steady-state mean-parallel (SSMP) simulations with synthetic fields. For SSMP flow, the entire domain was the relevant flow field, rather than only that portion affected by the pumping well. Variability, such as that shown in Figure 6-8 and Figure 6-9, was much reduced for the SSMP cases, due to the larger relative scale of the flow problem. Variance of p_{eq} did increase as the correlation scale increased even for the SSMP cases, but it remained minor. For example, the SSMP p_{eq} standard deviations ranged from only 0.02 to 0.04 over the same range in correlation scales as shown for the transient stream depletion cases in the previous figures.

In the South Platte River alluvial aquifer examples considered in this chapter, the transmissivity correlation range was generally larger than the scale of the problem as measured by the distance between the well and the river (i.e., $\lambda_{max}/a > 1$). In that case, it appears ergodicity cannot be reliably assumed, meaning that overall expected behavior may not be a good predictor of individual site behavior. Large-scale correlation was not the only limitation to applying the stochastic approach—statistical inhomogeneity and non-stationarity were also complicating factors for these non-synthetic transmissivity fields—but the transmissivity correlation scale is an important consideration.

CHAPTER 7 DISCUSSION and SUMMARY

7.1 Research Summary

The goal of this research was to evaluate effective transmissivity for the particular case of transient stream depletion caused by pumping a groundwater well in two-dimensional, heterogeneous aquifers. The results apply equally to streamflow gains caused by managed aquifer recharge. Quantifying the timing of such hydraulic interactions between rivers and groundwater pumping and recharge is necessary in the conjunctive management of groundwater and surface-water resources.

The scenario examined was an aquifer bounded on one side by a straight constant-head boundary representing a river and on the other side by a parallel zero-flux boundary representing an impermeable boundary such as the edge of a valley-fill aquifer. The effect of heterogeneous transmissivity on the timing of stream depletions was the focus. Numerical flow simulation and Gaussian geostatistical models of heterogeneity were used with the Monte Carlo approach and power averaging to identify effective transmissivities for a variety of heterogeneous conditions.

Geometric anisotropy was the primary geostatistical parameter evaluated for its influence on effective transmissivity. Other factors assessed were field variance (i.e., degree of heterogeneity), correlation model shape, nugget effect, correlation scale, and the orientation of principal correlation directions with respect to the river. Sensitivity to zonal anisotropy and field non-stationarity were also tested.

7.2 Primary Findings

In this study, effective transmissivity for transient stream depletion simulations spanned nearly the full range between the bounding arithmetic and harmonic means, varying with the spatial correlation structure of the transmissivity field. For cases with low to moderate

heterogeneity, the shape of the effective power-mean exponent curve as a function of geometric anisotropy was roughly similar to analytical estimates developed by Gelhar and Axness (1983) and Ababou (1991) for steady-state, mean-parallel flow in unbounded domains. And, no statistical difference was observed between effective transmissivity for transient stream depletion simulations and for simulations of steady-state, mean-parallel flow conducted with the same model domain and methods, although ensemble variability was notably larger for stream depletion compared to steady-state mean-parallel flow simulations.

The depletion response curve realizations, and the ensemble-mean curves in particular, were well-fit with a single-valued, time-invariant effective transmissivity. This is in contrast to studies that have found transient changes in effective transmissivity for transient flow conditions.

The similarity observed in this research between effective transmissivities for steady-state, mean-parallel flow and transient stream depletion is notable considering that different flow regimes were involved. It is also a potentially useful result given the large body of literature available analyzing steady-state, mean-parallel flow conditions for various domain and boundary configurations. For example, the Paleologos (1994) coordinate-direction transform for effective transmissivity in geometrically anisotropic fields under steady-state, mean-parallel flow conditions was found to also reasonably approximate effective transmissivity for transient stream depletion when the fields' principal correlation directions were oriented at oblique angles with respect to the river boundary (Figure 4-7, Page 77).

Despite general similarity to the analytical curves, there were differences in the simulated effective transmissivities. In certain anisotropy ranges, effective transmissivity was markedly shifted toward the geometric mean relative to the analytical curves, and overall results approached the bounding means more slowly than the analytical functions. The effective power-

average exponents compiled in this work in Figure 4-1 (Page 63) can be used to refine effective transmissivity estimates to be used in analytical models of stream depletion.

7.3 Sensitivity to the Geostatistical Model

The choice of which of the standard spatial-correlation models to use (spherical, Gaussian, or exponential) had no apparent effect on effective transmissivity results. It is concluded that the effect, if present, was too small to be significant within the experimental methods used for this work. The sensitivity of effective transmissivity to higher field variance was mild, but was clear for cases with moderate to high geometric anisotropy.

Using a nugget effect in the spatial correlation function had a significant impact. An increasing nugget shifted effective transmissivity away from the Weiner bounds and away from otherwise-expected values toward the geometric mean. The impact was proportional to the scale of the nugget relative to the total field variance. Regardless of whether a nugget in the spatial correlation model is a real property of a given transmissivity field, or merely an artifact of measurements and a tool of geostatistical models, this work demonstrated that a decision to include a nugget effect in the field model is consequential.

Stationarity was a condition for most of the transmissivity fields simulated in this work. The non-stationary cases that were tested had a linear trend in mean transmissivity oriented toward the river boundary. For fields with strong heterogeneity, the stationary-field results still provided a reasonable estimate for the effective transmissivity. For mildly heterogeneous fields, however, the presence of the trend skewed results partly toward a harmonic mean. Results in those cases were also highly sensitive to the choice of spatial averaging area.

Simulations were conducted in which field variance was anisotropic in addition to the field having an anisotropic correlation range (i.e., zonal anisotropy). Effective transmissivity in those

cases was shifted toward a harmonic mean and was independent of the geometric anisotropy ratio.

7.4 Other Findings

Ensemble variability increased with increasing scale of spatial correlation. Such behavior is expected in stochastic simulation, but considering that aquifer transmissivity typically exhibits regional-scale correlation ranges, this scale-dependant variability has particular significance for stream depletion. Expected behavior in a stochastic approach may frequently be a poor predictor of site-specific behavior for transient stream depletion cases when the correlation scale, in any direction, is large relative to the distance between the well and the river.

Geostatistical evaluation and groundwater flow modeling were performed for five areas of an alluvial valley-fill aquifer in northeast Colorado. For various locations within these five regions, equivalent transmissivity results varied widely, ranging anywhere between, and beyond, the arithmetic and harmonic means of local transmissivities. Equivalent transmissivity was close to the geometric mean in several locations, but results were highly sensitive to well location, and to the selected averaging area around those locations, such that a typical value was not identified. Overall, this research suggests that equivalent transmissivity may be quite location-specific due to the large correlation scales and statistical inhomogeneity of those fields.

7.5 General Applications

Groundwater hydrologists commonly use analytical models of stream depletion timing. Given a set of local-scale transmissivity values, this research helps address the basic question of how to best upscale (i.e., average) that data. It also suggests that the averaging process needs to represent a relatively large portion of the aquifer along the depleted stream reach.

At a higher level, the research results ask if spatial structure of the aquifer transmissivity is characterized sufficiently to support a choice of one type of average over another. In that respect, this work also provides a guide to the scale of uncertainty involved in choosing the appropriate transmissivity. For example, for the scenario where an aquifer's structure may not be well understood, but perhaps the aquifer is characterized well enough to estimate the arithmetic and harmonic means, and by knowing those means are the bounds of the effective transmissivity (given certain assumptions), then the hydrologist may decide whether further structural characterization is warranted to refine the effective transmissivity estimate.

The variability observed among statistically equivalent realizations in some cases highlights that averaging heterogeneity may not be reliable when the correlation scale of heterogeneity is large relative to the flow problem. The use of effective transmissivity values and analytical models in that case involves significant uncertainty even when there is knowledge of the aquifer's structure. This should factor into the decision of when to use numerical models.

Applications are not limited to analytical modeling. Even when a numerical model is used that can explicitly represent heterogeneity, this work shows how the representation of aquifer structure in the numerical model is consequential to the results. For example, even ostensibly small decisions made in the geostatistical modeling, such as including a nugget effect, can impact the numerically modeled results.

Finally, improving our understanding of how to account for heterogeneity under different circumstances is a fundamental topic in groundwater hydrology that has numerous and perhaps intangible benefits. For example, in ideal practice, a numerical model should merely be a calculator and not a so-called black box, meaning the model user has an informed expectation of the groundwater behavior they are simulating. Such knowledge aids in the development of

conceptual models, aids in the construction of a numerical model, and aids in reviews and critiques of numerical modeling results.

7.6 A Regional Example of Potential Impacts

Quantitative assessments of stream depletion timing are conducted frequently in the South Platte River basin of Colorado. This is not unique to Colorado—the need exists and is increasing in several states in the western U.S.A., as well as in other semi-arid regions around the world. But, the South Platte basin provides an example with potentially large-scale impacts, and with an emphasis on depletion timing, given the large extent of irrigated agriculture and aquifer recharge in that basin, both of which are regulated under Colorado's relatively long and evolved history of managing streamflow depletions caused by groundwater pumping.

Impacts of groundwater pumping on surface water rights have been regulated in Colorado since the 1970s. To pump a groundwater well for irrigation from the South Platte River alluvial aquifer, the user must compute the timing, location, and amount of surface water depletions caused by their pumping, and replace those streamflow depletions that occur out-of-priority in the context of Colorado's prior-appropriation water rights system.

The out-of-priority stream depletions are typically covered by constructing recharge ponds that are operated outside of the irrigation season or when surface water supplies are more available, such as in winter and during spring runoff from the Rocky Mountain snowpack. This system takes advantage of the time delay between recharge operations at the pond and consequent flow accretions to the river.

Managed aquifer recharge in the South Platte basin has increased over the last 30 years approximately from 20,000 acre-feet/year to over 200,000 acre-feet/year. Hundreds of recharge facilities now cover depletions from several thousand irrigation wells. Analytical stream

depletion models are frequently used to plan and operate these recharge systems, both to assess the timing of pumping depletions to be covered and to choose the locations and operating schedules for the recharge ponds. Often, hydrologists use reference transmissivity maps to compute a harmonic mean transmissivity along a narrow line between the well (or recharge pond) and the river and use this value in analytical models.

The stochastic treatment of heterogeneity, as is used in this research, acknowledges that the effective transmissivity will have a degree of error at any particular location, but the goal is to identify an expected-average transmissivity that is the most probable value over a large number of locations given the characteristics of the aquifer and the flow problem. Considering the large volumes of water pumped from and recharged to the South Platte alluvial aquifer annually, a seemingly small bias in the averaging method used in water management models might have significant cumulative impacts if that bias is applied uniformly across the basin. The results of this research (e.g., Figure 4-1, Page 63) could be used to improve those effective transmissivity estimates and to evaluate uncertainties and potential biases in water management practices for the basin.

7.7 Considerations for Future Research

Non-Stationarity. Effective transmissivity was demonstrated to be sensitive to the stationarity condition. This raises questions about applying the effective transmissivity trends identified in this research to certain stream-aquifer systems. Defining a transmissivity averaging area is a necessary step for tackling non-stationarity, but defining the area of influence in the case of transient stream depletion is a complex problem, with the affected area being at least partly dependent on the configuration of the transmissivity field.

Spatial Weighting for Averaging Transmissivity. It was observed in this research that the choice of spatial averaging area could reduce ensemble variability even though it did not influence ensemble-mean behavior. Further work on defining an area of influence and on spatial weighting within that area may further reduce ensemble variability and thereby improve utility of the effective transmissivity concept in stream depletion applications. As noted above, however, this appears to be a complex problem given that the area of influence is itself a function of the aquifer configuration.

In non-stationary fields, simulations suggested that computing a valid equivalent transmissivity may require greater spatial weighting in the areas near the river, with less weight near the pumping well. A similar indication for more general heterogeneous conditions may come from the derivation of the Glover solution (Glover and Balmer 1954). The first derivation steps involve diffusivity (T/S) of the wider aquifer which governs the rate of expansion of the drawdown cone. The final steps involve multiplying the pumping-induced gradient at the river, and only in the direction perpendicular to the river, by the transmissivity at the river. Although the Glover solution assumes a homogenous aquifer, the latter step suggests potential utility in identifying an effective transmissivity in the perpendicular direction near the river separately from the effective transmissivity of the wider area of influence.

Analytical Curves. The differences observed between simulations and the analytical curves plotted in this work are an avenue for additional study. The differences were greatest under anisotropies in the 10:1 range, and were larger for higher field variance. The overall shape of the relationship between geometric anisotropy and effective transmissivity was notably different for the higher variance cases. The differences may arise due to effects from having a bounded domain (e.g., Paleologos et al. 1996) or due to approximations in the analytical

estimates. Practical impacts are that the bounding means are approached more slowly than estimated by the analytical functions, and that even with relatively distinct anisotropy oriented parallel to the river (such as the 10:1 to 20:1 range) the effective transmissivity remained higher than a harmonic mean.

Non-Gaussian Simulators. This research was conducted using Gaussian simulators of heterogeneity which have the trait of connecting mid-range transmissivity values and leaving the higher and lower transmissivity zones with lower connectivity (Deutsch and Journel 1998). It may be of interest to repeat these simulations using geostatistical models that provide preferential connectivity of high-transmissivity zones. Path connectivity has been examined for other flow regimes and has been a notable direction of stochastic simulation research, but to this author's knowledge such scenarios have not been examined for transient stream depletion from a pumping well in a Glover-type analysis. Applicability of the Glover analytical solution, however, should be carefully reconsidered for cases with channelized heterogeneity.

REFERENCES

Ababou, R. and E.F. Wood. (1990). Comment on 'Effective groundwater model parameter values: Influence of spatial variability of hydraulic conductivity, leakance, and recharge' by J.J. Gomez-Hernandez and S.M. Gorelick". *Water Resources Research* 26(8):1843-1846.

Ababou, R. (1991). Identification of effective conductivity tensor in randomly heterogeneous and stratified aquifers. In Proceedings of the 5th Canadian/American Conference on Hydrogeology: Parameter Identification and Estimation for Aquifer and Reservoir Characterization (pp. 155-157).

Aiken, A., E. Anderman, E. Harmon, S. Paschke, D. Plazak, and M. Riemann (editors). (2000). Colorado Ground-Water Atlas. Colorado Ground-Water Association, P.O. Box 150036, Lakewood, Colorado 80215. (available at www.coloradogroundwater.org).

Anderson, M. P. (1997). Characterization of geological heterogeneity. In Subsurface Flow and Transport: A Stochastic Approach, edited by G. Dagan and S.P. Neuman. Cambridge University Press, Cambridge, UK. 241 pp.

Anderson, M. P., and W.W. Woessner. (2002). Applied groundwater modeling: simulation of flow and advective transport. Academic Press. 381 pp.

Bachmaier, M., and M. Backes. (2008). Variogram or semivariogram? Understanding the variances in a variogram. *Precision Agriculture* 9(3):173-175.

Bachmaier, M., and M. Backes. (2011). Variogram or Semivariogram? Variance or Semivariance? Allan Variance or Introducing a New Term? *Mathematical Geosciences* 43(6): 735-740.

Barnes, R. J. (1991). The variogram sill and the sample variance. *Mathematical Geology* 23(4): 673-678.

Barlow, P.M., and S.A. Leake. (2012). Streamflow depletion by wells—Understanding and managing the effects of groundwater pumping on streamflow: U.S. Geological Survey Circular 1376, 84 p. (http://pubs.usgs.gov/circ/1376/)

Bjorklund, L. J., and R. F. Brown. (1957). Geology and Ground-Water Resources of the Lower South Platte River Valley Between Hardin Colorado, and Paxton Nebraska. U.S. Geological Survey Water Supply Paper 1378.

Bredehoeft, J. (2011). Hydrologic Trade-Offs in Conjunctive Use Management. *Ground Water* 49(4):468–475. doi: 10.1111/j.1745-6584.2010.00762.x

Bredehoeft, J. and E. Kendy. (2008). Strategies for Offsetting Seasonal Impacts of Pumping on a Nearby Stream. *Ground Water* 46(1):23–29. doi: 10.1111/j.1745-6584.2007.00367.

Butler, J.J., Jr., X. Zhan, and V.A. Zlotnik. (2007). Pumping-induced drawdown and stream depletion in a leaky aquifer system. *Ground Water* 45(2):178–186.

Cardwell, W. T., and R. L. Parsons. (1945). Average Permeabilities of Heterogeneous Oil Sands: *Trans. AIME,* v. 160, p. 34–42.

Chang, C. M., and H.D. Yeh. (2012). Investigation of solute transport in nonstationary unsaturated flow fields. *Hydrology and Earth System Sciences* 16(11), 4049-4055.

CSU (Colorado State University). (2006). GIS file *river.shp* dated 05/23/2006, accessed 06/12/2011 at ids.colostate.edu.

Dagan, G. (1979). Models of groundwater flow in statistically homogeneous porous formations. *Water Resources Research* 15(1):47-63.

Dagan, G. (1982). Analysis of flow through heterogeneous random aquifers: 2. Unsteady Flow in Confined Formations. *Water Resources Research* 18(5):1571-1585.

Dagan, G., S.C. Lessoff, and A. Fiori. (2009). Is transmissivity a meaningful property of natural formations? Conceptual issues and model development. *Water Resources Research* 45, W03425, doi: 10.1029/2008WR007410.

Danquigny, C., P. Ackerer, and J. P. Carlier. (2004). Laboratory tracer tests on three-dimensional reconstructed heterogeneous porous media. *J. Hydrology* 294: 196–212.

de Dreuzy, J. R., P. de Boiry, G. Pichot, and P. Davy. (2010). Use of power averaging for quantifying the influence of structure organization on permeability upscaling in on-lattice networks under mean parallel flow. *Water Resources Research* 46(8), W08519.

Desbarats, A.J. (1992a). Spatial Averaging of Transmissivity in Heterogeneous Fields with Flow towards a Well. *Water Resources Research* 28(3):757–767.

Desbarats, A.J. (1992b). Spatial Averaging of Hydraulic Conductivity in Three-Dimensional Heterogeneous Porous Media. *Mathematical Geology* 24(3):249-267.

Desbarats, A.J. (1993). Geostatistical analysis of interwell transmissivity in heterogeneous aquifers. *Water Resources Research.* 29(3):1239-1246.

Desbarats, A.J. 1994. Spatial averaging of hydraulic conductivity under radial flow conditions. *Mathematical Geology* 26(1):1-20.

Deutsch, C.V., and A.G. Journel. (1998). GSLIB—Geostatistical Software Library and User's Guide. New York: Oxford University Press. Second Edition.

El-Kadi, A.I., and W. Brutsaert. (1985). Applicability of effective parameters for unsteady flow in nonuniform aquifers. *Water Resources Research* 21(2):183-198.

Engdahl, N. B., E. T. Vogler, and G. S. Weissmann. (2010). Evaluation of aquifer heterogeneity effects on river flow loss using a transition probability framework. *Water Resources Research* 46(1). W01506, doi:10.1029/2009WR007903.

Faunt, C. C., A.M. Provost, M.C. Hill, and W.R. Belcher. (2011). Comment on "An unconfined groundwater model of the Death Valley Regional Flow System and a comparison to its confined predecessor" by RWH Carroll, GM Pohll and RL Hershey. *Journal of Hydrology* 397(3): 306-309.

Fleckenstein, J. H., and G.E. Fogg. (2008). Efficient upscaling of hydraulic conductivity in heterogeneous alluvial aquifers. *Hydrogeology Journal* 16(7): 1239-1250.

Fleckenstein, J.H., R.G. Niswonger, and G.E. Fogg. (2006). River-aquifer interactions, geologic heterogeneity, and low-flow management. *Ground Water* 44(6):837-852.

Fogg, G. E. (1986). Groundwater Flow and Sand Body Interconnectedness in a Thick, Multiple-Aquifer System, *Water Resources Research* 22(5): 679–694.

Freeze, R.A. (1975). A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. *Water Resources Research* 11(5):725-741.

Freeze, R. A., and J.A. Cherry. (1979). Groundwater, 604 pp.

Gringarten, E. and C.V. Deutsch. (2001). Teacher's aide variogram interpretation and modeling. *Mathematical Geology* 33(4):507-534.

Gelhar, L.W. (1993). Stochastic Subsurface Hydrology. Prentice Hall. 390 pp.

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

Glover, R.E. (1977). Transient Ground Water Hydraulics. Water Resources Publications, Fort Collins, Colorado. Third Printing (1985).

Glover, R.E., and C.G. Balmer. (1954). River depletion resulting from pumping a well near a river. AGU Transactions 35(3): 468-470.

Gomez-Hernandez, J.J. and S.M Gorelick. (1989). Effective groundwater model parameter values: Influence of spatial variability of hydraulic conductivity, leakance, and recharge. *Water Resources Research* 25(3):405-419.

Goode, D. J., and C.A. Appel. (1992). Finite-Difference Interblock Transmissivity for Unconfined Aquifers and for Aquifers having Smoothly Varying Transmissivity. USGS Water-Resources Investigations Report 92-4124.

Goovaerts, P. (1997). Geostatistics for natural resources evaluation. Oxford university press.

Gringarten, E., and C.V. Deutsch. (2001). Teacher's aide variogram interpretation and modeling. *Mathematical Geology* 33(4): 507-534.

Gutjahr, A. L., L.W. Gelhar, A.A. Bakr, and J.R. MacMillan. (1978). Stochastic analysis of spatial variability in subsurface flows: 2. Evaluation and application. *Water Resources Research* 14(5): 953-959.

Harbaugh, A.W. (2005). MODFLOW-2005, The U.S. Geological Survey modular ground-water model—the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16.

Harbaugh, A.W., E.R. Banta, M.C. Hill, and M.G. McDonald. (2000). MODFLOW-2000, the U.S. Geological Survey modular ground-water model—User guide to modularization concepts and the Ground-Water Flow Process. USGS Open-File Report 00-92.

Hoeksema, R. J., and P.K. Kitanidis. (1985). Analysis of the spatial structure of properties of selected aquifers. *Water Resources Research* 21(4):563-572.

Huang, C. S., W.S. Lin, and H.D. Yeh. (2014). Stream filtration induced by pumping in a confined, unconfined or leaky aquifer bounded by two parallel streams or by a stream and an impervious stratum. *Journal of Hydrology*, 513, 28-44.

Hunt, B. (2014). Review of Stream Depletion Solutions, Behavior, and Calculations. *Journal of Hydrologic Engineering*, 19(1): 167-178.

Hurr, R.T., and P.A. Schneider, Jr. (1972a). Hydrogeologic characteristics of the valley-fill aquifer in the Brighton reach of the South Platte River Valley, Colorado. Water Resources Division, Colorado District, USGS Open-File Report.

Hurr, R.T., and P.A. Schneider, Jr. (1972b). Hydrogeologic characteristics of the valley-fill aquifer in the Julesburg reach of the South Platte River Valley, Colorado. Water Resources Division, Colorado District, USGS Open-File Report.

Hurr, R.T., and P.A. Schneider, Jr. (1972c). Hydrogeologic characteristics of the valley-fill aquifer in the Sterling reach of the South Platte River Valley, Colorado. Water Resources Division, Colorado District, USGS Open-File Report.

Jenkins, C.T. (1968). Techniques for computing rate and volume of stream depletion by wells. *Ground Water* 6(2): 37–46.

Indelman, P. (2003). Transient pumping well flow in weakly heterogeneous formations. *Water Resources Research* 39(10):1287.

Indelman, P., and B. Abramovich. (1994). A higher-order approximation to effective conductivity in media of anisotropic random structure. *Water Resources Research* 30(6): 1857-1864.

Indelman, P., and G. Dagan. (2004). A note on well boundary condition for flow through heterogeneous formations. *Water Resources Research* 40, W03601, doi: 10.1029/2003WR002602.

Indelman, P., and Y. Rubin. (1996). Average flow in heterogeneous media of trending hydraulic conductivity. *Journal of Hydrology* 183(1): 57-68.

Isaaks, E. H., and R. M. Srivastava. (1989). Applied Geostatistics. Oxford University Press.

Kendy, E., and J. D. Bredehoeft. (2006). Transient effects of groundwater pumping and surface-water-irrigation returns on streamflow. *Water Resources Research* 42(8), W08415, doi: 10.1029/2005WR004792.

Knight, J.H., M. Gilfedder, and G.R.Walker. (2005). Impacts of irrigation and dryland development on groundwater discharge to rivers—a unit response approach to cumulative impacts analysis. *Journal of Hydrology* 303:79-91.

Lachassagne, P., E. Ledoux, and G. de Marsily. (1989). Evaluation of hydrogeological parameters in heterogeneous porous media. Groundwater Management: Quantity and Quality, Proceedings of the Benidorn Symposium, October 1989. IAHS Pub. No.188:3-18.

Lindsey, D. A., W. H. Langer, and D. H. Knepper. (2005). Stratigraphy, lithology, and sedimentary features of Quaternary alluvial deposits of the South Platte River and some of its tributaries east of the Front Range, Colorado. U.S. Geological Survey Professional Paper 1705. 70p.

Matheron, G. (1967). Composition des Perméabilités en Milieu Poreux Hétérogène: Méthode de Schwydler et Règles de Pondération. Revue de l'Institut Français du Pétrole. 22(3):443–466.

Mathworks. (2013). Curve Fitting Toolbox™ User's Guide for MATLAB Version 2013a. The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098.

McWhorter, D.B., and D.K. Sunada. (1977). Ground-Water Hydrology and Hydraulics. Water Resources Publications, Highlands Ranch, Colorado.

McDonald, M.G., and A.W. Harbaugh. (1988). A modular three-dimensional finite-difference ground-water flow model. U.S. Geological Survey Techniques of Water-Resources Investigations, book 6, chap. A1, 586 p.

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

Miller, C.D., D.S. Durnford, M. Halstead, J. Altenhofen, and V. Flory. (2007). Stream Depletion in Alluvial Valleys using the SDF Semi-Analytical Model. *Ground Water* 45(4):506-514.

Molz, F.J., J. Gaun, and J. Wang. (2005). Spatial Weighting Functions: Transient Hydraulic Tests and Heterogeneous Media. *Ground Water* 43(2):215-221.

Naff, R. I. (1991). Radial Flow in Heterogeneous Porous Media: An Analysis of Specific Discharge. *Water Resources Research* 27(3): 307–316.

Neuman, S. P., and S. Orr. (1993). Prediction of Steady State Flow in Nonuniform Geologic Media by Conditional Moments: Exact Nonlocal Formalism, Effective Conductivities, and Weak Approximation. *Water Resources Research* 29(2):341–364.

Neuman, S.P.; V. di Federico. (2003). Multifaceted nature of hydrogeologic scaling and its interpretation. *Reviews of Geophysics*. 41(3):1014. doi: 10.1029/2003RG000130

Paleologos, E. K., S.P. Neuman, and D. Tartakovsky. (1996). Effective hydraulic conductivity of bounded, strongly heterogeneous porous media. *Water Resources Research* 32(5): 1333-1341.

Paleologos, E.K. (1994). Effective hydraulic conductivity of bounded, strongly heterogeneous porous media. Ph.D. dissertation, Dept. of Hydrology and Water Resources, University of Arizona, Tucson.

Remy, N., A. Boucher, and J. Wu. (2009). Applied geostatistics with SGeMS: A user's guide. Cambridge University Press.

Renard P., and G. de Marsily. (1997). Calculating equivalent permeability: a review. *Advances in Water Resources* 20(5):253-278.

Riva, M., A. Guadagnini, S.P.Neuman, and S. Franzetti. (2001). Radial flow in a bounded randomly heterogeneous aquifer. *Transport in Porous Media* 45:139-193.

Romeu, R. K., & B. Noetinger. (1995). Calculation of internodal transmissivities in finite difference models of flow in heterogeneous porous media. *Water Resources Research* 31(4): 943-959.

Ronayne, M. J., and S.M. Gorelick. (2006). Effective permeability of porous media containing branching channel networks. *Physical Review E* 73(2), 026305.

Rubin, Y. (2003). Applied Stochastic Hydrogeology. Oxford University Press.

Rubin, Y., and G. Dagan. (1989). Stochastic analysis of boundaries effects on head spatial variability in heterogeneous aquifers, 2, Impervious boundary. *Water Resources Research* 25(4): 707-712.

Rubin, Y., and K Seong. (1994). Investigation of flow and transport in certain cases of nonstationary conductivity fields. *Water Resources Research* 30(11): 2901-2911.

Sanchez-Vila, X. (1997). Radially convergent flow in heterogeneous porous media. *Water Resources Research* 33(7):1633-1641.

Sanchez-Vila, X., J. Carrera, and J.P.Girardi. (1996). Scale effects in transmissivity. *Journal of Hydrology* 183:1-22.

Sanchez-Vila, X., A. Guadagnini, and J. Carrera. (2006). Representative hydraulic conductivities in saturated groundwater flow. *Reviews of Geophysics*, 44(3). RG3002, doi: 10.1029/2005RG000169

Sarris, T. S., and E.K. Paleologos. (2004). Numerical investigation of the anisotropic hydraulic conductivity behavior in heterogeneous porous media. *Stochastic Environmental Research and Risk Assessment*, 18(3): 188-197.

Sheets, R. A., M.C. Hill, H.M. Haitjema, A.M. Provost, and J.P. Masterson. (2015). Simulation of Water-Table Aquifers Using Specified Saturated Thickness. *Groundwater*, 53(1): 151-157.

Smith, L. and R.A. Freeze. (1979). Stochastic analysis of steady state groundwater flow in a bounded domain. 2. Two-dimensional simulations. *Water Resources Research* 15(6):1534-1559.

Smith, R.O., P.A. Schneider, Jr., and L.R. Petri. (1964). Ground-Water Resources of the South Platte River Basin in Western Adams and Southwestern Weld counties, Colorado. USGS Water-Supply Paper 1658. 132pp plus plates.

SPDSS (South Platte Decision Support System). (2006a). SPDSS Groundwater Component, Phase 3, Task 34 Streambed Conductance Testing. Prepared for the CWCB and DWR by CDM Smith. June 9, 2006.

SPDSS. (2006b). SPDSS Ground Water Component Phase 3, Task 42.3 South Platte Alluvium Region Aquifer Configuration Technical Memorandum. Prepared for the CWCB and DWR by CDM Smith. November.

SPDSS. (2006c). SPDSS Ground Water Component Phase 3, Task 43.3 South Platte Alluvium Region Aquifer Property Technical Memorandum. Prepared for the CWCB and DWR by CDM Smith. December.

SPDSS. (2009). GIS files "tgrid0309" dated 03/19/2009. Accessed in file div1_gw_rasters.zip at http://cdss.state.co.us/GIS/Pages/Division1SouthPlatte.aspx on 11/22/2013.

SPDSS. (2013). South Platte Decision Support System Alluvial Groundwater Model Report. Prepared for the CWCB and DWR by CDM Smith. April.

Spector, M., and P. Indelman. (1998). On second-order correction to the effective conductivity of two-dimensional anisotropic media. *Water Resources Research*. 34(5):1357-1359.

Tartakovsky, D.M., A.Guadagnini, and L. Gaudagnini. (2000). Short Note: Effective Hydraulic Conductivity and Transmissivity for Heterogeneous Aquifers. *Mathematical Geology* 32(6):751-759.

Theis, C.V. (1935). The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. AGU Transactions 16(2):519-524.

Theis, C.V. (1941). The effect of a well on the flow of a nearby stream. AGU Transactions 22(3): 734-738.

Topper, R.E., and J.C. Wilson. (2003). Ground water atlas of Colorado. Colorado Geological Survey.

Tweto, O., and R.E. Schoenfeld. (1979). Geologic map of Colorado. Department of the Interior, United States Geological Survey.

Vandenberg, A. (1977). Pump Testing in Heterogeneous Aquifers. *Journal of Hydrology* 34(1/2): 45–62.

Vermeulen, P.T.M., C.B.M. te Stroet, and A.W. Heemink. (2006). Limitations to upscaling of groundwater flow models dominated by surface water interaction. *Water Resources Research* 42, W10406, doi: 10.1029/2005WR004620.

Warren, J. E., and H.S. Price. (1961). Flow in Heterogeneous Porous Media. *Society of Petroleum Engineering Journal*. 1:153–169.

Welhan, J.A. and M.F. Reed. (1997). Geostatistical analysis of regional hydraulic conductivity variations in the Snake River Plain aquifer, eastern Idaho. GSA Bulletin 109(7):855-868.

Welhan, J.A., T.M. Celmo, and E.L. Gégo, (2002). Stochastic simulation of aquifer heterogeneity in a layered basalt aquifer system, eastern Snake River Plain, Idaho. Geological Society of America Special Paper 353, p. 225-247, in Link, P.K. and L.L.

Mink, L.L., eds., Geology, Hydrogeology, and Environmental Remediation: Idaho National Engineering and Environmental Laboratory, Eastern Snake River Plain, Idaho.

Wen, X.H. and J.J.Gomez-Hernandez. (1996). Upscaling hydraulic conductivities in heterogeneous media: An overview. *Journal of Hydrology* 183:ix-xxxii.;

Williams, R. E. (1988). Comment on "Statistical theory of groundwater flow and transport: Pore to Laboratory, Laboratory to Formation, and Formation to Regional Scale" by G. Dagan. *Water Resources Research* 24(7):1197-1200.

Xu, C., and P.A. Dowd. (2012). The edge effect in geostatistical simulations. In Geostatistics Oslo 2012 (pp. 115-127). Springer Netherlands.

Yeh, H. D., Y.C. Chang, and V.A. Zlotnik. (2008). Stream depletion rate and volume from groundwater pumping in wedge-shape aquifers. *Journal of Hydrology* 349(3): 501-511.

Zlotnik, V.A. and D. Tartakovsky. (2008). Stream Depletion by Groundwater Pumping in Leaky Aquifers. *Journal of Hydrologic Engineering* 13(2):43–50.

APPENDIX - EXAMPLE REALIZATIONS OF HETEROGENEOUS FIELDS

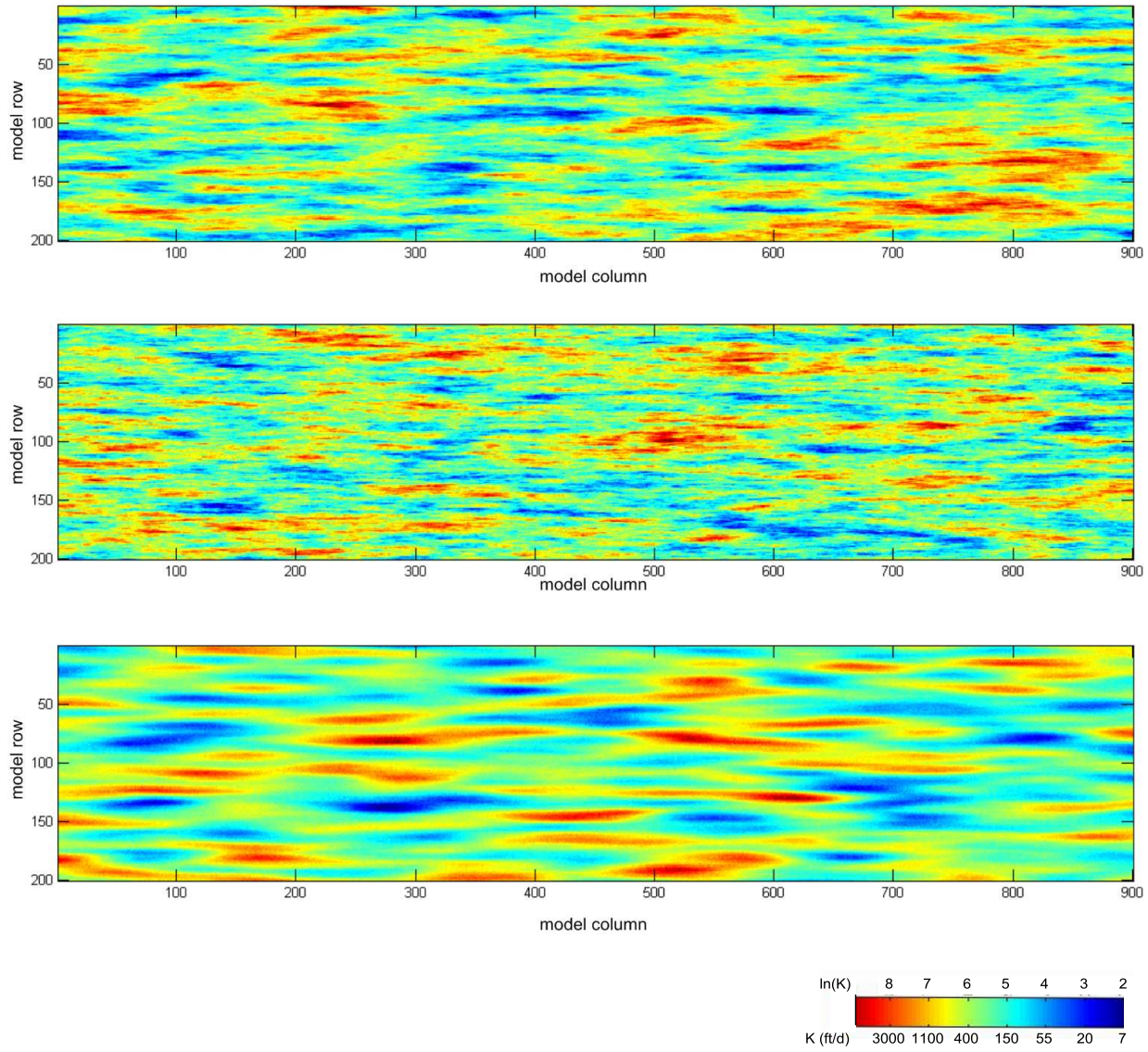

Figure A-1. Comparison of geometrically anisotropic (10:1) transmissivity fields generated with spherical (top), exponential (middle) and Gaussian (bottom) correlation functions. Correlation range is 100 columns x 10 rows.

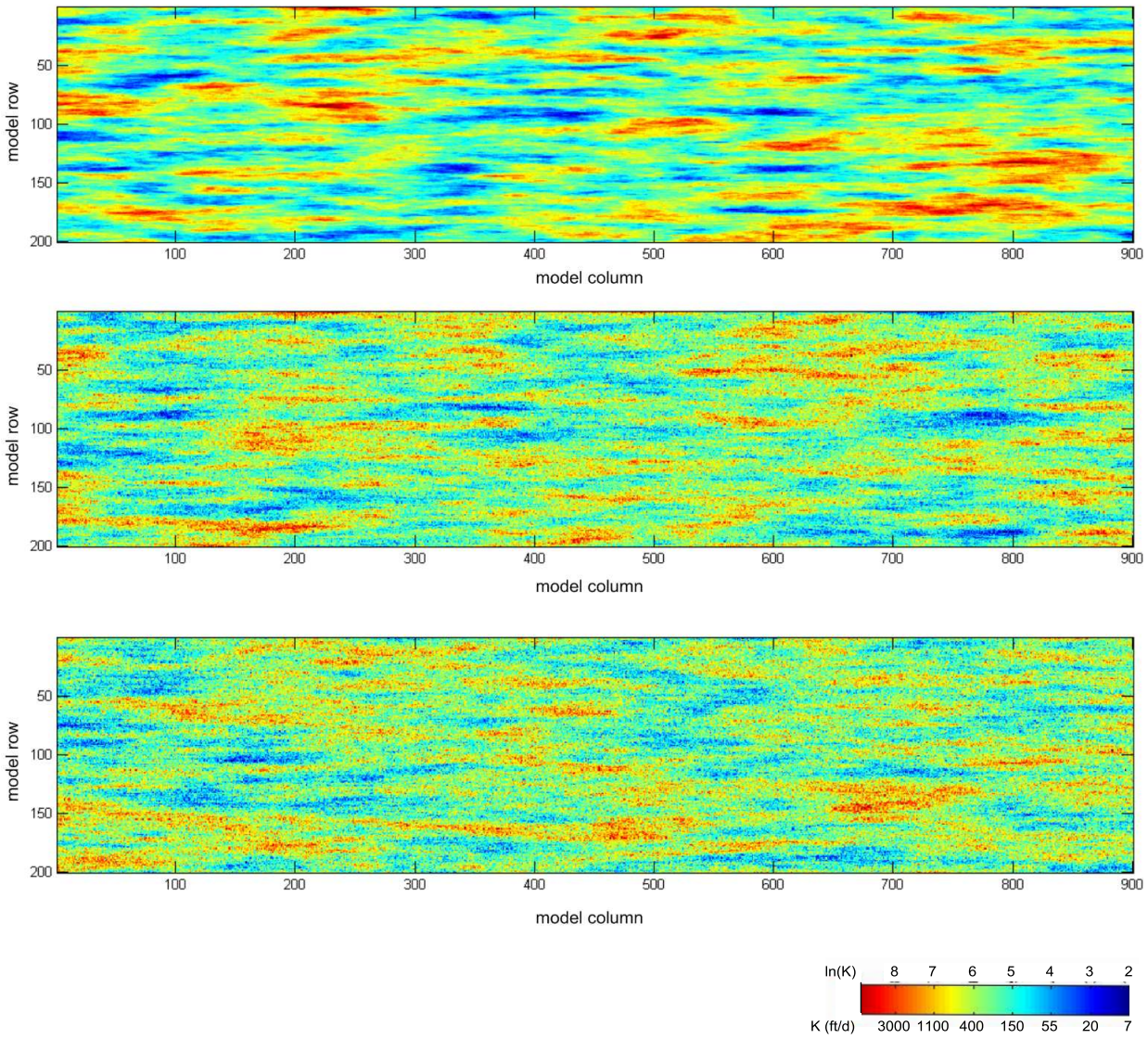

Figure A-2. Comparison of geometrically anisotropic (10:1) transmissivity fields generated with zero nugget effect (top), and nugget contributions of 33% (middle) and 50% (bottom). Correlation range is 100 columns x 10 rows.

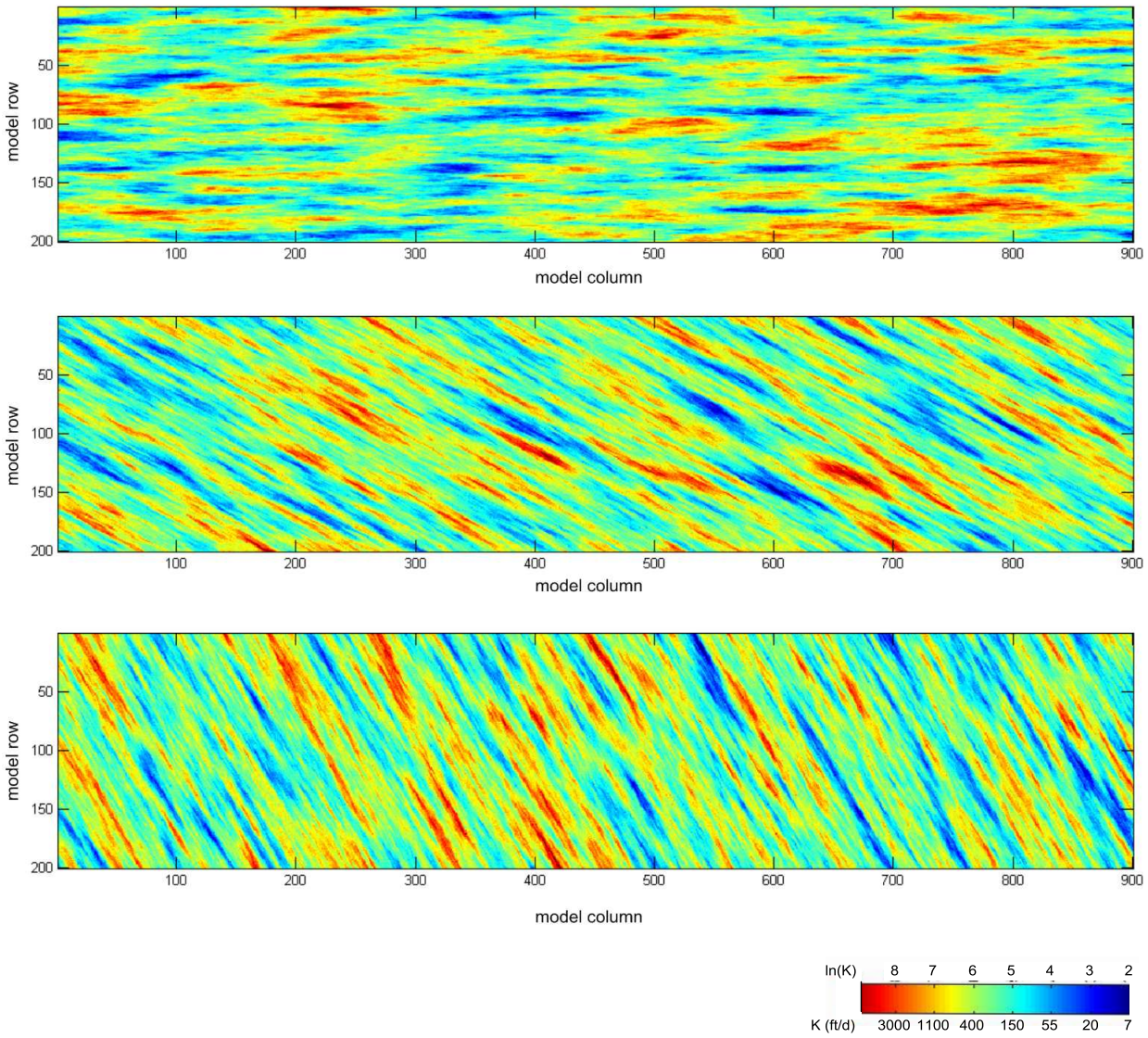

Figure A-3. Comparison of geometrically anisotropic (10:1) transmissivity fields generated with principal correlation direction at azimuth equal to 90° (top) -60° (middle), and -30°(bottom). Correlation range set equivalent to 100 x 10 model grid nodes.

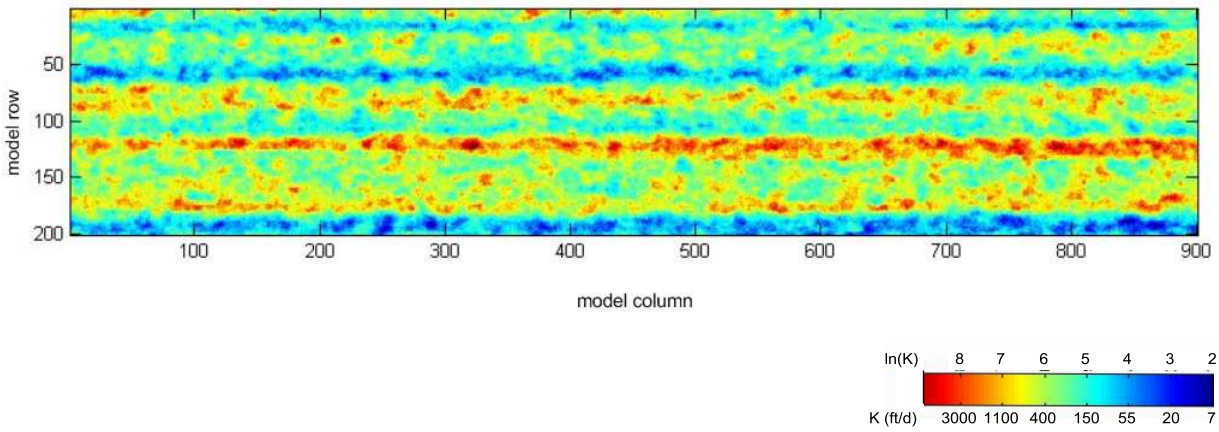

Figure A-4. Transmissivity field with zonal anisotropy. Isotropic correlation range equal to 15 model grid nodes, but with variogram sill in the y-direction (along columns) three times larger than in the x-direction.

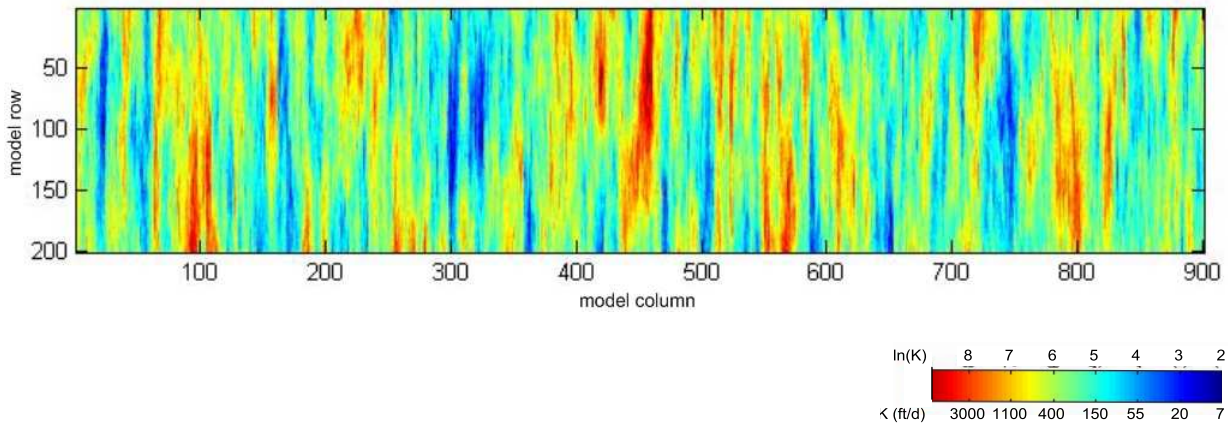

Figure A-5. Example of a transmissivity realization producing an extreme-outlier p_{eq} result. A high-contrast, high-transmissivity zone (red) functioned as a distinct channel by directly connecting a pumping well in the center (Row 101, Column 450) to the river (Row 1). The resulting depletion response was fast and described by an effective transmissivity much higher than obtained by spatial averaging of point values over an area wider than the channel.

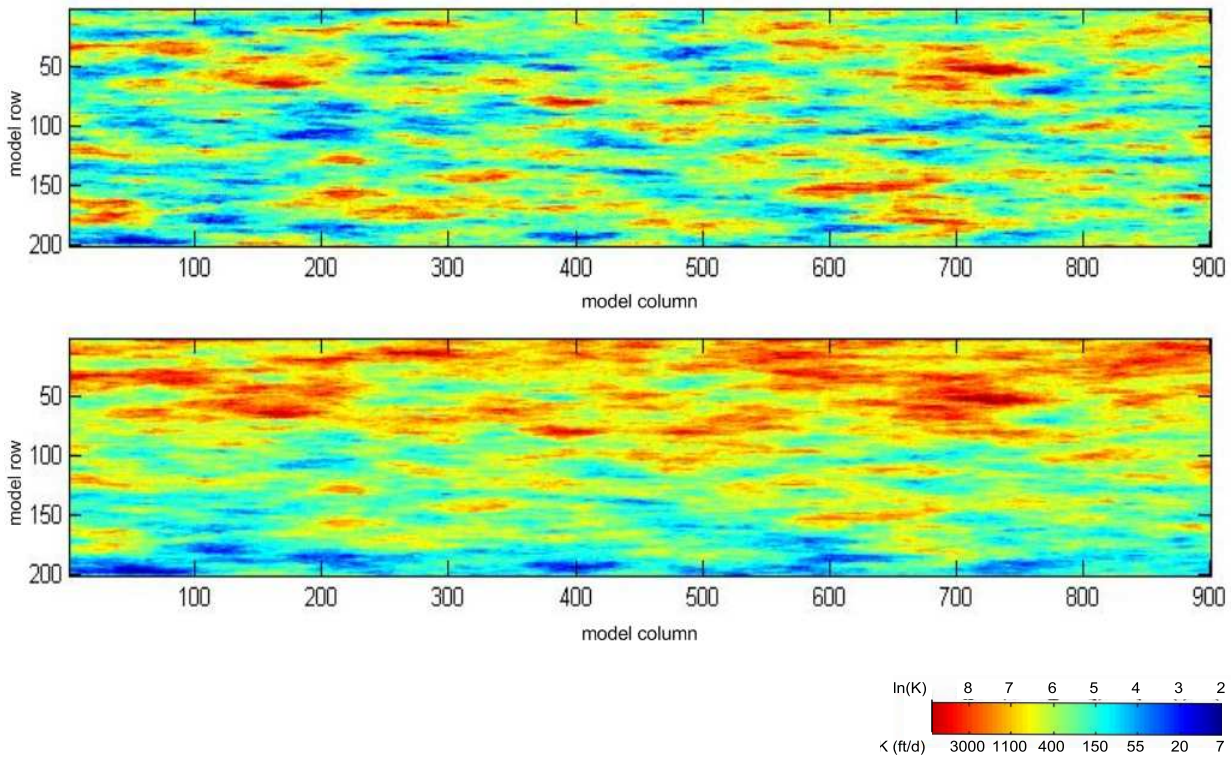

Figure A-6. Example of stationary field (top) and that same field with a trend in mean transmissivity added (bottom).