THESIS

HYDROGEOPHYSICAL INVESTIGATION OF UNCONFINED AQUIFER DRAINAGE BEHAVIOR USING TEMPORAL MICROGRAVITY AND WATER LEVEL DATA

Submitted by

Matthew Sturdivant

Department of Geosciences

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2020

Master's Committee:

Advisor: Michael Ronayne

William Sanford
Thomas Sale

ABSTRACT

HYDROGEOPHYSICAL INVESTIGATION OF UNCONFINED AQUIFER DRAINAGE BEHAVIOR USING TEMPORAL MICROGRAVITY AND WATER LEVEL DATA

Unconfined aquifers are commonly characterized by an analysis of water level changes in response to groundwater pumping during an aquifer test. Traditional analytical models predict the rate and extent of water level changes based on transmissive and storage properties of the aquifer. These models commonly assume instantaneous and complete dewatering of the pore space above a falling water table, which neglects time-dependent storage changes in the unsaturated zone. By sensing pumping-induced water mass changes in both the saturated and unsaturated zones, gravity surveys provide an opportunity for improved characterization of unconfined aquifers.

In this study, a time-lapse microgravimetric survey was performed during pumping from a shallow unconfined aquifer in northern Colorado. Water level data were collected at four monitoring wells located along a radial transect at 6.34, 15.4, 30.7. and 61.2 meters from the pumping well. Gravity measurements were collected adjacent to the second well at 15.4 meters. Pumping from the aquifer resulted in a water level decline ranging from 0.35 meters at the distant well to 1.5 meters at the closest well. A total of $3.89 \cdot 10^6$ kg of water mass was pumped during the test, resulting in a decline in gravitational acceleration of 27.2 microGals at the fixed measurement location. The gravity data are not adequately explained by traditional analytical models that predict negligible mass changes as the water table beings to stabilize. This

highlights potential inaccuracies in drawdown model assumptions that are not readily discernible with water level data alone.

ACKNOWLEDGEMENTS

I am tremendously grateful to all who helped me during my time at Colorado State University. I would first like to thank my advisor, Dr. Michael Ronayne, who never gave up on me and graciously continued to work with me even after I left campus. He is a truly kind mentor and completing my project would not have been possible without his continuous support.

I would like to thank Dr. William Sanford who supported me through instruction, guidance with lab and field equipment, and a lasting excitement for my project. I thank Dr. Thomas Sale for generously and without hesitation offering his time to join my committee. I am also thankful for Dr. Dennis Harry who introduced me to my project and was instrumental in the collection and interpretation of gravity data.

I thank Bill and those at the Colorado State University Horticulture Farm for allowing me to perform multiple aquifer tests, guiding me in operating the pumping well, and rearranging much of their set-up for our tests' needs.

Lastly, I would like to thank my wife Danielle and our son Noah for being the reason and the inspiration behind all that I do. Thank you for being with me at the times I needed it most.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 - INTRODUCTION

1.1 Background

Groundwater comprises almost one-third of freshwater resources. Further, approximately 80 percent of groundwater resources in the US are retrieved from unconfined aquifers [Maupin and Barber, 2005]. Evaluating water level changes and characterizing aquifer physical properties allows for better groundwater resource management. However, monitoring these changes requires expensive well installation, and determining the associated aquifer properties often depends on various simplifying assumptions that may not adequately reflect field conditions.

One or more portable gravimeters may be deployed to further constrain the physical behavior of aquifers by monitoring temporal gravity changes. Gravimeters placed on the surface measure an integral response to nearby subsurface mass variations associated with transient water storage changes. This technique has been effective in detecting large storage changes in aquifers [e.g. Pool and Eychaner, 1995; Howle et al., 2003; Gehman et al., 2009] and is applicable in various hydrologic models [e.g. Creutzfeldt et al., 2010; Van Camp et al., 2017]. Gravity data has shown utility in testing model assumptions and simplifications for hydrogeologic systems by integrating large-scale storage changes to quantify overall system characteristics [Creutzfeldt et al., 2010].

Gravity data has also been considered a tool for further constraining aquifer parameters. However, synthetic modeling of the gravitational response to pumping of a shallow unconfined aquifer predicts detectable anomalies of only tens of microGals for tests lasting one or more days, which may be a concerningly small magnitude relative to the precision of gravimetric equipment [Damiata and Lee, 2006]. Most studies of the utility of gravity measurements during

aquifer tests have relied on synthetic modeling, often without comparison to measured data sets. Consequently, these modeling studies produce gravity data sets subject to the assumptions of the model that is used to compute groundwater mass changes throughout time and space. Most previous studies have utilized analytical drawdown models [e.g. Damiata and Lee, 2006; Blainey et al., 2007; Herckenrath et al., 2012]

The drawdown-gravity model link is important because there has been significant development in analytical drawdown models with adjustments to theories describing flow mechanisms. These various models, which utilize different flow theories and aquifer parameterizations, impact the interpretation of gravity data. The commonly used Theis [1935] and Neuman [1972] drawdown models have been effective in estimating aquifer transmissivity values but have limitations constraining and describing storage properties. Curve fitting of drawdown data with the Neuman [1972] drawdown model consistently under-estimates specific yields of unconfined aquifers compared with expected values from laboratory testing [Nwankwor et al., 1984; Nwankwor et al., 1992; Narasimhan and Zhu, 1993]. Drawdown-gravity synthetic modeling experiments indicate that simultaneous fitting of both drawdown and gravity data may better constrain specific yield when estimating parameters for the Neuman [1972] solution compared to drawdown data alone [Blainey et al., 2007; Herckenrath et al., 2012]. This suggests that gravity data is a viable tool for more accurately estimating specific yield when fitting Theis [1935] or Neuman [1972] drawdown models.

Synthetic modeling studies produce gravity data that is dependent on the assumptions of the drawdown model that defines the shape from which water mass is removed. This dependency of the gravity model on the drawdown model is especially important to consider as there are competing explanations for the observed drawdown in unconfined aquifers. The

phenomenon of "delayed yield," exhibited through a flattening of the temporal drawdown gradient during intermediate stages of pumping, was first explained as a delayed water release above the saturated zone [Boulton, 1954]. Although the physical cause for this delayed release of water storage was undefined, it made conceptual sense and was mathematically effective to model the delay in drawdown using an empirically-controlled exponential function [Boulton, 1954]. This conceptual model was further adapted based on laboratory tests carried out in 1955 to include more parameters controlling uniform anisotropy for more realistic flow dynamics, fitting data well, but relying on empirical parameters with no physical meaning [Boulton and Pontin, 1971]. Because these models did not adequately describe the physical dynamics or governing equations controlling flow, it was necessary to adjust boundary conditions for a more meaningful drawdown model to explain delayed yield.

A numerical approach was developed by Cooley [1971] in which unsaturated-saturated flow dynamics reasonably characterized the shape of the time-drawdown curve, giving some physical basis for the unexplained storage mechanism in the Boulton and Pontin [1971] model. Neuman [1972] was critical of the previous approaches [Boulton, 1954; Boulton, 1970; Boulton and Pontin, 1971] for the use of empirical parameters and assumption of significant flow in the unsaturated zone. He suggested that water release is restricted to the saturated zone, and that aquifer compressibility and the expansion of water control the early stage of the time-drawdown curve whereas the gravity drainage of water controls the late stage [Neuman, 1972]. Despite the Boulton [1954; 1970] and Neuman [1972] models utilizing different explanations of aquifer storage release, there was no significant difference in either model's ability to fit drawdown data and no certainty which model more accurately described unconfined aquifers.

Nwankwor et al. [1984] argued that curve fitting of the Neuman [1972] model provided unrealistically low specific yield values, suggesting that delayed water storage release above the water table is significant and aquifer compressibility is over-estimated. Working at the Borden, Ontario unconfined sand aquifer site, Nwankwor et al. [1984] tested a volume balance method to estimate specific yield against curve fitting of Neuman [1972] drawdown. They found that specific yield increased over time, asymptotically approaching 0.3 which agreed with laboratory specific yield tests. Curve fitting using the Boulton [1963] and Neuman [1972] drawdown models resulted in best fitting specific yields of 0.08 and 0.07, respectively [Nwankwor et al., 1984].

Neuman [1987] suggested the volume balance method of estimating specific yield is inaccurate as it does not account for flow beyond the cone of depression. Nwankwor et al. [1992] addressed this by utilizing piezometers, a tensiometer-pressure transducer system, and gamma moisture probe during a pumping test. The results corroborated the earlier hypothesis [Nwankwor et al., 1984] that delayed drainage above the water table results in the S-shaped time-drawdown curve as the water table drawdown gradient flattens while water storage is released above.

Nwankwor et al. [1992] detailed an important relationship in which the rate of water-table decline (i.e., drawdown) at early times is much greater than the rate of decline of water content, indicating that a substantial amount of water may be temporarily stored above the water table. He further found that capillary fringe drawdown gives better estimates of specific yield than water table drawdown, which provided an explanation for the volume balance method asymptotically approaching the expected specific yield as the thickness of the tension-saturated zone decreases with time. Most excess, or delayed, storage in the pumping test was released

from the zone of tension saturation rather than the vadose zone, and the total water content lost from all zones roughly equaled the pumped volume. Thus, it was not necessary to include contributions of flow beyond the cone of depression [Nwankwor et al., 1992]. Bevan [2002] later performed a 7-day aquifer test at the same Borden, Ontario unconfined sand aquifer site and found that the capillary fringe continued to elongate through the duration of the aquifer test, substantiating the work of Nwankwor et al. [1984; 1992] and expressing the importance of characterizing unsaturated-saturated flow dynamics.

Moench [1995] combined unsaturated and saturated flow mechanisms by introducing Boulton's [1954] empirical function as a boundary condition in Neuman's [1972] drawdown model. Moench [1995] determined the inclusion of Boulton's [1954] to be an improvement upon previous assumptions of instantaneous drainage despite Narasimhan and Zhu [1993] remarking that the exponential relation does not accurately describe the drainage process. Narasimhan [1999] however later gave credence to Boulton's exponential model by evidencing creep behavior in unconfined aquifers, in which deformation kinematics were time-dependent and adequately described by an exponential function. Moench [2004] found that his model strongly fit late-drawdown data with a wide range of specific yield values, but non-instantaneous drainage from above the water table must be considered to better fit early and intermediate water level data.

Mathias and Butler [2006] and Tartakovsky and Neuman [2007] also developed analytical models that allowed storage above the initial water table and assumed volumetric soil moisture and relative hydraulic conductivity vary as exponential functions of pressure head. Moench [2008] tested his analytical model as well as the Mathias and Butler [2006] and Tartakovsky and Neuman [2007] analytical models with Bevan's [2002] data set. Moench

[2008] found that his model under-estimated the specific yield of the saturated zone when assuming instantaneous drainage. Moench [2008] concluded each model accurately yielded saturated zone hydraulic parameters when allowing non-instantaneous drainage but expressed the need for flow anisotropy in the unsaturated zone to characterize unsaturated parameters. Mishra and Neuman [2010] produced an analytical model with consideration to each of these hydrogeologic complexities, rationalizing early-time drawdown discrepancies by well-bore storage, and especially noting the need for a better governing relationships between pressure head, relative hydraulic conductivity, and water content for improvements upon the empirical nature of some of the saturated-unsaturated flow dynamics.

Introducing more boundary conditions and time-variant boundary conditions into analytical models has shown moderate impact in better fitting the early, intermediate, and late time stages of time-drawdown simultaneously, at the cost of increased parameterization and necessitating a network of monitoring wells to constrain the estimated aquifer parameters [Mao et al., 2011]. Mao et al. [2011] also expressed the need for better governing equations to produce a flow model that considers compressibility as well as tension-saturated and vadose storage, and accounts for local heterogeneity and saturation-dependent conductivities.

Modeling and data-fitting efforts have shown that different flow dynamics are dependent on aquifer characteristics and develop over time within the same aquifer. As drawdown has been shown to be insensitive to many of the parameters in analytical solutions, further study of the dynamics of saturated-unsaturated flow of unconfined aquifers would likely benefit from data sets sensitive to the parameters which drawdown alone cannot easily constrain.

Temporal relative gravity may be a useful geophysical tool during aquifer tests for better understanding and modeling of the aquifer's behavior. Field experiments involving gravity data

collection during aquifer tests are needed to test the magnitude of the gravity signal and whether this signal may validate hydrogeologic models for flow in unconfined aquifers. Additional field testing is also necessary to evaluate the ability of combined gravity-drawdown data sets to better constrain the parameters controlling flow and aquifer storage changes.

1.2. Field Experiment and Research Objectives

Despite numerous synthetic modeling studies evidencing the capacity of microgravimetry for further constraining aquifer storage properties, field applications are rare in the groundwater literature. This study provides a 24-hour field test of time-lapse microgravimetry and drawdown during controlled pumping from a shallow unconfined aquifer. The study area, located in northern Colorado, was the site of one previous aquifer test and gravity survey conducted by Woodworth [2011]. In that study, gravity data were collected over several multi-hour intervals and used to estimate aquifer properties assuming idealized (i.e., radially symmetric) drawdown behavior in the aquifer. Data interpretation was confounded by repositioning of the gravimeter multiple times during the test. Each change in gravimeter location was associated with a significant shift in magnitude of the absolute gravity anomaly, even upon revisiting the same location, so the magnitudes of the anomalies were disregarded and distinct relative gradients were analyzed [Woodworth, 2011].

The objective of this study was to perform a relatively long duration (24-hour) field test with a gravimeter maintained at a single location to allow for continuous measurement of the gravity anomaly. We investigate whether this scale of aquifer test results in a discernible anomaly that is significantly large in magnitude compared to the instrument precision, and if so, whether the behavior of the gravitational anomaly is consistent with

previously published synthetic data sets [e.g. Damiata and Lee, 2006; Blainey et al., 2007; Herckenrath et al., 2012; González-Quirós and Fernández-Álvarez, 2016].

CHAPTER 2 - METHODS

2.1. Site Description

The aquifer test was conducted at the Horticulture Field Research Center located about 8 kilometers northeast of the Colorado State University campus in Fort Collins, Colorado (Figure 1). The shallow, unconfined aquifer is used for irrigation throughout the 65-acre horticulture farm which grows a variety of organic and specialty crops for university research projects from May to October. The test took place on April 13, 2016, approximately one month before pumping would begin for irrigation. The initial water table was therefore unaffected by irrigation pumping and was, for our purposes, at an approximate steady-state condition.

Figure 1. Test Site location relative to Colorado State University [*Google Maps*, 2018].

The Pleistocene Broadway Alluvium, which forms the primary aquifer material at the site, consists of poorly sorted sands and gravels and thin lenses of clays and silts [Lindsey et al., 2005]. The aquifer exists below a roughly one-meter thick layer of Nunn clay loam soil [Moreland, 1980]. The saturated thickness of the alluvial aquifer is approximately 7.5 meters, and the aquifer is confined underneath by the Pierre Shale (Figure 2) [USGS, 1979]. The previous aquifer-test analysis by Woodworth [2011] indicated a transmissivity of 0.018 m^2/s and specific yield of 0.0415.

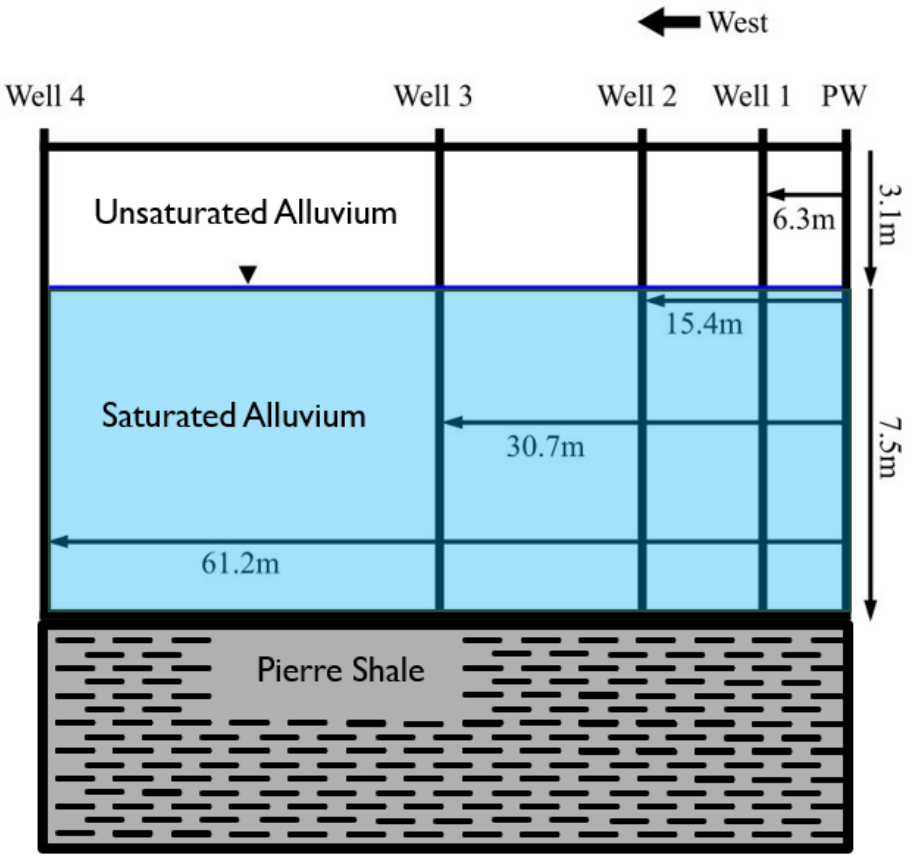

Figure 2. Schematic of pumping (PW) and observation wells at test site [modified from Woodworth, 2011]. Average initial water table depth is 3.1m, and the aquifer thickness is 7.5m, underlain by Pierre Shale. The gravimeter was placed at the same radial distance as observation well 2.

The pumping well selected for the aquifer test is a stainless steel, slotted well with a diameter of 50 cm and located adjacent to a transect of four monitoring wells. The monitoring wells are to the west of the pumping well at distances of 6.3, 15.4, 30.7, and 61.2 meters (Figure 2). Each of the monitoring wells is fully penetrating to the underlying shale and screened throughout the saturated zone.

2.2 Aquifer Test Measurements

2.2.1. Water Level and Discharge Measurements

The aquifer test was conducted over a 24-hour period, starting at 10:35 am April 13, 2016, and ending at 10:35 am on April 14, 2016. Water levels were measured at all four monitoring wells throughout the duration of the aquifer test. Level TROLL® [*In-Situ Inc.*, 2007] pressure transducers were installed in monitoring wells 1 and 2 to collect more frequent data at these locations closest to the pumping well. At well 1, water levels were measured every minute. At well 2, water levels were measured at a variable rate, decreasing from 5 collections per second initially to 1 collection per minute after 15 minutes of pumping to capture the rapid changes that occur quickly upon pumping.

Manual depth-to-water measurements were collected at wells 3 and 4 throughout the 24-hour aquifer test, with a brief gap in data collection during the night, from hours 10 to 15 of pumping. Additionally, manual depths were measured at each well prior to pumping, with an average initial depth to water table of 3.02 meters below ground surface. The depth to aquifer bottom was measured to be 10.6 meters, giving an average initial saturated thickness of 7.58 meters.

Pumped water was piped and discharged roughly 230 meters southeast of the pumping well. This was a sufficient distance to avoid any gravitational disturbance due to recharge from the pumped water. The water was discharged through a flowmeter that measured the cumulative number of gallons pumped during the aquifer test. The pumping rate of the well was calculated to be 0.045 m^3/s throughout the test, as determined by the measured flowmeter discharge (Figure 3). We collected discharged volume data 12 times and measured 3812 cubic meters of total water drawn from the aquifer. The discharge rate was consistent throughout the test, as indicated by the linear increase in discharged volume (Figure 3).

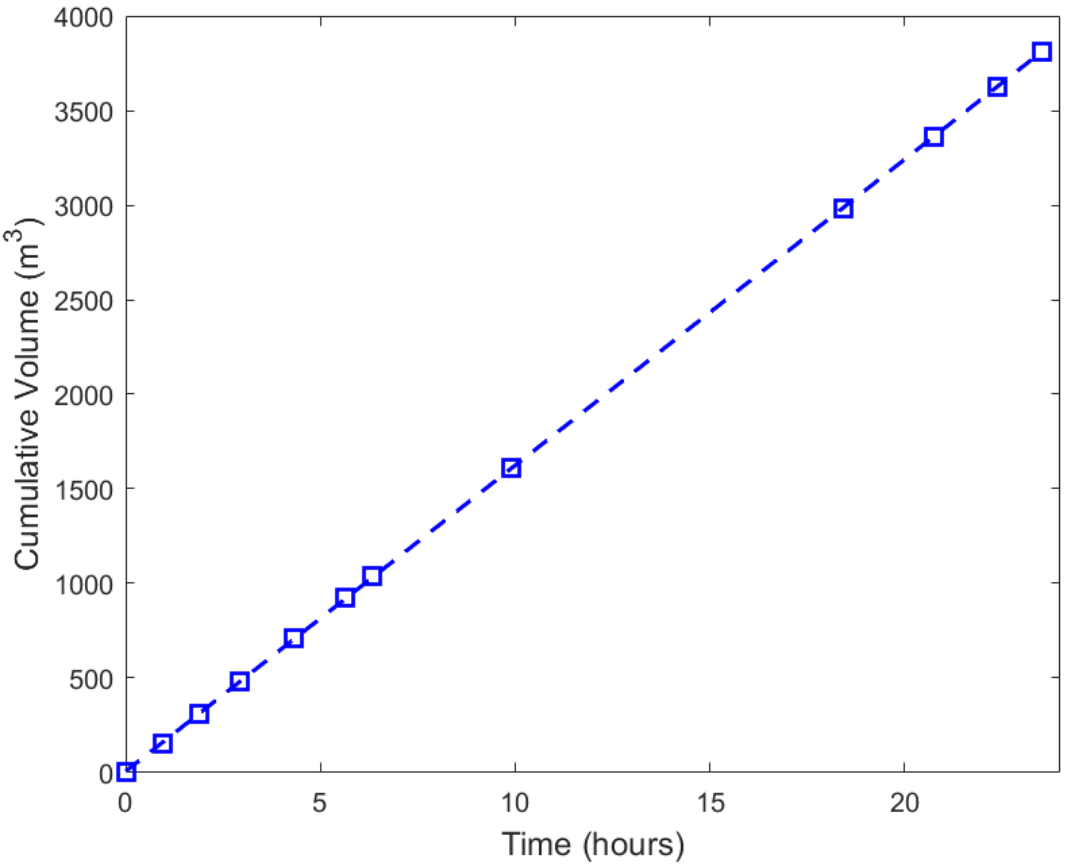

Figure 3. Flowmeter volume measured at the discharge point located 230 m southeast of pumping well. The dotted line represents the best-fitting discharge rate of 162 m^3/hr.

2.2.2. Gravity Data Collection

The Scintrex CG-5® [*Scintrex*, 2009] gravimeter was deployed to measure vertical gravity acceleration throughout the 24-hour pumping test. Previous work at this site by Woodworth [2011] indicated that moving the gravimeter during pumping required several minutes of wait time to allow the gravimeter's inner mechanism to settle, or data disturbances would be too great to isolate microGal precision anomalies. We therefore chose to maintain the gravimeter at a single location for the entire test duration. We placed the gravimeter at the same radial distance as monitoring well 2 (Figure 2) to associate the drawdown data with the change in vertical gravitational attraction at well 2. The gravimeter was placed on a 1-cm thick aluminum plate with three notches drilled for the three legs of the gravimeter's tripod. Additionally, aluminum sheets were set up along the sides of the gravimeter as a wind block to minimize tilt. We relied on the gravimeter's internal tilt correction for other minor tilt adjustments.

2.2.3. Evaluation of Potential Subsidence

The aquifer at the Horticulture farm site is pumped regularly for irrigation, so we did not anticipate significant subsidence due to pumping. However, changes in elevation of the gravimeter would potentially influence the measured gravity anomaly. Therefore, we monitored elevation changes to assess whether subsidence-related corrections were necessary for microGal-precision results.

Land surface elevations were measured using a total station survey instrument (Topcon Positioning Systems Inc., Livermore, CA). The total station was placed 95 meters away from the pumping well, and the total station's prism was buried into the ground adjacent to well 2. We

measured elevation changes of the prism 43 times throughout the 24-hour test. The measured change in elevation never exceeded 5 mm during the aquifer test (Figure 4).

A free-air correction is necessary to adjust for potential vertical movement of the gravimeter. Elevation changes on the order of centimeters correspond to multiple microGals [Telford et al., 1990]. If the ground level were to subside during the aquifer test, the gravimeter would move closer to the underlying mass resulting in an increase in gravity. Based on the surface elevation data collected with the total station, there were no observable changes in elevation beyond the millimeter scale, which results in anomalies on the order of 1 microGal with a gradient of 3.086 microGal/cm at 45° N latitude. Therefore, a free air correction due to elevation change is not necessary for this aquifer test (Figure 4).

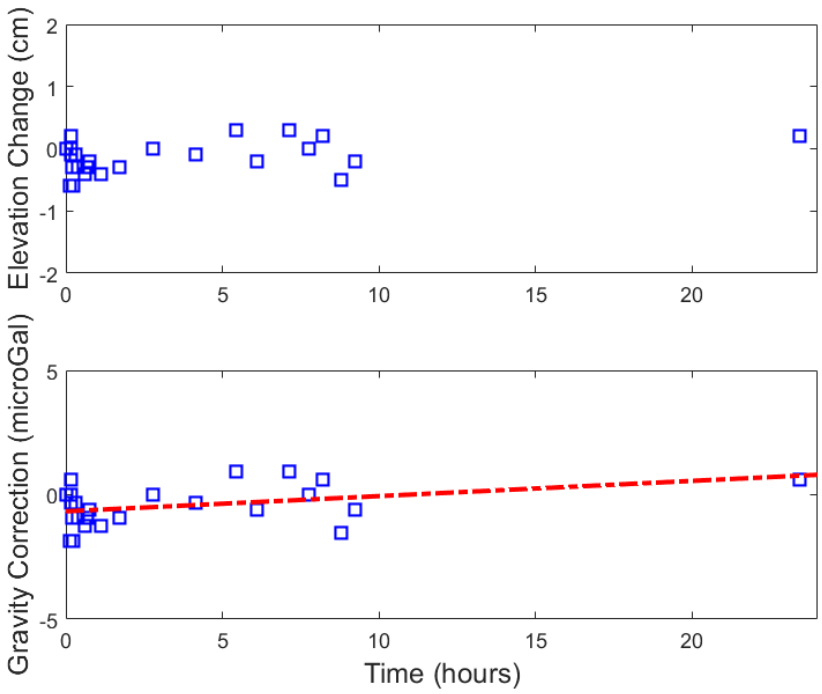

Figure 4. Change in ground surface elevation at monitoring well 2 and the associated free-air gravity correction. Negative elevation change indicates lowering (subsidence). The red dotted line indicates a negligible (<1 microGal) trend through the duration of the survey.

2.3. Gravity Data Reductions

To isolate the gravity anomaly associated with groundwater removal, we performed three data corrections, listed in order of increasing correction magnitude: (i) barometric correction to account for changes in air pressure, (ii) tide correction to account for gravitational tides and ocean loading, and (iii) drift correction to remove inherent instrument drift. We made no Bouguer elevation or terrain corrections to the data set, as we were only interested in the relative gravitational anomaly. Consequently, the absolute values of gravity are not known.

2.3.1. Barometric Air Pressure Correction

Changes in air pressure throughout the day represent mass changes above the gravimeter, so it is necessary to measure barometric air pressure and apply the associated corrections. The air pressure throughout the test was obtained from the nearest weather station at the Colorado State University campus, which is maintained by the Colorado Climate Center [CCC, 2016]. The weather station is approximately 8.5 km from the horticulture farm. Air pressure dropped by 6.49 mbar throughout the 24-hour aquifer test period (Figure 5). The associated correction nearby in Boulder, Colorado has been found to be -0.356 microGal/mbar [van Dam and Francis, 1998], resulting in a maximum correction of 2.31 microGals (Figure 5).

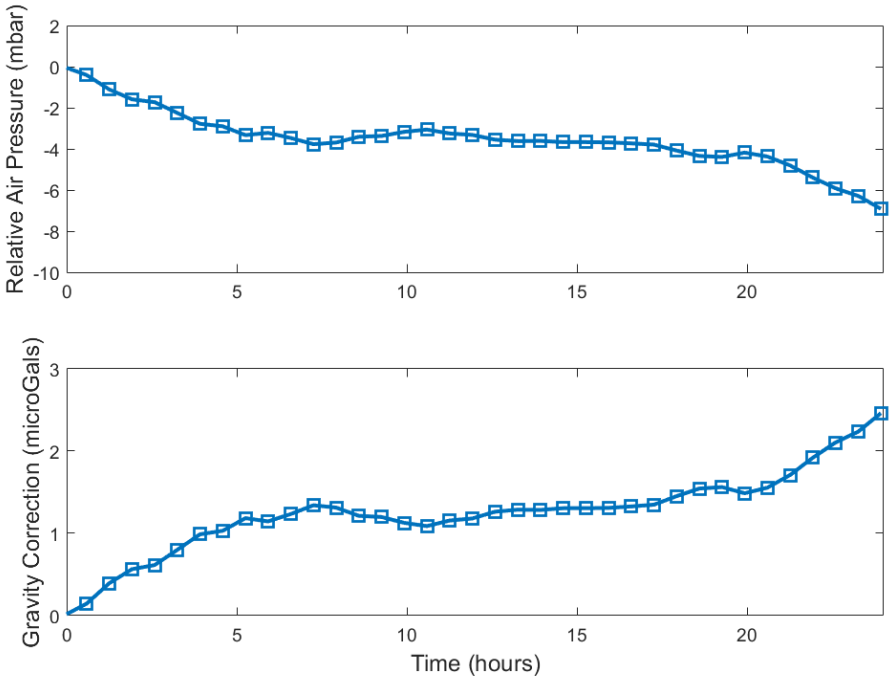

Figure 5. Relative atmospheric pressure versus time and the associated barometric correction. The initial air pressure at time 0 was 847.02 mbar. Barometric pressure data were collected at the Colorado State University Fort Collins Campus weather station [CCC, 2016].

2.3.2. Tidal and Ocean Loading Correction

As Earth tides change both with time and global positioning, it is necessary to calculate the time-varying gravitational impact of tides and remove this effect. The effects of tides are on the order of tens to hundreds of microGals, depending on location, and must be accounted for in microGal precision data sets. We verified the gravimeter's internal corrections based on user-specified position and time against computed values using the TSOFT software [Van Camp and Vauterin, 2005]. The values were found to agree within a microGal, so we utilized the gravimeter's internal tidal corrections (Figure 6). TSOFT was also used to evaluate the gravitational anomaly associated with ocean loading. As expected for a test site distant from oceans and large water masses, this effect was insignificant (less than a microGal).

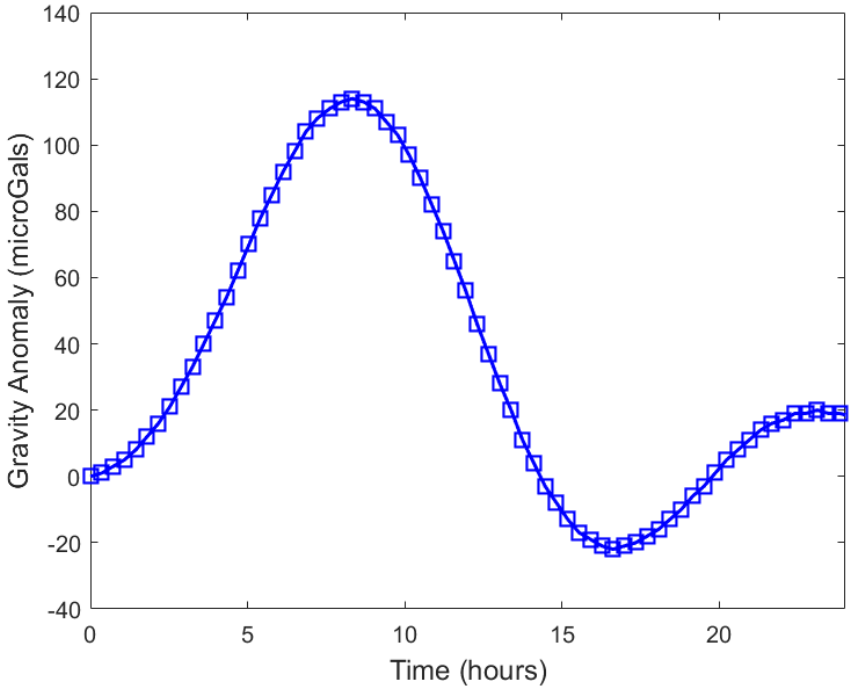

Figure 6. Gravity anomaly (microGal) due to tidal effects during the survey.

2.3.3. Manual Drift Correction

The largest magnitude correction for a multiple-hour gravimetry survey is for instrument drift. Drift rates are dependent on the properties of specific devices. For Scintrex CG-5® [*Scintrex*, 2009] gravimeters, the drift is manifested as a linear increase in measured gravity over periods of several hours. In surveys lasting more than a few hours, the drift rate may not be sufficiently modeled as constant, necessitating higher order functions to remove the effect to microGal-precision. To account for this, we measured the effect of drift for 12 hours continuously before and after data collection. We collected these drift data overnight at the bottom floor of the Warner College of Natural Resources at Colorado State University, where we did not anticipate seismic effects or localized changes in gravity. Although the device has built-in functionality to collect a drift value from such surveys, we manually determined drift values for a higher precision correction (Figure 7).

The magnitude of instrument drift was measured at 26 microGal/hr prior to field data collection and 24.6 microGal/hr after the field test (Figure 7). We infer that the drift gradient is therefore 26 microGal/hr at the start of the survey and 24.6 microGal/hr at the end of the survey and introduce a higher order function to linearly decrease drift during the survey at a rate of -1.4 microGal/hr/day (Figure 7).

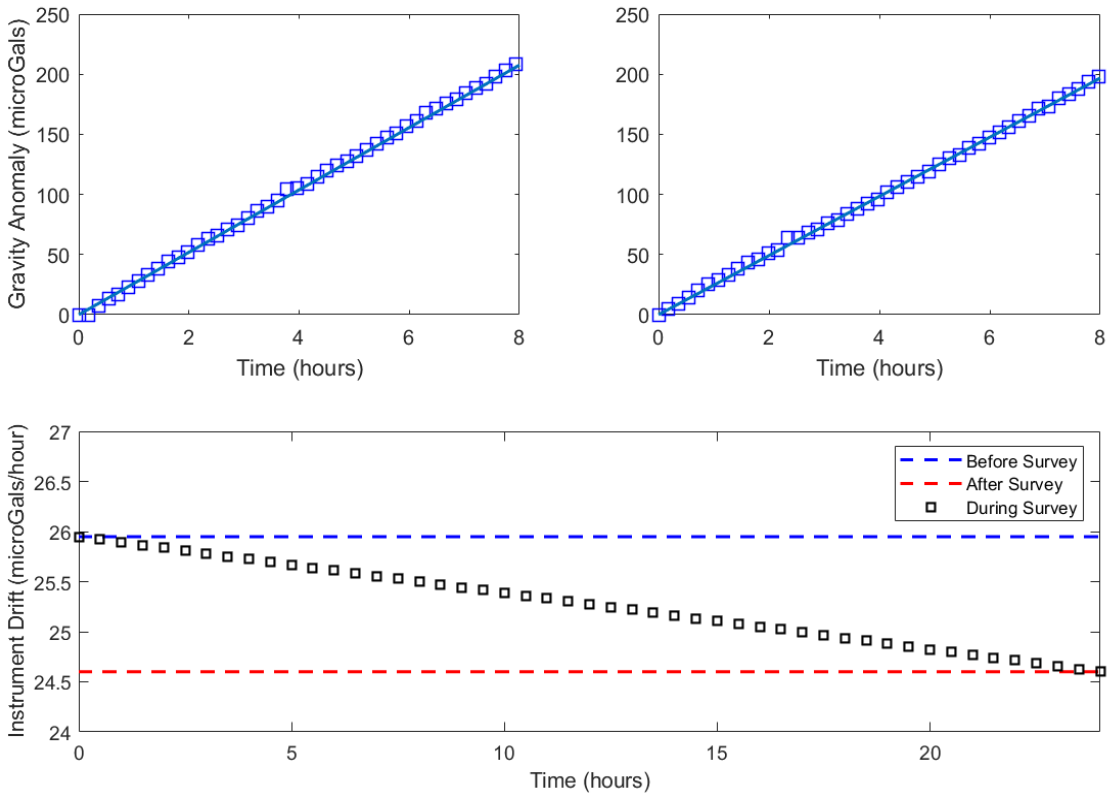

Figure 7. Measurement of instrument drift one day prior to survey (top left) and one day after survey (top right) and the drift function used for gravity reduction (bottom). The drift rate linearly decreases from the pre-survey rate to post-survey rate throughout the duration of the 24-hour survey.

The relative importance of all data corrections is shown in Figure 8. The uncorrected gravity anomaly increases linearly, as instrument drift is the largest magnitude effect over long surveys. After removing the effect of instrument drift, the data appear oscillatory through the

survey due to the large impact of tidal effects. After the removal of both drift and tides and less importantly, barometric effects, it is possible to isolate the gravitational anomaly associated with the removal of underlying water mass (Figure 8, bottom panel).

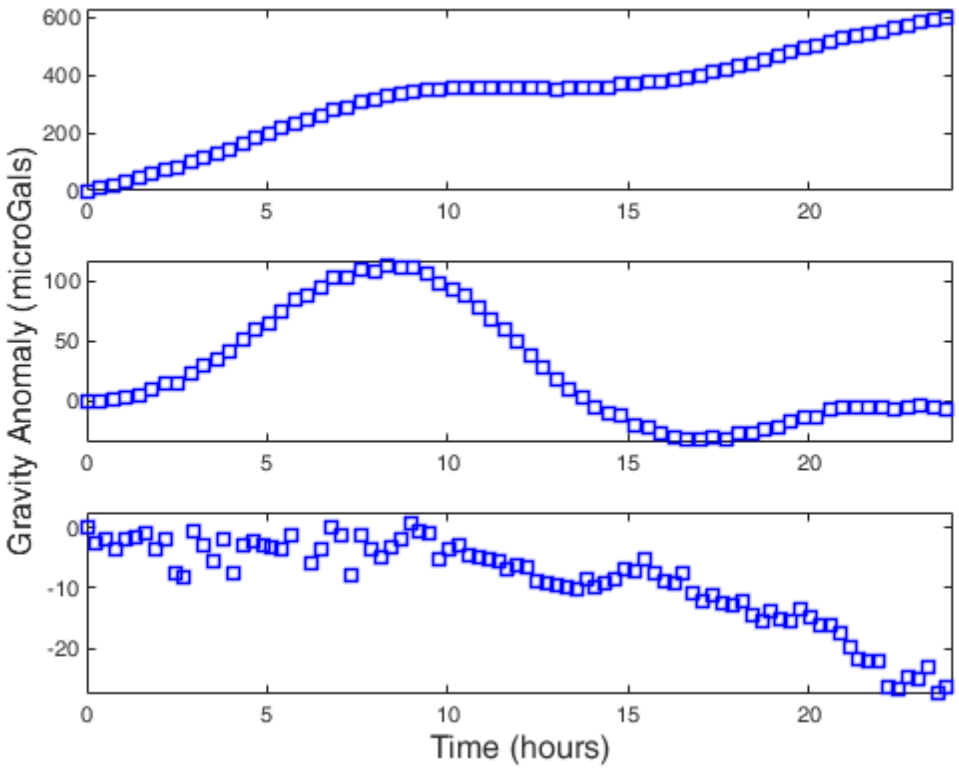

Figure 8. Measured gravity anomaly (microGals) during the 1-day aquifer test. (Top) uncorrected data. (Middle) gravity data corrected for instrument drift (see Figure 7). (Bottom) gravity data with corrections for drift, Earth tides (see Figure 6), and barometric effect (see Figure 5).

2.4. Aquifer Property Estimation Using Drawdown Data

We chose to estimate the aquifer parameters using the Theis [1935] analytical drawdown model, which is summarized in equations 2-1, 2-2, and 2-3 below. The modeled drawdown, d, at a given time, t_m, and radius, r_m, is dependent on the pumping rate, Q, the aquifer transmissivity, T, and the aquifer storativity, S.

$$u = \frac{r_m^2 S}{4Tt_m} \quad \text{(Eq. 2-1)}$$

$$W(u) = \int_u^\infty \frac{e^{-y}}{y} \partial y \quad \text{(Eq. 2-2)}$$

$$d(r_m, t_m) = \frac{Q}{4\pi T} W(u) \quad \text{(Eq 2-3)}$$

It is important to note that the Theis [1935] model uses a single storage parameter which is commonly intended to describe elastic release of water storage in confined aquifers. The Theis [1935] model will therefore not be able to characterize early transitions of storage release mechanisms occurring in our study's unconfined aquifer system. By fitting later-time data, however, the Theis [1935] model is mathematically equivalent to an incompressible Neuman [1972] model with the Theis storativity as an analog to Neuman specific yield. It is also worth noting that the transmissivity parameter occurs in two places in the drawdown solution (Eqs. 2-1 through 2-3), so the modeled drawdown is more sensitive to transmissivity than storativity, which is another anticipated challenge when using the Theis [1935] drawdown model to determine the best fitting storage parameter.

The best-fitting parameters (T and S) were computed by minimizing the sum of squared residuals, where an individual residual is the difference between the measured drawdown and the modeled drawdown at the associated time after pumping and well distance. The residual is defined in equation 2-4 below, and the associated sum of squared errors for each well is defined in equation 2-5.

$$res_{w_i} = d_{w_i} - f(r_w, t_i, P) \quad \text{(Eq. 2-4)}$$

$$err_w = \sum_{i=1}^{n_w} (res_{w_i})^2 \quad \text{(Eq.2-5)}$$

We chose two different schemes, WS1 and WS2, to assign weights to individual measurements for determination of best-fitting parameters. The weighting scheme WS1 straightforwardly gives the same relative weighting to each data point. Wells 1, 2, 3 and 4 contain 1441, 1538, 18, and 18 data points, respectively. Although WS1 results in a low absolute error, wells 1 and 2 dominate the estimation procedure given the abundance of data from those two observation locations.

As such, we developed weighting scheme WS2 to consider the number of elements in the error vectors of wells 1 and 2 and weighted the data at each well accordingly. Specifically, we weight the residual of each datum within each well by the length of the well's data vector, n_w, as shown in equation 2-6. When this residual vector is used in the non-linear inversion algorithm, the sum of the square of residuals in equation 2-7 is minimized.

$$res_{w_i} = \frac{d_{w_i} - f(r_w, t_i, P)}{\sqrt{n_w}} \quad \text{(Eq. 2-6)}$$

$$err_w = \sum_{i=1}^{n_w} \frac{(res_{w_i})^2}{n_w} \quad \text{(Eq.2-7)}$$

Equations 2-5 and 2-7 show how the overall error, or sum of squared residuals, is computed for each weighting scheme. WS2 thus reduces the bias associated with different data collection methodologies (i.e., drawdown measurement frequency) by decreasing the relative weight of well 1 and 2 data points.

2.5. Interpretation of the Gravity Anomaly using a Drawdown Model

We modeled the gravitational anomaly using a methodology similar to Damiata and Lee [2006]. This modeling approach approximates the gravitational anomaly associated with water mass removal by discretizing the drawdown cone as a series of stacked disks. We use the Theis

[1935] model described above with WS2 best-fitting aquifer parameters to obtain the cone of depression at discrete times. A 1-minute time step was used, and the model domain extended out to 400 meters. An example of the spatial discretization of the cone of depression using stacked disks is shown in Figure 9.

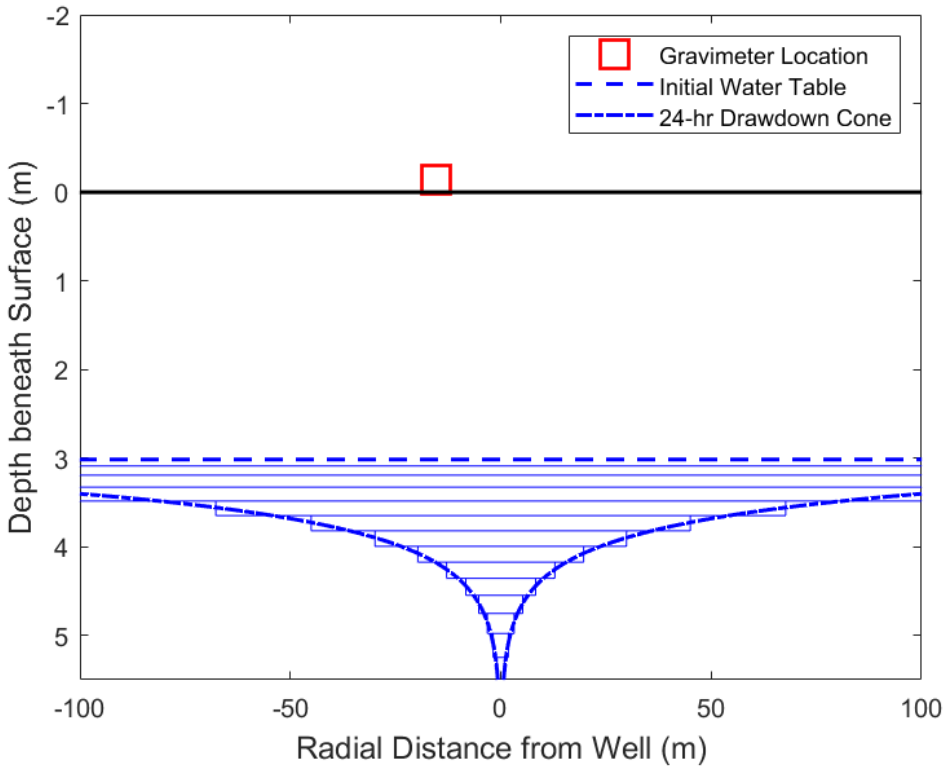

Figure 9. Disk stacking model for estimating gravity anomaly, following approach of Damiata and Lee [2006].

The disks are generated for the cone of depression at each time step in the model. Each disk's gravitational anomaly is the difference between the anomaly of two semi-infinitely extensive cylinders with offset heights, creating an inner solid disk. The gravitational anomaly associated with an infinitely extending cylinder at a given radial offset and depth relative to gravimeter location is given by Telford et al. [1990]. A conceptual model for this analysis is

shown in Figure 10, which displays the gravimeter position and the associated parameters to determine the anomaly of a single infinitely extending cylinder.

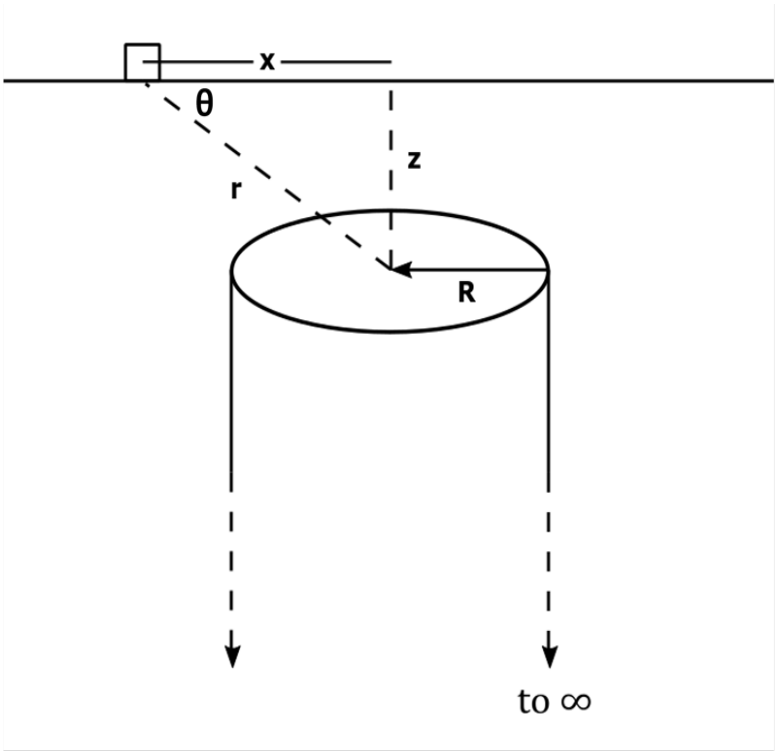

Figure 10. Model parameters for the gravitational attraction of a single offset infinitely extending cylinder.

Dependent on the position within the cone of depression, we use one of two Taylor series approximations (equations 2-8 and 2-9) to calculate the anomaly of given cylinders, representing the tops and bottoms of disks. The parameters r, R, and θ are described visually in Figure 10 and represent the distance from the gravimeter to the top and center of the cylinder, the cylinder radius, and the angle between the ground surface and the path from the gravimeter to the top and center of the cylinder, respectively. Additionally, γ is the gravitational constant (see Table 1), ρ_d is the disk density, $P_n(\mu)$ are Legendre polynomials of order n, and $\mu = \cos(\theta)$ [Telford et al., 1990].

$$while\ r > R$$

$$g(r,\theta) = 2\pi\gamma\rho_d R\left\{\left(\frac{R}{2r}\right) - \left(\frac{R}{2r}\right)^3 P_2(\mu) + \left(\frac{R}{2r}\right)^5 P_4(\mu) - \cdots\right\} \quad \text{(Eq. 2-8)}$$

$$while\ R > r$$

$$g(r,\theta) = 2\pi\gamma\rho_d R\left\{1 - 2\left(\frac{r}{2R}\right) P_1(\mu) + 2\left(\frac{r}{2R}\right)^2 P_2(\mu) - 2\left(\frac{r}{2R}\right)^4 P_4(\mu) + \cdots\right\} \quad \text{(Eq. 2-9)}$$

The disk density (ρ_d) is dependent on the pore space from which water mass drains, which we assume to be the specific yield (S_y) of the aquifer material, and the density of water (ρ_w) (see Table 1), in which $\rho_d = S_y \cdot \rho_w$. In our models we evaluate the disk density using specific yields ranging from 0.12 to 0.6 as described in Chapter 3.

Table 1. Table of model parameters for drawdown and gravity models.

Symbol	Parameter	Value(s)
Q	Pumping Rate (m^3/s)	0.045
t_m	Modeled Times (min)	1 – 1440
r_m	Modeled Radial Distances (m)	0.01 – 400
T	Best-Fitting Transmissivity (m^2/s)	0.014
S	Best-Fitting Storativity (--)	0.12
S_y	Specific Yield for Disk Model (--)	0.12 & 0.6
d	Modeled Drawdown (m)	See Eq. 2-3
x	Gravimeter Position (m)	15.4
Z	Depth to Disk (m)	3.10 – 11.45
r	Distance from Gravimeter to Disk Center (m)	$\sqrt{x^2 + z^2}$ = 15.71 – 19.19
R	Disk Radii (m)	r_m = 0.01 – 400
γ	Gravitational Constant ($m^3/kg{\cdot}s^2$)	6.672 x 10^{-11}
ρ_w	Water Density (kg/m^3)	999.75
ρ_d	Disk Density (kg/m^3)	$\rho_w \cdot S_y$ = 119.97 & 599.85
g	Gravitational Anomaly (m^2/s)	See Eqs. 2-8 & 2-9

CHAPTER 3 – RESULTS AND DISCUSSION

3.1. Drawdown Data Analysis

Groundwater extraction from the pumping well induced water table decline at all four observation wells (Figure 11). After 24 hours of pumping, the extent of water table decline ranged from 1.5 m at monitoring well 1 to 0.35 meters at the farthest monitoring well 61.2 meters from the pumping well (Figure 11).

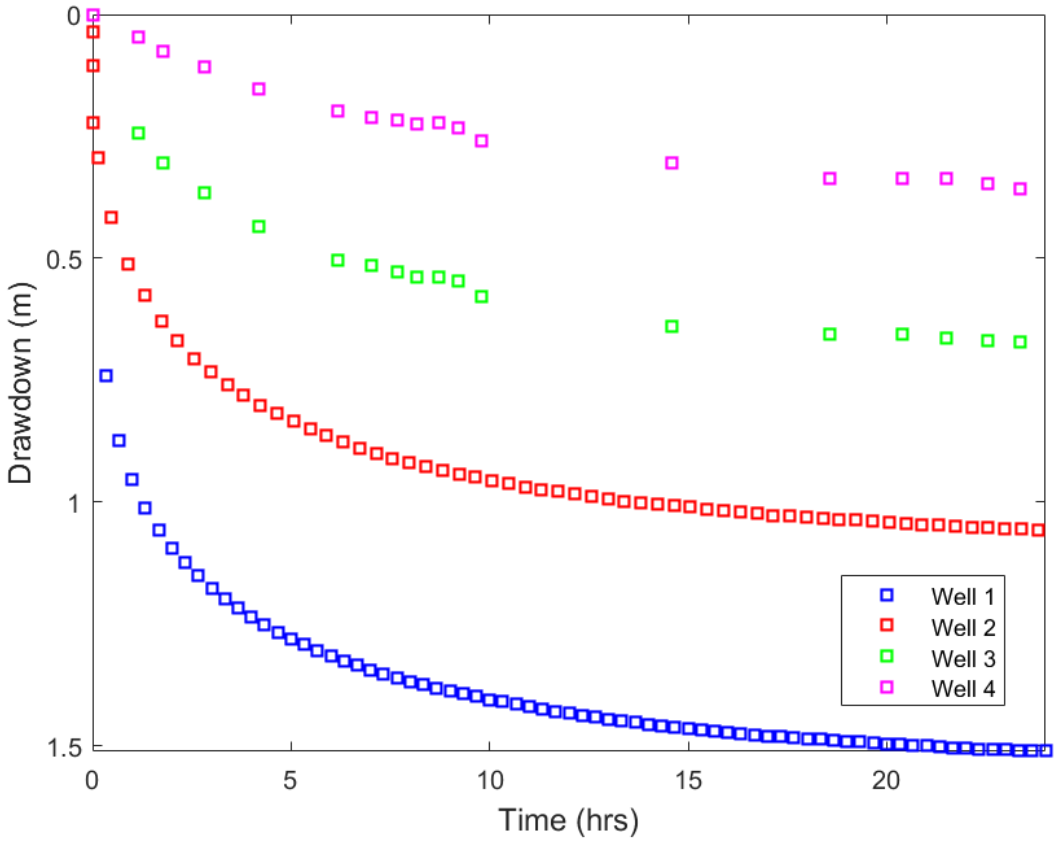

Figure 11. Measured drawdowns at all wells. Pressure transducers collected continuous drawdown data at wells 1 and 2, and manual measurements were made at wells 3 and 4.

We collected continuous pressure transducer data at wells 1 and 2, including 1441 total data points at well 1 and 1537 total data points at well 2. At observation wells 3 and 4, we

collected 18 manual depth-to-water measurements. When weighting all data points equally with the WS1 scheme, the parameter estimation solution is biased towards fitting drawdown at wells 1 and 2. If the aquifer is not homogeneous or has time-dependent properties, the parameters resulting in a good fit at wells 1 and 2 will not necessarily be a good fit farther from the pumping well (at wells 3 and 4), hence the need to develop and analyze the results from two weighting schemes WS1 and WS2.

We find that WS1 results in a best-fitting Theis model transmissivity of 0.017 m^2/s, and WS2 results in a transmissivity of 0.014 m^2/s. We also find that WS1 results in a best-fitting storativity of 0.059, whereas WS2 results in a substantially larger storativity of 0.122. Previous aquifer testing indicated best-fitting parameters of 0.018 m^2/s and 0.041 for transmissivity and specific yield, respectively, when fitting drawdown data with the Neuman [1972] drawdown model [Woodworth, 2011]. Woodworth [2011] also noted that when simultaneously fitting drawdown and gravity data the best fitting transmissivity decreased 56% to 0.008 m^2/s and specific yield increased 534% to 0.26, further indicating that the storage parameters are not well constrained by the analytical drawdown models.

As indicated by previous investigators [e.g., Nwankwor et al., 1984, 1992; Mao et al., 2011], analytical drawdown models are inherently insensitive to storage parameters. That is, changes to the storage parameters have less of an impact on the magnitude of drawdown than changes to transmissivity. We specifically look at the sensitivity of the Theis [1935] model by determining error values associated with a range of transmissivity and storativity values for both WS1 (Figure 12, top panel) and WS2 (Figure 12, bottom panel). For this analysis, the selected transmissivity values ranged from 5×10^{-4} m^2/s to 5×10^{-2} m^2/s and storativity values ranged from 5×10^{-3} to 0.5, increasing by equal increments in 100-element vectors. We evaluated the

model fits for 10,000 sets of modeled parameters from each combination of the 100-element vectors of transmissivity and storativity values. The summed error associated with each set of aquifer parameters was calculated by adding the error contribution of each well in accordance with Eq. 2-5 for WS1 and Eq. 2-7 for WS2. These errors were then divided by the maximum error from the matrix of simulated parameters as a form of normalization, to compare the relative errors associated with each weighting scheme.

As seen in Figure 12, the general behavior of the error function contours do not change significantly between weighting schemes. Because both schemes are based on the Theis [1935] drawdown model, the error function contours are skewed and elongated along the storativity axis, indicating small changes in model fit with changes to the storage parameter. In fact, the best-fitting parameter sets, identified as the location of minimum error in Figure 12, both fit within the same error contour, despite predicting storativity values differing by 107%. This analysis of error minimization alone does not make it clear which weighting scheme more appropriately describes the properties of the aquifer system, so it is necessary to compare the model fits of each.

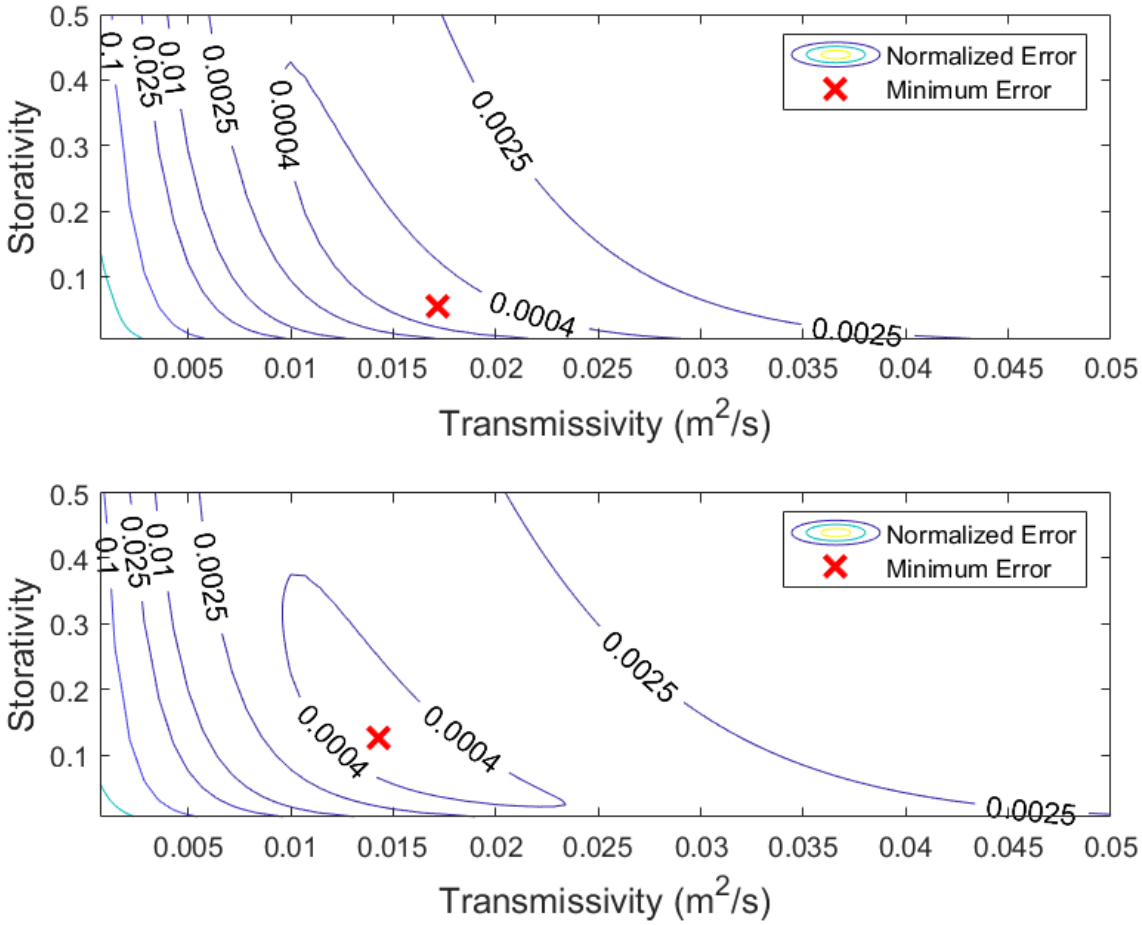

Figure 12. Normalized objective function for Theis model parameter (transmissivity, storativty) estimation. (Top) Results with data weighting scheme 1 (WS1); best-fitting T and S values are 0.017 m^2/s and 0.056, respectively. (Bottom) Results with data weighting scheme 2 (WS2); best-fitting T and S values are 0.014 m^2/s and 0.12, respectively.

To evaluate the effectiveness of WS1 and WS2 in describing drawdown, we compare the best fitting time-drawdown models of each weighting scheme in Figure 13. We first examine the fits at each well using WS1 (Figure 13, top panel) in which wells 1 and 2 are more strongly fit than wells 3 and 4. Specifically, the model significantly over-predicts drawdowns at wells 3 and 4. As anticipated, the model gives a best fit to the well 2 data; although, it under-predicts early-time drawdown (before about 2 hours) and over-predicts later-time drawdown (after about 9

hours). Similarly, well 1 drawdown is under-predicted before about 16 hours and over-predicted after about 18 hours. We know the capacity of the model to fit unconfined aquifer drawdown to be limited based on its use of a single storage parameter, but the discrepancies between measured and observed drawdown at wells 1 and 2 despite the dense data available could also be an indicator of spatial heterogeneity and/or time-dependent aquifer parameters.

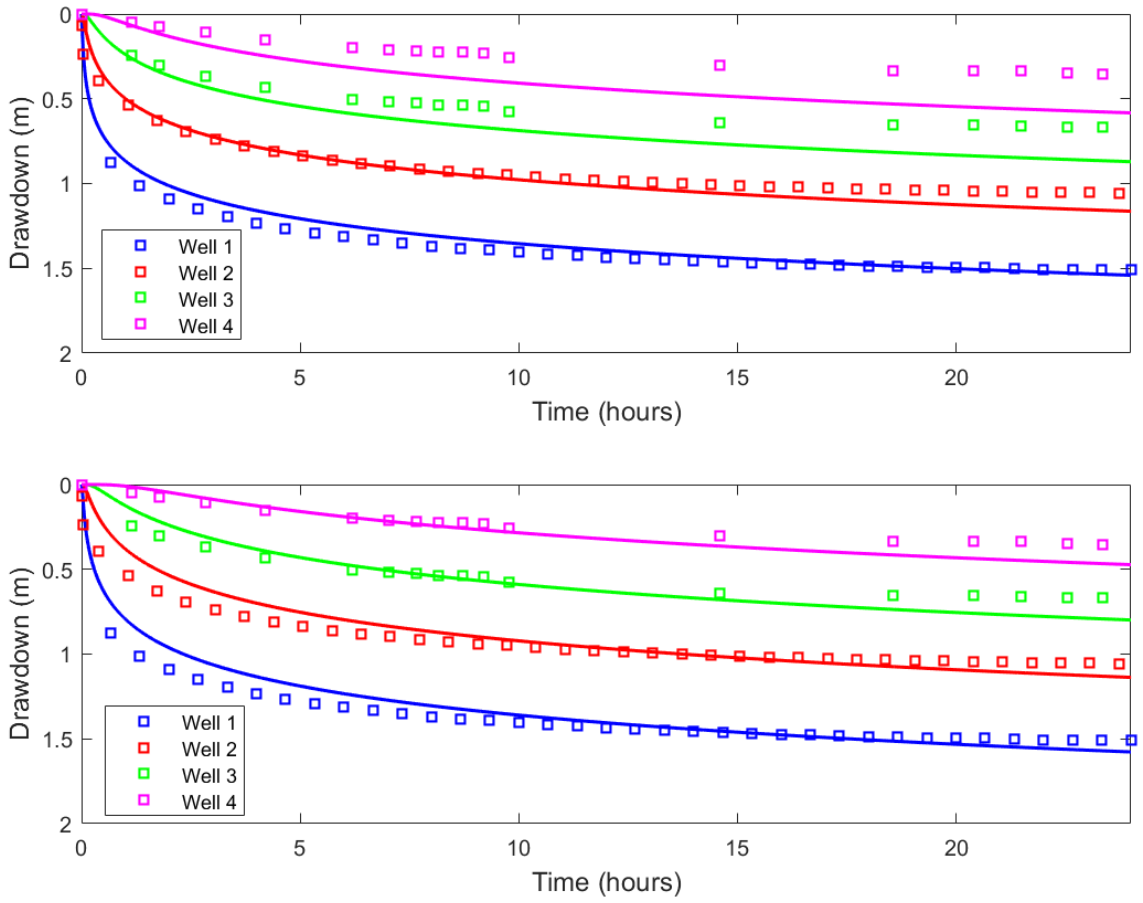

Figure 13. Comparison of Theis-solution modeled drawdown to the measured drawdown at each monitoring well. (Top) modeled response is based on best-fitting parameters from WS1. (Bottom) modeled response is based on best-fitting parameters from WS2.

The fits at each well using WS2 (Figure 13, bottom panel) exhibit a similar general behavior of under-predicting the observed drawdown at early time, then over-predicting the drawdown at later time. In fact, WS2 parameters result in a good fit of well 4 drawdown until over-predicting after about 8 hours. Well 3 drawdown is under-predicted until 8 hours then over-predicted. The same effect exists in wells 1 and 2 except over-predicting after 15 hours. Despite neither weighting scheme matching all data, WS1 consistently over-predicts drawdown at wells 3 and 4, whereas WS2 provides a good fit at each well for at least some portion of the time series.

Figure 14 displays the modeled drawdowns from both weighting schemes spatially at early, intermediate, and late times during the field test. The tendency of the models to under-predict then over-predict drawdown is apparent for each weighting scheme; however, the WS1 model (top) fits well at early times only and begins over-predicting drawdown more significantly than WS2 (bottom) afterwards. The WS2 model under-predicts drawdowns at early times but fits well intermediately, and more closely resembles measured drawdown at late times, especially at wells 3 and 4. As our drawdown-gravity model is derived from a single set of aquifer parameters, we assume a solution that most closely fits all wells simultaneously is more representative of the whole system than a solution that fits the wells with more available data. Accordingly, we henceforth utilize WS2 best-fitting parameters.

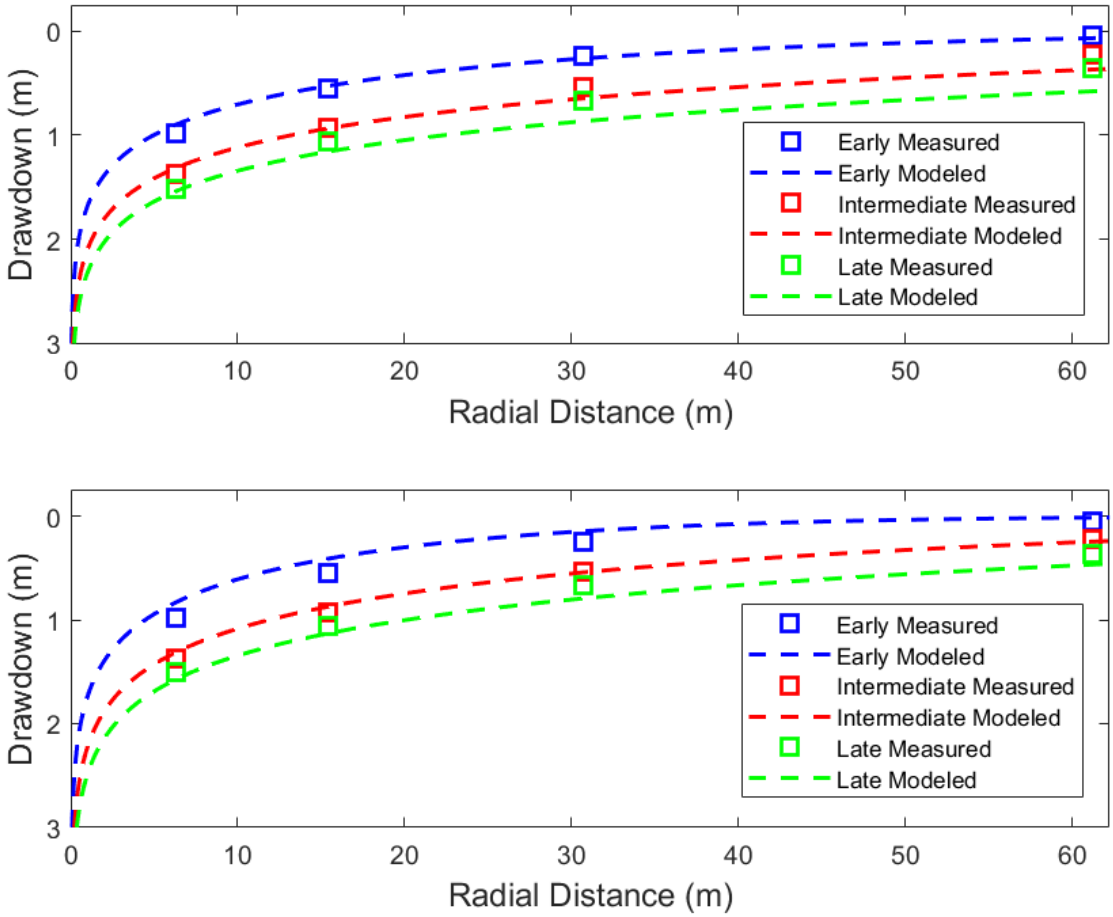

Figure 14. Comparison of measured and modeled drawdown versus distance from pumping well. Three times are considered: 1.15 hrs (early), 8.2 hrs (intermediate), and 23.4 hrs (late). (Top) Modeled response is based on best fitting parameters from WS1. (Bottom) Modeled response is based on best-fitting parameters from WS2.

To isolate the transition of storage mechanisms consistent with the pumping of unconfined aquifers, we collected data more rapidly at early times at well 2. The apparent shift of release mechanisms and associated S-shaped drawdown curve is illustrated in the log-log plot in Figure 15. A slowdown in drawdown gradient is observed from roughly 100 to 1000 seconds at well 2. Because data were not collected sufficiently rapidly at wells 1, 3 or 4, there is no apparent S-shaped drawdown curve, but this is not an indication that there is not an associated

shift of storage release mechanisms at these wells. As previously noted, the drawdown model is unable to characterize this effect and, consequently, rapid drawdown early at well 2 is significantly under-predicted.

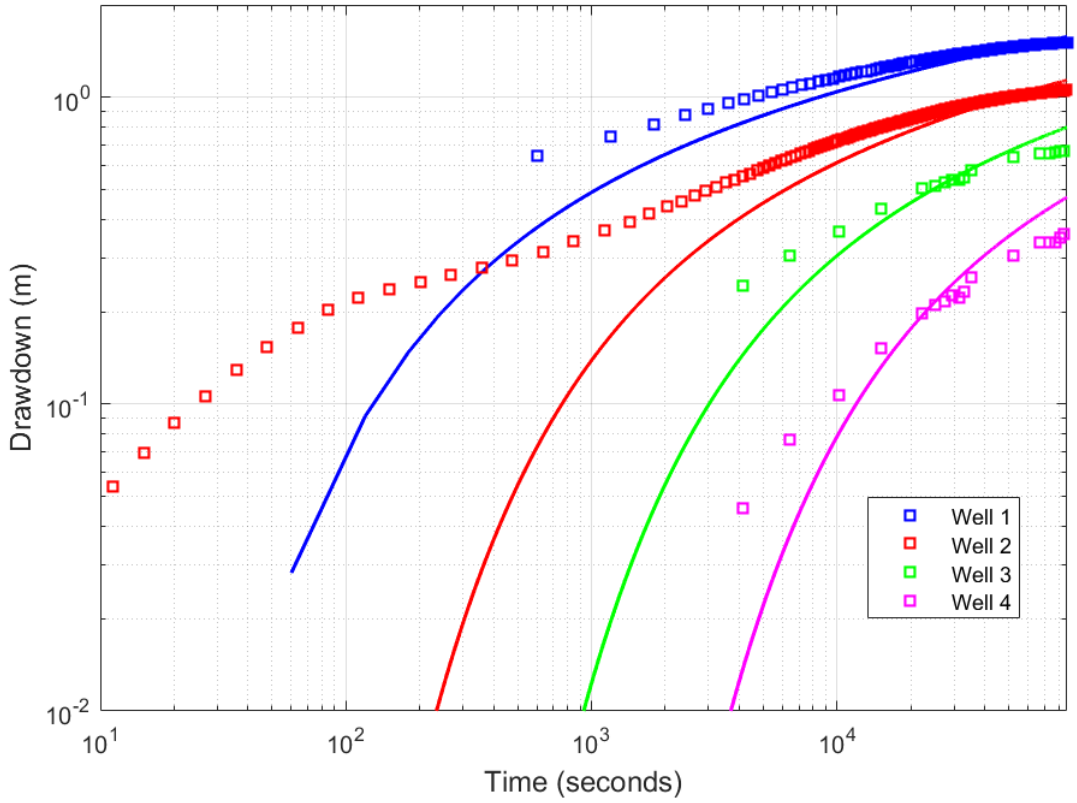

Figure 15. Log-log time-drawdown plots with measured drawdown at all wells and modeled drawdowns based on best-fitting Theis model parameters from WS2.

There is uncertainty in the storage release mechanisms of the aquifer, and there is complexity the model is unable to characterize. As an example, the tendency of the Theis [1935] model to under-predict then over-predict drawdown is consistent with the phenomenon of delayed drainage [Mao et al., 2011]. However, the drawdown data alone does not clarify whether it is anisotropy, heterogeneity, storage release transitions, or time-dependent parameters that the model fails to characterize and which of these site-specific factors may be most

impactful. We therefore seek to use the gravity data, which were collected simultaneously with drawdown, to further characterize the controls on storage change for the test site's shallow unconfined alluvial aquifer.

3.2. Gravity Data Analysis

Groundwater extraction resulted in a loss of water mass below the ground surface. We determined the total mass removed to be $3.89 \cdot 10^6$ kg using 3890 m^3 as the cumulative volume of water pumped as determined from the totalizer (Figure 3) and 999.75 kg/m^3 as the average water density. The loss of mass creates a density contrast between the saturated and drained layers of the aquifer, resulting in a decline in gravitational attraction. The 24 hours of pumping resulted in a large enough removal of underlying water mass to detect an observable decline in vertical gravitational attraction (Figure 16). The gravity data in Figure 16 are the corrected data previously shown in the bottom panel of Figure 8.

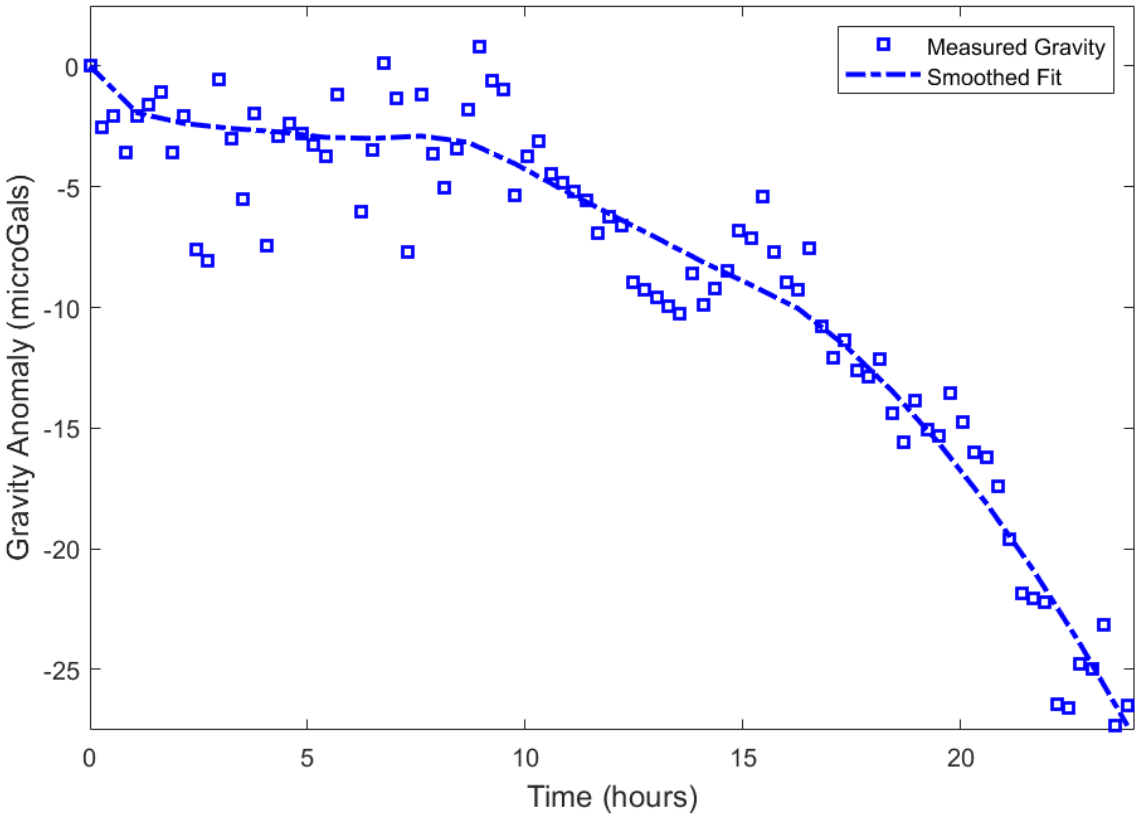

Figure 16. Gravity anomaly, following data reductions, measured at well 2.

The gravitational attraction decreased by 27.2 microGals after 24 hours of pumping (Figure 16). During the first 8 hours of pumping, the temporal change in gravity was roughly -0.3 microGals/hour, declining about 2.5 microGals. The measured gravity anomaly exhibits an inflection after 8 hours. In the remaining 16 hours the gravity gradient steepens to -1.5 microGals/hour, remaining relatively stable, and declining an additional 24.7 microGals.

These results support previous suggestions, primarily from synthetic modeling studies, that microgravimetric data can be used to detect and interpret groundwater mass changes during aquifer testing [Pool and Eychaner, 1995; Howle et al., 2003; Gehman et al., 2009; Creutzfeldt et al., 2010; Christiansen et al., 2011b; Van Camp et al., 2017]. Previous modeling studies of

gravitational response to aquifer pumping disagree on the magnitude and behavior of the response, largely due to the drawdown model attached to the gravity model [Damiata and Lee, 2006; Herckenrath et al., 2012; González-Quirós and Fernández-Álvarez, 2014; González-Quirós and Fernández-Álvarez, 2016]. As such, we evaluate the potential effectiveness of collected gravity data as a tool to better identify aquifer characteristics than drawdown data alone.

Previous synthetic modeling studies focused on unconfined aquifers have argued that simultaneously fitting gravity and drawdown data can better constrain specific yield [Damiata and Lee, 2006; Blainey et al., 2007; Herckenrath et al., 2012]. However, there is some question regarding whether gravity anomalies in aquifer tests will exceed the precision of the instrument. The instrument precision has been estimated multiple times at values of 4.76 microGals [Gehman et al., 2009], 2.5 – 5 microGals [Jacob et al., 2010], and 2 – 3 microGals [Christiansen et al., 2011a]. These values depend on the operation of the instrument as well as the methods used to quantify uncertainty. Even with conservative estimates of instrument precision at 5 microGals, this study documented a discernible and meaningful decline in gravitational anomaly due to pumping relative to the instrument precision.

Most synthetic modeling studies analyzing the gravitational response to pumping have estimated the anomaly spatially rather than temporally after a long duration of pumping, beginning with the Damiata and Lee [2006] model. That study considered a synthetic aquifer with a transmissivity of 0.05 m^2/s and a specific yield of 0.25. The simulated pumping rate was 0.063 m3/s, 40% larger than our study. At a radial distance of 15 m and after 1 day of pumping, the model predicted a gravity anomaly of roughly 18 microGals. This value is smaller than our measured anomaly of 27.2 microGals largely because they considered an initial water table depth of 10 m, compared to our initial water table depth at only 3.1 m below ground surface.

Interestingly, Damiata and Lee [2006] found that the gravitational anomaly after 1 day of pumping was relatively constant 1 – 10 meters from the well.

It is important to note that the Damiata and Lee [2006] gravity model is based on the Neuman [1972] drawdown model and consequently does not consider water storage release above the water table. Recognizing this, Herckenrath et al. [2012] developed a similar gravity-drawdown model which uses the Boulton [1970] and Moench [1997] approaches for modeling delayed drainage with an exponential function. For direct comparison with Damiata and Lee [2006] and Blainey et al. [2007], gravity data were only modeled spatially at the end of pumping, so the development of the gravity anomaly with time given delayed drainage is not apparent [Herckenrath et al., 2012]. However, the results of Herckenrath et al. [2012] show that the water-table decline (i.e., drawdown) is almost negligibly affected by the delayed drainage coefficient, whereas the behavior of gravity versus distance is much different with and without delayed drainage.

Recently, numerical models have been developed which consider a more robust description of flow near the water table. The model developed by González-Quirós and Fernández-Álvarez [2014] confirmed the gravity model's sensitivity to specific yield by plotting the simulated gravitational anomaly versus radius after 7 days of pumping under Damiata and Lee [2006] conditions over an array of specific yield values, noting that the magnitude of the anomaly at a given location increased roughly 2 microGals per 0.01 increase in specific yield. González-Quirós and Fernández-Álvarez [2016] evaluated the gravitational anomaly over time, and importantly, determined that the gravitational acceleration continuously decreased for the duration of the 7-day pumping test, steepening after 100 to 1000 minutes. That is, despite

drawdown rates eventually slowing near the pumping well, the gravitational anomaly steadily increased. This agrees with our measured drawdown and gravity data.

To test whether the gravity anomaly may be sufficiently predicted from our drawdown model, we evaluated the drawdown-gravity model using the best fitting storativity of 0.12 as specific yield and compared to measured values (Figure 17). The drawdown-gravity model at a specific yield of 0.12 resulted in a 24-hour modeled gravity anomaly just exceeding 5 microGals (Figure 17). Despite being lower in magnitude than the instrument precision, this model agrees well with the measured data in the first 8 hours (Figure 16). The approximate fit of the gravity model with a realistic specific yield at early times could suggest that the drawdown-gravity model which assumes no unsaturated or capillary fringe flow describes the system behavior in the early-intermediate times before storage release mechanisms begin to differ from modeled assumptions. As noted previously, however, the time-gravity inflection changes after 8 hours and the gravity declines more rapidly henceforth.

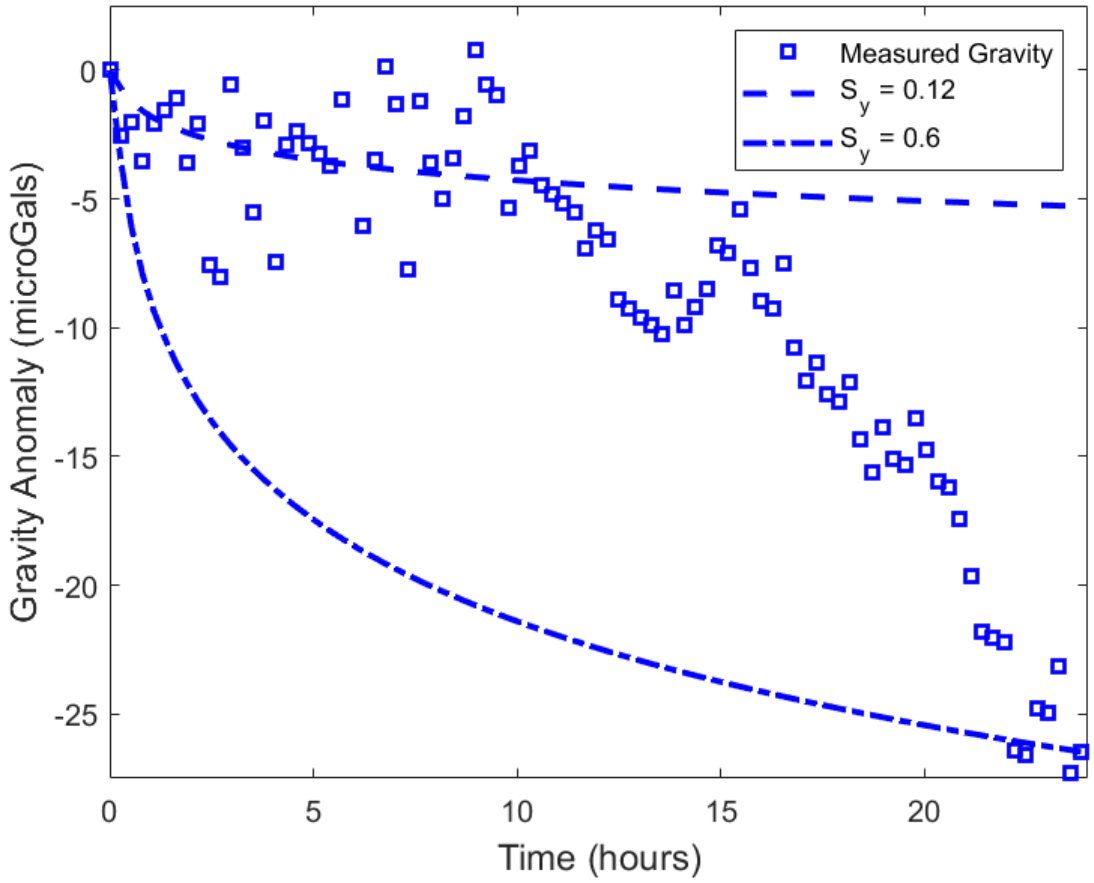

Figure 17. Modeled gravity anomaly using disk stacking methodology (Figure 9) with disk shape determined from Theis-modeled drawdowns (WS2 optimal parameters) and disk specific yields of 0.12 and 0.6.

As shown in Figure 17, a physically unrealistic specific yield of 0.6 is necessary to fit the same late-time total anomaly of 27.2 microGals. However, the shape of the gravity anomaly from the drawdown-gravity model is analogous to the shape of the drawdown curve. As such, the modeled gravity declines rapidly in the first 3 hours. According to this model with high specific yield, the magnitude of modeled gravitational anomaly in the first 3 hours is roughly the same as the anomaly in the following 21 hours. This is in direct conflict with the measured anomaly which reaches its steepest decline at the end of the field test. In fact, there is no specific yield

value or combination of Theis [1935] parameters in the drawdown-gravity model that reflects this change in inflection of gravitational anomaly.

Figure 18 shows combined plots of the measured drawdown and gravitational anomaly at well 2, which allows for an evaluation of the relationship between the drawdown behavior and the change in inflection in the gravity anomaly. The results show that the gravity gradient sharply increases as the rate of drawdown at well 2 begins to slow. This indicates that despite the water table progressing towards steady-state conditions, there is still rapid drainage of water in a region near enough the gravimeter to result in a sharp gravity decline. This is consistent with the findings of Mao et al. [2011], who noted that as drawdown approaches steady state conditions, water is released almost entirely from unsaturated and previously saturated zones. The log-log drawdown and gravity plots in Figure 18 show that the gravity gradient steepens about 2 hours after the transition of flow mechanisms that results in a reduction in drawdown rates from 100 to 1000 seconds. The magnitude of drainage due to compression may therefore be much smaller than the magnitude of drainage above the water table, and importantly, neither is characterized by the Theis [1935] model.

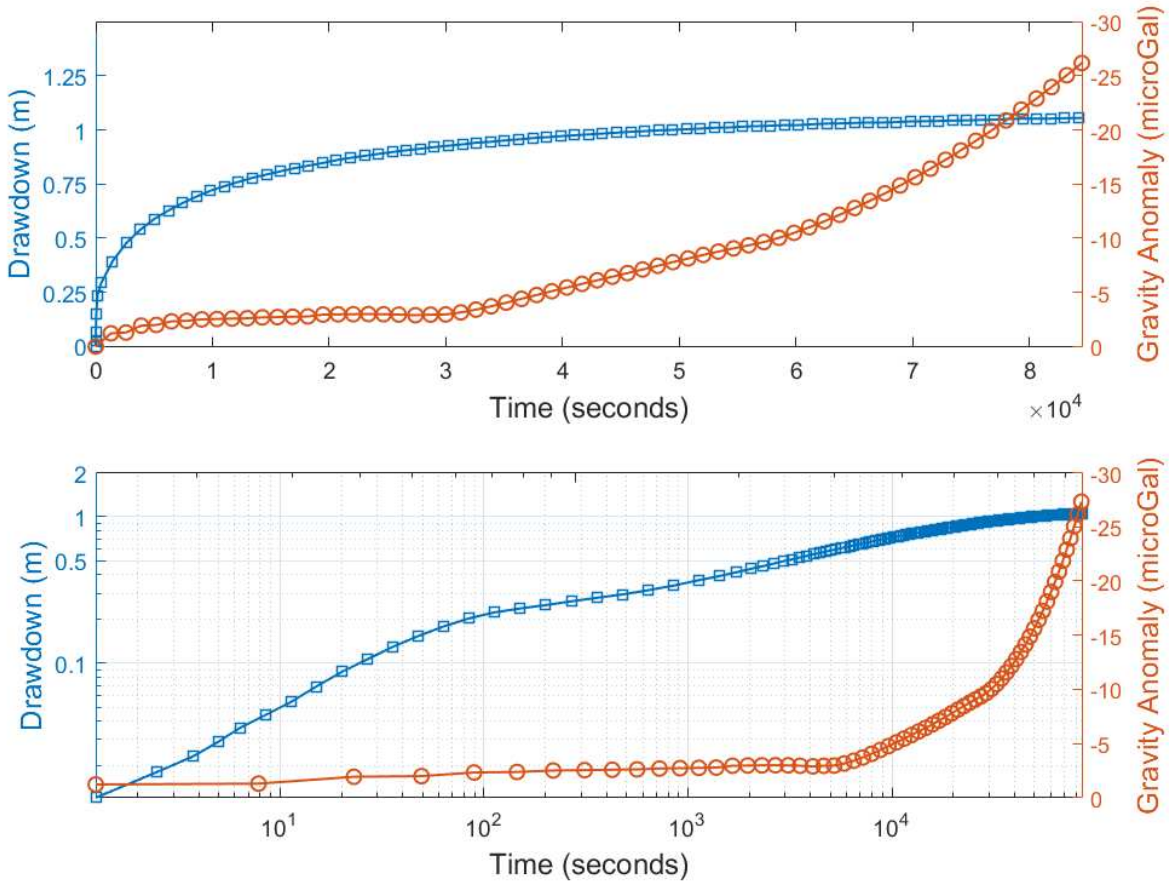

Figure 18. Dual plots of the measured drawdown and gravity anomaly at well 2 on (Top) arithmetic scale and (Bottom) logarithmic scale.

Because the modeled gravity with Theis assumptions differs significantly from the measured gravity, we assume that the analytical drawdown model does not accurately characterize flow mechanisms, and that the attached gravity model is too simplified to be used for constraining drawdown model parameters. Consequently, we explore the model assumptions and why the measured gravity data are not reflective of the modeled gravity data.

As discussed in the gravity data analysis, a specific yield of 0.6 in the drawdown-gravity model is necessary to match the measured gravity anomaly at the end of our 24-hour field test.

Given the assumption of instantaneous and complete drainage restricted to the initially saturated zone, this constitutes an unrealistic pore space as well as significantly more mass than was actually removed. It shows that no equal distribution of water removal within the initially saturated zone will adequately fit the measured gravity data without relying on unrealistically large specific yields. However, as shown in Eqs. 2-8 and 2-9, the gravitational anomaly measured at a location due to removal of a mass decreases with the distance and the angle to the mass. If water mass was removed closer to the gravimeter than predicted by the drawdown-gravity model, the measured anomaly would be larger than the modeled anomaly.

This result is consistent with hydrogeologic studies which evidence the significant magnitudes of vadose and tension-saturated drainage [e.g. Nwankwor et al., 1984, 1992; Narasimhan and Zhu, 1993; Bevan, 2002; Mao et al., 2011]. Nwankwor et al. [1992] specifically found that as water table drawdown slowed in the late stage of pumping, the water content profile began to drawdown, with rapid capillary fringe drawdown after reaching a critical tension in the tension-saturated zone. Most water storage was removed from the zone of tension-saturation during the late stage stages of pumping. This water content drawdown behavior was further supported by Narasimhan and Zhu [1993] in their analysis of soil columns.

Numerical models evaluating the spatiotemporal changes in aquifer storage during pumping further evidence vadose and tension-saturated storage release. Mao et al. [2011] finds that at early times during pumping tests, there are negligible changes in storage above the water table, with most storage removed near the pumping well in the saturated zone. At intermediate times, release of storage from the vadose zone began to increase, and drainage in the unsaturated zone became the primary source for well discharge. Later the storage change rates became minimal in the saturated zone, with storage release from unsaturated and previously saturated

zones reaching the rate of discharge as water table drawdown approached steady state conditions [Mao et al., 2011]. This is consistent with the numerical drawdown-gravity models of González-Quirós and Fernández-Álvarez [2016], which predict steadily decreasing gravitational acceleration despite slowing drawdown rates. In our study, this is evidenced by the simultaneous drawdown-gravity plots (Figure 18) that show gravity gradients to be largest late in the pumping test when the water table has begun to stabilize.

CHAPTER 4 – CONCLUDING REMARKS

4.1. Conclusions

A temporal gravity survey was conducted during a 24-hour aquifer pumping test of a shallow, relatively thin unconfined aquifer. Groundwater removal yielded a total gravity decline of 27.2 microGals at a distance of 15.4 meters from the pumping well. This result demonstrates that gravity anomalies larger than the instrument precision are detectable and interpretable. To obtain a clear gravity signal that is associated with groundwater pumping, we note the importance of carefully characterizing the instrument drift before and after the field survey, especially for surveys lasting 1 or more days. Additionally, the hydrogeologic setting is important to consider, as the gravity anomaly is attenuated at depth and may be undetectable in deep aquifer. Similarly, the ground surface location of the gravimeter is important in shorter duration surveys, as most water mass removal occurs radially closer to the pumping well.

In our study we determined that the temporal gravity anomaly exhibits an intermediate time change in inflection that is uncharacterized by many previous analytical drawdown-gravity models. Our analysis reinforces previous studies noting the limitations of analytical drawdown models for predicting water mass changes in unconfined aquifers. We conclude that the difference between measured and modeled gravity is therefore due to the assumptions of the drawdown model. Temporal gravity data shows utility as a tool to determine whether important model assumptions such as vadose storage release, tension-saturated zone storage release, and hydraulic anisotropy are important to consider. The parameters controlling these properties and their development through time may be detectable by one or more gravimeters collecting temporal gravity data during pumping tests.

4.2. Recommendations for Future Work

This study verifies the capacity of gravimeters to detect significant gravitational anomalies during aquifer pumping tests and therefore lends itself to various future work as there are unanswered questions concerning the exact nature of the water mass changes that produce the measured decline in gravity. We recommend testing more robust drawdown-gravity models considering non-steady state conditions which model gravity without using the disk stacking method while allowing for flow above and within the initial saturated zone, compressibility, and anisotropy, such as the coupled model developed by Gonzales-Quiros and Fernandez-Alvarez [2016]. Calibrating a more rigorous model using the drawdown and gravity data from this field test would further quantify gravity's utility in hydrogeophysical surveys and whether it is possible to estimate multiple aquifer parameters and/or quantify each parameter's sensitivity to gravity data.

Additionally, we recommend a longer survey to allow late-time flow mechanics to develop at distances farther from the pumping well. It would be useful to document the behavior of gravity as these flow regimes shift and develop. Further, a repeated study with multiple gravimeters may quantify gravity's ability to characterize aquifer heterogeneity. More data sets, analyses, and model developments are necessary to develop robust drawdown-gravity models which can accurately predict gravity anomalies during pumping as well as obtain digestible and meaningful information from pumping tests that better identifies aquifer characteristics than the drawdown data alone.

REFERENCES

Bevan, M. J. (2002), A detailed study of water content variation during pumping and recovery in an unconfined aquifer, M.Sci. thesis, Univ. of Waterloo, Waterloo, Ont., Canada.

Blainey, J. B., T. P. A. Ferre´, and J. T. Cordova (2007), Assessing the likely value of gravity and drawdown measurements to constrain estimates of hydraulic conductivity and specific yield during unconfined aquifer testing, *Water Resources Research, 43*, W12408, 1-9, doi: 10.1029/2006WR005678

Boulton, N. S. (1954), Unsteady radial flow to a pumped well allowing for delayed yield from storage, *Int. Assoc. Sci. Hydrol. Publ.*, 2, 472–477.

Boulton, N. S. (1963), Analysis of data from non-equilibrium pumping tests allowing for delayed yield from storage, *Proc. Inst. Civil Engr*., v. 26, 469-482.

Boulton, N. S. (1970), Analysis of data from pumping tests in unconfined anisotropic aquifers, *J. Hydrol.*, 10(4), 369-378.

Boulton, N. S. and J. M. A. Pontin (1971), An extended theory of delayed yield from storage applied to pumping tests in unconfined anisotropic aquifers, *Journal of Hydrology*, 14, 53-65.

Christiansen L., Lund S., Andersen O.B., Binning P.J., Rosbjerg D., and P. Bauer-Gottwein (2011), Measuring gravity change caused by water storage variations: performance assessment under controlled conditions. *J Hydrol*, 402(1–2):60–70.

Christiansen, L., P.J. Binning, D. Rosbjerg, O. B. Anderson, and P. Bauer-Gottwein (2011), Using time-lapse gravity for groundwater model calibration: An application to alluvial aquifer storage. *Water Resources Research*, 47, W06503, doi: 10.1029/2010WR009859.

Colorado Climate Center (CCC) (2016) Colorado State University Fort Collins Campus Weather Station Data. https://ccc.atmos.colostate.edu/~autowx/. Accessed May 2016

Cooley, R. L. (1971), A finite difference method for unsteady flow in variably saturated porous media: Application to a single pumping well, *Water Resources Research, 7(6), 1607.*

Creutzfeldt, B., A. Güntner, S. Vorogushyn, and B. Merz (2010), The benefits of gravimeter observations for modelling water storage changes at the field scale. *Hydrology and Earth Systems Sciences*, 14, 1715-1730, doi: 10.5194/hess-14-1715-2010.

Damiata, B. N. and T. C. Lee (2006), Simulated gravitational response to hydraulic testing of unconfined aquifers, *Journal of Hydrology*, *318*, 348-359, doi: 10.1016/j.jhydrol.2005.06.024.

Gehman, C. L., D. L. Harry, W. E. Sanford, J. D. Stednick, and N. A. Beckman (2009), Estimating specific yield and storage change in an unconfined aquifer using temporal gravity surveys, *Water Resources Research, 45*, W00D21, doi: 10.1029/2007WR006096.

González-Quirós, A. and J.P. Fernández-Álvarez (2014), Simultaneous Solving of Three-Dimensional Gravity Anomalies Caused by Pumping Tests in Unconfined Aquifers, *Math Geosci*, 46(6), 649-664, doi: 10.1007/s11004-016-9539-9

González-Quirós, A. and J.P. Fernández-Álvarez (2016), Forward coupled modeling and assessment of gravity anomalies caused by pumping tests in unconfined aquifers under unsteady-state conditions, *Math Geosci*, 49, 603-617, doi: 10.1007/s11004-016-9634-1

Google Maps (2018). Retrieved from https://www.google.com/maps/@40.5996182,-105.0234433,11.99z/data=!5m1!1e4

Herckenrath, D., E. Auken, L. Christiansen , A.A. Behroozmand, and P. Bauer-Gottwein (2012) Coupled Hydrogeophysical inversion using time-lapse magnetic resonance sounding and time-lapse gravity data for hydraulic aquifer testing: will it work in practice? *Water Resour Res*. doi:10.1029/2011WR010411

Howle, J. F., S. P. Phillips, R. P. Denlinger, and L. F. Metzger (2003), Determination of specific yield and water-table changes using temporal microgravity surveys collected during the second injection, storage, and recovery test at Lancaster, Antelope Valley, California, November 1996 through April 1997, *U.S. Geological Survey Water-Resources Investigations Report 03-4019*, 1-28

In-Situ Inc. (2007), *Level TROLL operator's manual*, In-Situ Inc., Fort Collins, Colorado, 1-75.

Jacob T, Bayer R, Chery J, Le Moigne N (2010), Time-lapse microgravity surveys reveal water storage heterogeneity of a karst aquifer. *J. Geophys. Res.*, 115, B06402. doi:10.1029/2009JB006616

Lindsey, D. A., W. H. Langer, and D. H. Knepper, Jr. (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*, 1-70.

Mathias, S. A., and A. P. Butler (2006), Linearized Richards' equation approach to pumping test analysis in compressible aquifers, *Water Resour. Res.*, 42, W06408, doi:10.1029/2005WR004680.

Mao, D., L. Wan, T. J. Yeh, C. Lee, K. Hsu, J. Wen, and W. Lu (2011), A revisit of drawdown behavior during pumping in unconfined aquifers, *Water Resources Research*, 47, W05502, doi:10.1029/200WR009326.

Maupin, M. A. and N. L. Barber (2005), Estimated withdrawals from principal aquifers in the United States, 2000, *U.S. Geological Survey Circular 1279*, 1-47

Mishra, P. K. and S. P. Neuman (2010), Improved forward and inverse analyses of saturated-unsatured flow toward a well in a compressible unconfined aquifer, *Water Resources Research*, 46, W07508, doi:10.1029/2009WR008899.

Moench, A. F. (1995), Combining the Neuman and Boulton models for flow to a well in an unconfined aquifer, *Ground Water*, 33(3), 378-384.

Moench, A. F. (1997), Flow to a well of finite diameter in a homogeneous, anisotropic water table aquifer, *Water Resour. Res.*, 33, 1397-1407, 593-596.

Moench, A. F. (2004), Importance of the Vadose Zone in Analyses of Unconfined Aquifer Tests, *Ground Water*, 42(2), 223-233.

Moench, A. F. (2008), Analytical and numerical analyses of an unconfined aquifer test considering unsaturated zone characteristics, *Water Resour. Res.*, 44, W06409, doi:10.1029/2006WR005736.

Moreland, D. C. (1980), *Soil survey of Larimer County area*, Colorado, National Cooperative Soil Survey, U.S.

Narasimhan, T. N., and M. Zhu (1993), Transient flow of water to a well in an unconfined aquifer: Applicability of some conceptual models, *Water Resour. Res.*, 29(1), 179–191, doi:10.1029/92WR01959.

Narasimhan, T. N. (1999), Boulton's Delayed Yield: A Different Context, *Ground Water*, 37(3), 323-326.

Neuman, S. P. (1972), Theory of flow in unconfined aquifers considering delayed response of the water table, *Water Resources Research*, 8, 1031-1045.

Neuman, S. P. (1987), On methods of determining specific yield, *Ground Water*, 25, 679-684.

Nwankwor, G. I., J. A. Cherry, and R. W. Gillham (1984), A comparative study of specific yield determinations for a shallow sand aquifer, *Ground Water*, 22, 764-772.

Nwankwor, G. I., R. W. Gillham, G. van der Kamp, and F. F. Akindunni (1992), Unsaturated and Saturated Flow in Response to Pumping of an Unconfined Aquifer: Field Evidence of Delayed Drainage, *Ground Water*, 30(5), 690–700, doi:10.1111/j.1745-6584.1992.tb01555.x.

Pool, D. R. and J. H. Eychaner (1995), Measurements of aquifer-storage change and specific yield using gravity surveys, *Ground Water*, 33, 425-432.

Scintrex (2009), *CG-5 operation manual*, Scintrex Ltd, Concord, Ontario, Canada.

Tartakovsky, G. D., and S. P. Neuman (2007), Three-dimensional saturated-unsaturated flow with axial symmetry to a partially penetrating well in a compressible unconfined aquifer, Water Resour. Res., 43, W01410, doi:10.1029/2006WR005153.

Telford, W. M., L. P. Geldart, and D. A. K. Sheriff (1990), *Applied Geophysics*, Cambridge University Press, Array Cambridge, England.

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, *Am. Geophys. Union Trans.,*16, 519-524.

United States Geological Survey (1979), Geologic map of Colorado, Interior - Geologic Survey, Reston, V.A.

Van Camp, M. and P. Vauterin (2005). Tsoft: graphical and interactive software for the analysis of time series and Earth tides, Computers & Geosciences, 31, 631-640. https://doi.org/10.1016/j.cageo.2004.11.015

Van Camp, M., O. de Viron, A. Watlet, B. Meurers, O. Francis, and C. Caudron (2017). Geophysics from terrestrial time-variable gravity measurements. *Review of Geophysics*, 55, 938-992. https://doi.org/10.1002/2017RG000566

van Dam, T. M. and O. Francis (1998), Two years of continuous measurements of tidal and nontidal variations in Boulder, Colorado, *Geophysical Research Letters*, 25, 393-396.

Woodworth, J. (2011), Estimation of unconfined aquifer hydraulic properties using gravity and drawdown data, M.Sci. thesis, Colorado State University, Fort Collins, Colorado.