Journal of Hydrology 597 (2021) 126285

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Use of microgravity for identification of delayed gravity drainage and conceptual model selection in unconfined aquifers

Andrés González-Quirós [a,b,*], José Paulino Fernández-Álvarez [b,c]

[a] School of Geosciences, University of Aberdeen, Aberdeen, UK

[b] Hydro-Geophysics and NDT Modelling Unit. Polytechnic School of Mieres, University of Oviedo, Mieres, Spain

[c] Department of Mining Exploitation and Prospection, School of Mines, University of Oviedo, Oviedo, Spain

ARTICLE INFO

This manuscript was handled by C. Corradini, Editor-in-Chief, with the assistance of William P. Gardner, Associate Editor

Keywords:
Unconfined aquifers
Unsaturated zone
Hydrogeophysics
Conceptual model
Microgravity

ABSTRACT

The estimation of hydraulic parameters at the aquifer scale typically relies in the analysis of pumping tests. Drawdown data, from a frequently limited number of boreholes and piezometers, is used in conjunction with analytical formulas derived for conceptually simplified models in a history matching procedure. In unconfined aquifers it is well known that the drainage process is controlled by the unsaturated zone. Several models have been proposed in the literature for the calculation of drawdown caused by pumping. Approaches range from the assumption of instantaneous and complete drainage to the inclusion of a delay term or the full implementation of the unsaturated zone to represent the delayed drainage in the unsaturated zone. Because borehole drawdown data is not informative about the unsaturated zone processes, it is common for practitioners to rely in the simplest models available, usually in some pumping test interpretation software. For example, Neuman's instantaneous drainage model is still widely used in the industry even when it has been demonstrated that under some circumstances provides incorrect estimations of key aquifer parameters. Additionally, as boreholes are scarce, spatial information is ignored and conceptual assumptions of axial-symmetry are a normal practice. In this work we show how the use of microgravity instruments, sensitive to storage variations in the near sub-surface, could be a simple, cheap and convenient tool for the identification of delayed drainage processes in unconfined aquifers. The joint use of drawdown and gravity data can be utilized to select the most preferable conceptual model for the parameter estimation problem.

1. Introduction

This paper focuses on the problem of choosing a conceptual model among those compatible with a given set of hydraulic observation data from pumping tests in unconfined aquifers.

Estimations of hydraulic parameters at the aquifer scale come usually from pumping tests by comparison of stationary/transient drawdown measurements in observation wells and piezometers with theoretical predictions from the solution of the forward problem (Theis, 1935; Cooper and Jacob, 1946; Neuman, 1972; Moench, 1995). These predictions may rely on different conceptual models which mathematically codify certain hypothesis regarding regime of flow, types of active physical and chemical processes, size (boundary conditions) and symmetries of the domain geometry as well as assumptions about the spatial parameter distribution.

In setting up this required conceptual model to solve the forward problem, the appropriateness of each choice will be established using professional judgement about two types of information: preexisting (prior) knowledge of any kind about the hypothesis mentioned above and knowledge of the behavior of measured flows and observed piezometric data (posterior information) including the measurement errors (e.g., Neuman and Wierenga, 2003).

As to the prior information, hydrogeological studies are famous for the scarcity of the subsurface data in real practice applications. Enough borehole logs are mostly not available, past geological/geotechnical reports on the site, if ever existing, may well be inconsistent and provide only qualitative guidelines and broad bounds about such important issues as the vertical and horizontal distribution of facies, essential to posit any worthy prior hypothesis about the spatial distribution of the hydraulic parameters or the expected boundary conditions.

Regarding the posterior information, the number of wells or piezometers is –on economic grounds- as diminished as possible, even one

* Corresponding author at: School of Geosciences, University of Aberdeen, Aberdeen, UK.
E-mail address: andres.quiros@abdn.ac.uk (A. González-Quirós).

https://doi.org/10.1016/j.jhydrol.2021.126285
Received 23 December 2020; Received in revised form 29 March 2021; Accepted 2 April 2021
Available online 14 April 2021

or none sometimes, aside from the pumping well. The construction procedure of the pumping/observation wells and the measurement process itself introduce perturbations and uncertainties (degree of partial penetration, identification/characterization of the skin effect, dynamic behavior of the pump, etc.), seldom characterized or even considered.

Therefore, even taking into consideration both the prior and the posterior knowledge together, there is room to wield freedom of choice in selecting different sets of hypotheses, mathematically translated into different conceptual models, deemed to underlie the process. An incorrect identification of the conceptual model for the forward problem would manifest itself in disguise as a bias in the values of the hydraulic parameters estimates obtained through an inversion procedure (e.g., Carrera et al., 2005).

Unconfined aquifers are subject to stress in terms of water extractions and contamination processes, frequently at or near the terrain surface. Their hydraulic characterization is relevant, therefore. For them, the modelling and interpretation of pumping tests is mathematically more challenging than for confined aquifers because the water table is both boundary condition and part of the solution (Mishra and Kuhlman, 2013). To circumvent this nonlinearity numerous assumptions have led to conceptual models that cannot fully represent the processes that occur in the aquifer. Pioneering works on the matter were those of Boulton (1954), Boulton (1963), Prickett (1965) and Dagan (1967). Neuman (1972), Neuman (1974) developed a model that imply instantaneous and complete drainage and has been for decades a standard in the industry. When processes in the vadose zone are deemed to be absent or negligible, it is the preferred option for its simplicity.

Some very well documented examples, though, as those from Cape Cod, Massachusetts, US (Moench et al., 2001) and Borden, Ontario, Canada (Bevan, 2002; Bevan et al., 2005) -dealt with later in this Section 3 of this paper- have shown how the estimates of some parameters may be biased if the unsaturated zone does indeed play a role, should it remain undetected or underestimated. For instance, the values of specific yield estimated via type-curve analysis with the Neuman model have been observed to yield too low values in some studies when compared with laboratory tests, volume balance methods and field experiments (e.g. Nwankwor et al., 1984, 1992; Endres et al., 2007; Moench et al., 2001; Moench, 2004, 2008). Moench (1995), Moench (1997) proposed a model that would provide the needed gradual delay in the drainage process. His model merges Boulton and Neuman models, introducing Boulton's convolution integral, with the delay index α, into the upper boundary of Neuman (1972), Neuman, (1974) model. Although an improvement in performance, Moench's model parameters do not have easy interpretation in physical terms and their relation to vadose zone parameters is unknown.

That neither instantaneous nor delayed drainage models do account for the water from the unsaturated zone has been corroborated in the last decades with the analysis of data from largely monitored such as the aforementioned pumping tests experiments of Cape Cod and Borden. This evidence suggests the importance of the vadose zone (Moench, 2008) and the inability of those approaches so far to truly capture the aquifer response. More advanced analytical solutions with incorporation of the unsaturated zone have been proposed in the last decade (Mathias and Butler, 2006; Mishra and Neuman, 2010, 2011), but their use and application in the groundwater industry is still very limited as their mathematical treatment is more complex. Additional mechanisms (e.g. hysteresis), flow behavior (e.g. horizontal gradients on the capillary fringe) and specific site characteristics (e.g. heterogeneities) (Bunn et al., 2010) have been called upon to explain discrepancies between the analytical and numerical models and field data.

However, the conditions met in such highly densely surveyed pilot sites, regarding completeness and precision of the data sets and level of professional scrutiny, are never met in real practice. Instead, the lack of information promotes simplicity as the criterion of choice for real everyday practice.

For unconfined aquifers two extreme types of conceptual models could be considered. Those explicitly fully including the hydrodynamics of the non-saturated zone -will be referred to as Hydraulic Vadose Zone (HVZ) models-; and those assuming instantaneous (Neuman) or delayed (Boulton-Moench) drainage of the porous media with the decline in the phreatic surface but avoiding explicit treatment of the unsaturated zone. As it has been (and will be) observed, both may well be compatible with a given set of piezometric observations and prior information available from a pumping test.

However, the implementation of the HVZ concept would require the extra cost of dealing both with added non linearities and with an increased complexity, by virtue of the additional hydraulic parameters (which ought to be characterized) describing the non-saturated behavior. If included without really existing, the results are loaded with unnecessary complexity and with incorrect estimates. If not included when really present, the assumed dynamics will be wrong and so will also be the estimates. Whether to implement or not the unsaturated zone in the modelling framework is a key decision for which any additional source of information, other than the hydraulic observations, is of great value. We see in this article how microgravity geophysics can help.

1.1. The role of Geophysics: Hydrogravity

Geophysical techniques can provide relevant information about groundwater processes and flow-constraining geological structures at multiple scales (Hubbard and Linde, 2011; Binley et al., 2015). One of the classical geophysical techniques is gravity, which provides information about spatial and temporal density (mass) changes in the subsurface (Crossley et al., 2013). In recent years, gravity has been used for hydrological applications to measure water storage changes at different spatial and temporal scales (Pool and Eychaner, 1995; Pool and Schmidt, 1997; Gehman et al., 2009; Naujoks et al., 2010; Creutzfeldt et al., 2010; Güntner et al., 2017) or water recharge processes (Chapman et al., 2008; Kennedy et al., 2014; Kennedy and Ferré, 2016) using absolute, relative or the very sensitive superconducting gravimeters with submicrogal precision, or at the global scale as exemplified by successful experiences such as the GRACE satellites (e.g., Cazenave and Chen, 2010).

Improved microgravity instruments with field precisions around or below the microGal ($1\ \mu\mathrm{Gal} = 10^{-8}\ \mathrm{m \cdot s^{-2}}$) level allow the mass changes linked to water injection/extraction processes to be tracked not only at global or regional scale, but also at the much smaller pumping scale. At the aquifer scale some works have assessed the utility of gravity to provide support in the groundwater model conceptualization by informing on aquifer properties such as heterogeneities (Poeter, 1990; Jacob et al., 2010) and anisotropy (Fernández-Álvarez et al., 2016). In a more quantitative fashion, gravity has been posited to constrain aquifer parameter estimation in unconfined aquifers when combined with hydraulic data (Damiata and Lee, 2006; Blainey et al., 2007; Herckenrath et al., 2012). A limitation for application in some of these works was the use of analytical models for the computation of the groundwater and/or the related gravity anomaly. To overcome this difficulties, numerical codes can be used (e.g. Leirião et al., 2009) to link the flow problem with the gravity problem, but the need to employ separate codes may rendered them cumbersome.

González-Quirós and Fernández-Álvarez (2014), González-Quirós and Fernández-Álvarez (2017) used the software COMSOL Multiphysics to build a coupled, single-code, flexible solution of the gravitational anomaly caused by pumping tests in unconfined aquifers under the assumption of complete and instantaneous drainage (known hereinafter as HGID), valid both for stationary and transient (Neuman) models. Though flexible as code, for those cases where the Vadose Zone processes are important the tool was still not available.

1.2. Motivation and objectives

The main aim of this paper is to show whether microgravity can be used, over unconfined aquifers, to discriminate between a conceptual model which explicitly includes the vadose zone (HVZ) and a conceptual model which assumes instantaneous drainage (HID), in those cases where the hydraulic observations are inconclusive.

First, as the right coupled numerical tool was yet not available, we have extended the capabilities of the coupled hydrogravity modelling code -known here as HGID and explained in González-Quirós and Fernández-Álvarez (2017)- to include the effects of the unsaturated or vadose zone both in the flow and gravity problems (hereafter known as HGVZ).

To test the model, we applied the approach (Section 3) to the widely known Borden test (Bevan, 2002; Bevan et al., 2005) with two goals: first, to validate the behavior of the flow code with data from a real test and, second, to simulate, coupled, the associated gravitational anomaly expected if gravity were to be used during the field experiment. In doing this, it became clear that the actual pumping rate in the Borden test was unfortunately very low compared to normal situations.

To evaluate the performance of gravity, we propose a synthetic case (Section 4) in which drainage in the unsaturated zone cannot be unambiguously detected from hydraulic data observations alone.

The procedure will be as follows:

1- First, the code will be used to generate a piezometric head field, sampled at the positions of some observation wells/piezometers and perturbed with noise. This synthetic field plays the equivalent role to a set of piezometric observations from a pumping test, performed in a real unconfined aquifer where the vadose zone is both present and relevant. Therefore, compliant with the conceptual model HVZ.
2- Second, the code will be run again but this time using the conceptual model HID (Neuman model) and it will be proved that, even without vadose zone, the same observed piezometric heads can be honoured. In doing so, a set of hydraulic parameters will be obtained which provides the best fit, under this ID model, to the data generated in step 1. Now, therefore, a synthetic set of piezometric measurements consistent with two conceptual models is available. No distinction between the two conceptual models is possible, by construction, from this measurement set alone.
3- Third, the gravity module -which we call HGVZ – will be used with the HVZ model to calculate the evolution over time of the gravity anomaly at different offsets from the well. Once adequately perturbed with noise, these calculations will mimic those that a relative gravity meter would have measured over time if stationed at those points since the beginning of the water extraction: the gravity anomaly above a terrain where the vadose zone do matters. It is important to bear in mind that these curves can be obtained in the field, although here they are simulated. They will be referred to as field gravity curves.
4- Fourth, the gravity module -this time HGID, from instantaneous drainage- will be used with the HID model using the hydraulic parameters obtained in step 2, to grasp under what conditions the difference with the field gravity anomaly (step 3) becomes significant, that is, not explained by the random error level achievable in the field.
5- Finally, in Section 4.4 we show that not only for this particular set of parameters but also for no combination of hydraulic parameters, the gravity anomaly calculated over the HID conceptual model can approach the field gravity curve. This means that the field gravity curve remains ostensibly distant, from a certain time after the start of the pumping onwards, from any possible simulation outcome that could be performed using the Neuman conceptual model which, therefore, would be discarded on the grounds of gravity measurements.

This strategy provides a way to discard the instantaneous drainage concept when the vadose zone is really present but remains undetected by hydraulic data alone.

2. Materials and methods

2.1. Coupled Hydro-Gravity model with vadose zone effects (HGVZ).

The first tool built in this paper is a numerical code able to simulate pumping tests including the full behavior of vadose zone. The possibilities of the code are very flexible. For the specific application of pumping from a single well, as it is the purpose in this article, we restrict the explanation to models in COMSOL Multiphysics® with 2D axisymmetric domains and a partially penetrating well placed on the axis of symmetry. 3D model with multiple pumping or injection wells and inclusion of aquifer characteristics such as heterogeneity or anisotropy can be easily included in the modelling flowchart.

For simulation of the hydrogravity model we followed a sequential approach for implementation. First, the flow model was defined and implemented with inclusion of vadose zone characteristics (code HVZ). Later, a coupled gravity extension was added to obtain the associated gravity anomaly (code HGVZ). Both models were run simultaneously and fully coupled, and time-dependent solutions were obtained at the points of interest. A more detailed description is presented below.

2.1.1. Flow model (HVZ)

The code solves Richards' equation for unsaturated zone flow

$$K_r \frac{1}{r}\frac{\partial}{\partial r}\left(K(\psi) r \frac{\partial \sigma}{\partial x}\right) + K_z \frac{\partial}{\partial z}\left(K(\psi)\frac{\partial \sigma}{\partial z}\right) = C(\psi)\frac{\partial \sigma}{\partial t} \tag{1}$$

where ψ is the pressure head, σ is the drawdown in the vadoze zone and $C(\psi)$ is the soil moisture capacity defined as

$$C(\psi) = \frac{\partial \theta}{\partial \psi} \tag{2}$$

$\theta[-]$ is the volumetric water content.

The model supports direct implementation of the Brooks and Corey and van Genuchten parametric models, but extension to other model in the literature is possible within the flexible options provided by COMSOL. In this work we restricted our simulations to the Brooks and Corey (1964) model with which the volumetric moisture content, $\theta[-]$ is computed as

$$\begin{aligned} &\theta = \theta_r + (\theta_s - \theta_r)\left(\frac{h_b}{h_c}\right)^n && h_c < h_b \\ &\theta = \theta_s && h_c \geq h_b \end{aligned} \tag{3}$$

$\theta_r[-]$ and $\theta_S[-]$ are the relative and saturated volumetric moisture content, h_c [m] is the pressure head, h_c [m] the air entry (or bubbling) pressure and n[-] the pore size distribution index.

The relative hydraulic conductivity is computed as

$$\begin{aligned} &K_{rel} = \left(\frac{h_b}{h_c}\right)^{2+3n} && h_c < h_b \\ &K_{rel} = 1 && h_c \geq h_b \end{aligned} \tag{4}$$

The modelling included partial penetration and borehole storage effects.

2.1.2. Hydrogravity code (HGVZ)

The flow code HVZ was coupled with a gravity module to compute the vertical acceleration at a point of mass m caused by the gravitational field measured on the surface, given by the expression:

$$\boldsymbol{g}_z = -G\frac{\Delta m}{r^3} cos\alpha \boldsymbol{r} \tag{5}$$

$G = 6.673 \times 10^{-11}\ \mathrm{m^3\ kg^{-1}\ s^{-2}}$ is the gravitational constant; $\boldsymbol{r}$ is the vector from the measurement instrument to the target mass and α is the angle between the vertical line and $\boldsymbol{r}$. The value of $\mathbf{g}$ is expressed in this work in microGal (1 μGal = 10^{-6} Gal = 10^{-8} m s^{-2}).

There is hydrogravity literature to compute the gravity anomaly linked to different drainage processes to the changing water table -assumed uniform over a broad extension- at regional and global scales (e.g. Pool, 2008; Gehman et al., 2009). Or, more interesting here, related to pumping wells under different hypothesis (e.g. Damiata and Lee, 2006; Blainey et al., 2007; Herckenrath et al., 2012; González-Quirós and Fernández-Álvarez, 2014, 2017). They determined the geometry of a domain, completely drained, between the initial and final water table. The mass change per unit volume, proportional to the specific yield S_y (6), is now to be substituted in expression (5), which, when integrated over the whole drained domain gives the gravity acceleration produced at the point of interest on the surface, as desired.

$$\Delta m_u(x,y,z,t) = S_y(x,y,z,t)\rho_w \tag{6}$$

When the drained domain includes the unsaturated zone, the change in water content per unit volume of aquiferθ [−] is not the S_y anymore, but $\Delta\theta$ defined as in (7), which varies in space and time and has to be obtained as part of the solution from the flow model.

$$\Delta\theta(x,y,z,t) = \theta(x,y,z,t) - \theta(x,y,z,t_0) \tag{7}$$

Expression (6) giving the target mass change linked to drainage now turns into (8), variable in time throughout the domain.

$$\Delta m_u(x,y,z,t) = \Delta\theta(x,y,z,t)\rho_w \tag{8}$$

Equation (8) is implemented in COMSOL Multiphysics to build the gravity model extension, here called (HGVZ), capable of fully considering the vadose zone dynamics substituting Eq. (6) in González-Quirós and Fernández-Álvarez (2017). It must be note that the main difference with the previous approach relies in that (8) applies to the full aquifer, while in the previous model mass variation was restricted to a "drained" volume, defined between the initial and final water table and known as drawdown or depression cone. The additional strategy allows flexibility for both the choice of the conceptual hydrogeological base model and the gravity associated computation.

Detailed descriptions of the modules and boundary conditions required to build the gravity module in COMSOL Multiphysics are explained in Butler and Sinha (2012) and in our previous references for combination with the flow module.

3. The Borden Case: An example where two conceptual models are acceptable

We aim here to validate the predictions of the flow part (HVZ) of our code using the observed data from a well-known groundwater pilot site deeply analysed in the literature: the Borden pumping test, performed in the Canadian Airforce base in Ontario, Canada (Bevan et al., 2005). The Borden case is also one where different conceptual models have been explicitly tested against measured hydraulic data.

3.1. Test description and flow model

The Borden site has hosted different field experiments well-known in the groundwater community (MacFarlane et al. (1983), Nwankwor et al. (1984), Nwankwor et al. (1992), Endres et al. (2000), Bevan et al. (2003), Bevan et al. (2005), Endres et al. (2007), Bunn et al. (2010), Sudicky and Illman (2011)). In particular, the test simulated in this paper was conducted in August 2001 (Bevan, 2002; Bevan et al., 2005). During the test, water was extracted from the unconfined aquifer for 7 days using a constant pumping rate of 40 ± 1 l min^{-1}. The aquifer of 9 m thickness was mainly comprised of horizontal lenses of medium and fine-grain sand and silty fine-grain sand with infrequent silt, silty-clay, and coarse sand layers (MacFarlane et al., 1983; Sudicky, 1986). The Initial water table was located at 2.75 m below ground surface (see Fig. 1). The pumping well had 0.13 m inner diameter and was screened 3.65 m over the bottom of the aquifer. Data was collected in 25 piezometers installed in different locations around the pumping well and six neutron probes (Fig. 1) (Bevan et al., 2005; Moench, 2008).

The extensive data sets have produced, under various authors and with *different conceptual models*, a range of estimated values of hydraulic properties.

For validation purposes, and to evaluate the capabilities of the numerical tool, we chose the parameters estimated by previous authors using two conceptual models in the literature: the instantaneous drainage (ID) assumption using Neuman (1974) analytical model as estimated by Endres et al. (2007), and the vadose zone (VZ) model using Richards' equation with the parameters estimated by Moench (2008) using numerical modelling (Table 1).

Fig. 2 shows the fit to the measured hydraulic heads in 2 piezometers at approximately the same distance from the well but at different height in the aquifer, P15 and WD1B. The mismatch in the first 2-minute stage of the drawdown curve is caused by wellbore storage and the skin effect as observed by Moench (2008). The mismatch at the end of the test has been explained as the levels reaching a steady-state regime through contact with a lateral constant level water body (see Endres et al., 2007).

For the deep piezometers, such as P15, the fit is good for the second and third stage of the drawdown curve and both conceptual models are indistinguishable. For shallow piezometers, represented here by WD1B, the conceptual model including the vadose zone is better but just a few centimetres different from the levels predicted by the alternative conceptual model with instantaneous drainage. This difference is undetectable in most real-world pumping rests.

These results are consistent with analyses from previous works (Endres et al., 2007) and therefore we consider validated the HVZ module. It will be used in the following to calculate the forward flow problem both with vadose zone effects and under the instantaneous drainage hypothesis.

3.2. Coupled hydrogravity model

To assess whether a gravity campaign would be able to see any difference between the two conceptual flow scenarios, we simulated a time-lapse microgravity survey performed on the ground surface.

The gravity extension (HGVZ) was used to calculate the gravity anomalies associated to both conceptual models (instantaneous drainage and full vadose zone) and test whether they are both measurable by state-of-the-art gravity instrumentation and distinguishable on the basis of those gravity measurements alone. The parameters used are those in Table 1 with the actual pumping rate. We postulate three gravity stations on the surface (not existent in the real Borden case) where relative gravity values are taken every half day during the whole duration of the pumping test (seven days). The field gravity curves simulated for the case including vadose zone can be seen in Fig. 3. The computed gravity anomaly happens to be, even for these very low pumping rates (lower than 1 l·s^{-1}, 40 ± 1 l min^{-1}) above the limits of detectability of modern gravity meters (see the 5 μGal line).

Finally, we simulated the gravity anomaly for the Borden test at the same three gravity stations using the instantaneous drainage (HGID) conceptual model (circle markers every 0.5 days in Fig. 3). The gravity differences for the two alternative conceptual models are represented in Fig. 3. A systematic pattern of increasing difference in time -associated with higher water storage changes in the unsaturated zone in the VZ model (González-Quirós and Fernández-Álvarez, 2020)- tends to develop but nonetheless remains under 2 μGal. Undetectable therefore for practical purposes unless sophisticated high precision gravity meters are carefully employed (e.g., Kennedy et al. 2016).

The Borden pumping rates were too low for a practical scenario. However, the results suggest that higher pumping rates, 10^1 l s^{-1} or above, should then lead to easily measurable gravity anomalies. But, for

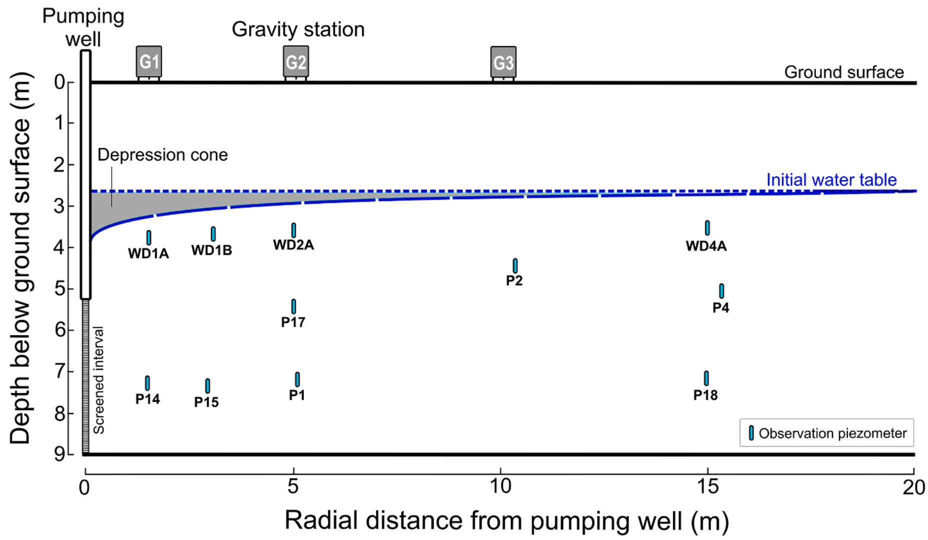

Fig. 1. Characteristics of the Borden pumping test and location of some of the observation piezometers (modified from Bevan et al., 2005) and locations of the simulated gravity stations at the surface.

Table 1
Borden aquifer parameters estimated by different authors and used in the modelling in this work.

Parameter name	Symbol	Unit	ID[a]	VZ[b]
Horizontal hydraulic conductivity	K_r	m s^{-1}	6.27E−5	6.6E−5
Vertical hydraulic conductivity	K_z	m s^{-1}	2.79E−5	3.11E−5
Specific storage	S_s	m^{-1}	5.78E−5	5.67E−5
Specific yield	S_y	–	0.201	0.301
Residual volumetric moisture content	θr	–		0.03
Saturated volumetric moisture content	θs	–		0.331
Air entry (or bubbling) pressure head	h_b	m		−0.330
Pore size distribution index	n	–		5.85

[a] Endres et al. (2007) using Neuman (1974) instantaneous drainage (ID) analytical model. [b]Moench (2008) using numerical variable saturated model performed with VS2DT and using Brooks and Corey (1964) equations.

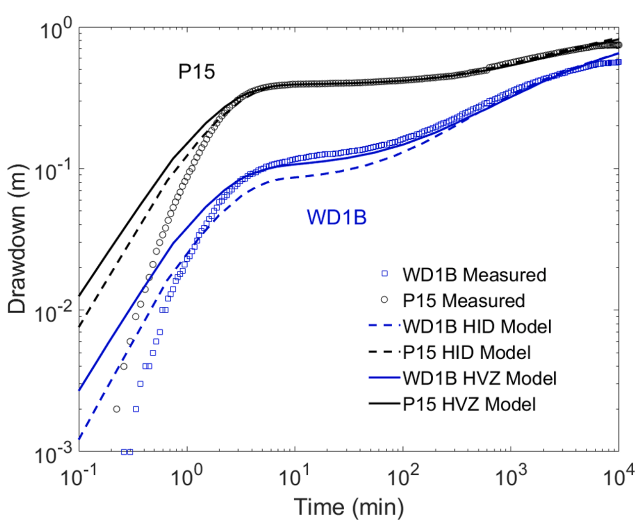

Fig. 2. Comparison of measured data from two of the piezometers (P15 and WD1B) with results of models for Instantaneous Drainage (HID) and vadose zone (HVZ). Early-stage differences (<1 min) are due to delayed piezometer response.

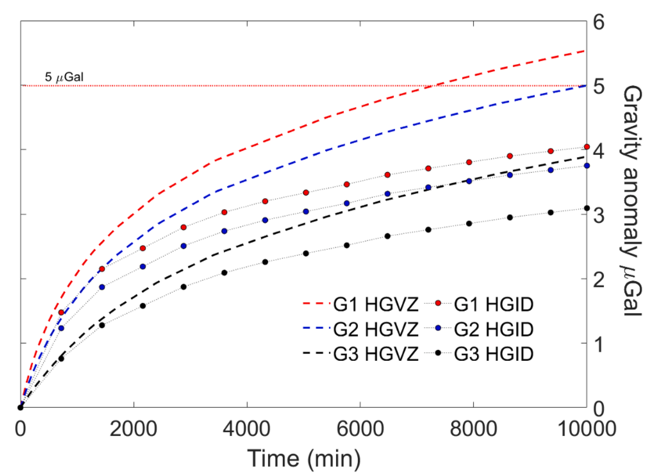

Fig. 3. Simulated gravity anomaly (absolute value) at the three stations for the HGID and HVVZ models.

these materials and geometry the aquifer gets dry. In the next section, we pose a hypothetical -although very plausible in practice- case, able to bear more normal pumping rates, which effectively shows the contribution of gravity to distinguish between the cases with and without vadose zone.

4. Hydrogeophysical conceptual model discrimination

4.1. Case description

The hypothesized model has two layers (Fig. 4). The upper layer is an unconfined aquifer, composed of medium-grained sand of glacio-deltaic or glacio-fluvial origin similar to the one of the Borden test that was previously defined in Section 3. The lower second layer, with higher permeability, is a silty to clean sand and with properties obtained from Damiata and Lee (2006). The parameter values are compiled in Table 2.

The model is conceived to mimic the behavior of a real case where the shallower formation, less permeable but still of a sandy nature, as also is the lower formation, provides a delayed drainage once the heads in the lower aquifer decline due to pumping. The pumping rate now is higher than that in the Borden case, 0.06309 $m^3\ s^{-1}$ (63.09 l s^{-1}), common in real applications, susceptible of having a gravity effect at the

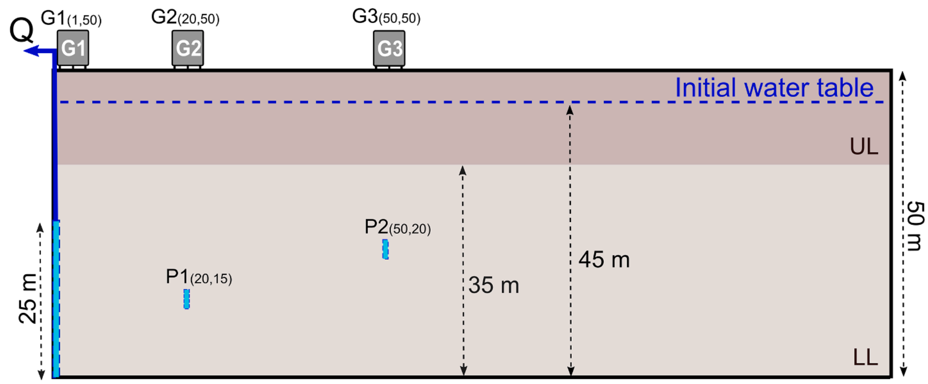

Fig. 4. Characteristics of the two-layer model utilized in the analysis. See Table 2 for details.

Table 2
Model Parameters assigned to the two-layer reference model. See also Fig. 4.

Parameter name	Symbol	Value	Unit
Upper Layer (UL)			
Thickness	b_u	10	m
Horizontal Hydraulic Conductivity	K_r	5.89×10^{-5}	m s^{-1}
Vertical Hydraulic Conductivity	K_v	3.14×10^{-5}	m s^{-1}
Specific Storage	S_s	3.32×10^{-5}	m^{-1}
Pore size index	n	2.5	–
Residual volumetric moisture content	θ_r	0.07	–
Saturated volumetric moisture content	θ_s	0.37	–
Specific Yield	S_y	0.3	–
Lower Layer (LL)			
Thickness	b_l	35	m
Hydraulic Conductivity	K	1×10^{-4}	m s^{-1}
Specific Storage	S_s	1×10^{-5}	m^{-1}

surface and used in previous works with application to the hydrogravity problem (Blainey et al., 2007; Damiata and Lee, 2006; Herckenrath et al., 2012; Leirião et al., 2009; González-Quirós and Fernández-Álvarez, 2014, 2017; Maina and Guadagnini, 2018).

The pumping well is partially penetrating, with 0.1 m radius and a 25 m screen length at the aquifer bottom aiming to extract water from the lower formation. The initial water table is 5 m below ground surface. Two observation piezometers are assumed. The three gravity stations are located 1, 20 and 50 m off the pumping well.

4.2. Forward flow simulation

For this model, the drawdowns measured at the two piezometers at the end of the total duration of the pumping test (7 days) are calculated (with HVZ, thus considering the full vadose zone) and perturbed using an uncorrelated Gaussian noise of zero mean and standard deviation of

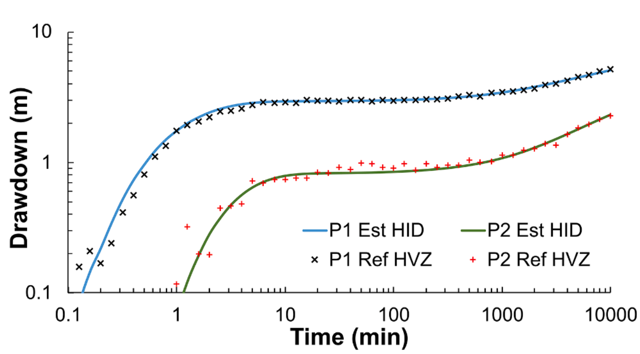

Fig. 5. Results of calibration with ID conceptual model for the two piezometers.

0.03 m to mimic sensible hydraulic heads as would have been measured in the field (see markers in Fig. 5).

4.3. HID parameter estimation

The WTAQ software (Barlow and Moench, 2011) is employed in combination with PEST (Doherty, 2020), to inverse-estimate the hydraulic parameters of such single "homogeneous" aquifer using the HID model. The results are compiled in Table 3. The fitting between the "*observed field data*", simulated with vadose (HVZ) zone assumptions and the parameter values compiled in Table 2 (discrete markers), and the predictions for the single aquifer HID Neuman model estimations (continuous lines) are shown in Fig. 5.

With those parameters, the field data can then be also nicely fitted. Therefore, with these observed data, a conceptual model which includes the vadose zone may not be differentiated from a conceptual model which assumes instantaneous drainage. Both are compatible with the data. The only hint in the estimated values is the abnormally low value for the Sy (0.20) instead of the real value of 0.30 (Table 2).

4.4. Conceptual model discrimination using gravity data

Using both sets of parameters that lead to the same drawdown data response with both conceptual models, we simulated the gravity anomaly for the instantaneous drainage assumption (HGID code) and the gravity (as explained in Section 2.1 -HGVZ code-) for the model that fully incorporates unsaturated zone effects. This has been done in three gravity stations located at the surface at 1, 20 and 50 m from the pumping well (Fig. 4).

The two conceptual models, represented each by its own hydraulic parameter set, are able to fit the piezometric head field curves. Now the question is whether the gravity response measured at the surface is different for the two conceptual models and if this difference can be measured. Both for this single layer instantaneous drainage estimates, and for the parameters of the "real" model case with vadose zone, gravity predictions at the stations are produced.

Table 3
Results of the parameter estimation using the Neuman conceptual model.

Single Layer Parameter Estimation Results with HID model.					
Parameter	Symbol	Estimated value	Unit	95% confidence	
				Upper	Lower
Horizontal Hyd. Conductivity	K_r	1.1×10^{-4}	m s^{-1}	1.064×10^{-4}	1.137×10^{-4}
Vertical Hyd. Conductivity	K_v	5.4×10^{-5}	m s^{-1}	5.176×10^{-5}	5.633×10^{-5}
Specific Storage	S_s	1×10^{-5}	m^{-1}	9.299×10^{-6}	1.075×10^{-5}
Specific Yield	S_y	0.1988	–	0.182	0.216

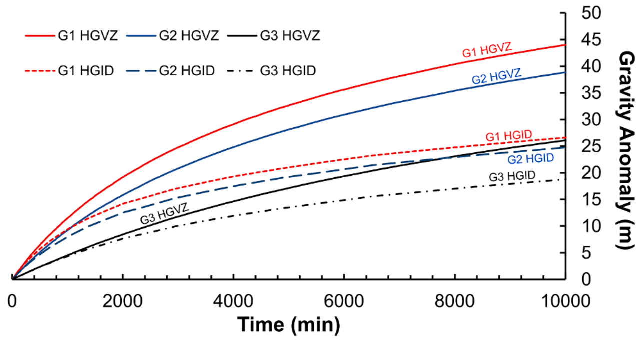

Fig. 6. Simulated gravity anomaly (absolute value) at the three stations for the two-layer (true) model with vadose zone (HGVZ) and the one-layer model with assumption of instantaneous drainage (HGID).

Fig. 6 shows the answer. The gravity anomaly as measured in stations G1, G2 and G3 for the model with vadose zone (HGVZ) is plotted in continuous line. The same is plotted with dashed lines, for the same stations, for the instantaneous drainage conceptual model (HGID). The difference between the continuous and dashed line, for each station, has to be looked upon for interpretation. In all the three stations this difference increases with time and becomes measurable from a certain time onwards. For the G1 station (in red), closer to the pumping well, the difference at the end of the seven-day period of pumping rises to around 17.2 μGal. For Station G2 (in blue) to 13.9 μGal and for station G3 (in black), the farthest, to 7.1 μGal.

The possibility to measure the difference then, depends both on the time after the pumping starts and on the offset of the gravity station. In Fig. 7 this difference is plotted against both variables. The lines of 5 μGal, 10 μGal and 15 μGal are drawn for reference. Taking 5 μGal as an achievable precision for a careful field work (Scintrex Limited, 2009), the difference could be seen in G1 after 1.5 days, slightly later (2 days) for G2 and after 4 days in the more distant G3. That is, with the pumping rate of this example, the measured gravity change would be enough to discriminate between the two hypotheses after 2 days at a distance closer to 20 m from the pumping well. On the contrary, with distance from the pumping well both models tend to be hardly distinguishable, with differences lower than 5 μGal farther from 65 m from the pumping well, even after 7 days.

Up to now it has been shown that the gravity predictions obtained with *that* particular set of hydraulic parameters shown in Table 3, which provides a good fit of the measured heads under the Neuman conceptual model, are measurably different from the gravity predictions of the real field model presenting vadose zone behavior. This result is to be taken as a proof of concept. But It remains now to make sure that this is true also for *no other* set of hydraulic parameters. That is, that there is no (sensible) set of hydraulic parameters among those fitting the measured piezometric heads under both conceptual models, which is also -simultaneously-able to do so for the gravity measurements.

To explore this issue, the list of relevant parameters included in the non-dimensional groups characterizing the Neuman response has been chosen. Two objective functions to quantify the goodness of fit have been defined and plotted in the same graph: one for the hydraulic head misfit and the other for the gravity misfit, bearing the same structure. Formally:

$$\phi = \sum_{i=1}^{n} \left[d_{VZ} - F_{ID}(m)\right]^2 \tag{9}$$

d_{VZ} is the observed data (here the datasets simulated for the reference model with unsaturated zone using the parameters compiled in Table 2) both for head and gravity; and $F_{ID}(m)$ is the simulated response (both for head and gravity) using the ID model assumption. n is the number of datapoints in each observation dataset.

We performed parametric sweeps to compute the objective function for the drawdown and gravity response varying within admissible ranges for combination of the four parameters estimable with Neuman model, that is K_r, K_z, S_s *and* S_y.

The results are summarized in Fig. 8. Weighted objective functions, $\overline{\phi} = \sqrt{\phi/n}$, for each set of observations (gravity in blue and hydraulic

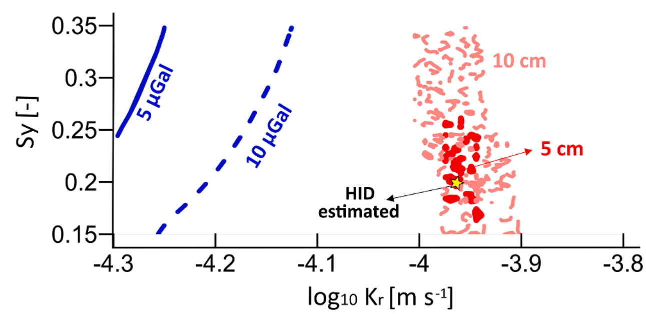

Fig. 8. Plot of the objective admissible contours for gravity (blue) and hydraulic data (red). With a star we show the model that best fits the hydraulic data using the ID model (HID estimated) with parameters in Table 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

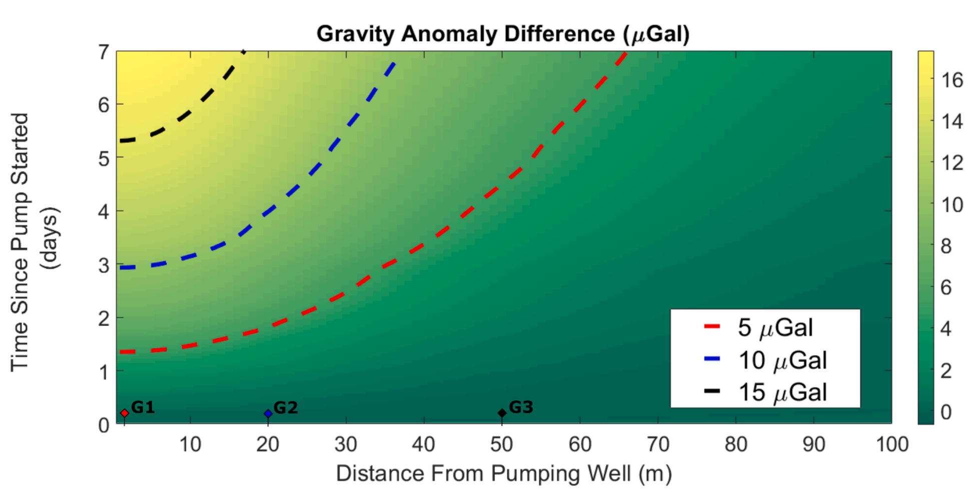

Fig. 7. Difference, in μGal, between simulated HGVZ and HGID responses with time (y-axis) and distance from the pumping well (x-axis). Location of the 3 gravity stations is indicated in the bottom axis.

head in red) are plotted as representative admissible contours against the two key parameters: specific yield and hydraulic conductivity. The model that best fits the hydraulic data (HID) from Table 3 is also superimposed with a star symbol.

It seems clear, then, that every sensible set of parameters, giving a good hydraulic fit under both conceptual models will give different gravity responses for this specific conditions. Different enough to be measured and, therefore, different enough to be discriminated upon that basis. Nonetheless, it must be noted that the differences in gravity signal are specific for the aquifer characteristics, pumping rate and depth to water table of this specific example. For analysis and comparisons of models with different pumping rates and hydraulic parameters the results should be analyzed in a case-by-case basis, each leading to different hydraulic and gravity curves with relative difference between two plausible models that could be lower or higher than the shown in this work.

5. Discussion and conclusions

In this article we show how simple microgravity observations, that can be done while pumping at a certain distance of the well, may allow to discard one from two hypothesis or hydraulic conceptual models when the set of piezometric measurements is inconclusive for that matter. The conceptual models considered regard the inclusion or not of the dynamic behavior of the vadose zone. The first conceptual model includes the full behavior of the vadose zone and the second conceptual model neglects it completely, assuming complete and instantaneous drainage while phreatic levels are declining. This second conceptual model, the Neuman hypotheses, is well known and employed in industry.

The fact that gravity measurements proves useful to help choosing the right conceptual model avoids or at least softens the problem of biased estimated properties values obtained under wrongly chosen conceptual models.

To show this, we have built new, more flexible codes able to simulate joint flow and gravity problem, with extended the capabilities compared to existing previous codes. The possibility to fully account for the unsaturated zone for flow and gravity purposes allows the joint use of hydraulic and gravity data/information in the same inversion problem. The code has been validated -for its flow part- using the data from the Borden aquifer which, unfortunately, has proven not to be a good candidate for the use of gravity due to the very small pumping rates admissible.

CRediT authorship contribution statement

Andrés González-Quirós: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Software, Visualization, Writing - original draft, Writing - review & editing. **José Paulino Fernández-Álvarez:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Supervision, Writing - review & editing.

Declaration of Competing Interest

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

Acknowledgments

Andrés González Quirós is the recipient of a Royal Society – Newton International Fellowship (NIF\R1\182210). We thank Professor Ty Ferre and an anonymous reviewer for their constructive comments which contributed to improve the manuscript.

References

Barlow, P.M., Moench, A.F., 2011. WTAQ Version 2: A Computer Program for Analysis of Aquifer Tests in Confined and Water-table Aquifers with Alternative Representations of Drainage from the Unsaturated Zone. US Department of the Interior, US Geological Survey.

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

Bevan, M.J., Endres, A.L., Rudolph, D.L., Parkin, G., 2003. The non-invasive characterization of pumping-induced dewatering using ground penetrating radar. J. Hydrol. 281 (1-2), 55–69.

Bevan, M.J., Endres, A.L., Rudolph, D.L., Parkin, G., 2005. A field scale study of pumping-induced drainage and recovery in an unconfined aquifer. J. Hydrol. 315 (1-4), 52–70.

Binley, A., Hubbard, S.S., Huisman, J.A., Revil, A., Robinson, D.A., Singha, K., Slater, L. D., 2015. The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales. Water Resour. Res. 51 (6), 3837–3866.

Blainey, J.B., Ferré, T.P.A., Cordova, J.T., 2007. Assessing the likely value of gravity and drawdown measurements to constrain estimates of hydraulic conductivity and specific yield during unconfined aquifer testing. Water Resour. Res. 43 (12), W12408.

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. Civ. Eng. 26 (3), 469–482.

Brooks, R.H., Corey, A.T., 1964. Hydraulic properties of porous media. Hydrology papers (Colorado State University) no. 3.

Bunn, M.I., Jones, J.P., Endres, A.L., Rudolph, D.L., 2010. Effects of hydraulic conductivity heterogeneity on vadose zone response to pumping in an unconfined aquifer. J. Hydrol. 387 (1-2), 90–104.

Butler, S.L., Sinha, G., 2012. Forward modeling of applied geophysics methods using Comsol and comparison with analytical and laboratory analog models. Comput. Geosci. 42, 168–176.

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., Slooten, L.J., 2005. Inverse problem in hydrogeology. Hydrogeol. J. 13 (1), 206–222.

Cazenave, A., Chen, J., 2010. Time-variable gravity from space and present-day mass redistribution in theEarth system. Earth Planet. Sci. Lett. 298 (3-4), 263–274.

Chapman, D.S., Sahm, E., Gettings, P., 2008. Monitoring aquifer recharge using repeated high-precision gravity measurements: A pilot study in South Weber. Utah. Geophys. 73 (6), WA83–WA93.

Cooper Jr, H.H., Jacob, C.E., 1946. A generalized graphical method for evaluating formation constants and summarizing well-field history. Eos, Trans. Am. Geophys. Union 27 (4), 526–534.

Creutzfeldt, B., Güntner, A., Thoss, H., Merz, B., Wziontek, H., 2010. Measuring the effect of local water storage changes on in situ gravity observations: Case study of the Geodetic Observatory Wettzell. Germany. Water Resour Res 46, W08531.

Crossley, D., Hinderer, J., Riccardi, U., 2013. The measurement of surface gravity. Rep. Prog. Phys. 76 (4), 046101. https://doi.org/10.1088/0034-4885/76/4/046101.

Dagan, G., 1967. A method for determining the permeability and effective porosity of unconfined anisotropic aquifers. Water Resour. Res. 3, 1059–1071.

Damiata, B.N., Lee, T.-C., 2006. Simulated gravitational response to hydraulic testing of unconfined aquifers. J. Hydrol. 318 (1-4), 348–359.

Doherty, J., 2020. PEST v17. Model-independent Parameter Estimation, User Manual, Watermark Numerical Computing.

Endres, A.L., Clement, W.P., Rudolph, D.L., 2000. Ground penetrating radar imaging of an aquifer during a pumping test. Ground Water 38 (4), 566–576.

Endres, A.L., Jones, J.P., Bertrand, E.A., 2007. Pumping-induced vadose zone drainage and storage in an unconfined aquifer: a comparison of analytical model predictions and field measurements. J. Hydrol. 335 (1-2), 207–218.

Fernández-Álvarez, J.-P., González-Quirós, A., Rubio-Melendi, D., 2016. Assessment of the value of microgravity to estimate the principal directions of the anisotropic transmissivity of aquifers from pumping tests: A study using a Hough transform based automatic algorithm. J. Appl. Geophys. 134, 172–182.

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

González-Quirós, A., Fernández-Álvarez, J.P., 2014. Simultaneous solving of three-dimensional gravity anomalies caused by pumping tests in unconfined aquifers. Math. Geosci. 46 (6), 649–664.

González-Quirós, A., Fernández-Álvarez, J.P., 2017. Forward coupled modeling and assessment of gravity anomalies caused by pumping tests in unconfined aquifers under unsteady-state conditions. Math. Geosci. 49 (5), 603–617.

González-Quirós, A., Fernández-Álvarez, J.P., 2020. In: Incorporation of Unsaturated Zone Effects in Coupled Hydrogeophysical Modelling of Gravity Anomalies Caused by Pumping Tests. European Association of Geoscientists & Engineers. https://doi.org/10.3997/2214-4609.202020134.

Güntner, A., Reich, M., Mikolaj, M., Creutzfeldt, B., Schroeder, S., Wziontek, H., 2017. Landscape-scale water balance monitoring with an iGrav superconducting gravimeter in a field enclosure. Hydrol. Earth Syst. Sci. 21 (6), 3167–3182.

Herckenrath, D., Auken, E., Christiansen, L., Behroozmand, A.A., Bauer-Gottwein, P., 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. 48, W01539. https://doi.org/10.1029/2011WR010411.

Hubbard, S.S., Linde, N., 2011. Hydrogeophysics. In: Wilderer, P. (Ed.), Treatise on Water Science, vol. 2. Academic Press, Oxford, pp. 401–434.
Jacob, T., Bayer, R., Chery, J., Le Moigne, N., 2010. Time-lapse microgravity surveys reveal water storage heterogeneity of a karst aquifer. J Geophysical Res 115, B06402. https://doi.org/10.1029/2009JB006616.
Kennedy, J., Ferré, T.P.A., Güntner, A., Abe, M., Creutzfeldt, B., 2014. Direct measurement of subsurface mass change using the variable baseline gravity gradient method. Geophys. Res. Lett. 41 (8), 2827–2834.
J.R. Kennedy T.P.A. Ferré Accounting for time-and space-varying changes in the gravity field to improve the network adjustment of relative-gravity data Geophysical Journal International 204 2 2016 2016 892 906.
Leirião, S., He, X., Christiansen, L., Andersen, O.B., Bauer-Gottwein, P., 2009. Calculation of the temporal gravity variation from spatially variable water storage change in soils and aquifers. J. Hydrol. 365 (3-4), 302–309.
Maina, F.Z., Guadagnini, A., 2018. Uncertainty quantification and global sensitivity analysis of subsurface flow parameters to gravimetric variations during pumping tests in unconfined aquifers. Water Resour. Res. 54 (1), 501–518.
Mathias, S.A., Butler, A.P., 2006. Linearized Richards' equation approach to pumping test analysis in compressible aquifers. Water Resour. Res. 42, W06408.
MacFarlane, D.S., Cherry, J.A., Gillham, R.W., Sudicky, E.A., 1983. Migration of contaminants in groundwater at a landfill: A case study: 1. Groundwater flow and plume delineation. J. Hydrol. 63 (1-2), 1–29.
Mishra, P.K., Neuman, S.P., 2010. Improved forward and inverse analyses of saturated-unsaturated flow toward a well in a compressible unconfined aquifer. Water Resour. Res. 46 (7), W07508.
Mishra, P.K., Neuman, S.P., 2011. Saturated-unsaturated flow to a well with storage in a compressible unconfined aquifer. Water Resour. Res. 47 (5), W05553.
Mishra, P.K., Kuhlman, K.L., 2013. In: Advances in Hydrogeology. Springer New York, New York, NY, pp. 185–202. https://doi.org/10.1007/978-1-4614-6479-2_9.
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 (6), 1397–1407.
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 analysis of an unconfined aquifer test considering unsaturated zone characteristics. Water Resour. Res. 44, W06409.
Moench, A.F., Garabedian, S.P. and LeBlanc, D.R. (2001). Estimation of hydraulic parameters from an unconfined aquifer test conducted in a glacial outwash deposit, Cape Cod, Massachusetts. U.S. geological Survey Profesional Paper 1629, 69 pp.
Naujoks, M., Kroner, C., Weise, W., Jahr, T., Krause, P., Eisner, S., 2010. Evaluating local hydrological modelling by temporal gravity observations and a gravimetric three-dimensional model. Geophysical. J. Int. 182, 233–249.
Neuman, S.P., 1972. Theory of flow in unconfined aquifers considering delayed gravity response. Water Resour. Res. 8 (4), 1031–1045.
Neuman, S.P., 1974. Effects of partial penetration on flow in unconfined aquifers considering delayed aquifer response. Water Resour. Res. 10, 303–312.
Neuman, S. P., & Wierenga, P. J. (2003). Comprehensive Strategy of Hydrogeologic Modeling and Uncertainty Analysis for Nuclear Facilities and Sites.
Nwankwor, G.I., Cherry, J.A., Gillham, R.W., 1984. A comparative study of specific yield determinations for a shallow sand aquifer. Groundwater 22 (6), 764–772.
Nwankwor, G.I., Gillham, R.W., Kamp, G., Akindunni, F.F., 1992. Unsaturated and saturated flow in response to pumping of an unconfined aquifer: Field evidence of delayed drainage. Groundwater 30 (5), 690–700.
Poeter, E.P., 1990. A new tool: delineation of textural heterogeneities in unconfined aquifers, using microgravity surveys during pumping. Bullet. Assoc. Eng. Geologists xxvii (3), 315–325.
Pool, D.R., 2008. The utility of gravity and water-level monitoring at alluvial aquifer wells in southern Arizona. Geophysics 73 (6), WA49–WA59.
Pool, D.R., Eychaner, J.H., 1995. Measurements of aquifer-storage change and specific yield using gravity surveys. Ground Water 33 (3), 425–432.
Pool, D.R., Schmidt, W., 1997. Measurement of groundwater storage change and specific yield using temporal gravity method near Rillito Creek, Tucson, Arizona US Geological Survey. Water Resour. Investig. Rep. 97–4125.
Prickett, T.A., 1965. Type-curve solution to aquifer tests under water-table conditions. Ground Water 3 (3), 5–14.
Limited, S., 2009. Scintrex Autograv System Operation Manual. Ontario, Canadá, p. 304p.
Sudicky, E.A., 1986. A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process. Water Resour. Res. 22 (13), 2069–2082.
Sudicky, E.A., Illman, W.A., 2011. Lessons learned from a suite of CFB Borden experiments. Groundwater 49 (5), 630–648.
Theis, C.V., 1935. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground-water storage. Eos, Trans. Am. Geophys. Union 16 (2), 519–524.