

---

# On the meaning of the transmissivity values obtained from recovery tests

M. Willmann · J. Carrera · X. Sánchez-Vila · E. Vázquez-Suñé

**Abstract** Recovery tests are based on estimating transmissivity,  $T$ , from the heads that rebound after pumping has stopped. Recovery tests can be performed in wells where conventional constant-rate pumping tests would not be possible. Test interpretation is based on the simple Theis recovery method, related to late time drawdown in an infinite homogeneous aquifer. Yet, field data often cannot be explained by the homogeneous theory. Because  $T$  is heterogeneous over an evolving range of scales, it is important to evaluate the support scale of hydraulic tests. Numerical simulations are performed to show that heterogeneity in  $T$  can explain these field observations. It is also shown that the local  $T$  value around the well can be inferred from early time-recovery data, assuming ideal conditions, whereas late time data yield a large-scale (regional) representative value. Even when recovery is observed for a short time, indirect information about the regional value can also be obtained. A method for the interpretation of recovery tests is proposed based on the Theis recovery method that takes into account the heterogeneity of aquifers. Finally, some guidelines are provided for best test performance depending on the scale of the problem.

**Résumé** Les tests de remontée sont basés sur l'estimation de la transmissivité,  $T$ , d'après les niveaux piézométriques qui se redressent suite à l'arrêt du pompage. Les tests de remontée peuvent être réalisés dans des puits où les tests conventionnels à pompage constant ne pourraient pas être possible. L'interprétation du test est basé sur la simple méthode de Theis à la remontée, relative aux dernières données de rabattement dans un aquifère infini homogène. Cependant, les données de terrain ne peuvent pas souvent être expliquées par la théorie homogène. Parce que  $T$  est

hétérogène sur une échelle évolutive d'échelles, il est important d'évaluer l'échelle qui s'accommode aux tests hydrauliques. Des simulations numériques sont réalisées pour montrer que l'hétérogénéité de  $T$  peut expliquer ces observations de terrain. Il est également montré que la valeur de  $T$  autour du puits peut être interprétée grâce aux premières données du tests en supposant les conditions idéales à cette interprétation, tandis que les derniers rabattement représentent une valeur de  $T$  à une échelle plus large (régional). Même lorsque la remontée est observée sur une courte durée, une information indirecte sur la valeur régionale peut également être obtenue. Une méthode pour l'interprétation des tests de remontée est proposée basée sur la méthode de la remontée de Theis qui prend en compte l'hétérogénéité des aquifères. Finalement, des règles générales sont apportées pour améliorer la performance du test en fonction de l'échelle du problème.

**Resumen** Las pruebas de recuperación se basan en estimar la transmisividad,  $T$ , a partir de las cabezas que se recuperan después que el bombeo se ha detenido. Las pruebas de recuperación pueden realizarse en pozos donde las pruebas de bombeo convencionales, a caudal constante, no son posibles. La interpretación de la prueba se basa en el método simple de recuperación de Theis, relacionado a los últimos valores del abatimiento, en un acuífero homogéneo infinito. No obstante, los datos del campo no pueden ser explicados a menudo por la teoría homogénea. Debido a que  $T$  es heterogéneo, por encima de un rango creciente de escalas, es importante evaluar la escala de apoyo de las pruebas hidráulicas. Se realizan las simulaciones numéricas, para mostrar que esa heterogeneidad en  $T$  puede explicar las observaciones de campo. También se muestra que los valores de  $T$  locales, alrededor del pozo, pueden inferirse a partir de los datos de recuperación iniciales, asumiendo condiciones ideales, mientras que los datos finales entregan un valor representativo de escala grande (regional). Aún cuando la recuperación se observe durante un tiempo corto, también puede obtenerse una información indirecta sobre el valor regional. Se propone un método, que tiene en cuenta la heterogeneidad de los acuíferos, para la interpretación de pruebas de la recuperación, con base en el método de recuperación de Theis. Finalmente, se dan algunas pautas, para la mejor ejecución de la prueba, que dependen de la escala del problema.

---

Received: 12 December 2005 / Accepted: 1 December 2006  
Published online: 1 February 2007

© Springer-Verlag 2007

---

M. Willmann (✉) · J. Carrera · X. Sánchez-Vila · E. Vázquez-Suñé  
Department of Geotechnical Engineering and Geosciences,  
Technical University of Catalonia,  
Jordi Girona 1-3, 08034, Barcelona, Spain  
e-mail: matthias.willmann@upc.edu  
Fax: +34-934017251

**Keywords** Hydraulic testing · Heterogeneity · Recovery tests · Scale effects

## Introduction

A large body of work in hydrogeology has been devoted to the design of hydraulic tests that can provide good estimates of aquifer parameters. Specifically, recovery tests are easy to perform and provide reliable estimates of transmissivity,  $T$ . Recovery tests consist of observing the build-up (or recovery) of hydraulic heads after pumping has stopped. If possible, measurements should continue until the head has recovered to its prior-to-pumping value.

The advantages of a recovery test stem from its simplicity: (1) a recovery test follows naturally from a pumping test, because it only requires the recording of heads after pumping has ceased; (2) it can be used even when pumping rates are difficult to control; (3) it is fairly inexpensive and no equipment or additional observation wells are required apart from a water-level measuring device; and (4) results are usually not sensitive to well losses.

The analysis of a recovery test follows the Theis solution to pumping a fully penetrating well of zero radius in an infinite, homogeneous aquifer. The most common and easiest way to interpret a recovery test is to use the Theis recovery method (Theis 1935), an approximation to late time data which leads to:

$$s = \frac{2.303Q}{4\pi T} \log\left(\frac{t+t_p}{t}\right) = 0.183 \frac{Q}{T} \log(t^*), \quad (1a)$$

where  $s$  is the residual drawdown,  $Q$  (constant) is the pumping rate,  $T$  is the transmissivity,  $t_p$  is the pumping time, and  $t$  is the elapsed time since pumping stopped. The variable  $t^* = (t+t_p)/t$  is termed equivalent time. In dimensionless form, Eq. (1a) can be rewritten as:

$$s_D = \log\left(\frac{t_D + t_{PD}}{t_D}\right) = \log(t^*) \quad (1b)$$

where  $s_D = 4\pi Ts/2.303Q$  and  $t_D = Tt/SL^2$  with  $L_c$  representing some characteristic length of the problem and  $S$  being storage coefficient. The variable  $t_{PD}$  corresponds to the dimensionless pumping duration. Equations (1a) and (1b) are valid for large times, meeting the condition  $u = r^2 S/4Tt < 0.01$  (e.g. Freeze and Cherry 1979), where  $r$  is the distance between pumping and observation points. In terms of equivalent time, this condition is  $t^* < 1 + 0.04t_{PD} \frac{L_c^2}{r^2}$ . In recovery tests, heads are usually observed at the pumping well, so that  $r$  is the well radius, which tends to be small, making this condition easy to meet.  $Q$  can be approximated as  $V_w/t_p$ , where  $V_w$  is the total volume of water pumped. This result is useful when  $Q$  is not constant, even if this variability might affect early time recovery.

The interpretation of a recovery test is performed by plotting residual drawdown against equivalent time on a

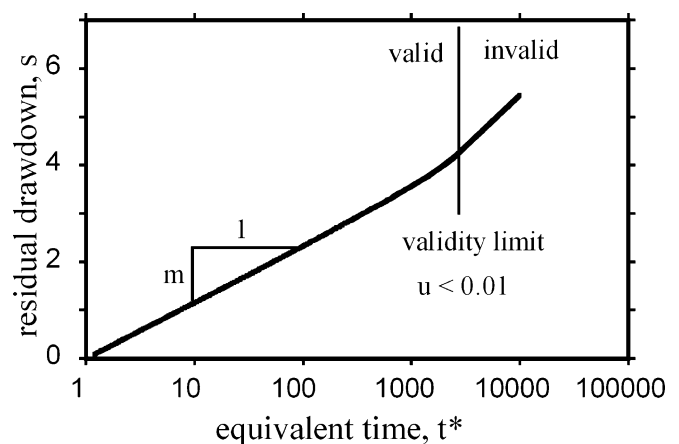
semi-logarithmic plot (Fig. 1). The use of equivalent time causes late time data to be displayed on the left side, corresponding to small residual drawdown. Equation (1a) shows that late time data displays a straight line passing through the origin, provided that no residual drawdown remains when the aquifer reaches equilibrium. The slope,  $m$ , of this line is the coefficient in Eq. (1a). Knowing  $m$  allows estimating transmissivity as:

$$T = 0.183 \frac{Q}{m} \quad (2)$$

In principle, the interpretation of one recovery test using a semi-log plot would not allow estimating  $S$ . Recent research on recovery tests has concentrated on overcoming this limitation (Agarwal 1980; Banton and Bangoy 1996; Goode 1997; Chenof and Chapuis 2002; Zheng et al. 2005). They focus on finding an alternative interpretation to estimate the storage coefficient by also using heads at observation points different from the pumping well.

The Theis solution and, thus, the Theis recovery method, are based on some assumptions that are not met in reality, but may affect test results: infinite aquifers, negligible well radius, perfectly confined aquifer, constant pumping rate, and homogeneous hydraulic parameters. Actually, aquifers in the field are bounded, wells do have a finite radius, flow to a pumping well may not be purely horizontal but have a vertical component, pumping rate may be kept constant with difficulty, and transmissivity and storativity are not homogeneous. Some of these departures from the ideal hydrologic conditions have been thoroughly studied for hydraulic tests and several methodologies to account for them are available in the literature (e.g. Streltsova 1988; Bourdet 2002).

Heterogeneity remains a major difficulty in the interpretation of field hydraulic testing. Hydraulic conductivity measured at some small representative scale is known to vary by orders of magnitude from one point to another,



**Fig. 1** “Ideal” Theis recovery plot (also called Horner plot), displaying residual drawdown versus equivalent time, defined as  $t^* = (t+t_p)/t$ . Transmissivity can be calculated from the slope  $m$  through Eq. 2. The validity limit of the Theis recovery method  $u$  is  $r^2 S/4Tt < 0.01$

even in apparently homogeneous aquifers (e.g. Gelhar 1993). This result makes it important to know the individual support scale of the measurements. If the sampled volumes of an aquifer are too large, estimated parameters would be artificially homogenized. In large projects, it is not unusual to perform different types of tests involving several support scales in order to find the adequate input value for numerical models (Shapiro and Hsieh 1998; Martinez-Landa and Carrera 2005).

A few results are available for the response of heads to pumping in a heterogeneous aquifer. Analytical solutions are available for simplified heterogeneities like a disk of transmissivity  $T_1$  embedded in a homogenous matrix of transmissivity  $T_2$  (Moench and Hsieh 1985; Butler 1990), geometries with three different zones, each one with a different  $T$  value shaping as disks (Butler and Liu 1993), or with complicated, but, in any case, prefixed geometries (Chu and Grader 1999). Schad and Teutsch (1994) analyzed the support scale of pumping tests by looking at an extensive set of field tests, and thus, were able to characterize an effective heterogeneity length scale. Meier et al. (1998) found that the late time behavior of heads, measured by the slope of drawdown vs. log time was: (1) the same for all points independent of their relative location with respect to the well, and (2) characterized by the effective transmissivity of the aquifer, thus, making Jacob's method also valid for heterogeneous media. These findings were confirmed analytically by Sánchez-Vila et al. (1999) using a small perturbation approach. Meier et al. (1999) showed that specific-capacity measurements depend strongly on the local transmissivity next to the well. More recently, Coptý and Findikakis (2004) and Neuman et al. (2004) have focused on the possibility of obtaining the integral scale and the variance of local log-conductivity from a small number of pumping tests in a given aquifer. None of the above results was derived for recovery tests but these results are directly applicable here.

Heterogeneity affects recovery data at all times. Heterogeneity also affects lateral boundaries that can be conceptualized as extreme cases of heterogeneity. There are, as well, certain effects that influence early time data. The most important effects considered here are vertical leakage, wellbore storage, and skin effect. In case of a large amount of available drawdown data, it is possible to analyze these effects in both pumping and recovery tests, and there is extensive literature on this subject (e.g. Streltsova 1988; Bourdet 2002). In any case, the point is that they are limited in time. Using the results of Papadopoulos and Cooper (1967) and Agarwal et al. (1970), wellbore storage and skin effect can be ruled out if:

$$t > \frac{S_r^2}{2.25T} \quad (3a)$$

$$t > 10 \frac{S_w e^{2\sigma}}{T} \quad (3b)$$

where  $r_s$  is the radius of the region sampled by the test,  $S_w$  is the well storage (water surface area in open wells), and  $\sigma$  is the skin factor. Using rough estimates of  $S$  and  $T$ , pumping and recovery time can be defined so as to eliminate these perturbations from the ideal case. Oppositely, following Eq. (3a) for interpreting a given test, the sampling radius,  $r_s = 1.5\sqrt{Tt/S}$ , can be solved. Depending on the scale of the sampling radius and additional data, the analyst of the pumping test has to judge whether all these early effects can be ruled out and, thus, the remaining data are governed by the underlying heterogeneous  $T$  field.

In order to grasp the behavior of recovery tests in aquifers in the field, four selected plots of recovery tests are shown in Fig. 2. It is clearly visible that, in most cases, the shapes of the curves are far from that of Fig. 1 (that is, far from ideal). First, drawing a single straight line is in most cases difficult. Second, the extrapolation of these straight lines rarely crosses the origin as predicted by the homogeneous theory. Local scale heterogeneity of transmissivity can explain these two effects (variable slope and extrapolation line not passing through the plot origin).

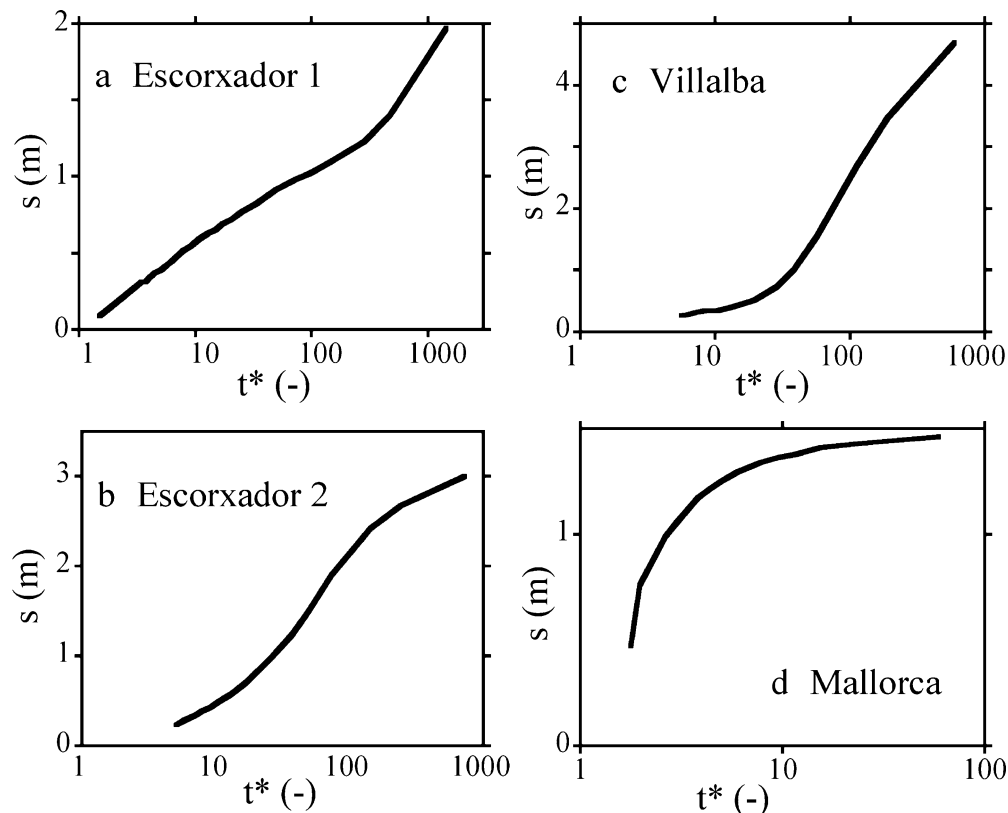
Proper accounting for heterogeneity needs to be done in the context of geostatistical inversion. In practice, however, this accounting is only done when several pumping tests and/or several observation wells are available (Yeh and Liu 2000; Meier et al. 2001; Vesselinov et al. 2001). The scope of recovery tests is usually much more limited than when several pumping and observation wells are available. Interpretation of recovery tests is usually done by using the simple homogeneous model. The immediate question arising is what is the meaning of the value obtained using this simple interpretation?

The objective of the work described here is to answer the above question and specifically to study if the classical interpretation of recovery tests yields meaningful and useful hydraulic parameter values in heterogeneous aquifers. If yes, as conjectured by extrapolating the results of Meier et al. (1998), then the practical question would be how long to pump or to observe recovery in order to obtain meaningful results. To address these issues, numerical simulations are performed first and then the results to the tests of Fig. 2 are applied.

## Methodology for the numerical simulations

The methodology is adapted from that of Meier et al. (1998). The methodology can be summarized into four basic steps: (1) generation of heterogeneous transmissivity fields; (2) numerical simulation of recovery tests within these fields; (3) interpretation and (4) comparison of interpretation results with hydraulic parameters representative of the original transmissivity field. These steps are discussed below.

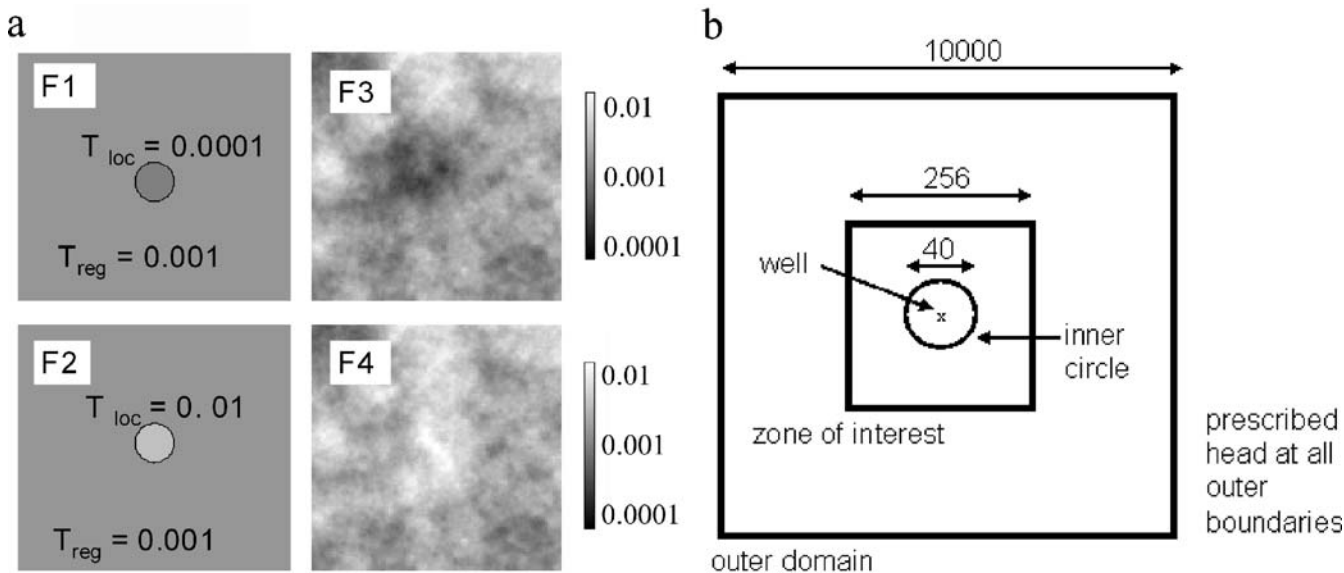
**Fig. 2** Examples of field recovery tests performed within the city of Barcelona, Spain. The Theis recovery plots are plotted as residual drawdown vs. dimensionless equivalent time. Escorxador 1 and 2 are tests performed at Escorxador Park, Villalba is a test at Villalba Street and Mallorca is a test at Mallorca Street. Only Escorxador 1 (a) displays a shape similar to that in a homogeneous aquifer. Notice that even in this test, early time data ( $t^* > 100$ ) departs from the ideal condition, reflecting well effects such as skin, wellbore storage, etc. (negative skin in this case)



### Transmissivity fields

The problem domain is identical for all transmissivity fields. The domain consists of a  $256 \times 256$  zone of interest. An inner disc of radius 20 around the pumping

well is also defined to specify heterogeneity patterns. Four different heterogeneous transmissivity fields (F1–F4) were generated (see Fig. 3a). F1 and F2 are deterministic with  $T = T_{loc}$  for all elements located within



**Fig. 3** a Transmissivity fields F1–F4. F1 and F2 show a disc around the well with  $T$  being one order of magnitude lower (F1) or higher (F2) than the regional value. F3 and F4 are simulated heterogeneous  $T$  fields conditioned at the well location to a value one order of magnitude higher (F3) or lower (F4) with respect to the geometric mean of the  $T$ -point values. b Set-up of the numerical simulations. The fields in a correspond to the zone of interest. In the outer domain, the grid becomes increasingly coarse towards the boundary, whereas it is refined at the well. Plot b is not to scale. The characteristic length,  $L_c$ , is 20, equal to the radius of the inner disc in fields F1 and F2 and equal to the correlation length in fields F3 and F4. All numbers are unit length



the inner disc, and a different value,  $T=T_{\text{reg}}$ , outside of the disc. In field F1,  $T_{\text{loc}}$  is one order of magnitude lower than  $T_{\text{reg}}$ , whereas the opposite holds for F2. F3 and F4 are defined as individual realizations of a random stationary field with a spherical covariance function with a correlation length,  $\lambda$ , of 20 and a variance of 1 in terms of  $\ln T$ . Both fields are conditional to the  $T$  value around the well,  $T_{\text{loc}}$ , which is one order of magnitude lower (F3) or larger (F4) than the effective large scale transmissivity,  $T_{\text{reg}}$ . The random  $\ln T$  fields were generated using the Gaussian random simulator GCO-SIM3D (Gomez-Hernandez and Journel 1993). The correlation length was chosen equal to the radius of the disc in F1 and F2 for later comparison. Notice that only two realizations are used. The underlying idea is not to follow a Monte Carlo procedure, but to generate individual realizations to see whether information could be obtained about recovery tests in a given aquifer.

### Simulation of recovery tests

Pumping and recovery was simulated using the finite-element flow-modeling code FAITH (Sánchez-Vila et al. 1992). The zone of interest (Fig. 3b) consists of a grid containing  $256 \times 256$  heterogeneous regular square cells. The boundary conditions applied to this ideal grid are, first, zero drawdown at the outer boundaries and, second, a constant pumping rate prescribed at the well (located in the center of the grid). The pumping rate is set to zero after pumping stops. These boundary conditions allow a total drawdown recovery for very large times.

The grid is modified in order to avoid boundary effects (Fig. 3b). First, the grid is enlarged outside the zone of interest to ensure that the outer boundary is far enough and does not affect the results. In this outer domain,  $T$  is homogeneous and set to value outside the disc ( $T_{\text{reg}}$ ) in the case of F1 and F2 and to the geometric mean ( $T_{\text{reg}}$ ) in F3 and F4. Second, the grid was refined around the well to improve simulation accuracy. The level of refinement was verified by comparison with the analytical solution for a homogeneous medium (for  $u < 0.01$ ). This refinement does not affect the regular shape of the transmissivity fields generated. The same approach was shown by Meier et al. (1998) to provide drawdown results that are not affected by boundary conditions.

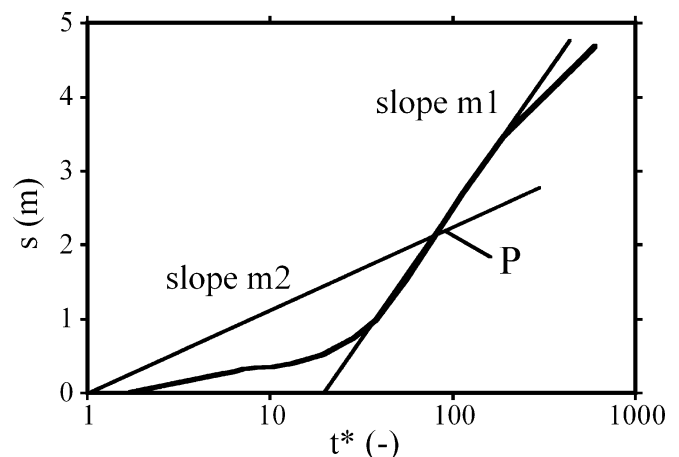
One of the objectives is to obtain insight into the recommended pumping duration. Therefore, three tests durations were simulated for each field: a short-term test with a dimensionless pumping time  $t_{\text{PD}}=0.1$ , an intermediate-term test with  $t_{\text{PD}}=1$ , and a long-term test with  $t_{\text{PD}}=10$ . Here,  $t_{\text{PD}}$  is defined as in Eq. 1b, using the disc radius (F1 and F2) or the correlation length (F3 and F4) as characteristic length,  $L_c$ , and as the transmissivity value the (constant)  $T_{\text{loc}}$ . Because  $T_{\text{loc}}$  is either  $10 T_{\text{reg}}$  (F1 and F3) or  $0.1 T_{\text{reg}}$  (F2 and F4), actual pumping is different for each field: (1)  $t_{\text{PD}}=0.1$  corresponding to a local (short) test with a very small radius of influence, so that the

perturbation in terms of drawdown has barely reached the limit of the inner disc or the correlation length. Therefore, the support volume sampled by the test during pumping is a small area around the well; (2)  $t_{\text{PD}}=1$ , an intermediate test where pumping lasts long enough to bring the drawdown signal to a distance comparable to the characteristic length,  $L_c$ ; and (3)  $t_{\text{PD}}=10$ , a very long test that samples areas well beyond the characteristic distance, thus, affecting the area corresponding to  $T=T_{\text{reg}}$ . In all cases, the recovery time was set to five times the pumping time. Notice that measurements in field tests are seldom performed after elapsed time of one to two times the pumping time. Usually, the recovery time is often shorter (a fraction of the pumping time).

### Analysis of the simulated tests

A number of methods are available for interpretation of recovery tests. Possibly, diagnostic plots (displaying the derivative of residual drawdown with respect to  $\log t_{\text{PD}}/t^*$ ) are the most informative method. A large body of literature is available in the oil industry, where recovery tests are termed “shut-in” tests, about ideal features that can be identified with diagnostic plots. Unfortunately, diagnostic plots require pumping rates to be well monitored. Moreover, existing tools are better suited for analyzing early time behavior. For the study of large-scale effects, the variations of the Theis recovery method described below were applied.

The interpretation consists of three steps. First, dimensionless drawdown is plotted against the logarithm of the equivalent time ( $t^*$ ). Second, two different slopes are defined at any given time. Slope  $m_1$ , which is defined as the tangent of the residual drawdown data, is its derivative with respect to  $t^* = (t + t_p)/t$  (Fig. 4) is considered here. This slope is computed by using moving windows in order



**Fig. 4** Two different calculations of slope  $m$  in a heterogeneous aquifer at the example test site Villalba. At any point  $P$ , the slope can be calculated as that of the tangent ( $m1$ ) or as that of the secant, the line between the origin and  $P$  ( $m2$ ). Note that  $m1$  and  $m2$  should be identical for late time data in an ideal homogeneous aquifer

to avoid numerical artifacts. A second slope,  $m_2$ , is defined as that corresponding to the secant that joins any given point of the semi-log plot with the origin ( $s=0$ ,  $t^*=1$ ). Third, slopes are converted into normalized estimates of transmissivities  $T^*=T/T_{loc}$  by means of

$$T_{m1}^* = 0.183 \frac{Q}{T_{loc} m_1} \quad (4)$$

$$T_{m2}^* = 0.183 \frac{Q}{T_{loc} m_2} \quad (5)$$

which are simple extensions of Eq. (2) but normalized by the well (local)  $T$  value. Note that in the case of a homogeneous medium (Fig. 1),  $T_{m1}^* = T_{m2}^* = 1$  should be obtained in the range of validity of the method. Notice also that  $T_{loc}^* = 1$  for all fields but  $T_{reg}^* = 10$  in fields F1 and F3, whereas  $T_{reg}^* = 0.1$  in fields F2 and F4.

## Discussion of results

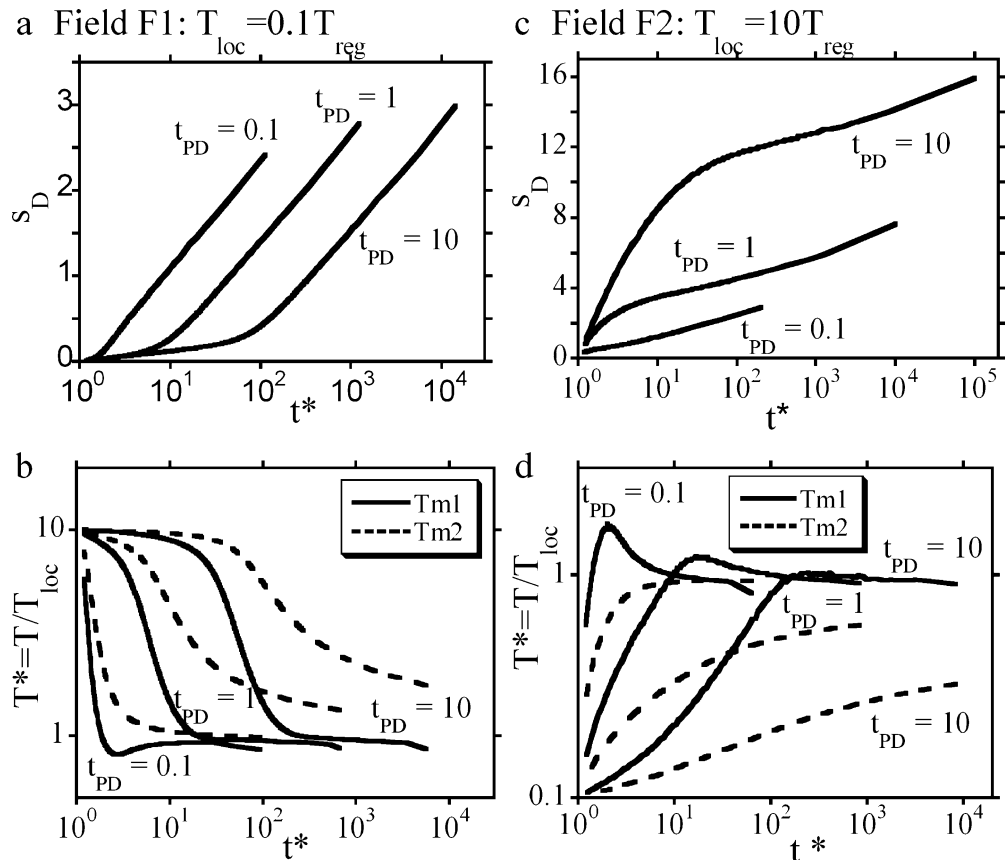
Results are presented in terms of both residual dimensionless drawdown and equivalent transmissivities vs. equivalent time for all fields and pumping durations.

The short-term test with  $t_{PD}=0.1$  in field F1 ( $T_{reg}^* = T_{reg}/T_{loc} = 10$ ) looks basically like the homogeneous

case (Fig. 5a). The resulting late time straight line crosses the axes close to the origin of the semi-log plot, as would happen in a homogeneous field. Furthermore, the estimated transmissivity is very close to the local value. Only for very long times (several times the pumping period) would a deviation from this behavior be found. Therefore, the estimated transmissivity value would increase with recovery time. In short, it would not be possible to resolve (find) the regional value with such a short test. The reason is that in such a short pumping time, the signal has barely gone beyond the inner circle and the test behaves mostly as if performed in a homogeneous medium.

With a longer test,  $t_{PD}=1$ , drawdown at the well is affected by the transmissivity beyond the inner circle. Estimated transmissivities obtained from  $T_{m1}^*$  at early times are close to  $T_{loc}^*$  (equal to 1). As time increases ( $t^*$  tends to 1), transmissivity tends to  $T_{reg}^* T_{m2}^*$ ; on the other hand, transmissivity always lies between  $T_{loc}^*$  and  $T_{reg}^*$ , converging to  $T_{reg}^*$  only for long recovery times. Extrapolating the early time data crosses the time axis at  $t^*>1$ . Similar results, but even more pronounced, can be seen for  $t_{PD}=10$  (Fig. 5a). The extrapolation of early time data crosses the equivalent time axis at an even higher value than for the case  $t_{PD}=1$  and the one for late time data crosses  $t^*=1$  exactly. For these latter two cases,  $t_{PD}=1$  and 10, a very short recovery test would lose all the relevant information regarding  $T_{reg}$ . Further, in both cases,  $T_{m2}^*$  is at all points between  $T_{loc}^*$  and  $T_{reg}^*$  and larger than  $T_{m1}^*$ .

**Fig. 5** Numerical results for fields F1 and F2. **a** Dimensionless drawdown for three different pumping durations for F1, and **b** corresponding estimated transmissivities ( $T^*=T/T_{loc}$ ) derived from either (4) or (5). **c** Dimensionless drawdown for three different pumping durations for F2, and **d** corresponding estimated transmissivities. Shown are only the portions of the curves meeting the condition of validity of the Theis recovery method ( $u<0.01$ )



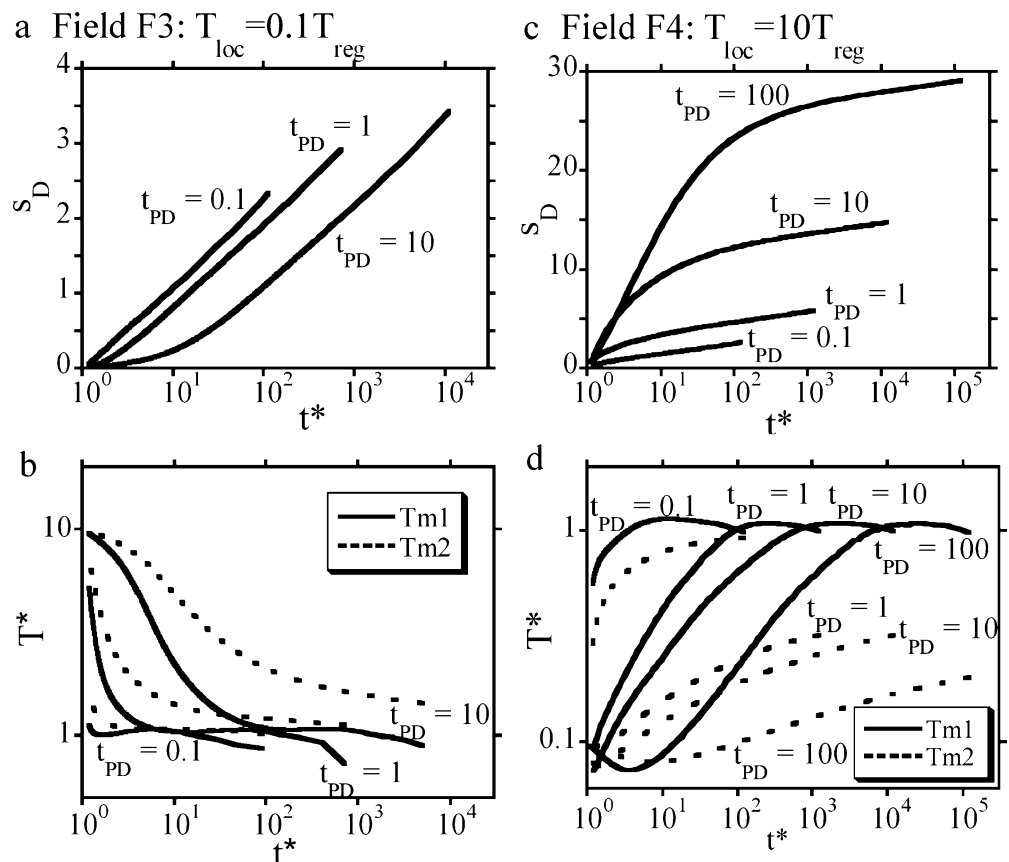
In field F2,  $T_{reg}^*$  is 0.1. The short-term test ( $t_{PD}=0.1$ ) is essentially identical to the corresponding test with  $T_{reg}^* = 10$  (the difference between Fig. 5a,c for  $t_{PD}=0.1$  is limited to changes in vertical scales, except for very late recovery, i.e. small  $t^*$ .  $T_{loc}$  is well resolved and the extrapolation of the data passes through the origin (zero drawdown at infinite time). This test again behaves like a test performed in a homogeneous medium as the heads are insensitive to the  $T_{reg}$  value. The intermediate test,  $t_{PD}=1$  (see Fig. 5c, d), displays a behavior very different from the corresponding test in field F1.  $T_{loc}$  is well defined from the short-time data. However,  $T_{m1}^*$  tends to the regional value, but it does not stabilize as in the F1 case. This difference can be attributed to the fact that  $T_{reg}$  is now smaller than  $T_{loc}$ . As a result, drawdowns progress slowly beyond the inner circle. Therefore, only a relatively small area of  $T_{reg}$  is actually sampled. The extrapolation of the slope corresponding to early time data would cross the zero drawdown line at  $t^*<1$ . In the homogeneous interpretation of a field test, this extrapolation is usually interpreted as if the original water level were never to be recovered. If the recovery is observed for very long times, the line bends, and the slope of the late time data would really cross the  $y$ -axis at a value close to 1. This result is again a reason for extending the recovery time as much as possible. These effects are similar but more pronounced in the case of  $t_{PD}=10$  (Fig. 5) with both  $T_{loc}^*$  and  $T_{reg}^*$  resolved visibly by  $T_{m1}^*$ . Again, the extrapolation of the early time slope leads to a crossing point well beyond

$t^*=1$ . The late time slope crosses exactly through the origin.  $T_{m2}^*$  lies, in all the cases, between  $T_{loc}^*$  and  $T_{reg}^*$ , but converges to  $T_{reg}^*$  faster than  $T_{m1}^*$ . This implies that  $T_{m2}^*$  can be used as an estimate for the not resolved  $T_{reg}^*$  when recovery is short. When applying Eq. (3a) to the turning points in the semi-log plots for fields F1 and F2, values of  $r=14$  and 25, respectively, are obtained. As the real distance is 20, this value shows that Eq. (3a) does indeed lead to a rough estimate of heterogeneity scale.

Transition from  $T_{loc}$  to  $T_{reg}$  estimated from  $T_{m1}$  is preceded by a small trough when  $T_{loc}<T_{reg}$  (Fig. 5b) and by a small peak when  $T_{reg}>T_{loc}$  (Fig. 5d). These results are caused by the change in transmissivity (they occur for  $t_{DP}\approx 0.1$ , regardless of the pumping duration) that causes recovery initially to speed up (when  $T_{reg}<T_{loc}$ , Fig. 5d). The effect is similar to double porosity, linear flow, or boundary effects.

Fields F1 and F2 can be viewed as a simplistic representation of realistic heterogeneous media. Field sites will rarely display a single geometric feature separating local and regional  $T$  values. The spatial distribution of  $T$  values will not be radially symmetric. Recovery tests in fields F3 and F4 (Fig. 6) display some similarities to those in F1 and F2, but there are a number of distinct differences. For the shortest test ( $t_{PD}=0.1$ ), only the local value can be well resolved in both fields, even though  $T_{m2}^*$  is slightly bent towards the regional value (recall that now  $T_{reg}$  is the effective transmissivity, the geometric mean of point values). Notice that the local value corresponds here

**Fig. 6** Numerical results for fields F3 and F4. **a** Dimensionless drawdown for 3 different pumping durations for F3, and **b** corresponding estimated transmissivities. **c** Dimensionless drawdown for three different pumping durations for F4 with additional pumping time  $t_{PD}=100$ , and **d** corresponding estimated transmissivities. Shown are only the portions of the curves meeting the condition of validity of the Theis recovery method ( $u<0.01$ )



to the one used for conditioning the fields. Because the correlation length is large compared to the cell size, this local value would correspond also to some integrated value of the local  $T$  values around the well.

Results from the intermediate test ( $t_{PD}=1$ ) in fields F3 and F4 are similar to those of the short-term test, but late time data provide an improved estimation of the regional value with respect to the shortest test. However, the behavior of the two fields is qualitatively different in that both  $T_{m1}$  and  $T_{m2}^*$  identify regional transmissivity much better in field F4 (high  $T_{loc}$ , Fig. 6d) than in F3 (low  $T_{loc}$ , Fig. 6c). It is clear that the low  $T$  zone is “screening” the high  $T$  regional value in field F3. Contrarily,  $T_{reg}$  becomes well defined if the pumping duration is long enough (see Fig. 6 for  $t_{PD}=10$ ) when  $T_{loc}$  is small. In order to register a sharp definition of  $T_{reg}$  when it is smaller than  $T_{loc}$ , a larger  $t_{PD}$  value is required. The results corresponding to  $t_{PD}=100$ , also shown in Fig. 6, also confirms this requirement.

In summary,  $T_{m1}$  provides good estimates of  $T_{loc}$  with early time data and eventually tends to  $T_{reg}$  if both pumping time and recovery durations are long enough.  $T_{m2}$ , on the other hand, helps to identify departure from ideality, as it consistently lies between  $T_{loc}$  and  $T_{reg}$ . At late times, both values converge for sufficiently long pumping and recovery times.

### A reinterpretation of the field tests

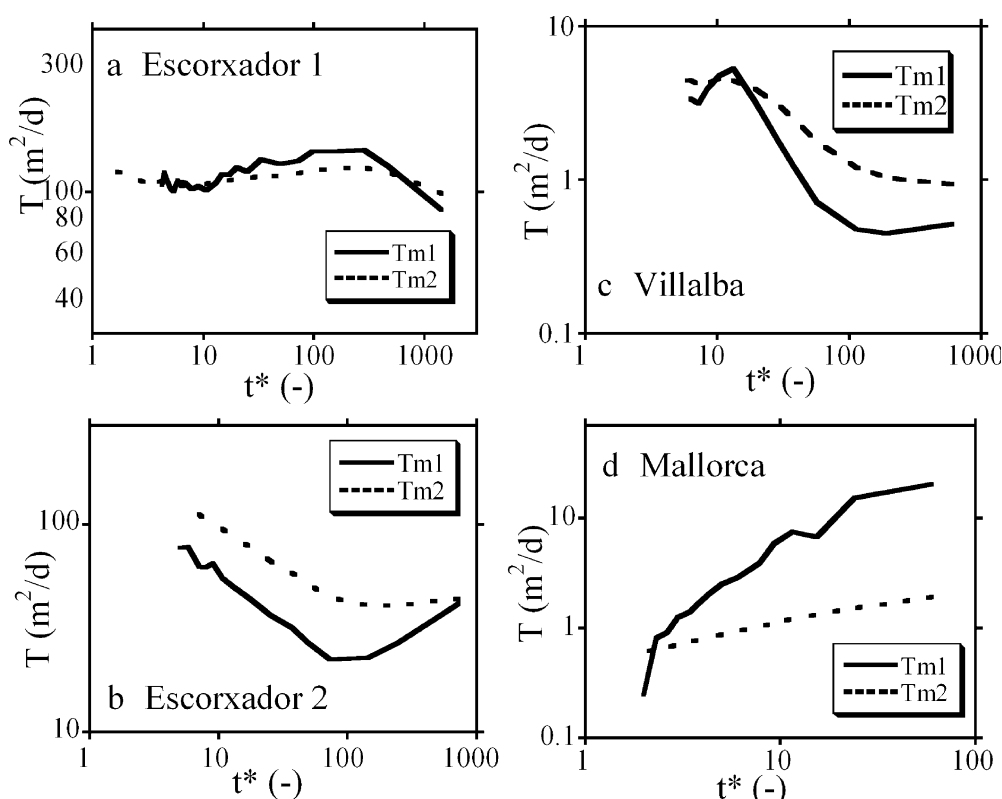
In this section, the above results are used to discuss the interpretation of the field tests shown in Fig. 2. Curves of

estimated transmissivity vs. equivalent time resulting from Eqs. (4) and (5) are shown in Fig. 7. The only difference with respect to the numerical simulations from synthetic fields presented in Figs. 5 and 6 is that now the values of  $T_{m1}$  and  $T_{m2}$  shown are dimensional, because the local value of  $T$  at the well is not known.

### Tests at Escorxador Park

Two long-term recovery tests ( $t_p=12$  h, constant rate) were performed at different times at two wells located 74 m apart (Fig. 2a,b). The shapes of the recovery curves are appreciably different. Whereas test 1 follows a straight line for  $t^*<300$  and looks like the expected result for a homogeneous aquifer, the same is not true for test 2, where the slope increases with time (decreases with  $t^*$ ) without reaching a plateau. Despite these differences, very similar late time transmissivity estimates were determined from both tests. These estimates are consistent with  $T$  values derived from long-term pumping tests (Vázquez-Suñé and Ondiviela 2003). Early time data lead to a difference in estimated transmissivities,  $T_{loc}$ , of about one order of magnitude between the two tests. From the methodology already explained, skin effect and effect of wellbore storage can be ruled out. This result would mean that well 2 is located in an area less conductive than well 1, which is consistent with the fact that observed draw-downs are larger in well 2.

**Fig. 7** Reinterpretation of the field test data presented in Fig. 2. The tests results are plotted as transmissivities vs. equivalent time.  $T_{m1}$  and  $T_{m2}$  are defined in Eqs. 4 and 5 but not normalized by  $T_{loc}$  as it is not known. **a**  $T_{m1}$  and  $T_{m2}$  are basically the same but in **b–d** they differ particularly for early times





### Test at Villalba Street

Recovery data of test Villalba ( $t_p=9$  h, constant rate) are displayed in Fig. 2c. As for the Escorxador 2 test, the shape of the curve does not display a homogeneous behavior. Using Eqs. (3a) and (3b), skin effect and wellbore storage can be neglected. The apparent transmissivity as a function of time is plotted in Fig. 7c. From the plot, it can be concluded that  $T_{reg}$  is larger than  $T_{loc}$ . In this test, an external confirmation of the estimation of  $T_{reg}$  is confirmed as it is very close to the value calibrated in a regional groundwater model of the area (Vázquez-Suñé and Ondiviela 2002).

### Test at Mallorca Street

The short recovery test ( $t_p=0.32$  h, constant rate) in Fig. 2d displays clearly two different slopes. The corresponding estimated transmissivities in Fig. 7d suggest an increase from  $T_{loc}$  to  $T_{reg}$  of almost two orders of magnitude. The check with Eq. (3a) renders a sampling radius,  $r_s$ , of around 10 cm for the time where the change in slope takes place. This result clearly indicates that both the pumping and recovery times were too short. The different  $T_{loc}$  is most likely caused by skin effect or wellbore storage. This result was confirmed by interpreting the test using an automatic fitting tool (Ondiviela et al. 2001) that provides about the same result for  $T$  as our estimated value  $T_{loc}$  and a very high skin factor.

### Conclusions and recommendations

Interpretation of recovery tests using the simple Theis recovery method can provide valuable information about representative parameters in heterogeneous aquifers, even though the method was developed for homogeneous media. Actually, it may be possible to discriminate between at least two representative transmissivity values, corresponding to the local  $T$  values ( $T_{loc}$ ) surrounding the well, and some regional representative value ( $T_{reg}$ ). The analysis presented here implies that the rate of early time recovery is informative of  $T_{loc}$ , whereas that of late time data yields information about  $T_{reg}$ . This result is valid as long as early time effects can be filtered out. The latter was shown for pumping tests by Butler (1988) for a radially symmetric structure and then confirmed by Meier et al. (1998) for general heterogeneity. The results presented here have several practical implications about the design and interpretation.

Regarding test design the following recommendations are advisable:

1. Pumping duration. Pumping duration should be chosen depending on the scale,  $L_c$ , to be characterized. At the very least,  $t_p$  should equal  $SL_c^2/T$  where  $S$  and  $T$  are local estimates of storativity and transmissivity, respectively. Preferably, pumping durations should be much longer, for example,  $10 SL_c^2/T$ , especially if  $T_{reg}$  is expected to be smaller than  $T$ .

2. Recovery period. The design should allow for a long recovery period, not shorter than twice the pumping duration, and, if possible, much longer than twice the duration. Notice that this increase in time can be achieved at very low cost because time scale is logarithmic and only a few measurements have to be added. Because drawdowns become small for late time recovery, the reliability of these drawdowns needs to be assessed by comparing them to measurement errors.
3. Pumping rate. Late time residual drawdown is sensitive to the total volume of water pumped,  $V_W$ , as  $s \sim V_W / 4\pi T t$  (that is, it does not depend on the time evolution of pumping rate). Average pumping rate,  $V_W/t_p$ , should be designed for a sizable value of  $s$ . Recovery tests are not highly sensitive to high frequency changes in flow rate, but early time recovery is sensitive. Still, using  $V_W/t_p$  may become the only realistic option when pumping rate cannot be controlled.
4. Natural head trends. Heads should be monitored prior to the test so as to ensure either that they have stabilized prior to pumping or that a natural trend can be fitted. It should be noticed that much information is contained in late time residual drawdowns that are small. Residual drawdowns are equal to the difference between heads during recovery, which can be measured accurately, and natural heads. Efforts should be made to estimate the latter natural heads as accurately as possible. These efforts include not only prior stabilization, but also monitoring boreholes unaffected by the test. Mistaken natural head evolution would lead to wrong  $T_{reg}$  estimation.

Regarding the interpretation, the following recommendations are advisable:

1. At any given time the head recovery plot can yield two slopes:  $m_1$ , tangent, and  $m_2$ , secant though the origin. Using these,  $T_{m1}$  and  $T_{m2}$  can be calculated.
2. When both pumping and recovery times are large, it is possible to obtain good estimates of both  $T_{loc}$  and  $T_{reg}$ . Better estimations are obtained when the local value at the well is smaller than the regional  $T$  value, provided that sizable pumping rates can be sustained.
3.  $T_{loc}$  is better represented by  $T_{m1}$ . Early time effects (for example, skin effect) should be ruled out by calculating the corresponding sampling radius. A small radius indicates that these effects are still relevant and mask  $T_{loc}$  that cannot be estimated.
4. If pumping duration and recovery are long enough, late time  $T_{m1}$  and  $T_{m2}$  will tend to  $T_{reg}$ . The advantage of using  $T_{m2}$  is that it converges to the large-scale value much faster than  $T_{m1}$ . If recovery is too short to resolve  $T_{reg}$  ( $T_{m1} \neq T_{m2}$ ),  $T_{m2}$  can be used to approximate  $T_{reg}$ . The resulting value should be suspected.  $T_{m2}$  lies between  $T_{loc}$  and  $T_{reg}$ . Therefore,  $T_{reg}$  will be larger than  $T_{m2}$  if  $T_{loc}$  is small, and vice versa.

These recommendations are valid for the interpretation of recovery in the pumping well. Recovery in observation

wells usually only renders an intermediate value of transmissivity as early time responses are delayed and late time responses are also usually not resolved. This lack of resolution results in that total drawdown is smaller at an observation well and measurement errors become more important.

**Acknowledgements** This work was partly supported by ENRESA (Empresa Nacional de Residuos Radioactivos, S. A.) and by the Departament d'Universitats, Recerca i Societat de la Informació of the Catalan Government, Spain.

## References

- Agarwal RG (1980) A new method to account for producing time effects when drawdown type curves are used to analyze pressure built-up and other test data. SPE 9289, SPE-AIME 55th Annual Technical Conference and Exhibition, Dallas, Texas, September 1980
- Agarwal RG, Al-Hussainy R, Ramey HJJ (1970) An investigation of wellbore storage and skin effect in unsteady liquid flow. 1. Analytical treatment. AIME Trans 249:279–290
- Banton O, Bangoy LM (1996) A new method to determine storage coefficient from pumping test data. Ground Water 35:772–777
- Bourdet D (2002) Well test analysis: the use of advanced interpretation models (Handbook of petroleum exploration and production), 1st edn. Elsevier, Amsterdam
- Butler JJ Jr (1988) Pumping tests in nonuniform aquifers: the radially symmetric case. J Hydrol 101:15–30
- Butler JJ Jr (1990) The role of pumping tests in site characterization: some theoretical considerations. Ground Water 28:340–394
- Butler JJ Jr, Liu WZ (1993) Pumping tests in nonuniform aquifers: the radially asymmetric case. Water Resour Res 29:259–269
- Chenof D, Chapuis RP (2002) Methods to determine storativity of infinite confined aquifers from a recovery test. Ground Water 10:385–389
- Chu LF, Grader AS (1999) Transient pressure and rate analysis for active and interference wells in composite systems. In Situ 23:297–332
- Coptly NK, Findikakis AN (2004) Stochastic analysis of pumping test drawdown data in heterogeneous geologic formations. J Hydraul Res 42:59–67
- Freeze RA, Cherry JA (1979) Groundwater. Prentice Hall, Eaglewood Cliffs, NJ
- Gelhar LW (1993) Stochastic subsurface hydrology. Prentice Hall, Eaglewood Cliffs, NJ
- Gomez-Hernandez JJ, Journel AG (1993) Joint sequential simulation of multiGaussian fields. In: Soares A (ed) Geostatistics Troja '92, Kluwer, Dordrecht, pp 85–94
- Goode DJ (1997) Composite recovery type curves in normalized time from Theis exact solution. Ground Water 35:672–678
- Martinez-Landa L, Carrera J (2005) An analysis of hydraulic conductivity scale effects in granite (FEBEX experiment, Grimsel, Switzerland). Water Resour Res 41, W03006. DOI 10.1029/2004WR003458
- Meier PM, Carrera J, Sánchez-Vila X (1998) An evaluation of Jacob's method for interpretation of pumping tests in heterogeneous formations. Water Resour Res 34:1011–1025
- Meier PM, Carrera J, Sánchez-Vila X (1999) A numerical study on the relationship between transmissivity and specific capacity in heterogeneous aquifers. Ground Water 37:611–617
- Meier PM, Medina A, Carrera J (2001) Geostatistical inversion of cross-hole pumping tests for identifying preferential flow channels within a shear zone. Ground Water 39:10–17
- Moench AF, Hsieh PA (1985) Evaluation of slug tests in wells containing a finite-thickness skin-comment. Water Resour Res 21:1459–1461
- Neuman SP, Guadagnini A, Riva M (2004) Type-curve estimation of statistical heterogeneity. Water Resour Res 40, W04201. DOI 10.1029/2001WR001072
- Ondiviela M, Vázquez-Suñé E, Carrera J, Plana D (2001) Estudio hidrogeológico para el túnel de la c/ Mallorca de la penetración de la línea de alta velocidad en Barcelona [Hydrogeological study for the tunnel at Mallorca street about the influence of the high velocity train]. Internal report, Hydrogeology group, UPC, Barcelona
- Papadopolus IS, Cooper HHJ (1967) Drawdown in a well of large diameter. Water Resour Res 3:241–244
- Sánchez-Vila X, Carrera J, Colominas I (1992) Directional effects on convergent flow tracer tests. In: Hoetzel H, Werner A (eds) Proc 6th International Symposium on Water Tracing, Karlsruhe, Germany, September 1002, pp 407–414
- Sánchez-Vila X, Meier PM, Carrera J (1999) Pumping tests in heterogeneous aquifers: an analytical study of what can be obtained from their interpretation using Jacob's method. Water Resour Res 35:943–952
- Schad H, Teutsch G (1994) Effects of the investigation scale on pumping test results in heterogeneous porous aquifers. J Hydrol 159:61–77
- Shapiro AM, Hsieh PA (1998) How good are estimates of transmissivity from slug tests in fractured rock? Ground Water 36:37–48
- Streltsova TD (1988) Well testing in heterogeneous formations. Wiley, New York
- Theis CV (1935) The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground water storage. Trans Am Geophys Union 16:519–524
- Vázquez-Suñé E, Ondiviela M (2002) Interpretació de l'assaig de bombament del pou del C/ Villalba dels Arcs [Interpretation of the pumping test at the well Villalba dels Arcs street]. Internal report, Hydrogeology group, UPC, Barcelona
- Vázquez-Suñé E, Ondiviela M (2003) Interpretació dels dos assaigs de bombament dels pous 1 i 2 del Parc de l'Escorxador [Interpretation of two pumping tests at wells 1 and 2 at the park Escorxador]. Internal report, Hydrogeology group, UPC, Barcelona
- Vesselinov VV, Neuman SP, Illman WA (2001) Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff. 2. Equivalent parameters, high-resolution stochastic imaging and scale effects. Water Resour Res 37:3019–3041
- Yeh TCJ, Liu SY (2000) Hydraulic tomography: development of a new aquifer test method. Water Resour Res 36:2045–2105
- Zheng L, Guo JQ, Lei Y (2005) An improved straight-line fitting method for analyzing pumping test recovery data. Ground Water 43:939–942