Improving parameter estimation in transient groundwater models through temporal differencing

L.J.M. Peeters [a], **D. Rassam** [a], **J. Lerat** [a]

[a]*CSIRO. Water for a Healthy Country National Research Flagship*
Email: luk.peeters@csiro.au

Abstract In the majority of transient groundwater modelling studies, parameters are estimated by matching calculated heads to timeseries of observed heads. In regional groundwater systems, the variation in head between bores can be $10s$ or $100s$ of meters, while variation of head over time seldomly is larger than a few meters. Reducing the mismatch between observed and calculated heads, especially with automated calibration procedures, will often lead to a good approximation of the average water level in the bore, while the temporal dynamics are less well reproduced. The temporal dynamics however are often of the utmost importance to understand and predict the aquifer system under study.

This study explores the potential of using the sign of the difference of two subsequent head measurements as an additional calibration target to improve parameter estimation. The sign of the difference between two measurements isolates the temporal signal in the data, while it reduces the amount of redundant information in the calibration dataset, compared to adding the difference between two observations.

The relative parameter error reduction, introduced by Doherty and Hunt (2009), is used to assess the added value of temporal differencing for parameter estimation. This measure is based on linear error propagation through singular value decomposition of the sensitivity matrix of the model. This implies that it can be calculated even before actually calibrating a model, as long as an estimate of the error for the potential new observations is provided.

The methodology is tested on a regional-scale, calibrated transient groundwater model of the Upper Namoi valley in New South Wales (McNeilage, 2006). The results show that the introduction of temporal differencing as additional calibration data reduces the error variance of the estimated parameters. The results however also highlight that choices in the conceptualisation of the model, especially the implementation of boundary conditions, can dominate the amount of information that can be extracted from a calibration dataset and negatively affect the succes of parameter estimation through inverse modelling.

Keywords: transient groundwater model calibration, parameter estimation, PEST, MODFLOW

1 INTRODUCTION

The use of groundwater head differences rather than absolute head values is often advised to improve the representation of temporal groundwater level dynamics in transient groundwater models (Hill and Tiedeman, 2007; Doherty et al., 2010). Knowles et al. (2007) used weighted temporal differencing of head data in the inverse modelling of the Willunga Basin in Australia to simultaneously estimate conductivity, storativity and recharge. In Bakker et al. (2008) the dynamics of the groundwater system were represented as temporal moments which were subsequently used in groundwater model calibration.

A more straight-forward approach is to use the difference between two subsequent head observations as calibration target. The SMPDIFF utility in PEST (Doherty, 2009) for instance provides a tool for this approach. If this kind of data is used solely in a calibration, the model will probably reproduce the dynamics of the system fairly well, but the risk exists that the piezometric surface and thus the spatial gradients are not well captured by the model. This can be alleviated by including both the observed values of head and the temporal differencing in the objective function. The main drawback of this approach is that, especially if the model does not capture the temporal variations of the system very well, large observed variations in head will result in large residuals for both the absolute head value and for the difference in head. These extremes can therefore dominate the objective function and the calibration process.

To avoid this effect, the use of the sign of the temporal difference is suggested as additional calibration target together with the observed head values as it captures the essence of the temporal dynamics. The potential effect of including this kind of temporal information on parameter estimation is evaluated for an existing, previously calibrated, regional groundwater model in New South Wales, Australia, using the relative parameter error reduction statistic (Doherty and Hunt, 2009).

2 METHODS

2.1 Temporal Differencing

The sign of the difference in head in the ith observation well at time t, $dh_{i,t}$ is defined as

$$dh_{i,t} = \frac{h_{i,t} - h_{i,t-1}}{|(h_{i,t} - h_{i,t-1}|} \tag{1}$$

where $h_{i,t}$ is the head at time t and $h_{i,t-1}$ is the head at time $t-1$.

The objective function for a model with n observation wells and $k(i)$ the number of observations for observation well i, can than be written as

$$O = \sum_{i}^{n} \sum_{t}^{k(i)} w_{i,t} \left(h_{i,t} - \widehat{h_{i,t}} \right)^2 + \sum_{i} \sum_{t} v_{i,t} \left(dh_{i,t} - \widehat{dh_{i,t}} \right)^2 \tag{2}$$

where $w_{i,t}$ and $v_{i,t}$ are weights and the simulated equivalents to $h_{i,t}$ and $dh_{i,t}$ are $\widehat{h_{i,t}}$ and $\widehat{dh_{i,t}}$ respectively. The weights of the observations are usually chosen to be inversely proportional to their measurement error.

2.2 Parameter Identifiability and Relative Parameter Error Reduction

The calibration data set will determine to what extend parameters can be estimated through inverse modeling. Relative parameter error reduction is a statistic (Doherty and Hunt, 2009) which can inform the modeller before, during or after calibration how many and which parameters can be identified from a given calibration data set. As this statistic is based on linear theory and differentiability of model outputs to adjustable model parameters, its calculation is straightforward from the Jacobian, the finite difference approximation of the sensitivity matrix, and estimates of the prior uncertainty in parameters and observations used for calibration. A detailed description of the theory and calculation of this statistic can be found in Moore and Doherty (2005) and Doherty (2010b).

Parameter space can be subdivided into solution space and null space through singular value decomposition (Moore and Doherty, 2005). The solution space contains linear combinations of parameters which

affect model outcomes for which observations are present. The null space contains parameter combinations that, if superimposed on a parameter set that already calibrates the model, do not or minimally affect model outcomes corresponding to observations.

If X represents the actions of a model on its parameters p, then the observations h can be written as:

$$h = Xp + \epsilon \tag{3}$$

with ϵ representing measurement and structural noise. For non-linear models, X is approximated by the Jacobian. It has to be noted that the Jacobian is computed through perturbation of an optimal parameter set. The selection of the optimal parameter set will have an influence on the sensitivity analysis.

If there are more elements in p than can be uniquely estimated from h, singular value decomposition (SVD) can be used to divide parameter space into calibration space and null space according to:

$$X^t Q X = [U_1 U_2] \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} V_1^t \\ V_2^t \end{bmatrix} \tag{4}$$

with Q the observation weight matrix, U an orthogonal matrix spanning observation space, S a diagonal matrix containing the singular values in decreasing order and V an orthogonal matrix spanning the parameter space. Subscripts 1 and 2 denote solution and null space respectively. The identifiability of a parameter i can be defined as (Doherty and Hunt, 2009):

$$PI_i = \left(V_1 V_1^T \right)_{i,i} \tag{5}$$

where subscript (i, i) indicates the ith diagonal element of the matrix. This value can be thought of as the projection of parameter i on the calibration space. If its value is 1, it lies completely within the calibration space and can therefore be uniquely estimated. If the value is 0, the parameter is in the null space, which means that the observation data set contains no information to constrain parameter i.

The PI-statistic does not take into account the effect of noise in the measurement data on the estimation of parameters. Under the assumption that parameter and observation error are independent and that the observation weight matrix Q is inversely proportional to the error in the observations from both measurement error and structural model error, the post calibration error variance of parameter $i, \left[\sigma_2^2\right]_i$, can be written as (Moore and Doherty, 2005; Doherty, 2010a):

$$\left[\sigma_2^2\right]_i = i^t V_2 V_2^t C(p) V_2 V_2^t i + \sigma_r^2 i^t V_1 S_1^{-1} V_1^t i \tag{6}$$

where i is a unit vector with all elements except element i equal to zero, σ_r^2 the reference variance and $C(p)$ the variance-covariance matrix that describes the pre-calibration uncertainty of the parameters. The reference variance σ_r^2 can be estimated as the objective function divided by the number of observations minus the number of parameters. The pre-calibration error variance of parameter i, $\left[\sigma_1^2\right]_i$, is therefore

$$\left[\sigma_1^2\right]_i = i^t C(p) i \tag{7}$$

The relative parameter error reduction can then be defined as:

$$ER_i = 1 - \frac{\left[\sigma_2^2\right]_i}{\left[\sigma_1^2\right]_i} \tag{8}$$

which has a maximum value of 1.

One of the most important aspects of the analysis, is the selection of the truncation point, the number of eigenvalues to include in the solution space. Doherty (2010b) suggests to select the truncation point at that singular value for which inclusion of the singular value would lead to an increase in error variance associated with the complementary column of the V matrix, rather than a decrease in error variance.

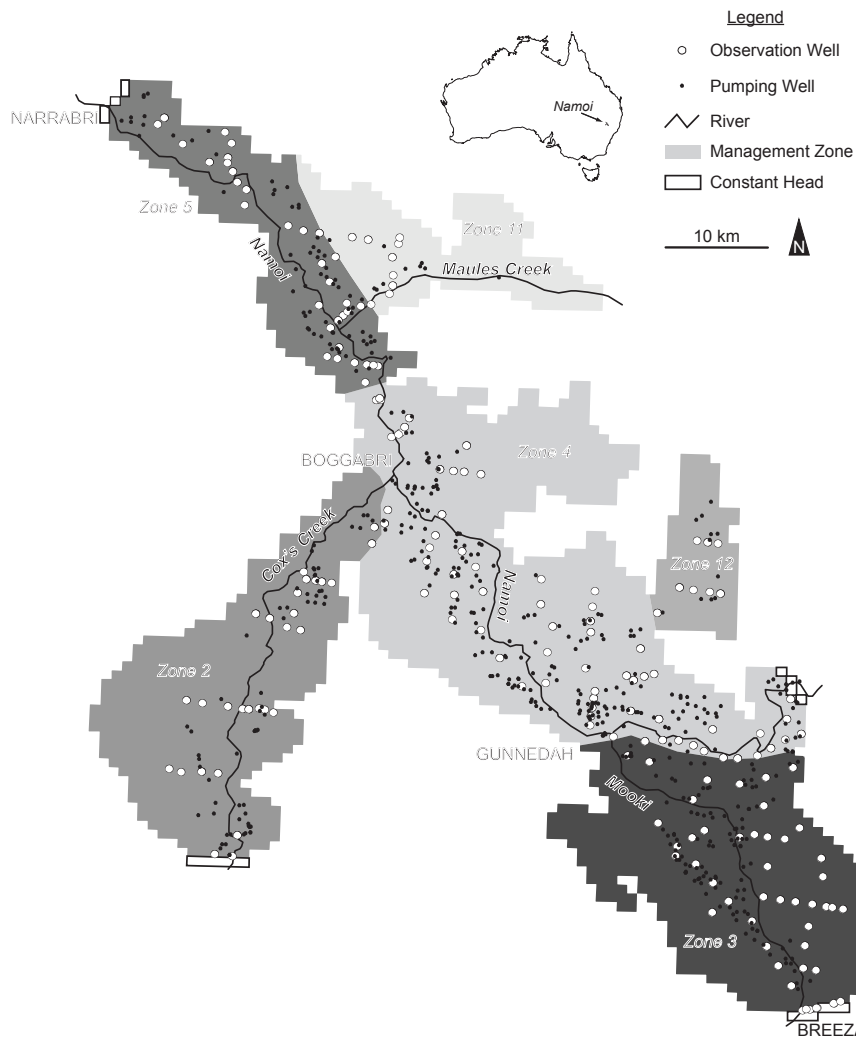## 3   CASE STUDY: UPPER NAMOI, NSW

Figure 1. Upper Namoi Groundwater Flow Model (after McNeilage, 2006)

The Upper Namoi Groundwater Flow Model (McNeilage, 2006) is a MODFLOW-model to simulate groundwater flow in the alluvial deposits of the Namoi River in New South Wales, Australia, upstream of Narrabri (Fig. 1) and was developed and is maintained by the New South Wales Office of Water. Average rainfall in the study area is $657\ mm$ and average annual pan evaporation is $1718\ mm$. Downstream Boggabri, the Namoi River gains water from the aquifer, while upstream, the river is considered losing. The Mooki River, Maules and Cox's Creek are considered to be dominantly losing streams, partially connected to the aquifer. The alluvial deposits of the Namoi River reach a maximum thickness of $170\ m$ and consist of moderately well sorted sands and gravels with inter-bedded clay in the Gunnedah Formation overlain by extensive over-bank clays with channel sands and gravels in the Narrabri Formation.

The transient, two-layer MODFLOW model has 82 columns and 106 rows with a resolution of $1000\ m$. The simulation period is from July 1985 to June 2001, divided in monthly stress periods. The lateral model boundaries are conceptualized as no-flow boundaries, while the major rivers are implemented as a head-dependent flux. 523 pumping wells (Fig. 1) are active in the model with a seasonally varying pumping rate. The same seasonal pattern is used for recharge rates from irrigation. Diffuse recharge is computed as unevaporated rainfall multiplied by a spatially varying recharge rate multiplier. Within the simulation period, two major floods occur, in July-August 1998 and October-November 2000. During these flood events, the recharge rate multiplier is increased in the Namoi floodplain. Interaction with upstream and downstream aquifers is implemented through constant head boundaries (Fig. 1). Based on

bore logs, zones of constant hydraulic properties are defined.

Table 1 lists all parameter groups considered in the calibration of the model. All parameters are log-transformed as the relationship between parameter and model output is more likely to become linear. The pre-calibration uncertainty in table 1 pertains to the log of the parameter. The off-diagonal terms of $C(\boldsymbol{p})$ are assumed to be zero, i.e. the parameters are assumed to be uncorrelated.

Table 1. Parameter groups used in calibration

Group	Description	n	range	$\sigma^2(\log)$
HK	Horizontal Conductivity	27	$0.05 - 20$ m/d	1
VK	Vertical Conductivity	27	$10^{-4} - 10^{-2}$ m/d	1
SS	Specific Storage	8	$10^{-7} - 10^{-3}$ 1/m	4
SY	Specific Yield	6	$10^{-3} - 5 \times 10^{-1}$	0.25
IRR	Irrigation	2	$30 - 72$ mm/yr	0.25
RCH	Recharge Rate Multiplier	27	$0.01 - 0.15$	1
FLOOD	Flood Recharge Rate Multiplier	24	$0.03 - 0.15$	1
KRIV	Riverbed Conductance	14	$10 - 500$ m/d	1

The calibration data set consists of 298 observation well screens with time series of groundwater head, indicated in Fig. 1. The measurement noise is assumed to be Gaussian with a standard deviation of 1.1 m, which leads to an observation weight of 0.826. This error estimate assumes an error related to the surveying of the reference level of the bore of 1 m and an error in measuring the actual water level in the bore of 0.1 m.

The observation weights associated with the sign of the temporal difference of the head observations are based on the 0.10 m estimate of the measurement error. A 1000 realisations of the observation dataset are created by adding random noise with a normal distribution with zero mean and standard deviation of 0.1. The sign of the temporal differencing is calculated and the observation weight is equal to the frequency for which the statistical simulation and the observation are the same.

The reference variance from the objective function is 16.29 m^2 for the dataset with only head measurements and is 17.66 m^2 for the head dataset with temporal differencing.

4 RESULTS AND DISCUSSION

The Jacobian, $\boldsymbol{X}$ is calculated using the PEST software (Doherty, 2010b) with a parameter increment of 1% for both cases. The singular value decomposition and subsequent calculation of parameter identifiability and relative parameter error reduction is carried out in Matlab. The truncation point between solution space and null space is determined by applying the method suggested in Doherty (2010a) and resulted in solution space with dimension 65 for the head only case and dimension 96 for the head and temporal differencing case.

The relative parameter error reduction for each parameter group is shown in Fig. 2. This figure clearly reveals the potential of using temporal differencing as additional calibration target. It reduces parameter error for almost all parameters. Most noteworthy however are the groups RCH, FLOOD and to a lesser extent SS. The original head dataset has no or very little information to constrain the recharge multiplier, while including temporal differencing provides at least limited information on this parameter group. The flood recharge multiplier shows a bimodal pattern. For the first twelve multipliers, corresponding to the 1998 flood, the improvement is marginal, while for the last twelve parameters, the 2000 flood, including temporal differencing reduces parameter error substantially. A possible explanation lies in the fact that, except for management zone 2, the flood of 1998 was more severe and the response more easily observed in hydrographs. The 2000 flood occurred during the pumping season and its signal in the hydrograph is therefore masked by other factors. The other two parameter groups that benefit the most from including temporal differencing are vertical conductivity and specific storage. The former controls leakage between upper and lower model layer and the former the storage in the lower layer. It is therefore not surprising that any additional information on the temporal dynamics will improve the estimate of these parameters.

The groundwater flow model has constant head boundary conditions at four lateral boundaries simulating interaction with neighbouring aquifers. To assess the effect of this conceptual model choice on the parameter estimation, the above described procedure is repeated, with inclusion of multipliers for each of the four constant head boundaries. A very low variance of 10^{-4} is associated with these parameters. Fig.
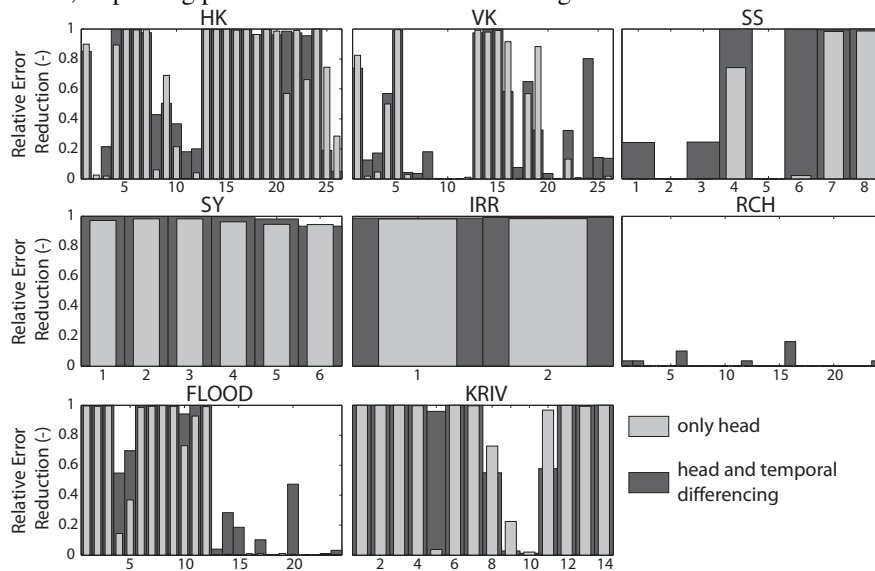
Figure 2. Relative parameter error reduction per parameter group for the head only case (blue) and the head and temporal differencing case (red)
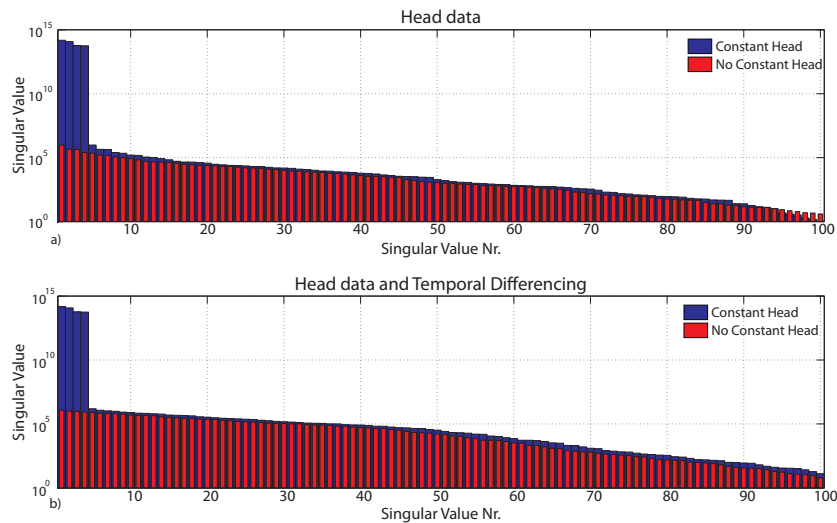

Figure 3. Singular values with and without inclusion of constant head boundary as parameter for head only (a) and head and temporal differencing (b)

3 shows the singular values for both cases with and without inclusion of the constant head boundaries in the sensitivity analysis. While the singular values computed without the constant head boundary decrease gradually, the singular values including constant head show a marked drop after the fourth singular value. Closer inspection of the V matrix shows that the first four singular values are dominated by the constant head boundaries. Fig. 4 shows the relative parameter error reduction when using only these four singular values and including the constant head boundaries as parameters. Although there are some slight differences between both cases, it is clear that the potential of parameter identification in the presence of constant head boundaries is very limited.

5   CONCLUSIONS

The relative parameter reduction error statistic is used to evaluate the added value of including the sign of temporal differences of head observations to improve parameter estimation for a transient, regional-scale groundwater model.

The results show that temporal differencing has a large potential to improve the estimation of parameter controlling the temporal dynamics of the system, such as recharge and flood multipliers, leakage and storage parameters.
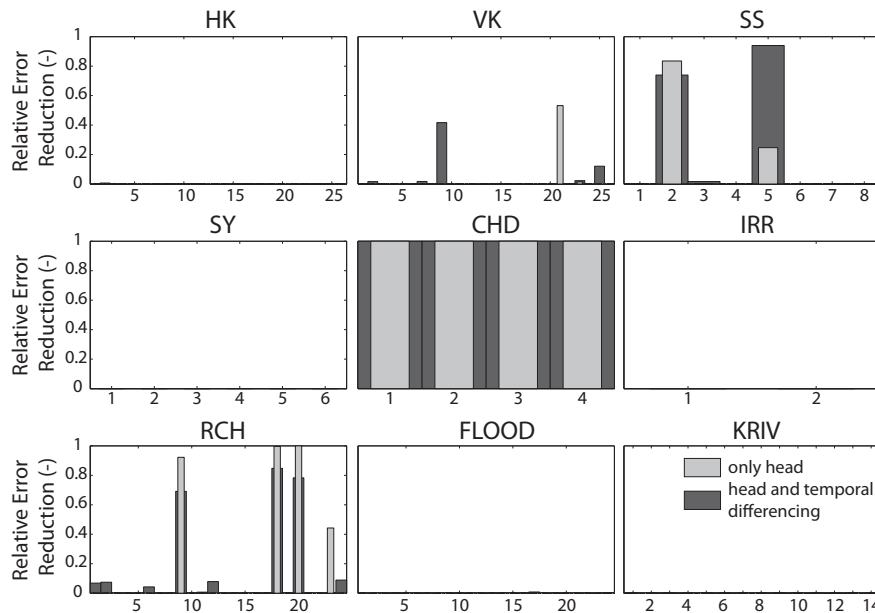
Figure 4. Relative parameter error reduction per parameter group for the head only case (blue) and the head and temporal differencing case (red) taking in account uncertainty in constant head boundaries

The results also show that conceptual model choices, such as constant head boundaries, may have a large effect on the model outcomes and can limit the identifiability of parameters through inverse modelling. In the current version of the model, which is under development at New South Wales Office of Water, the issue of the constant head boundaries is being addressed.

REFERENCES

Bakker, M., K. Maas, and J. R. Von Asmuth: 2008, 'Calibration of transient groundwater models using time series analysis and moment matching'. *Water Resour. Res.* **44**(4), W04420–.

Doherty, J.: 2009, 'Groundwater Data Utilities Part A: Overview'. Technical report, Watermark Numerical Computing.

Doherty, J.: 2010a, 'Addendum to the PEST manual'. Technical report, Watermark Numerical Computing.

Doherty, J.: 2010b, 'PEST: Model independent parameter estimation. User Manual 5th edition'. Technical report, Watermark Numerical Computing.

Doherty, J., M. Fienen, and R. Hunt: 2010, 'Approaches to Highly Parameterized Inversion: Pilot Point Theory, Guidelines, and Research Directions'. Technical report, USGS Scientific Investigations Report 2010-5168.

Doherty, J. and R. J. Hunt: 2009, 'Two statistics for evaluating parameter identifiability and error reduction'. *Journal of Hydrology* **366**(1-4), 119–127.

Hill, M. C. and C. R. Tiedeman: 2007, *Effective groundwater model calibration*. Wiley.

Knowles, I., M. Teubner, A. Yan, P. Rasser, and J. W. Lee: 2007, 'Inverse groundwater modelling in the Willunga Basin, South Australia'. *Hydrogeology Journal* **15**(6), 1107–1118.

McNeilage, C.: 2006, 'Upper Namoi groundwater model. Groundwater Management Area 004; Zones 2, 3, 4, 5, 11 and 12. Model development and calibration'. Technical report, NSW Department of Natural Resources Parramatta.

Moore, C. and J. Doherty: 2005, 'Role of the calibration process in reducing model predictive error'. *Water Resources Research* **41**(5).