

Article

Toward a Priori Evaluation of Relative Worth of Head and Conductivity Data as Functions of Data Densities in Inverse Groundwater Modeling

Nuan Sun [†], Kuai Fang  and Chaopeng Shen ^{*} 

Civil and Environmental Engineering, Pennsylvania State University, State College, PA 16801, USA; nuans1991@gmail.com (N.S.); kxf227@psu.edu (K.F.)

^{*} Correspondence: cshen@enr.psu.edu; Tel.: +1-814-863-5844

[†] Now at A. Morton Thomas and Associates, Inc., Rockville, MD 20850, USA.

Received: 27 March 2019; Accepted: 5 June 2019; Published: 8 June 2019



Abstract: Groundwater hydraulic head (H) measurements and point-estimates of hydraulic conductivity (K) both contain information about the K field. There is no simple, *a priori* estimate of the relative worth of H and K data. Thus, there is a gap in our conceptual understanding of the value of the K inversion procedure. Here, using synthetic calibration experiments, we quantified the worth of H and K data in terms of reducing calibrated K errors. We found that normalized K error e_K could be approximated by a polynomial function with first-order terms of H and K data densities μ_H and μ_K , which have been normalized by the correlation lengths of the K field, and a mutually inhibitive interaction term. This equation can be applied to obtain a rough estimate of the uncertainty prior to the inversion for a case with a similar configuration. The formulation suggests that the inversion is valuable even without K data. The relative worths of H and K depend heavily on existing data densities and heterogeneity. K can be ten times more informative when it is sparse. Noise perturbation experiments show that we should incorporate noisy K data when K is sparse, but not a large amount of low-quality K estimates. Our conclusions establish a crude, baseline estimate of the value of calibration. A similar assessment method for information content can be employed for more complex problems.

Keywords: modflow; PEST; uncertainty; information content

1. Introduction

The hydraulic conductivity (K) is a porous media property that is of great importance to various applications including integrated hydrologic modeling [1,2], contaminant fate and transport predictions, evaluation of groundwater resources [3], and analytical modeling [4]. On the one hand, K can be inferred from pumping tests [5] or lithological estimates. Typically, K values are not independently distributed but auto-correlated in space [6]. Therefore, one K measurement provides some information not only about the site, but also about the adjacent region. On the other hand, K can also be estimated inversely by calibrating a suitable groundwater flow model to an observed hydraulic head (H) using a prescribed recharge. There are mature inversion packages such as model-independent parameter estimation and uncertainty analysis (PEST) [7], which adjust K values, so that simulated H is close to observed data. Often, they make use of available (but sparse) known K data points and geostatistical models to constrain (or, in machine learning terminology, regularize) the inversion process. Therefore, both H and scattered K data help reduce the uncertainty, i.e., they both carry information content about the K field.

Despite a significant body of literature on the information content (or data worth) of K measurements, which we will summarize below, a conceptual knowledge gap exists regarding the relative value of H data and the K inversion process. Recently, integrated hydrologic modeling studies, e.g., [8–11], have relied on large-scale estimates of hydraulic conductivity, e.g., [12], for parameterizing K , but these estimates may not offer adequate resolution and accuracy for local groundwater modeling purposes. Integrated modeling studies often skip the step of K inversion. A part of the reason may be that, prior to conducting the calibration, there is no method to estimate how much uncertainty can be reduced by this process. On the other hand, it is common to find studies where the calibrated K fields are presented without associated uncertainty estimates. Given the amount of available data, how much uncertainty in the K field can be reduced by calibrating K to H ? What is the value of adding point-scale K measurements, which are inherently noisy, for improving the accuracy of calibration? What is the value of securing more K or H measurements? *A-priori* answers to these questions may help gauge the expected return of the calibration effort prior to carrying out the work.

Sifting through literature, we found that there are no simple answers to these inquiries. Previous studies on optimal experimental designs have provided valuable contributions to understanding the data worth of K measurements [13–15]. Their objectives were mostly to design the optimal way of obtaining new K data points for different purposes. For instance, James et al. [13] used a Bayesian data-worth framework to determine the location of data points that maximize their value in reducing the total remediation cost, where the pollutant plume was uncertain with regard to both location and extent. Neuman et al. [15] extended data-worth analysis by using the Bayesian model averaging approach to maximize the cost-benefit of K data points. Freeze et al. [16] also optimized the locations of measurements to minimize uncertainties with aquitard continuity and K distribution, while considering minimized economic regret. Tucciarelli et al. [17] employed a chance-constrained stochastic technique to find the best number and locations of additional measurements, which could result in a minimum total cost of data acquisition and groundwater reclamation. One similar chance-constrained model was coupled to an integer-programming sampling network design model in Wagner [14] in order to optimize pumping and sampling strategies. The worth of data can also be evaluated via sequential data assimilation methods such as ensemble Kalman filtering [18,19].

Despite the large body of literature, to the best of our knowledge, no work compared the value of H data in the calibration of K . In other words, the information content of H data was not given sufficient attention. However, the comparison between H and K values can be both challenging and nuanced due to the many strategies that were utilized to optimize spatial locations of data points, each with different objectives, constraints, and results. Therefore, we turn to a more modest question, “With a fixed, regularly spaced configuration of data locations, what is the value of H and K data, respectively?”

Centered around the abovementioned main question, this paper is organized to answer the following three sub-questions: (1) In a uniformly distributed data setting, can we describe calibrated K and H errors as functions of (preferably dimensionless) data densities, and are errors functions of recharge and boundary conditions? (2) How do the worths of K and H compare under different scenarios of H and K densities? (3) How do uncertainties with known K values impact our analyses? In the interest of reducing the dimensionality of the problem, we seek to non-dimensionalize data densities and errors with respect to the spatial heterogeneity of K . We will verify that the resulting numbers are truly non-dimensional characterizations of system features. Dimensionless numbers have been used in hydrological analysis with success. For example, see Haitjema and Mitchell-Bruker [20] for the water table ratio and Li et al. [21] for the six dimensionless numbers characterizing watershed hydrologic response types. These dimensional numbers, when verified, reveal underlying hydrologic dynamics and allow the conclusion to be migrated to different scales.

The advantage of the simple configuration is that the cause (data density and other factors) and effects (calibrated K error) can be clearly ascertained. While we acknowledge that such a configuration is rare, and the real-world is much more complex, our effort at least provides a baseline scenario that sheds insights and allows comparison with other strategies. As a result, it is a step forward in improving our understanding. Although they play important roles in groundwater modeling, the storage coefficient and transient modeling are outside of the scope of the present study. Furthermore, even though groundwater can be studied using Richards equation [22–26] as an integrated component of the hydrologic system [27], or the Boussinesq approximation [28], we focus on the saturated groundwater flow.

2. Methodology

2.1. Experimental Design

As discussed above, we used uniformly distributed H observations to make the study tractable. A prescribed fraction of the H data points were randomly sampled and assumed to have K estimates. Several realizations of such selections were prepared and tested in the calibrations. The collocation of H and K data points occurs because when a fine-scale K estimate is recorded, e.g., from pumping test, a water level reading can normally be obtained. We assumed a steady-state, 2D unconfined aquifer with uniform recharge.

The worth of data, in the context of this paper, is defined as the reduction of calibrated H and K error due to the inclusion of more data. As inverse modeling error is relative to the range of K variability, its absolute value has limited meaning, and is also not transferrable. Hence, we defined a dimensionless calibrated K error, e_K , as the normalized root-mean-squared logarithmic error (RMSLE) between log-transformed calibrated and synthetic K :

$$\begin{aligned} RMSLE &= \sqrt{\frac{\sum_{i=1}^N [\log(K_i^o) - \log(K_i^c)]^2}{N}} \\ e_K &= \frac{RMSLE}{\log(P_{90}) - \log(P_{10})} \end{aligned} \quad (1)$$

where K_i^o and K_i^c are the synthetic and calibrated hydraulic conductivity values at the i -th location in the domain, respectively. N is the total number of cells in the domain, and $\log(P_{90})$ and $\log(P_{10})$ are the 90th and 10th percentiles of synthetic conductivity values, with their differences characterizing the variability. Using these values instead of maximum and minimum values avoids skewing the error estimate by extreme values in the stochastically generated K fields. Other percentiles may also be considered. We will show that the normalization removes the dependence of error on the variability of the K field.

Similarly, the normalized error of H , e_H , was calculated for the collection of synthetic wells, as a normalized root mean squared error (RMSE).

$$\begin{aligned} RMSE &= \sqrt{\frac{\sum_{i=1}^n (H_i^o - H_i^c)^2}{n}} \\ e_H &= \frac{RMSE}{H_{max} - H_{min}} \end{aligned} \quad (2)$$

where H_i^o and H_i^c are the “observed” and simulated groundwater head at the i -th synthetic H observation locations, respectively, and H_{max} and H_{min} are the maximum and minimum groundwater head across the domain corresponding to the synthetic K field.

2.1.1. Synthetic Domain

Our computational domain is a rectangular unconfined aquifer. Six random K fields (Figure 1) distinguished by different spatial correlation lengths, λ , were generated with FIELDGEN, a supplemental utility for the model water flow and balance simulation model (WaSiM) [29]. We manually varied the mean, standard deviation, and correlation length when running FIELDGEN to generate the fields with different λ . The number of fields was chosen as a balance of the total work load and the representation of heterogeneity. The range of λ that we tested ensured that, at the largest λ , there are at least several clusters of high and low K values. Further increasing the range of λ will make it too labor intensive and computationally expensive to complete our numerical experiments. The K fields are all of log-normal distribution, with K values between 0.5 m day^{-1} to 200 m day^{-1} . It needs to be noted that the absolute value of this range is irrelevant if non-dimensionalization of the system is valid, which means the conclusions are only dependent on the dimensionless numbers. Then, we fitted a spherical model to the extracted empirical semi-variogram. The formula for the spherical model is:

$$\gamma(h) = \begin{cases} c \left[1.5 \frac{h}{\lambda} - 0.5 \left(\frac{h}{\lambda} \right)^3 \right], & \text{if } h \leq \lambda \\ c, & \text{if } h > \lambda \end{cases} \quad (3)$$

where $\gamma(h)$ is the empirical semi-variogram calculated from the generated field, h is the lag distance, c is the sill, and λ is the correlation length. We used a bound-constrained version of the “fminsearch” command from Matlab® [30] to find the c and λ that minimized the sum of squared differences between the theoretical model in Equation (3) and the empirical variogram. We cut off the empirical semi-variogram at 8000 m before fitting the variogram model.

The model domain spans 8000 m in both horizontal coordinate directions. The top and bottom elevations were 150 m and -100 m, respectively. Except for the ones that test the effects of recharge (Section 2.4), all experiments employed a uniform recharge of 500 mm/year. Again, if the dimensionless analysis is found to be valid, what is important for the groundwater flow system is the ratio of recharge to conductivity. By default, the eastern side was assigned a specified head (Dirichlet) boundary condition, with a water head of 130 m, while the three other sides were set as no-flow boundaries. However, the eastern and western boundary conditions were varied to test the effects of different recharge and boundary conditions.

2.1.2. Synthetic Observations of H and K

The calibration of a K field requires observed H values and optionally known K values [31,32]. The locations of synthetic H observations were evenly distributed throughout the domain, with an interval of 500 m leading to 256 virtual wells (Figure 2). While the data may look dense, if the correlation length is small, the data is relatively sparse. In terms of calibrated K error, as we will show later, the problem is dimensionless, in the sense that only the ratio between data density and correlation length matters. For a field with a small λ as in Figure 1a, the data is not dense after all. The water head at each observation well was extracted from a forward simulation with the aforementioned synthetic K fields, recharge, and boundary conditions. We randomly assigned known K values to $x\%$ of the observation wells.

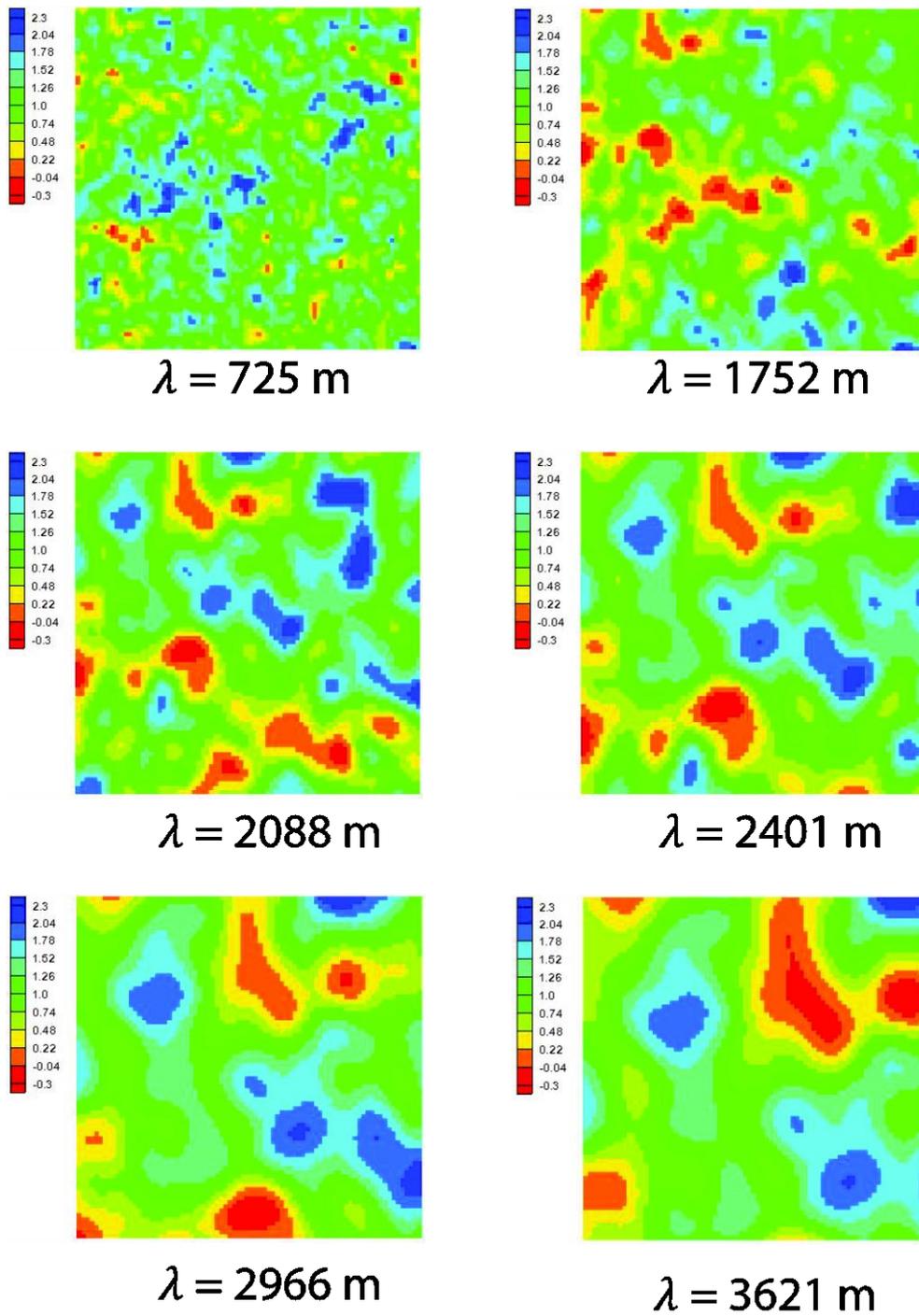


Figure 1. Random log(K) fields generated for our experiments. λ is the correlation length of the log(K) field. Smaller λ produces more heterogeneous fields and larger distances between data points.

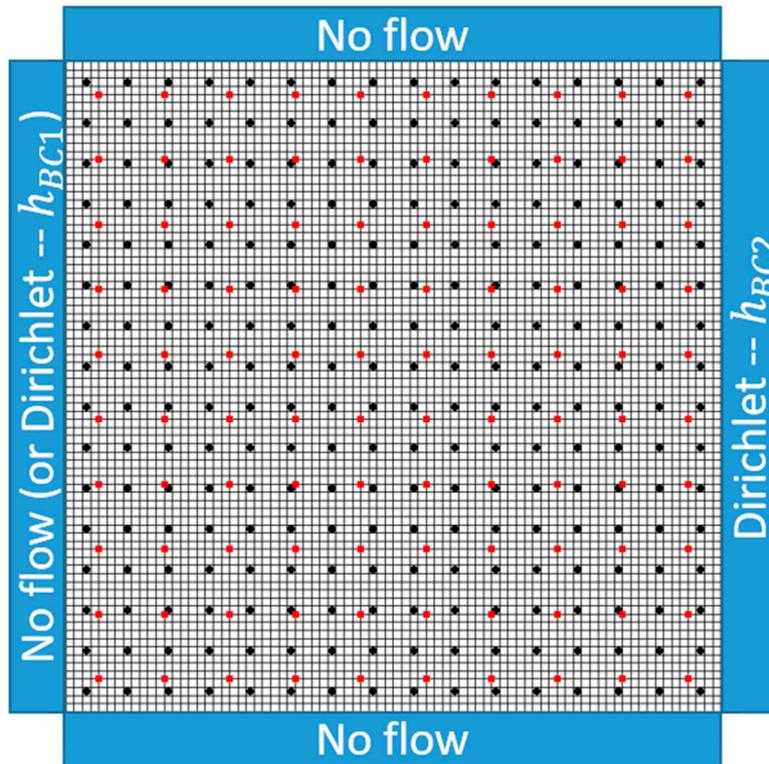


Figure 2. Locations of observation points (black) and pilot points (red). Boundaries of cells constituting the domain grid are shown with black grid lines. The elevation of the ground surface is uniformly set at 150 m. North and south sides of the domain are no-flow boundary conditions while the eastern side is Dirichlet with a value of h_{BC2} ; h_{BC2} is 130 m by default but varied during the tests for the effects of recharge and boundary conditions (Section 2.4). The Western side of the domain is no-flow by default but Dirichlet was also tested during in the experiments about boundary conditions.

2.2. Inverse Modeling

We used MODFLOW-2000 [33] as the groundwater flow model, and the model-independent parameter estimation and uncertainty analysis (PEST) [34] as the inverse modeling tool. PEST estimates K with the assistance of pilot points, each of which carries an initial K value that is used in calibrations and interpolated to all the grid cells in the domain. We set pilot points at the centroid of each 600×600 square box of the domain (see Figure 2). The inversion procedure in PEST adjusts K with an iterative approach to minimize the following objective function:

$$\begin{aligned} \Phi_m &= (\mathbf{H}^o - \mathbf{M}(\mathbf{K}))^T \mathbf{Q} (\mathbf{H}^o - \mathbf{M}(\mathbf{K})) \\ \Phi_g &= \Phi_m + \alpha \Phi_r \\ \mathbf{K}^c &= \underset{\mathbf{K}}{\operatorname{argmin}} (\Phi_g) \end{aligned} \tag{4}$$

where Φ_m is the unregularized objective function, Φ_r is a regularization term with α as a parameter, Φ_g is the global objective function, vector \mathbf{K} contains the conductivity values of the field, \mathbf{H}^o is the observation, and \mathbf{Q} is a weight matrix used to define greater contributions of certain pairs of observations. In the present simulation, a uniform weight is assigned to all K observations. \mathbf{M} is the model that predicts the system responses, given the parameter set \mathbf{K} , and \mathbf{K}^c is the calibrated K field that minimizes Φ_g .

Without $\alpha \Phi_r$, the inversion process can be non-unique, which means that different sets of parameter values may produce similar outcomes. K could be overfitted, which may lead to large errors when used in the predictive mode [31,34]. To reduce overfitting, $\alpha \Phi_r$ implements a penalization procedure called

Tikhonov Regularization [35], which introduces geospatial structure as a constraint in the calibration. ‘Regularization’ here is synonymous to ‘penalization.’ The regularization term is:

$$\Phi_r = (\mathbf{d} - \mathbf{R}(\mathbf{K}))^T \mathbf{Q}_r (\mathbf{d} - \mathbf{R}(\mathbf{K})) \quad (5)$$

where \mathbf{Q}_r is a diagonal matrix consisting of squared weights assigned to each observation, \mathbf{R} is a regularization operator that expresses a certain geostatistical constraint, e.g., the difference between a trial parameter value and the parameter value at a site, given neighbors’ values (and a variogram model), and \mathbf{d} is a ‘system-preferred’ state, which is 0 in this case.

In PEST, α is estimated together with calibration. With the help of a definition for an ‘acceptably good’ Φ_m value, Φ_m^1 , PEST found the α that minimizes Φ_r while satisfying that Φ_m is no greater than Φ_m^1 . This regularization penalizes parameter values that are far from the value expected of geostatistical models (built from data). The system can estimate more pilot points than there are observations, because each new pilot points is automatically accompanied by a spatially interpolated value. The algorithm attempts to ensure that the extra degrees of freedom, which carry little information, are discarded [36]. The interpolation method we selected was Kriging with a spherical model. The number of pilot points was set to 40% of H data points across all experiments. Considering the objective of the study, we did not test the fraction of pilot points as a control variable.

2.3. Non-Dimensionalization of Data Density and Errors

If we draw our conclusions as a function of dimensionless numbers, they are more broadly applicable than dimensional ones. The density of observation data needs to be examined relative to the spatial heterogeneity of the field, which is characterized by λ . Smaller λ indicates a more rapid varying K in space, which requires more observational data to constrain. To reduce the degrees of freedom, we propose a dimensionless number, the effective data density, μ_H , which quantifies the ratio between the correlation length and square root of data density:

$$\mu_H = \frac{\lambda}{d_H} = \frac{\lambda}{\sqrt{A/n}} \quad (6)$$

where A is the domain area, n is the number of wells where H is measured, and d_H measures the average distance between data points. μ_H can be interpreted as “the square root of the number of H data points in a square box with an area of λ^2 ”. A greater μ_H indicates a slower variation of K relative to the distance between measurement points, and thus, more information about the H field.

Similarly, we can quantify the relative density of wells with known conductivity values, which are implemented by randomly assigning synthetic K values to $x\%$ of observation wells. It gives rise to the average distance between known K values, d_K :

$$d_K = \frac{d_H}{\sqrt{x\%}} \quad (7)$$

where x is the fraction of H observation wells with known K values. Then, similar to Equation (6), the dimensionless factor to quantify conductivity data acquisition μ_K is derived as:

$$\mu_K = \frac{\lambda}{\sqrt{A/(n \times x\%)}} \quad (8)$$

The use of μ_H and μ_K allows us to greatly reduce the number of experiments and simplify the experimental design. If they can effectively and adequately characterize the system, we can avoid simultaneously varying the number of data points and spatial heterogeneity. Instead, we can adjust only the latter. To verify the validity of the dimensionless numbers, we compared the errors of calibrated K from a series of experiments, which contained different combinations of λ , n , and $x\%$

and produced similar dimensionless numbers. Five different μ_H values and five different μ_K were tested while keeping other variables constant. For each $x\%$ value other than zero, three sets of random locations were chosen for known K , leading to three separate calibrations. The results were then averaged to obtain the final calibration errors.

2.4. Recharge and Boundary Conditions

We examined the impacts of recharge and boundary conditions (BC) on calibration errors. Six recharge levels ranging between 100 mm/year and 1000 mm/year and two values of μ_H were tested. To prevent the recharge from raising the water table above the ground surface, the Dirichlet boundary was set to 115 m. Other model settings were identical to the default.

We also tested the impacts of the Dirichlet boundary conditions. For this test and this test only, we applied another Dirichlet boundary to the western domain boundary (Figure 2 right panel). We ran model calibrations with six μ_H values, while fixing μ_K at zero. We also ran five μ_K values while keeping μ_H at 4.18. We compared calibrated H and K errors from these different recharges and BCs.

2.5. Experimental Design and Multivariate Polynomial Curve Fitting

We tested a total of 31 pairs of (μ_H, μ_K) . For each pair, where μ_K is non-zero, we ran three random realizations of K fields. This experimental design resulted in a total of 79 calibration experiments. All experiments had the same domain geometries and locations of water head observations. Different μ_H was achieved by employing K fields with different λ (Figure 1), while keeping n constant.

We fitted error as a polynomial function (maximum second order) of μ_H and μ_K values:

$$e(\mu_H, \mu_K) = \mathbf{P} \times \vec{\mu} \quad (9)$$

where \mathbf{P} is a vector of coefficients in polynomial fitting, $\vec{\mu} = (1, \mu_H, \mu_H^2, \dots, \mu_K, \mu_K^2, \mu_H\mu_K)$ is the vector of predictors, and $e(\mu_H, \mu_K)$ is the calibration error. Our experiments were constrained within the range of μ values. \mathbf{P} was fitted using Matlab[®] curvefitting Toolbox. The term $\mu_H\mu_K$ in Equation (9) represents an interaction term. A probability value (p -value) was calculated for the null hypothesis that the coefficient is equation to zero based on t -tests for each of the curve fitting coefficients. Furthermore, the goodness of fit was evaluated using the coefficient of determination (R^2) and the root mean of squared error between the calibrated K/H errors and the values predicted by the polynomial function.

2.6. The Influence of Measurement Noise

The known K values that are supplied to the inversion algorithm can be estimated from pumping tests, lithology, specific capacity, and drawdown data, or conductivity test of samples [5]. In practice, regardless of which method is used, there will be errors. To study the impacts of K measurement noise on the calibrated K , we conducted perturbation experiments, where we added a synthetic noise to the observations. Since the K field is assumed to be log-normally distributed, we perturbed K values as:

$$\log(K) = \log(K^*) + \varepsilon \quad (10)$$

where K^* is the true conductivity value, K is what is supplied to the calibration algorithm, and $\varepsilon \sim N(0, \sigma_n)$ is a Gaussian noise with a standard deviation of σ_n . If σ_n is 0.2, it means 33% of the data points are perturbed to be either 50% larger or 37% smaller than K^* . If $\sigma_n = 1$, then 33% of the K data points are perturbed by more than an order of magnitude. The influence of the noise is quantified as an amplification factor:

$$\beta = \left(\frac{e_K}{e_K^*} \right) \quad (11)$$

where $e_K^* = e_K(\mu_H, \mu_K, \sigma_n = 0)$ is the average calibrated K error of the calibrations with noise-free K data, as defined in Equation (1), and e_K is the error for calibrations with added noise. The experiment

was very labor-intensive, as we tested M realizations of the noise and random selections of K positions. We could not regularly sample the three-dimensional parameter space of (μ_H, μ_K, σ_n) space as we did earlier. Instead, we explored a few lines in that space. In addition, since the readings of groundwater head (or depth to the water table) are generally more accurate compared to the K noise, we ignored H errors to reduce the dimensionality of the analysis. For comparing the results to the cases without any K data, we also defined β^0 :

$$\beta^0 = \frac{e_K(\mu_H, \mu_K = 0, \sigma_n = 0)}{e_K(\mu_H, \mu_K, \sigma_n = 0)} \quad (12)$$

β^0 is the ratio of errors between cases calibrated “without K data” and “with noise-free K data.” It can be compared with other β s because they have common denominators.

2.7. Relative Data Worth

We can calculate the reduction of K error with respect to one measurement point: $\frac{\Delta e_K}{\Delta n_H}$ and $\frac{\Delta e_K}{\Delta n_K}$. These values can be approximated numerically by calculating the increments in e_K or e_H (given by Equation (9)), as the number of data points increase. They can also be derived from the fitted polynomials in Equation (9). Then, we can examine two ratios that measure the relative worth of data:

$$\begin{aligned} R_\mu &= \frac{\partial e_K}{\partial \mu_H} : \frac{\partial e_K}{\partial \mu_K} \\ R_n &= \frac{\Delta e_K}{\Delta n_H} : \frac{\Delta e_K}{\Delta n_K} \end{aligned} \quad (13)$$

R_μ is the relative data worth ratio with respect to a unit increase in μ_H or μ_K and is only a function of these two factors. Since μ is a nonlinear function of n , far more data points are needed to increase a unit of μ when μ is high. This relationship can make R_μ difficult to interpret and use. R_n indicates whether it is more beneficial to add an H or a K data point, which has a direct practical meaning. However, R_n has three control variables: n_H , n_K and λ . We will show the influence of n_H , n_K under two different λ values.

We considered the following two scenarios when calculating R_n : (A) A new K data point is always accompanied by a new H data point, which is relevant when we plan to install new wells. In this scenario, $\frac{\Delta \mu_H}{\Delta n_K} = \frac{\Delta \mu_H}{\Delta n_H} = \frac{\lambda}{2\sqrt{\lambda n_H}}$; (B) we can add a K measurement without adding H data. This scenario is relevant when we can conduct a pumping test from an existing well, or extract K estimates from interpreting existing literature. Under this scenario, $\frac{\Delta \mu_H}{\Delta n_K}$ is 0. We denoted R'_n as the data worth ratio calculated under scenario (B).

3. Results and Discussion

3.1. Verification of the Effectiveness of Non-Dimensionalization

Given similar μ_H and μ_K values, the normalized conductivity error e_K is tightly clustered (Figure 3), although λ , n , and $x\%$ varied substantially (Table A1). This behavior verified that μ_H and μ_K are effective dimensionless numbers to characterize e_K , which allowed us in only altering λ in later experiments. In addition, Figure 3 suggested e_K can be described as a smooth function of μ_H and μ_K . In these preliminary experiments, we did not observe any non-monotonicity or fluctuations. Such smoothness and monotonicity serve as the basis of fitting a polynomial function to the relationships between normalized error and data densities.

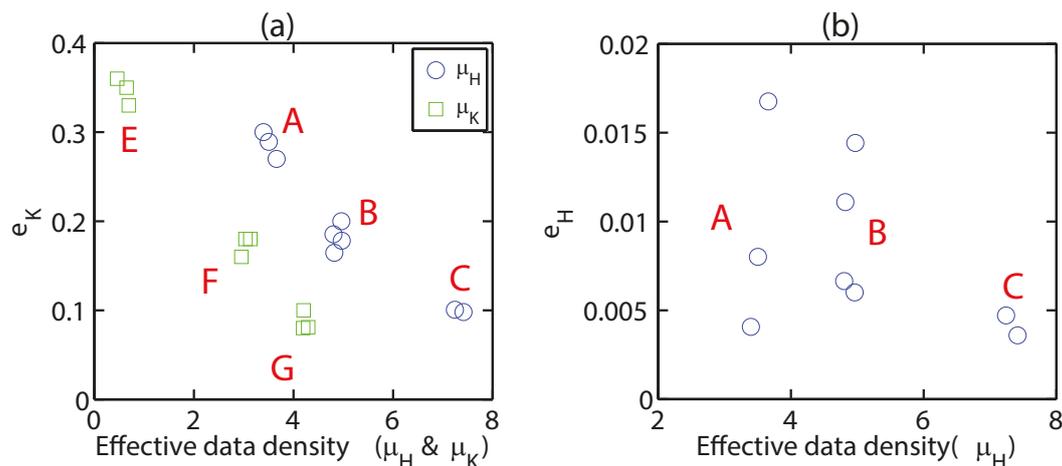


Figure 3. Verifications of normalized data densities as effective dimensionless numbers for e_K (a) and e_H (b). A–G annotate clusters. When we tested the effectiveness of μ_H (blue circles), μ_K was kept constant. Similarly, when we tested μ_K (green squares), μ_H was kept constant. In the three blue circle clusters, water head is the only synthetic observational data in the model. Thus, their μ_K are zero. From the clustering pattern in the left panel, it is clear that e_K is similar for similar (μ_H, μ_K) pairs, even though the correlation lengths and the number of data points are different. This figure suggests (μ_H, μ_K) are effective dimensionless parameters controlling e_K . However, it is not the case for e_H . The experimental configurations and results for the clusters on these figures are provided in Table A1.

However, we did not obtain a clustering pattern for water head error. The e_H values showed obvious scattering even with similar μ_H and μ_K . Therefore, it is meaningless to further test e_H . In summary, μ_H and μ_K are effective dimensionless numbers to characterize the system for e_K but not for e_H . As a result, the conclusions to be drawn later for e_K as a function of μ_H and μ_K were applicable to different λ , $x\%$, n combinations, while those for e_H were only valid for μ_H and μ_K values that we specified.

3.2. Impact of Recharge and Boundary Condition on Model Calibration Errors

Normalized errors are independent of recharge (Figure 4a,b). Various recharge inputs lead to different water head ranges and, consequently, different absolute error values. However, after the error is normalized on the ranges of H and K , they become flat and non-responsive to recharge. In summary, these experiments show that the influence of recharge is linear, and can be removed by normalizing the error with respect to the range of values in the domain. Therefore, we no longer considered recharge in our factorial experiments.

Comparing $e_K \sim \mu_H$ and $e_K \sim \mu_K$ curves obtained with two different boundary conditions, we noted that the boundary condition had little impact and the two sets of curves almost overlapped (Figure 4c). The two boundary conditions tested were Dirichlet and Neumann, which approximate lakes, rivers (Dirichlet), impervious mountain blocks (Neumann), and so on. However, the same cannot be said about e_H : There are gaps between Dirichlet and Neumann BCs and the gaps are not constant (Figure 4d). The addition of Dirichlet boundary increases the water head error e_H and makes calibration results more stochastic. These patterns mean that the conclusions to be drawn later in this paper with regard to e_K , but not e_H , can be generalized for many different situations, with little impact from the environmental settings. As we will primarily focus on e_K , we will remove BCs from further consideration.

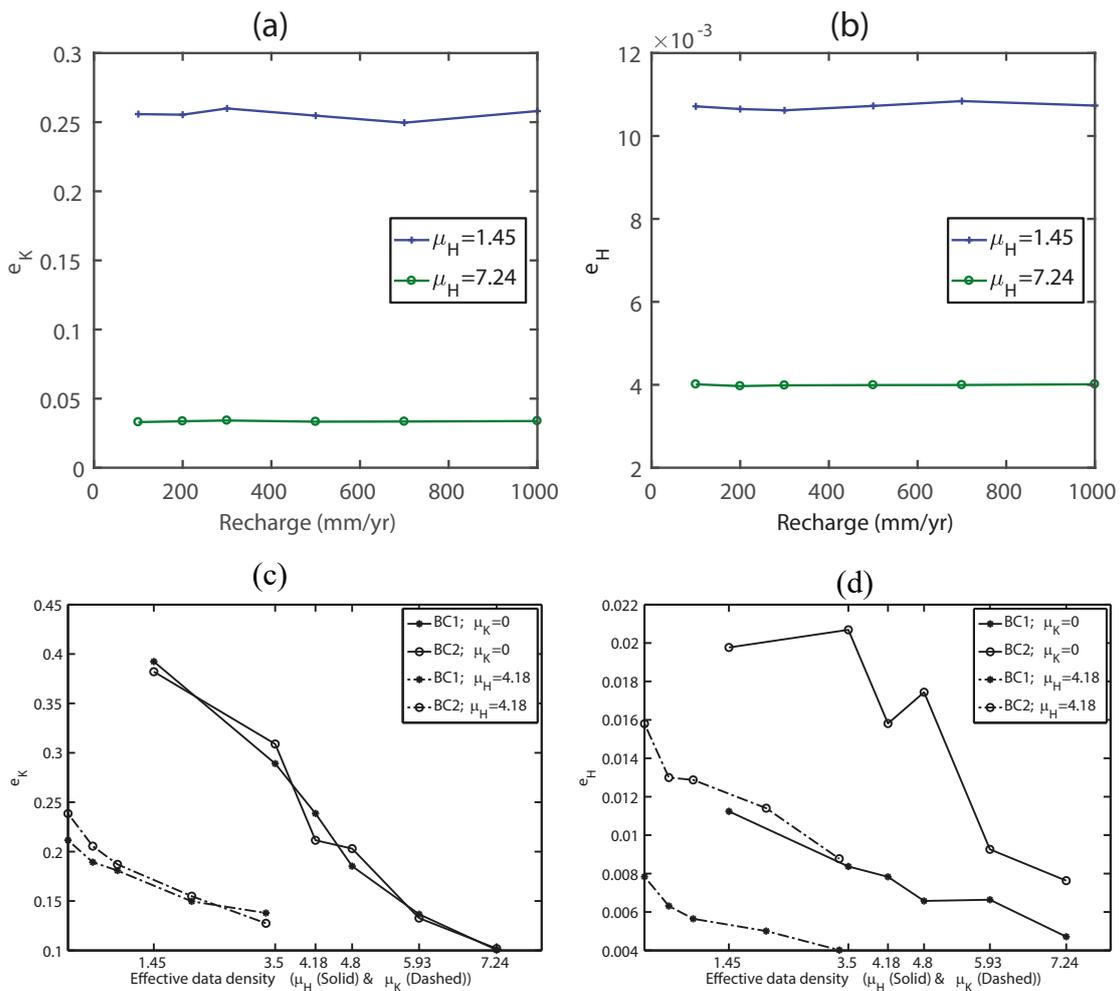


Figure 4. Rescaled water head error (a) and conductivity error (b) as functions of recharge under two different H data densities. The errors are obviously impacted by μ_H , but at each μ_H level, recharge does not influence e_K or e_H . This pattern allows us to remove recharge as a control variable from our experiments. (c,d) e_K (c) and e_H (d) as functions of normalized data densities. BC1 means domain with one Dirichlet boundary and BC2 stands for the domain with two Dirichlet boundaries. Since different boundary conditions generate the same curves, it indicates our analysis of e_K can be valid for different boundary conditions.

3.3. Errors as a Function of Normalized Data Densities

When we hold μ_H constant and increase μ_K , e_K gradually decreases as one expects (Figure 5a). The decline in e_K is almost linear. The slopes of the equi- μ_H lines decrease slightly for higher μ_H values, and the gaps between the lines become smaller at higher μ_K , indicating a moderate interaction between μ_K and μ_H . The $e_K \sim \mu_K$ curves become flatter when μ_H is higher, suggesting that when μ_H is higher, the marginal gain attained by the addition of μ_K decreases. e_H shows a generally similar trend, but there is a more noticeable quadratic trend (Figure 5b).

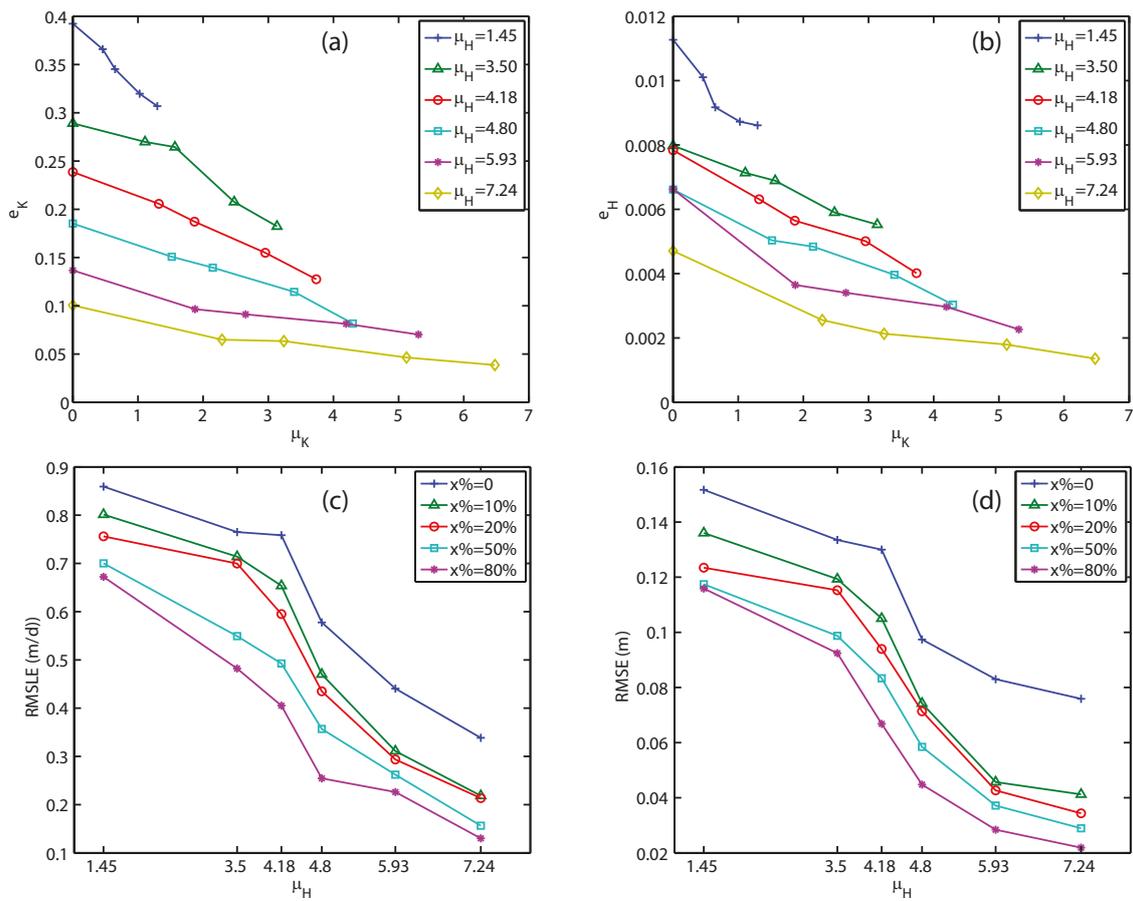


Figure 5. Normalized K error (a) and H error (b), and dimensionalized K error (c) and H error (d), as functions of normalized data densities. Each point on the plot represents the mean of three calibrations, except when $x\% = 0$. It is obvious that e_K decreases as μ_K or μ_H increases and $e_K \sim \mu_K$ at each μ_H level is almost linear.

Toward higher μ_K , we should be able to build a more accurate variogram model for the interpolation procedure during calibration. However, at least in the tested range between μ_H and μ_K , such a benefit is hardly observable. At the same time, as e_K is computed from comparing “observed” and calibrated K of the entire domain, the monotonously and smoothly varying e_K suggests the regularization approach is effective in reducing overfitting errors.

Viewing the data in a different way, when we keep a constant $x\%$, the error apparently decreases smoothly, as we increase the effective data densities (Figure 5c,d). In this Figure, as μ_H increases, μ_K also increases proportionally. When $x\%$ is increased from 0 m to 10%, the reduction of both RMSE and RMSLE are more significant than when it increased from 10% to 20%. As mentioned previously, this pattern is perhaps due to the moderate interaction between μ_H and μ_K .

3.4. Multivariate Polynomial Curve Fitting

Stepwise multiple regressions show that high-order terms (i.e., the second order) are statistically insignificant for e_K so that the system is mostly linear in the range of tested μ_K and μ_H (Table 1). The small value of the coefficient for (μ_H, μ_K) compared to the other terms confirms that the interaction between the two variables is mild. We created a 3D surface using three terms: μ_K , μ_H , and $\mu_K\mu_H$. Therefore, the final fitted equation can be written as:

$$e(\mu_H, \mu_K) = p_0 + p_1\mu_H + p_2\mu_K + p_5\mu_H\mu_K \tag{14}$$

where p_0 to p_5 are fitting parameters contained in \mathbf{P} of Equation (9). Meanwhile, the quadratic term, μ_K^2 , is statistically significant (p -value = 0.001) for e_H . However, the R^2 without the quadratic term is adequately high, and adding the term does not increase it notably. In the interest of parsimony, we chose not to include μ_K^2 in the fitted formula for e_H either. As there are no effective dimensionless numbers that characterize e_H , we focused on e_K .

Table 1. Multivariate polynomial curve fitting for the following equation: $e(\mu_H, \mu_K) = p_0 + p_1\mu_H + p_2\mu_K + p_3\mu_H^2 + p_4\mu_K^2 + p_5\mu_H\mu_K$. p -value is the probability of the null hypothesis that the corresponding coefficient is equal to 0 according to the t -statistic.

Calibration Errors		Coefficients in Fitting Equation						Fitting Goodness	
		p_0	p_1	p_2	p_3	p_4	p_5	R^2	e_f
e_K	p -Value	0	0	0	0.946	0.697	0	0.975	0.0162
	Value	0.464	-0.0533	-0.061	0	0	0.0075		
e_H	p -Value	0	0	0	0.939	0.001	0	0.973	0.00043
	Value	0.012	0.0011	-0.00156	0	0	0.00015		

Note: The p -value of p_4 coefficient for e_H is small but still set to be 0. The two values inside the parentheses are the values including the quadratic term μ_K^2 when fitting e_H .

For the first-order terms, the difference between p_1 and p_2 is ~15%. Therefore, in regions where the first-order terms dominate and μ_H and μ_K are similar, the data worth of new H and K data points are comparable. This finding suggests, at least across the range tested, groundwater head observations have great value in reducing uncertainties in the K field, and we do not necessarily require knowledge of K values to reduce e_K . Also, since e_K is normalized to the range of variability of K , we cannot transfer e_K to dimensionalized uncertainty estimates without some knowledge about the K field. Meanwhile, since p_1 and p_2 are both negative and p_5 is positive, the interaction terms exert a mutually inhibitive effect: The existence of each type of data reduces the marginal benefit of the other type of data. For example, when H observations are dense, because $\frac{\partial e_K}{\partial \mu_K} \approx -0.061 + 0.0075 \mu_H$, additional K observations do not help as much as when H data density is lower. This almost linear relationship with mutual inhibition is a novel finding. Although we might have an intuitive expectation of an inhibitive relationship, this study is the first to quantitatively determine its relative magnitude.

The fitted surface well describes the errors as a function of μ_H and μ_K (Figure 6), as most points are scattered closely to the surface. As we assumed μ_K cannot be greater than μ_H , the valid region is limited to the left lower triangle of the $\mu_H \sim \mu_K$ plane. We also provided a contour plot for the surface for a more numerically accurate representation (Figure 7). The contour patterns between e_H and e_K are similar. The contours are denser near the lower left corner, because the gradients of error with respect to μ are larger when μ is small as a result of smaller mutual inhibition effects.

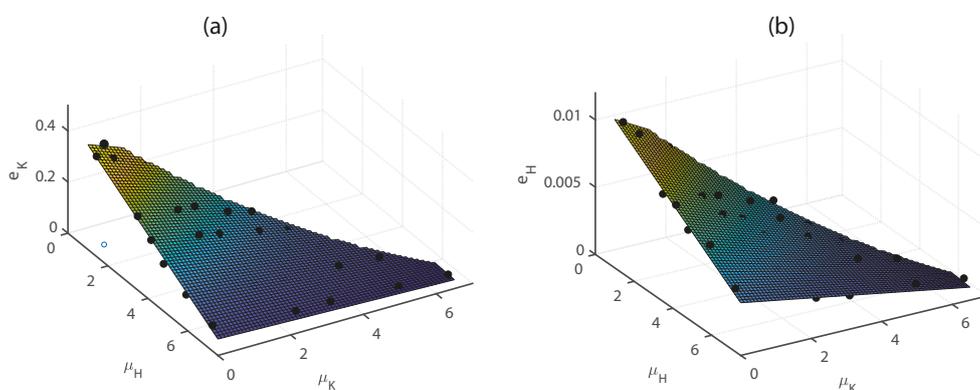


Figure 6. 3D visualization of adjusted multiple polynomial curve fitting of e_K (a) and e_H (b) as a function of normalized data densities. The data fall close to the surface. Some data points fall below the surface and are not visible at the shown angles.

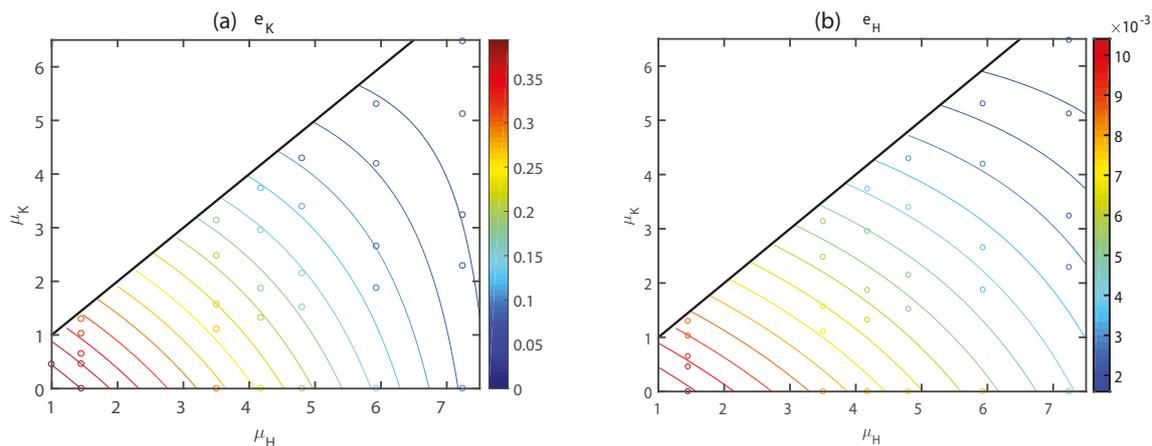


Figure 7. Contour representation of the surface fitted to e_K (a) and e_H (b) as functions of data densities. Scattered points indicate the data points used to construct the contours.

3.5. Relative Data Value

With either numerical approximation or analytical derivation, we can estimate the ratio between marginal error reduction rates to provide us some insights. For example, R_μ becomes $R_\mu \approx \frac{p_1 + p_5 \mu_K}{p_2 + p_5 \mu_H}$, which is clearly a function of both μ_K and μ_H . The μ_K tested scatter mostly within the range between 0 to 4.5 (Figure 5). Given a unit increase in the effective data density, H data appears to be more effective in reducing e_K (Figure 8). R_μ is >1 for the greater part of the plane, especially for K error. It rises quickly toward the high μ_H , low μ_K region near the upper left corner of the figure. This pattern results from the calculation of R_μ : μ is not a linear function of n . To increase one unit in μ when μ is high, many data points are required.

R_n is easier to interpret as it shows the ratio of information content brought in by the next data point of H vs. K , but to examine it, we must consider the nonlinear influence of λ . When λ is small (highly heterogeneous field, Figure 8c,d), R_n ranges between 0.1 and 0.8, which means a K data point will always bring in more information content than an H data point. This difference may be counter-intuitive, considering the magnitude of coefficients p_1 (-0.053) is only slightly smaller than p_2 (-0.061). An important factor is that n_K is always smaller than n_H in our tested ranges, which is normally the case with available groundwater data. R_n contours radiate out almost linearly in the shape of a fan and is dense near the left edge of the figure. The linear pattern of the contours in Figure 8a,b suggest R_n is almost a function of $x\%$ for this high-heterogeneity case ($\lambda = 725$ m). When $n_K < 10$, $R_n < 0.1$, meaning K measurement is sparse compared to H , there is 10 times more information value in new K data points than H . However, when λ is large (more homogeneous field), we are in a relatively data-rich environment, where the mutual inhibition effect becomes more important. R_n becomes markedly larger and more nonlinear (Figure 8e,f). Toward the left-edge, contours are dense and still mostly vertical. In that region, R_n mainly depends on the amount of K data, and H density has little impact. R_n then increases toward the upper-right corner. If other conditions are equal, new K points bring in more relative value in the small- λ case than in the large- λ case. As the fraction of K data points increases, new H data points become increasingly useful relative to K .

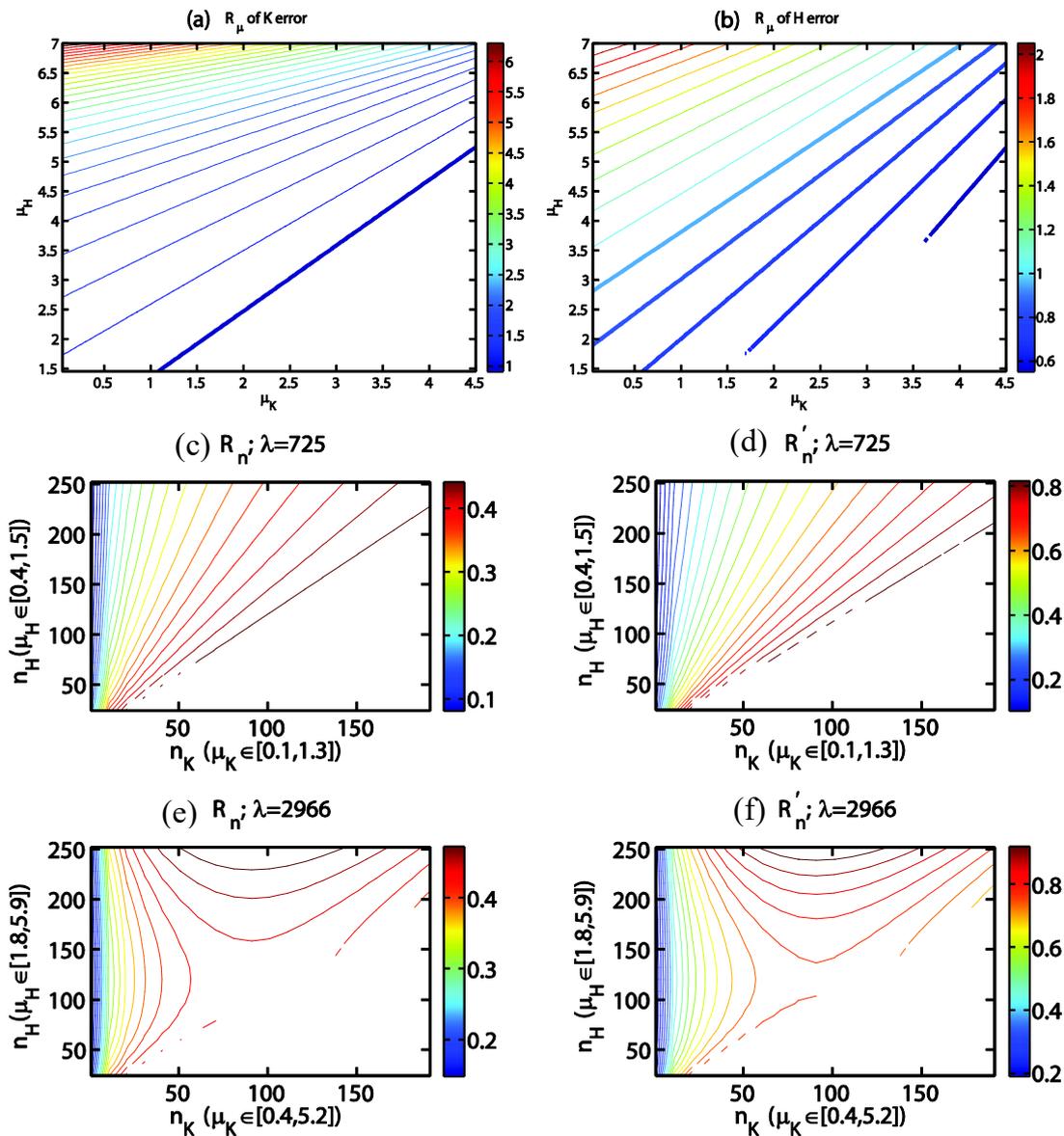


Figure 8. R_μ of K error (a) and H error (b) as functions of normalized data densities μ_H and μ_K . Bold lines highlight the less-than-one values among the contours. (c–f) R_n and R'_n as functions of n_H , n_K , and λ . R_n is calculated assuming each new K data point entails a new H data point. R'_n is calculated assuming new K and H data points are independent of each other.

3.6. The Influence of Noise with K Observations

Although we looked for predictive formulations to describe the relationship between β and its three control variables (μ_K , μ_H , σ_n), we have not been able to find a simple and reliable predictive formulation with $R^2 > 0.5$. However, we can draw some inferences from our results.

1. When μ_H is fixed, β , in general, grows as a function of increasing μ_K (Figure 9), but when $\mu_K < 3$, the error amplification is almost close to 1 and β is not very sensitive. The largest impact in this category is with the case ($\mu_K = 1.45, \mu_H = 1.3, \sigma_n = 1$). Recall in this case, 33% of the K data points have been perturbed more than an order of magnitude, but the impact on calibrated K nonetheless appears limited ($\beta = 1.29$). However, when μ_K is larger than 3, the errors grow significantly. At ($\mu_K = 6.48, \mu_H = 7.24, \sigma_n = 0.5$), $\beta = 11.6$ (the upper-rightmost point in Figure 9), which means the inversion essentially failed. Another such case is ($\mu_K = 5.12, \mu_H = 7.24, \sigma_n = 1$) and $\beta = 4.57$ (visible in Figure 9 blue line). At ($\mu_K = 6.48, \mu_H = 7.24, \sigma_n = 0.2$), even if the perturbation is moderate, it still causes a significant error amplification ($\beta = 1.47$, the rightmost circle on the lower solid green line).
2. Larger μ_H can help inhibit error amplification. When μ_K and σ_n are kept the same and μ_H is increased, β always decreases. However, this effect is small when $\mu_K < 3.5$ because the amplification is already small.
3. Even though larger μ_K increases the error amplification from a noise-free baseline (Figure 9), incorporating K data points nonetheless reduces error compared to corresponding cases with $\mu_K = 0$ (note β_0 in Figure 9), as long as μ_K is not too large. For example, under a sparse-data scenario ($\mu_K = 0.45$ or $\mu_K = 1.3$), even if $\sigma_n = 1$, which means significant noise, the error is still less than the case without incorporating K data.
4. Under the combined conditions of high $\mu_K (> 3.5)$ (again, this means there are on average $3.5^2 = 12.25$ conductivity data points in an area of λ^2) and high noise ($\sigma_n \geq 0.5$), the error amplification skyrockets and dominates over the information content of K . For example, ($\mu_K = 5.1, \mu_H = 7.2, \sigma_n = 1$), $\beta = 4.57$, and $e_K = 0.21$ which is greater than the case with the same μ_H but without K data ($\mu_K = 0$). At $\mu_K = 6.5$ and $\mu_H = 7.2$, even a σ_n of 0.5 is sufficient to bring the error amplification factor to 11.6. For these cases, the errors are too large, and the inversion failed.

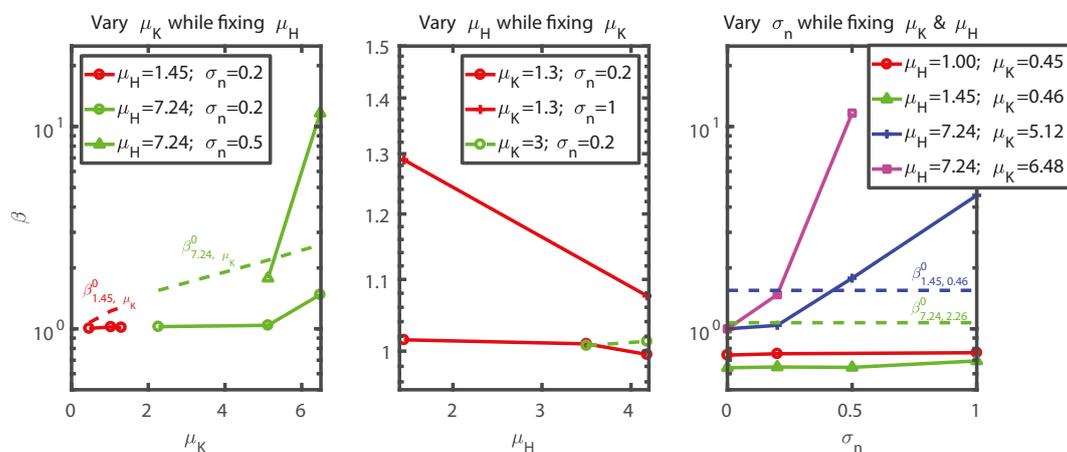


Figure 9. Error amplification factor β , the ratio of errors between “with noisy K data,” and “with noise-free data” as functions of μ_K , μ_H , and σ_n . As explained before, since this experiment is very expensive, we could only afford to explore a few lines. For comparison, dashed lines indicate β^0 .

Based on these observations, we conclude that, perhaps counter-intuitively, noise with K estimates are more malicious under high μ_K scenarios. When there are only a few K data points, we should incorporate them, even if we know they may have significant noise, provided that the log-standard deviation of that noise is not more than 1 order of magnitude. However, when there are a great number of low-quality K estimates, it is, in fact, better not to use them during the inversion and completely rely on H data. It is possible that the low-quality K has made it very difficult for the inversion process to infer a usable variogram.

4. Conclusions

The main contribution of the paper is to expose the functional form of inverse modeling error as dependent on data densities and the interaction between the density terms for the simplest case. The simple dependence of calibrated K error (but not H error) on the normalized densities have not been shown before. The conclusions of this study, i.e., Equation (14) with coefficients in Table 1, can be applied a priori to roughly estimate the value of groundwater model calibration. For consultants reading others' works, which resulted from a calibration but without details concerning uncertainty, the formula here can be helpful. We found that the calibrated K error, e_K , can be well described by the sum of a linear function of μ_H and μ_K , and a mild, mutually inhibitive interaction term, which indicates if H density is high, the marginal value of K is reduced and vice versa. This functional form fills a knowledge gap about the value of data and the inversion procedure itself. As the formula is derived in dimensionless forms, it can be applied to various scales and heterogeneity settings. However, the calibrated H error cannot be similarly described by dimensionless numbers. BCs and recharge are found to have little impact on normalized K error. The absolute value of the coefficients of the first-order terms are similar, but relative data worth ratio ($H:K$) for the next data point, R_n , is strongly dependent on the existing data densities. If we assume that a new K observation must entail a new H observation, across most of the tested parameter range, a new H measurement has less than 40% of the data worth of a new K measurement ($R_n < 0.4$). When K is sparse, this ratio can be less than 10% ($R_n < 0.1$). In a domain with higher heterogeneity, R_n is mostly determined by the fraction of wells with K measurements.

Considering the noise inherent with known K estimates, we should incorporate K data when there are relatively few K data points, or if the data quality is high. Especially, some knowledge about the range of K variability is required to convert the normalized error estimates to ones with units. However, if there are a great number of K data points with low quality, it is, in fact, better not to incorporate them because they make it difficult to estimate a valid variogram.

5. Limitations

The findings help to build a first-order, conceptual understanding, but we must realize real-world situations are far more complex. Certainly, the method employed in this work is simple and empirical. The geometric configuration of the domain and measurement points is simplistic. Therefore, it only represents a rough, a priori estimate, and it needs to be used with caution. More advanced methods should be used to determine the optimum location to place new data points. Finally, our experiments were carried out using the PEST algorithm in the environment of groundwater modeling system software, and the results are, thus, conditioned on some decisions of the program, e.g., the choice of the regularization parameter.

Author Contributions: Conceptualization, C.S.; methodology, K.F.; software, N.S.; validation, N.S. and C.S.; writing—original draft preparation, N.S.; writing—review and editing, C.S.; visualization, N.S.; supervision, C.S.; project administration, C.S.; funding acquisition, C.S.

Funding: This work was support by the U.S Bureau of Land Management under Interagency Agreement L11PG00354, as part of Work for Others funding from Lawrence Berkeley National Lab, provided by the U.S. Department of Energy, Office of Science, under Award Number DE-AC02-05CH11231.

Acknowledgments: Many thanks to the anonymous reviewers whose comments have helped to improve this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Experimental configurations and errors for the verifications of dimensionless numbers are shown in Figure 3.

Cluster	μ_H	μ_K	λ	n_H	n_K	e_H	e_K
A	3.4	0	725	1444	0	0.0041	0.3
	3.5	0	1752	256	0	0.008	0.29
	3.66	0	2088	196	0	0.017	0.27
B	4.8	0	2088	361	0	0.006	0.19
	5	0	2401	256	0	0.007	0.2
	4.8	0	2966	169	0	0.011	0.17
	5	0	3621	121	0	0.014	0.08
C	7.2	0	3621	256	0	0.101	0.005
	7.4	0	2966	169	0	0.980	0.004
E	1.5	0.45	725	256	25	N/A	0.36
	1.4	0.64	725	256	49	N/A	0.35
	1.5	0.69	1751	49	16	N/A	0.33
F	3.1	2.95	1751	256	196	N/A	0.18
	3.4	3.13	2088	169	144	N/A	0.18
	4.3	3.04	2088	256	121	N/A	0.16
G	4.8	4.3	2401	256	196	N/A	0.081
	5.9	4.194	2966	256	121	N/A	0.08
	4.7	4.2	2088	324	256	N/A	0.1

References

- Maxwell, R.M.; Putti, M.; Meyerhoff, S.; Delfs, J.-O.; Ferguson, I.M.; Ivanov, V.; Kim, J.; Kolditz, O.; Kollet, S.J.; Kumar, M.; et al. Surface-subsurface model intercomparison: A first set of benchmark results to diagnose integrated hydrology and feedbacks. *Water Resour. Res.* **2014**, *50*, 1531–1549. [[CrossRef](#)]
- Shen, C.; Niu, J.; Phanikumar, M.S. Evaluating controls on coupled hydrologic and vegetation dynamics in a humid continental climate watershed using a subsurface-land surface processes model. *Water Resour. Res.* **2013**, *49*, 2552–2572. [[CrossRef](#)]
- Freeze, R.A.; Cherry, J.A. *Groundwater*; Prentice-Hall: Upper Saddle River, NJ, USA, 1979.
- Atangana, A. Analytical solutions for the recovery tests after constant-discharge tests in confined aquifers. *Water* **2014**, *40*, 595. [[CrossRef](#)]
- De Marsily, G. *Quantitative Hydrogeology: Groundwater Hydrology for Engineers*; Academic Press: Ann Arbor, MI, USA, 1986.
- Rehfeldt, K.R.; Boggs, J.M.; Gelhar, L.W. Field study of dispersion in a heterogeneous aquifer: 3. Geostatistical analysis of hydraulic conductivity. *Water Resour. Res.* **1992**, *28*, 3309–3324. [[CrossRef](#)]
- Doherty, J. *PEST: Model Independent Parameter Estimation*; Watermark Numerical Computing: Brisbane, Australia, 2010. Available online: <http://www.pesthomepage.org/Downloads.php> (accessed on 7 June 2019).
- Shen, C.; Riley, W.J.; Smithgall, K.R.; Melack, J.M.; Fang, K. The fan of influence of streams and channel feedbacks to simulated land surface water and carbon dynamics. *Water Resour. Res.* **2016**, *52*, 880–902. [[CrossRef](#)]
- Maxwell, R.M.; Condon, L.E.; Kollet, S.J.; Maher, K.; Haggerty, R.; Forrester, M.M.; Maher, K. The imprint of climate and geology on the residence times of groundwater. *Geophys. Res. Lett.* **2016**, *43*, 701–708. [[CrossRef](#)]

10. Shen, C.; Chambers, J.Q.; Melack, J.M.; Niu, J.; Riley, W.J. Interannual Variation in Hydrologic Budgets in an Amazonian Watershed with a Coupled Subsurface–Land Surface Process Model. *J. Hydrometeorol.* **2017**, *18*, 2597–2617. [[CrossRef](#)]
11. Ji, X.; Lesack, L.F.W.; Melack, J.M.; Wang, S.; Riley, W.J.; Shen, C. Seasonal and Interannual Patterns and Controls of Hydrological Fluxes in an Amazon Floodplain Lake With a Surface-Subsurface Process Model. *Water Resour. Res.* **2019**, *55*, 3056–3075. [[CrossRef](#)]
12. Gleeson, T.; Moosdorf, N.; Hartmann, J.; Van Beek, L.P.H. A glimpse beneath earth’s surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity. *Geophys. Res. Lett.* **2014**, *41*, 3891–3898. [[CrossRef](#)]
13. James, B.R.; Gorelick, S.M. When enough is enough: The worth of monitoring data in aquifer remediation design. *Water Resour. Res.* **1994**, *30*, 3499–3513. [[CrossRef](#)]
14. Wagner, B.J. Evaluating Data Worth for Ground-Water Management under Uncertainty. *J. Water Resour. Plan. Manag.* **1999**, *125*, 281–288. [[CrossRef](#)]
15. Neuman, S.P.; Xue, L.; Ye, M.; Lu, D. Bayesian analysis of data-worth considering model and parameter uncertainties. *Adv. Water Resour.* **2012**, *36*, 75–85. [[CrossRef](#)]
16. Freeze, R.A.; James, B.; Massmann, J.; Sperling, T.; Smith, L. Hydrogeological Decision Analysis: 4. The Concept of Data Worth and Its Use in the Development of Site Investigation Strategies. *Ground Water* **1992**, *30*, 574–588. [[CrossRef](#)]
17. Pinder, G.; Tucciarelli, T. Optimal data acquisition strategy for the development of a transport model for groundwater remediation. *Water Resour. Res.* **1991**, *27*, 577–588. [[CrossRef](#)]
18. Zeng, L.; Chang, H.; Zhang, D. A Probabilistic Collocation-Based Kalman Filter for History Matching. *SPE J.* **2011**, *16*, 294–306. [[CrossRef](#)]
19. Dai, C.; Xue, L.; Zhang, D.; Guadagnini, A. Data-worth analysis through probabilistic collocation-based Ensemble Kalman Filter. *J. Hydrol.* **2016**, *540*, 488–503. [[CrossRef](#)]
20. Haitjema, H.M.; Mitchell-Bruker, S. Are Water Tables a Subdued Replica of the Topography? *Ground Water* **2005**, *43*, 781–786. [[CrossRef](#)]
21. Li, H.-Y.; Sivapalan, M.; Tian, F.; Harman, C. Functional approach to exploring climatic and landscape controls of runoff generation: 1. Behavioral constraints on runoff volume. *Water Resour. Res.* **2014**, *50*, 9300–9322. [[CrossRef](#)]
22. Van Dam, J.C.; Feddes, R.A. Numerical simulation of infiltration, evaporation and shallow groundwater levels with the Richards equation. *J. Hydrol.* **2000**, *233*, 72–85. [[CrossRef](#)]
23. Fahs, M.; Younes, A.; Lehmann, F. An easy and efficient combination of the Mixed Finite Element Method and the Method of Lines for the resolution of Richards’ Equation. *Environ. Model. Softw.* **2009**, *24*, 1122–1126. [[CrossRef](#)]
24. Farthing, M.W.; Ogden, F.L. Numerical Solution of Richards’ Equation: A Review of Advances and Challenges. *Soil Sci. Soc. Am. J.* **2017**, *81*, 1257. [[CrossRef](#)]
25. Richards, L.A. CAPILLARY CONDUCTION OF LIQUIDS THROUGH POROUS MEDIUMS. *Physics* **1931**, *1*, 318. [[CrossRef](#)]
26. Zhou, J.; Liu, F.; He, J.-H. On Richards’ equation for water transport in unsaturated soils and porous fabrics. *Comput. Geotech.* **2013**, *54*, 69–71. [[CrossRef](#)]
27. Shen, C.; Phanikumar, M.S. A process-based, distributed hydrologic model based on a large-scale method for surface–subsurface coupling. *Adv. Water Resour.* **2010**, *33*, 1524–1541. [[CrossRef](#)]
28. Su, N. The fractional Boussinesq equation of groundwater flow and its applications. *J. Hydrol.* **2017**, *547*, 403–412. [[CrossRef](#)]
29. Schulla, J. *Model Description (WaSiM)*; Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology: Zurich, Switzerland, 2015. Available online: <http://www.wasim.ch/en/index.html> (accessed on 7 June 2019).
30. D’Errico, J. Bound Constrained Optimization Using Fminsearch. Matlab Central [Internet]. 2012. Available online: <https://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd-fminsearchcon> (accessed on 7 June 2019).
31. Tonkin, M.; Doherty, J. Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. *Water Resour. Res.* **2009**, *45*, 1–17. [[CrossRef](#)]

32. Tonkin, M.; Doherty, J.; Moore, C. Efficient nonlinear predictive error variance for highly parameterized models. *Water Resour. Res.* **2007**, *43*, 1–15. [[CrossRef](#)]
33. Harbaugh, B.A.W.; Banta, E.R.; Hill, M.C.; McDonald, M.G. *MODFLOW-2000, The US Geological Survey Modular Ground-Water Model—User Guide to Modularization Concepts and the Ground-Water Flow Process*; U.S. Geological Survey: Reston, WV, USA, 2000. Available online: <https://pubs.usgs.gov/of/2000/0092/report.pdf> (accessed on 7 June 2019).
34. Doherty, J. Ground Water Model Calibration Using Pilot Points and Regularization. *Ground Water* **2003**, *41*, 170–177. [[CrossRef](#)]
35. Tikhonov, A.; Arsenin, V. *Solutions of Ill-Posed Problems*; V.H. Winston: Washington, DC, USA, 1977.
36. Schölkopf, B.; Smola, A.J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, USA, 2002.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).