Data Availability induced Geological Model Uncertainty in Groundwater Flow and Land Subsidence Simulations

Shih-Jung Wang[1]*, Quoc Cuong Nguyen[1,2], Yu-Chen Lu[3], Yonatan Garkebo Doyoro[4,5,6], Duc-Huy Tran[1]

1 Graduate Institute of Applied Geology, National Central University, Taoyuan, Taiwan, R.O.C.
2 DRAGON - Mekong Institute, Can Tho University, Vietnam
3 Institute of Materials Science & Engineering, National Central University, Taoyuan, Taiwan, R.O.C.
4 Earth System Science, Taiwan International Graduate Program (TIGP), Academia Sinica, Taipei, Taiwan, R.O.C.
5 Department of Earth Science, National Central University, Taoyuan, Taiwan, R.O.C.
6 Department of Applied Geology, School of Natural Sciences, Adama Science and Technology University, Adama, Ethiopia

sjwang1230@gmail.com

Abstract

The study adopted a synthetic geological model (SGM) based on the geological characteristics in Taiwan as the baseline. Considering various data availability based on multiple borehole numbers, the incorporation of geological knowledge, and combining geophysical data, several simulated geological models were obtained and used in groundwater flow and land subsidence simulations. The results show that 17 boreholes at a 300 m × 300 m site with incorporating geological knowledge provides good assessment of land subsidence. The model using geophysical data with correction from 13 boreholes provided good results. This study demonstrates that additional data can decrease the uncertainty in geological and numerical models. The results can be used by engineers to construct a suitable geological model for engineering projects based on the precision requirements and budget. This study was part of the content published in Bulletin of Engineering Geology and the Environment (Wang et al., 2022).

Keywords: Geological model uncertainty; Data availability; Geological knowledge; Borehole number; Geophysical data assimilation; Land subsidence.

Background

Uncertainty in hydrogeological modeling has received attention for decades (Lelliott et al. 2009; Benedek and Molnár 2013; Mahmoudpour et al. 2016; Juang et al. 2019; Tran et al. 2022) and it becomes an important topic for hydrogeological simulations (Marinoni 2003; Shi et al. 2008; Guillaume et al. 2012). Previous studies have shown that the sources of uncertainty in groundwater hydrology simulations come from numerical model settings, data input, and the geological model (Refsgaard et al. 2006; Hassan et al. 2008). Numerous studies have analyzed uncertainty in land subsidence and groundwater flow modeling (e.g., Ferronato et al. 2006, Wang et al. 2015). However, relatively few have considered the uncertainty in geological models (e.g., Tran et al. 2022). The present study thus focuses on evaluating geological model uncertainty based on the availability of input data during the construction of a geological model for groundwater flow and land subsidence modeling. Geological models with (1) various numbers of boreholes (to evaluate the influence of borehole density on the simulated geological model), (2) with and without the incorporation of geological knowledge (to evaluate the suitability of incorporating geological knowledge into a geological model), and (3) with and without geophysical data assimilation (to evaluate the effectiveness of combining different and complementary data types to minimize geological model uncertainty) were assessed. The results of this study can be used to reduce the uncertainty of geological models with various data availability. They can also minimize risk when modelers and stakeholders make decisions. This study was part of the content published in Bulletin of Engineering Geology and the Environment (Wang et al. 2022).

Methodology

This study investigates the effect of the uncertainty in geological models with various data availability on groundwater flow and land subsidence. The baseline is a synthetic geological model (SGM) developed in this study. MODFLOW and Aquifer-System Compaction (SUB) packages in groundwater modeling system (GMS) are adopted to simulate groundwater flow and land subsidence, sequentially. The numerical results from each geological model were compared with those of SGM. To evaluate the effect of the number of boreholes, data from 1 to 17 boreholes were taken from SGM to reconstruct a geological model (i.e., simulated geological model). More borehole numbers are not feasible in reality thus it is not considered here. The horizon ID method was used in the geological model simulation with correction based on geological knowledge. To evaluate the effect of geological knowledge, the horizon ID method was used in the geological model simulation without correction based on geological knowledge. To evaluate the effect of geophysical data assimilation, electrical resistivity tomography (ERT) was used in SGM to assess the clay thickness. The original estimated clay thickness and that corrected based on data from 9, 13, and 17 boreholes with the cokriging method were input into the SUB package in MODFLOW, respectively, to estimate the influence on land subsidence simulations.

Results and discussions

Geological model

The developed SGM with a multilayer system was built in the GMS, as shown in Fig. 1. The model consisted of nine layers but the silt layer at the top layer of the model was not used in the numerical simulation because it was too thin and discontinuous, making the grid setting difficult. Clay material in the model was considered the target of compressive medium, which typically results in high land subsidence.

Simulated geological models were constructed based on various data availability. Five step of borehole numbers (1, 5, 9, 13, and 17) are adopted to simulate the geological models. Data from 1 or 5 boreholes were insufficient for capturing the distribution of clay. Data from 9, 13 and 17 boreholes were more sufficient for constructing an approximate clay distribution. The distribution of clay thickness became closer to that of SGM with increasing borehole number. Geological knowledge was considered in the horizon ID method for various borehole numbers. The simulated geological models are more close to the SGM than those without the incorporation of geological knowledge. Geophysical data (ERT) and various borehole numbers are assimilated. The simulated geological models are much close to the SGM than those only using various boreholes or ERT data.

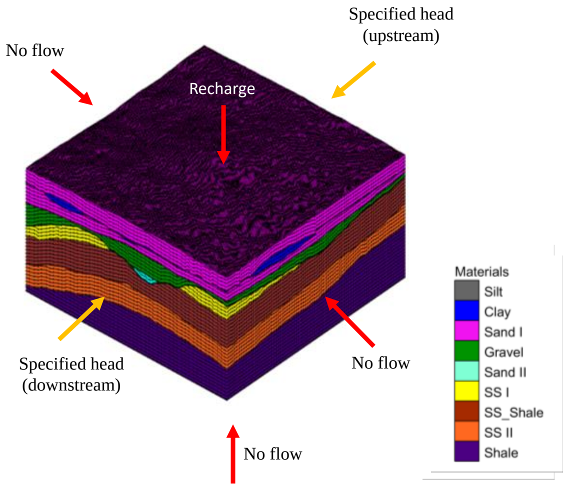

Figure 1 Synthetic geological model and boundary conditions. The two specified heads represent the upstream (north) and downstream (south), respectively. To control the groundwater flow from the upstream to the downstream, the east and west boundaries of the model were set as no-flow boundaries. Because of the low permeability bedrock, the bottom boundary of the model was set as a no-flow boundary. The top boundary was set to a constant recharge. SS = sand stone. (Wang et al. 2022)

Model comparisons

Since land subsidence assessment in this study is focused on clay material, the comparison on the distribution of clay thickness between the simulated geological model and SGM was conducted first. A comparison of clay thickness accuracy between the simulated geological models based on data from various numbers of boreholes and SGM is shown in Fig. 2. There is no correlation (R^2 = 0) between the model based on data from one borehole and SGM. The R^2 values increase with increasing number of boreholes (R^2 = 0.733 for 17 boreholes). RMSE decreases from 2.252 m for one borehole to 0.770 m for 17 boreholes. The quantification results show that increasing the number of boreholes in the study area for the geological model construction can decrease uncertainty. The trend is nonlinear and affected by the location of selected boreholes and the distribution of clay. Although increasing the borehole number to check if the results are convergent in this kind of synthetic model is possible, it is not feasible in reality thus more borehole numbers are not considered here.

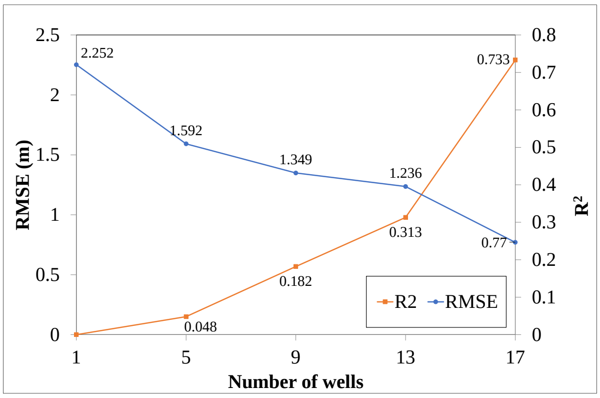

Figure 2 Comparison of clay thickness accuracy between SGM and simulated geological models based on data from various numbers of boreholes. Blue and orange lines indicate RMSE and R^2, respectively. RMSE decreases from 2.252 to 0.770 m and R^2 value increases from 0 to 0.733 when the borehole number increases from 1 to 17. (Wang et al. 2022)

Groundwater flow and land subsidence are simulated based on the simulated geological models and then the results are compared to those of SGM using an error assessment. Figure 3 shows a comparison of the numerical results of land subsidence between the simulated geological model based on data from various data availability and SGM. For the model based on data from 17 boreholes without the incorporation of geological knowledge, the R^2 value slightly increases to 0.316 and RMSE decreases to 7 mm. The results show that the accuracy of this method is very low. The use of data from 17 boreholes and the incorporation of geological knowledge improve the results. The results demonstrate that geological knowledge is necessary for simulating a geological model and land subsidence based on borehole data; its incorporation dramatically increases the accuracy of numerical results and reduces model uncertainty.

The results obtained from the model that used only ERT data show an improvement compared to those for data from boreholes without the incorporation of geological knowledge. The R^2 value is 0.640 and RMSE is 7 mm (Fig. 3). These values are worse than those for the model based on data from 17 boreholes with the incorporation of geological knowledge. Using only ERT data for land subsidence simulation embeds high uncertainty. The results of the simulated geological model that used ERT data with corrected data from 13 and 17 boreholes show a remarkable improvement in the land subsidence simulation. Specifically, the R^2 values are 0.812 and 0.894 and the RMSE values are 3 and 3 mm, respectively (Fig. 3). The results also show higher accuracy than that obtained using 17 boreholes with the incorporation of geological knowledge. The R^2 value increases from 0.787 to 0.894 and RMSE decreases from 4 to 3 mm. The results demonstrate that using data from a limited number of boreholes or only ERT data cannot provide reliable numerical results of land subsidence. A combination of borehole data and ERT data effectively reduces the required well number and uncertainty of geological models and increases numerical model accuracy.

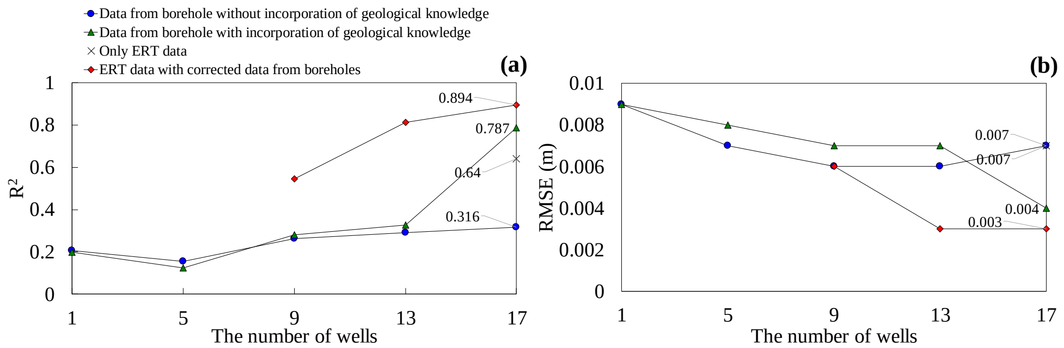

Figure 3 (a) R^2 and (b) RMSE values of comparison in land subsidence assessment between numerical results obtained from SGM and simulated geological models based on data from boreholes without incorporation of geological knowledge, data from boreholes with incorporation of geological knowledge, only ERT data, and ERT data with corrected data from boreholes. (Wang et al. 2022)

Conclusions

This study performed numerical simulations of groundwater flow and land subsidence for SGM as the baseline. Geological models with various data availability were built. The results from these models were compared with those of SGM to assess the geological model uncertainty. The quantification results show that the clay thickness, groundwater level, and land subsidence results for the model based on borehole data approach those of SGM with increasing borehole number. The model based on data from 17 boreholes that incorporates geological knowledge provided acceptable groundwater flow and land subsidence results for a 300 m × 300 m site. Applying the horizon ID method without the incorporation of geological knowledge yielded assessment results that were far from those of SGM. To increase accuracy and decrease the uncertainty of the geological model, borehole data can be combined with ERT data via the cokriging interpolation method. This combination decreases the required borehole number and yields a dramatic improvement compared with the results obtained using only ERT data or borehole data. The study results can be used by engineers or researchers to determine a suitable strategy for engineering geology projects based on the precision requirements and budget.

References

Benedek K, Molnár P (2013) Combining structural and hydrogeological data: Conceptualization of a fracture system. *Engineering Geology* 163:1-10.

Ferronato M, Gambolati G, Teatini P, Bau D (2006) Stochastic poromechanical modeling of anthropogenic land subsidence. *International Journal of Solids and Structures* 43(11-12): 3324-3336.

Guillaume JH, Qureshi ME, Jakeman AJ (2012) A structured analysis of uncertainty surrounding modeled impacts of groundwater-extraction rules. *Hydrogeology journal* 205:915-932.

Hassan AE, Bekhit HM, Chapman JB (2008) Uncertainty assessment of a stochastic groundwater flow model using GLUE analysis. *Journal of Hydrology* 3621-2:89-109.

Juang CH, Zhang J, Shen M, Hu J (2019) Probabilistic methods for unified treatment of geotechnical and geological uncertainties in a geotechnical analysis. *Engineering Geology* 249:148-161

Lelliott M, Cave M, Wealthall G (2009) A structured approach to the measurement of uncertainty in 3D geological models*. Quarterly Journal of Engineering Geology Hydrogeology* 421:95-105.10.1144/1470-9236/07-081

Mahmoudpour M, Khamehchiyan M, Nikudel MR, Ghassemi MR (2016) Numerical simulation and prediction of regional land subsidence caused by groundwater exploitation in the southwest plain of Tehran, Iran. *Engineering Geology* 201:6-28.

Marinoni O (2003) Improving geological models using a combined ordinary–indicator kriging approach. *Engineering Geology* 691-2:37-45 .

Refsgaard JC, Van der Sluijs JP, Brown J, Van der Keur P (2006) A framework for dealing with uncertainty due to model structure error*. Advances in Water Resources* 2911:1586-1597.

Shi X, Wu J, Ye S, Zhang Y, Xue Y, Wei Z, Li Q, Yu J (2008) Regional land subsidence simulation in Su-xi-Chang area and Shanghai City, China. *Engineering Geology* 1001-2:27-42.

Tran D-H, Wang S-J, Nguyen QC (2022) Uncertainty of Heterogeneous Hydrogeological Models in Groundwater Flow and Land Subsidence Simulations – A Case Study in Huwei Town, Taiwan. *Engineering Geology* 298:106543.

Wang S-J, Lee C-H, Hsu K-C (2015) A technique for quantifying groundwater pumping and land subsidence using a nonlinear stochastic poroelastic model*. Environmental Earth Sciences* 7312: 8111-8124.

Wang S-J, Nguyen QC, Lu Y-C, Doyoro YG, Tran D-H (2022) Evaluation of Geological Model Uncertainty Caused by Data Sufficiency Using Groundwater Flow and Land Subsidence Modeling as Example. *Bulletin of Engineering Geology and the Environment* 81(331).