JOURNAL OF WATER AND LAND DEVELOPMENT
e-ISSN 2083-4535
Polish Academy of Sciences (PAN) Institute of Technology and Life Sciences – National Research Institute (ITP – PIB)

JOURNAL OF WATER AND LAND DEVELOPMENT
DOI: 10.24425/jwld.2025.154268
2025, No. 65 (IV–VI): 238–246

Numerical simulation of pumping tests in heterogeneous aquifers using a Python-driven framework

Ibrahim S. Fekry*, Nagy A. Hassan, Ayman A. Nassar

Ain-Shams University, Faculty of Engineering, Irrigation and Hydraulics Department, P.O. Box 11566, El-Abbasia, Cairo, Egypt

* Corresponding author

RECEIVED 20.02.2025 ACCEPTED 23.04.2025 AVAILABLE ONLINE 24.06.2025

Abstract: Accurate interpretation of pumping test data in stratified aquifers requires approaches that account for vertical heterogeneity, a factor often neglected in conventional analytical solutions. This study presents a Pythodriven axisymmetric numerical modelling framework, built using MODFLOW 6 and FloPy, to simulate both pumping and recovery phases in vertically heterogeneous confined aquifers.

The model discretises the domain radially and vertically to allow layer-specific representation of hydraulic conductivity, while specific storage is assigned uniformly. An optimisation-based inverse modelling approach was used to estimate aquifer parameters by minimising the difference between observed and simulated drawdowns. Applied to a case study in Bahariya, Egypt, the results yielded hydraulic conductivity values consistent with the site's stratigraphy – ranging from approximately 10^{-5} m·d^{-1} in shale to over 27 m·d^{-1} in limestone – and a specific storage of $4{\cdot}10^{-8}$ m^{-1}. The simulated radius of influence was 133.67 m, and the root mean square error between the observed and simulated drawdown was 0.01 m.

Sensitivity analysis demonstrated that vertical discretisation had the greatest influence on model accuracy, with coarser grids increasing residual error by nearly 90% and reducing the radius of influence by 9%. The temporal resolution had minimal impact on accuracy but significantly affected computation time.

This framework offers an open-source, automated, and script-based tool for simulating pumping tests in layered aquifer systems, enabling more reliable estimation of hydraulic parameters for both scientific and applied groundwater studies.

Keywords: axisymmetric model, FloPy, hydraulic parameter estimation, MODFLOW 6, numerical groundwater modelling, pumping test, vertically heterogeneous aquifers

INTRODUCTION

Pumping tests are essential for determining aquifer hydraulic properties such as transmissivity (*T*), storativity (*S*), and hydraulic conductivity (*Hk*). These parameters govern groundwater flow and are fundamental for well design, groundwater resource management, and aquifer characterisation. Analytical solutions, such as those proposed by Theis (1935), Cooper and Jacob (1946), Hantush and Jacob (1955), Boulton (1963), and Neuman (1972), have been widely used to interpret pumping test data. However, these models rely on simplified assumptions, including aquifer homogeneity, isotropy, and infinite radial extent, which rarely hold true under field conditions (Gunawardhana *et al.*, 2021).

Although each of these solutions introduces refinements to account for specific conditions – such as unconfined behaviour, leaky boundaries, or delayed yield – they are inherently limited in layered (stratified) aquifer systems. In such settings, vertical contrasts in *Hk* can significantly affect drawdown behaviour between layers, leading to biased parameter estimates if vertical heterogeneity is not accounted for (Chen and Jiao, 1999; Wu *et al.*, 2017; Zhu *et al.*, 2020).

Numerical models provide a more flexible alternative by allowing for complex boundary conditions and the simulation of spatially variable parameters. MODFLOW, developed by the U.S. Geological Survey, is a widely adopted code for modelling groundwater flow in heterogeneous, anisotropic aquifers (Harbaugh *et al.*, 2000; Langevin *et al.*, 2017). Its integration with

Python-based tools such as FloPy enables automation, reproducibility, and user-defined control over simulation and calibration workflows (Bakker *et al.*, 2016; Hughes *et al.*, 2024). Recent studies have also explored the integration of MODFLOW with optimisation algorithms and unstructured grids to enhance model adaptability in arid and geologically complex regions (Saqr *et al.*, 2023; Saqr *et al.*, 2025).

However, most published applications rely on full three-dimensional discretisation, which significantly increases computational demand, particularly when high resolution is needed near pumping wells. Axisymmetric models offer an efficient alternative by simplifying the domain to two dimensions while preserving radial flow geometry. Earlier studies have demonstrated axisymmetric formulations using structured (Langevin, 2008) and unstructured (Bedekar, Scantlebury and Panday, 2019) MODFLOW grids. Yet, these models typically focus on idealised or homogeneous conditions and do not integrate automated calibration, sensitivity analysis, or dual-phase (pumping and recovery) simulation into a single reproducible workflow.

This study fills that methodological gap by developing a Python-driven, axisymmetric numerical model using MODFLOW 6 and FloPy (Harbaugh, 2005). The model is capable of simulating pumping tests in stratified aquifers with vertical heterogeneity, supports automated parameter estimation for individual layers, includes a sensitivity analysis module, and handles both pumping and recovery phases in a unified framework.

MATERIALS AND METHODS

STUDY AREA

The field pumping test analysed in this study was conducted at a deep, vertically stratified aquifer located along the Bani Mazar – El Bahariya Road, within the Bahariya Oasis, Western Desert, Egypt. This region forms part of a sedimentary basin characterised by thick, alternating sandstone, shale, and limestone hydrogeological units. Based on lithological interpretation and field observations, the aquifer is classified as confined, with the piezometric surface recorded above the upper confining shale unit.

The test involved two wells: a pumping well and a nearby observation well, positioned approximately 10.7 m apart. The pumping well was drilled to a depth of 378 m and screened from 212 m to 378 m, while the observation well was completed to 300 m, with a screen interval from 184 m to 300 m. Both wells are situated at nearly the same ground elevation. The depth to groundwater, measured from ground level, was approximately 122.05 m in the pumping well and 121.84 m in the observation well, confirming confined aquifer conditions.

The diameter of both wells is approximately 8.5 inches (0.216 m). The maximum recorded drawdown during the test was 12.75 m in the production well and 0.56 m in the observation well. Drawdown data were collected during both the pumping and recovery periods, allowing a complete evaluation of aquifer response. Well construction details are summarised in Table 1.

The lithological profile, derived from borehole logs and verified through geophysical logging, indicates an upper sandstone unit underlain by interbedded sandstone and limestone, followed by alternating shale and dolomitic limestone. The

Table 1. Pumping and observation well data

Parameter	Pumping well	Observation well
Well depth (m)	378	300
Screen interval (m)	212–378	184–300
Well diameter (m)	0.216	0.216
Distance from pumping well (m)	–	10.7
Pumping rate ($m^3 \cdot d^{-1}$)	5,280	–
Groundwater depth (m)	122.05	121.84
Maximum drawdown (m)	12.75	0.56

Source: own elaboration.

limestone hydrogeological units at depth represent the productive part of the aquifer, as confirmed by the location of the well screens and the aquifer's response during pumping. Although the base of the aquifer was not encountered in this borehole, its depth is inferred from nearby wells in the Bahariya region. The site is located in a desert setting, with no observed recharge, seepage, or leakage, and the stratified lithology strongly influences vertical flow and drawdown propagation.

A conceptual cross-section illustrating the aquifer stratification, screen placements, and well configuration are presented in Figure 1.

NUMERICAL MODEL FRAMEWORK

An axisymmetric numerical model was developed using MODFLOW 6 (Langevin *et al.*, 2017), with all input files and control scripts automated via FloPy (Bakker *et al.*, 2016). The model was designed to simulate the drawdown behaviour of a vertically heterogeneous confined aquifer under both pumping and recovery phases, using a radial-symmetric approach to reduce computational demands while maintaining spatial accuracy near the well. This approach follows the conceptual transformation introduced by Langevin (2008), which allows three-dimensional radial flow to be simulated using a two-dimensional vertical slice while preserving the key characteristics of cylindrical symmetry.

The model domain was conceptualised as a 2D vertical cross-section representing radial flow around a central pumping well. The radial extent was set to 500 m, and the vertical domain ranged from the saturated top elevation of −121.84 m to −500 m, representing the saturated portion of the aquifer. The pumping well screen spanned from −212 m to −378 m, and the observation well screen from −184 m to −300 m.

The domain was discretised into 188 vertical layers and 89 radial columns. Vertical discretisation was controlled using a dynamic cell thickness approach centred on the well screen. The top and bottom 0.01 m cells corresponded to the well screen boundaries. Above and below the screen, cell thickness was increased exponentially using multipliers of 1.6 and 1.3, respectively. Between the screen depths, a finer multiplier of 1.1 was used to capture vertical gradients. This structure allowed the screen intervals to be precisely represented in the numerical grid while preserving lithological fidelity from borehole logs.

In the radial direction, the first column matched the well radius (0.22 m), followed by a second column of 0.01 m to

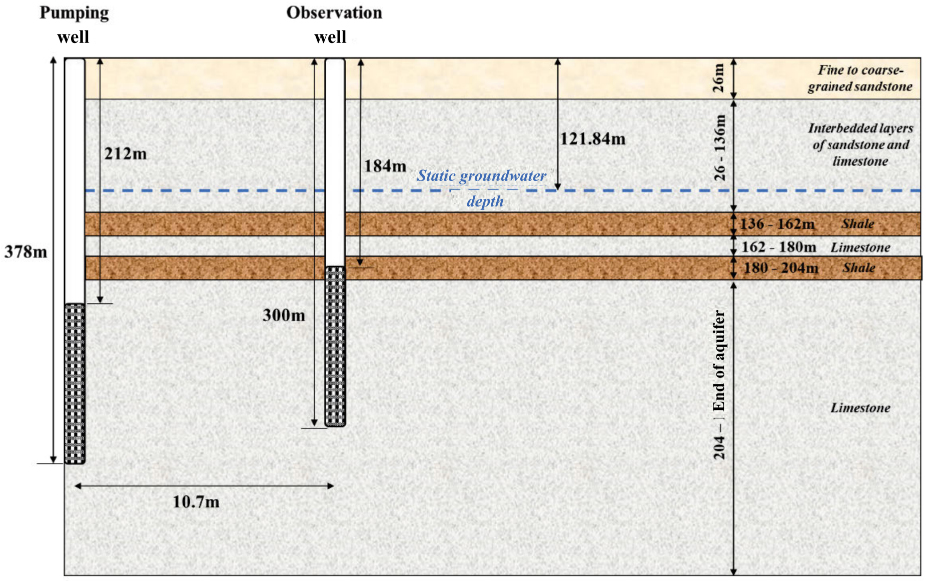

Fig. 1. Cross-sectional diagram showing the well configuration, lithological layers, and screen intervals; source: own elaboration

capture near-well gradients. Further columns expanded outward using a 1.1 multiplier, ensuring a balanced trade-off between accuracy and computational efficiency. The full grid refinement strategy and screen placement are illustrated in Figure 2.

Boundary conditions were selected based on the regional pre-pumping equilibrium conditions. A constant head boundary equal to −121.84 m was assigned at the outer radial edge. The top and bottom boundaries were defined as no-flow, consistent with the confined aquifer setting. The initial head throughout the model domain was set to −121.84 m.

The simulation included two stress periods: a pumping phase of 2,880 min, followed by a recovery phase of 2,160 min. The baseline time discretisation included 1,440-time steps with a multiplier of 1.05, ensuring sufficient temporal resolution to capture early-time drawdown behaviour.

The model used specific storage (*Ss*) to represent the confined aquifer conditions. The *Hk* values were assigned per layer based on field lithology. The output drawdown was extracted at the observation well for both pumping and recovery periods.

This configuration is hereafter referred to as the baseline simulation, and it served as the reference setup for model validation, sensitivity analysis, and parameter estimation. It was designed to achieve a balance between spatial–temporal resolution and computational efficiency, based on preliminary grid refinement trials.

MODEL VALIDATION USING ANALYTICAL SOLUTIONS

General information

To ensure that the developed numerical model accurately reproduces well-established analytical behaviour, a validation step was conducted using four classical analytical solutions: Theis (1935), Neuman (1975), Boulton (1963), and Cooper and Jacob

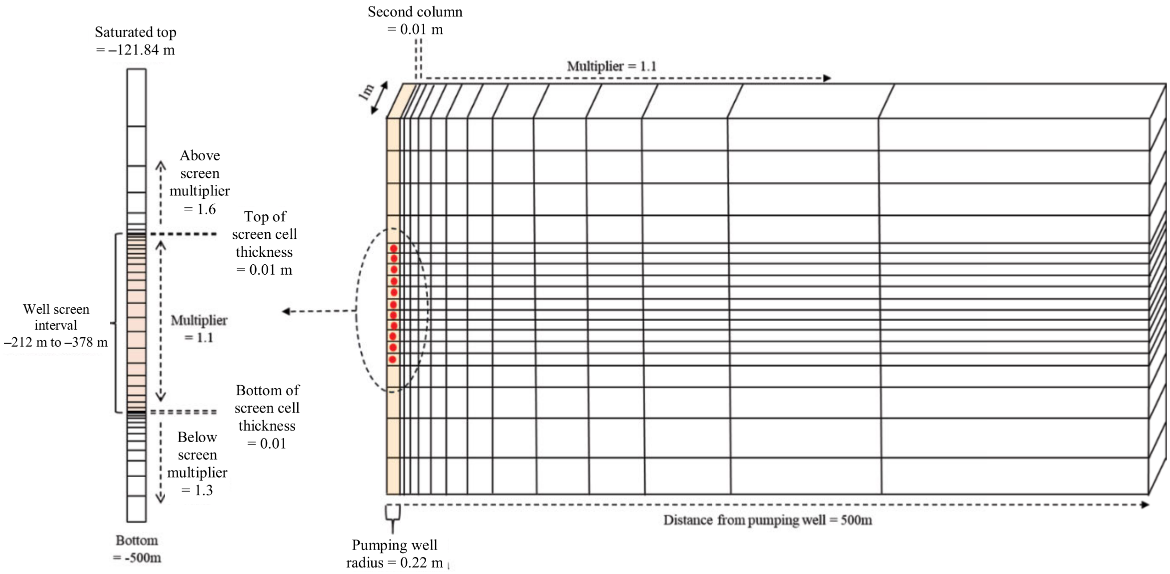

Fig. 2. Axisymmetric model domain showing vertical and radial grid refinement and well screen placement; source: own elaboration

(1946). This step aimed to verify that the numerical formulation, discretisation, and solver setup could simulate homogeneous and idealised aquifer conditions reliably before applying the model to the heterogeneous field case.

Input data and assumptions

Transmissivity (*T*) and *Ss* values for each analytical solution were estimated using AquiferTest software, which applies curve-fitting techniques to observed drawdown data. The software does not incorporate lithological data or layer-specific properties; it assumes homogeneous aquifer conditions and full penetration of the pumping well – limitations that are inherent to the analytical methods themselves. The resulting transmissivity values were then used to compute *Hk* by dividing *T* by the saturated thickness (*B*), which was assumed to be 378.16 m, extending from the observed water table elevation (−121.84 m) to the model bottom (−500 m). There is a summary of the estimated values of *T*, *Ss*, and the derived *Hk* for each method in the Table 2.

Table 2. Hydraulic parameters derived from analytical solutions

Method	Transmissivity ($m^2 \cdot d^{-1}$)	Specific storage (m^{-1})	Hydraulic conductivity ($m \cdot d^{-1}$)
Theis solution	15,380	10^{-7}	40.67
Neuman solution	26,438	10^{-7}	69.91
Boulton solution	26,438	10^{-7}	69.91
Cooper and Jacob solution	30,585	$5.6 \cdot 10^{-16}$	80.87

Source: own elaboration.

These values were then uniformly assigned across all model layers in the validation simulations to reflect the assumption of vertical homogeneity and full well penetration, as required by the analytical solutions. Both horizontal and vertical hydraulic conductivities were set equal (i.e., *Vk*/*Hk* = 1), consistent with isotropic conditions. The same grid refinement, domain extent, and boundary conditions used in the stratified model were retained for the validation simulations to ensure numerical consistency.

Implementation of analytical solution settings

Each analytical method was numerically implemented using a configuration that closely replicates its theoretical assumptions.
- Theis (1935): confined, homogeneous, isotropic aquifer under constant-rate pumping; uniform *Ss* was used; *Sy* was not applicable.
- Neuman (1975) and Boulton (1963): simulated as unconfined aquifers with delayed gravity drainage; in these cases, *Sy* was set based on analytical results, although the numerical model maintained full saturation (a limitation acknowledged for consistency with the solution form).
- Cooper and Jacob (1946): represented late-time confined aquifer conditions assuming quasi-steady behaviour, where the drawdown increases linearly with log(time); *Sy* was not used; *Ss* was applied uniformly. While the transmissivity value estimated by this method was within a reasonable range, *Ss* value obtained was significantly lower than typical field values, reaching $5.6 \cdot 10^{-16}$. This is a known limitation of the Cooper and Jacob method, which is based on late-time drawdown analysis where early-time storage effects are not well captured. As such, the derived *Ss* should not be interpreted as physically representative. It was retained here solely for validation purposes to match the theoretical drawdown behaviour of the analytical curve and highlight the limitations of applying such methods in practical aquifer characterisation.

The numerical model was then run using these analytical parameter sets under corresponding assumptions. This process did not aim to reflect field heterogeneity but to confirm that the numerical setup could replicate textbook behaviour under controlled theoretical conditions. While assigning a single *Hk* value across shale and limestone layers is unrealistic in actual settings, it was essential here to match the analytical framework and validate model behaviour under ideal assumptions.

The simulated drawdown from the numerical model with the analytical curves and observed data for each solution are compared in Figure 3. Minor discrepancies between numerical and analytical drawdown curves are primarily attributed to vertical discretisation and differences in well representation, which are known limitations of analytical approaches. The close agreement across all cases confirmed the model's ability to replicate classical drawdown behaviour, thus validating its reliability as a numerical simulation tool.

PARAMETER ESTIMATION THROUGH INVERSE MODELLING

After validating the model under homogeneous conditions, the next step involved estimating aquifer parameters under stratified conditions to reflect site-specific heterogeneity.

Aquifer parameters were estimated using an inverse modelling approach, which adjusted the model inputs to minimise the difference between observed and simulated drawdown during the pumping test. The focus was on estimating layer-specific *Hk* and *Ss* values consistent with the stratified nature of the aquifer and the observed drawdown response.

The optimisation process was performed using the "fmin" function from the SciPy Python library, which implements a Nelder–Mead algorithm to minimise an objective function based on the root mean square error and total residual error between observed and simulated drawdowns. The error function considered both the pumping and recovery phases simultaneously to improve calibration robustness.

Each lithological unit was assigned an independent *Hk* value that was iteratively adjusted until the best-fit simulation was achieved. The model maintained fixed screen locations, stress period durations, and boundary conditions during this process. The *Ss* was estimated for each layer based on calibration outcomes.

RESULTS

PARAMETER ESTIMATION AND AQUIFER BEHAVIOUR

The model was run in inverse modelling mode to estimate the hydraulic parameters of the vertically layered aquifer system. Each geological layer – sandstone, shale, or limestone – was assigned a unique hydraulic conductivity (*Hk*) value, and

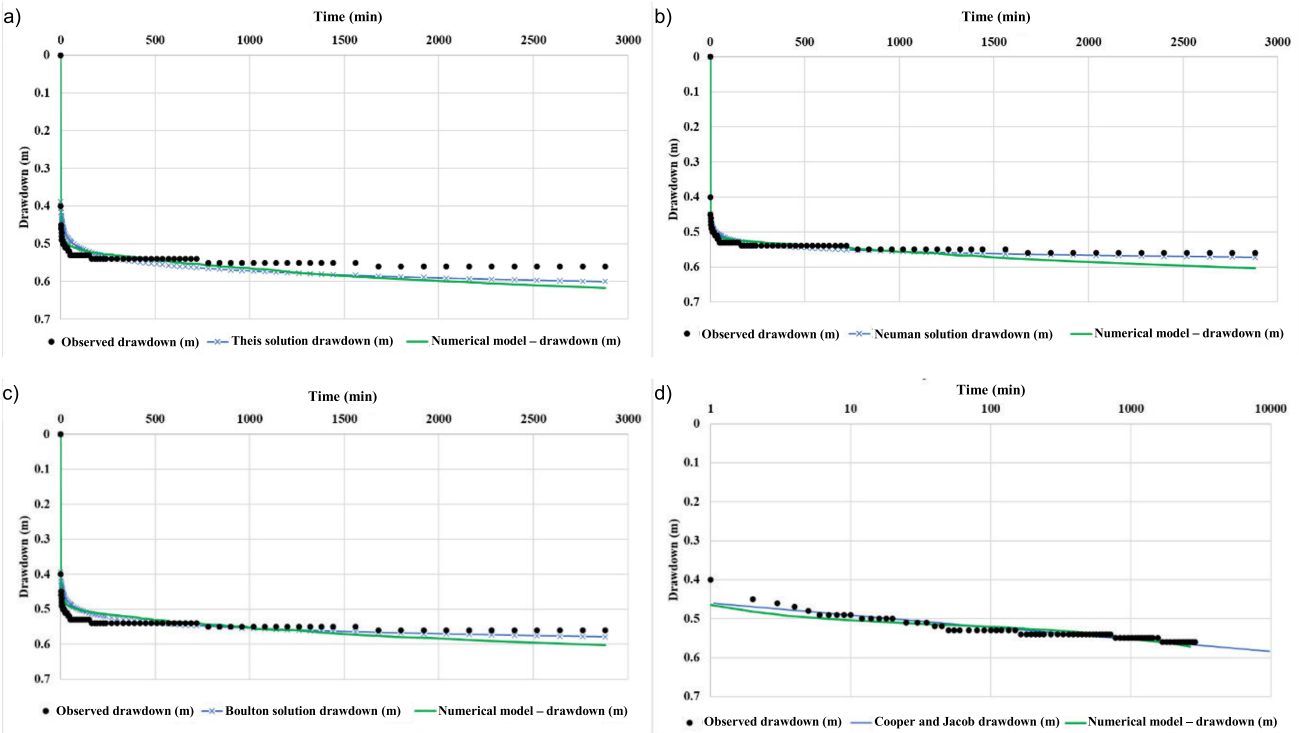

Fig. 3. Comparison of observed drawdown, analytical solutions, and numerical model results under homogeneous assumptions for such methods: a) Theis, b) Neuman, c) Boulton, d) Cooper and Jacob; source: own study

a uniform specific storage (*Ss*) was estimated for the entire domain. Specific yield (*Sy*) was not used, consistent with the confined conditions confirmed by the field data. The confinement of the aquifer was supported by low-permeability shale layers above and below the productive zones, as well as piezometric levels lying above the uppermost layer. Under such conditions, the aquifer's response is predominantly controlled by elastic storage rather than drainage from pores. The estimated *Ss* value of $4 \cdot 10^{-8}$ m^{-1} is within the typical range reported for confined aquifers (Freeze and Cherry, 1979).

In Table 3, the estimated layer-specific *Hk* values were summarised. Shale layers were found to have *Hk* values on the order of 10^{-5} m·d^{-1}, while sandstone and limestone layers exhibited significantly higher values, up to 27.2 m·d^{-1}. These values are in line with ranges reported for similar hydrogeological units in Bahariya and other parts of Egypt (Hamdan and Sawires, 2013)

The *Hk* estimates revealed pronounced vertical heterogeneity. The radius of influence, defined as the maximum radial distance affected by pumping, was estimated to be 133.67 m, which has direct implications for well spacing in design applications.

The agreement between observed and simulated drawdowns is shown in Figure 4. The model successfully captured both early-time drawdown and late-time recovery trends. Additionally, the model reproduced a steep drawdown gradient in the layers adjacent to the pumping well screen, which is consistent with expected near-well behaviour in confined stratified aquifers. The root mean square error (*RMSE*) was 0.01 m, which represents less than 2% of the maximum observed drawdown. The total residual error was 16.92 m, calculated as the sum of squared differences between observed and interpolated simulated drawdowns at each observation time. While the cumulative value may appear high, it results from the dense simulation structure comprising 1440 time steps. When distributed over the full duration, the average residual per step remains below 0.012 m, which is well within acceptable modelling tolerances given the observed maximum drawdown of 0.56 m.

Table 3. Estimated hydraulic parameters from inverse modelling

Soil material	Layer top level (m)	Layer bottom level (m)	Hydraulic conductivity (m·d^{-1})
Sandstone	−121.84	−136	2
Shale	−136	−162	$5.5 \cdot 10^{-5}$
Limestone	−162	−180	26.88
Shale	−180	−204	$5.8 \cdot 10^{-5}$
Limestone	−204	−500	27.2

Source: own study.

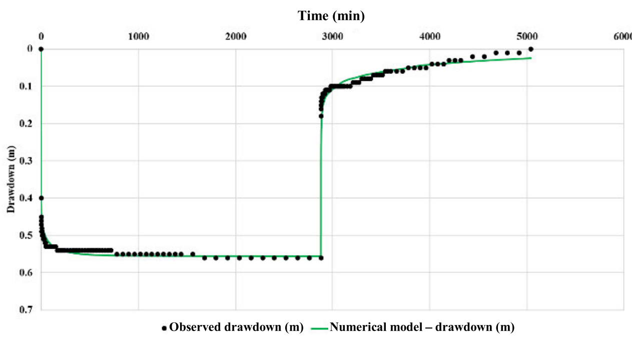

Fig. 4. Observed vs simulated drawdown for numerical model solution; source: own study

SENSITIVITY TO NUMERICAL DISCRETISATION

The influence of numerical discretisation on model accuracy and computational efficiency was evaluated through a structured sensitivity analysis. Three aspects were tested independently: vertical layering, radial grid spacing, and temporal resolution. For each, two modified configurations (finer and coarser) were assessed relative to the baseline model. The performance metrics included the *RMSE,* total residual error, radius of influence, and computation time.

Vertical layering exhibited the most significant impact on model accuracy. While both the baseline and finer grids maintained the same *RMSE* (0.01 m), the coarser configuration resulted in a higher *RMSE* (0.02 m), a two-fold increase in residual error (32.1 m vs 16.9 m), and a reduced radius of influence by 9.1% (121.53 m). These results confirm the importance of adequate vertical refinement for capturing early drawdown dynamics and vertical gradients in head distribution – particularly in stratified systems with alternating low- and high-permeability layers.

Radial grid spacing influenced the residual error and estimated radius of influence, despite *RMSE* values remaining stable across all configurations (0.01 m). The coarser grid led to a 7.9% reduction in the radius of influence (117.53 m), while the residual error increased substantially (from 16.9 to 31.1–32.1 m). This highlights the sensitivity of flow convergence around the pumping well to horizontal resolution and supports maintaining finer grids in the near-well zone where steep hydraulic gradients dominate.

Time step resolution had a negligible effect on *RMSE*, which remained constant at 0.01 m across all configurations. However, residual error increased slightly in the coarser setup (by approximately 8.9%), while computation time decreased markedly – from 58.2 s in the baseline case to just 9 s. These findings suggest that coarse temporal resolution may be acceptable for preliminary assessments, but finer resolution is recommended for applications requiring detailed transient behaviour, such as the estimation of early-time aquifer response.

The results of all tested configurations are summarised in Table 4. In Figure 5, the observed vs simulated drawdown curves for all scenarios are presented, allowing a visual comparison of the transient response. Overall, vertical discretisation was found to be the most influential factor affecting model accuracy, while temporal refinement offered the greatest efficiency gains without compromising model fidelity.

DISCUSSION

GENERAL DISCUSSION

This study highlights the effectiveness of a numerical modelling framework in simulating the drawdown response of a vertically layered confined aquifer. Layer-specific hydraulic conductivity (*Hk*) values were estimated through inverse modelling, while a uniform specific storage (*Ss*) value was applied across all layers. The model successfully reproduced both early-time drawdown and late-time recovery phases observed during the pumping test, confirming the value of incorporating stratigraphic heterogeneity into the model design.

Unlike conventional analytical solutions that assume aquifer homogeneity, the numerical model captured the distinct hydraulic behaviour of each hydrogeological unit. Estimated *Hk* values ranged from approximately $5.5 \cdot 10^{-5}$ m·d^{-1} in low-permeability shale layers to over 27 m·d^{-1} in high-permeability limestone units – values that are consistent with ranges reported for Bahariya and other Egyptian hydrogeological units (Hamdan and Sawires, 2013). The estimated *Ss* of $4 \cdot 10^{-8}$ m^{-1} is within the typical range for confined aquifers (Freeze and Cherry, 1979), supporting the dominance of elastic storage under confined conditions.

PRACTICAL IMPLICATIONS

The model-estimated radius of influence, approximately 133.67 m, provides a practical indicator for well spacing and pumping interference analysis. This value reflects vertical heterogeneity in the aquifer system and actual boundary conditions, unlike simplified analytical estimates. Such site-specific outputs offer added value in the design of pumping schemes and groundwater monitoring networks. To provide a comparative analytical benchmark, the radius of influence was also estimated using the empirical formula $r = 3000\sqrt{T}$, as proposed by Driscoll (1986). Using the Theis-derived transmissivity (T) (15,380 m^2·d^{-1}), the empirical estimate yields approximately 126 m, which is in close agreement with the numerical result. However, such empirical formulas neglect lithological layering, aquifer boundaries, and transient storage effects. The observed agreement reinforces the model's reliability, while also demonstrating the added value of numerical modelling in vertically layered confined aquifer systems.

Sensitivity analysis revealed that vertical discretisation had the most notable effect on model outputs. Coarsening the vertical

Table 4. Summary of sensitivity analysis results for baseline and modified discretisation configurations

Test	Configuration	*RMSE* (m)	Residual error (m)	Radius of influence (m)	Computation time (s)
Baseline		0.01	16.9	133.67	58
Vertical discretisation	finer layers	0.01	17.71	133.67	71.4
	coarser layers	0.02	32.1	121.53	14
Radial grid refinement	finer near pumping well	0.01	32.05	133.60	65.4
	coarser near pumping well	0.01	31.11	117.53	51
Time step configuration	finer time steps	0.01	18.13	133.67	88.2
	coarser time steps	0.01	18.4	133.67	9

Explanations: *RMSE* = root mean square error.
Source: own study.

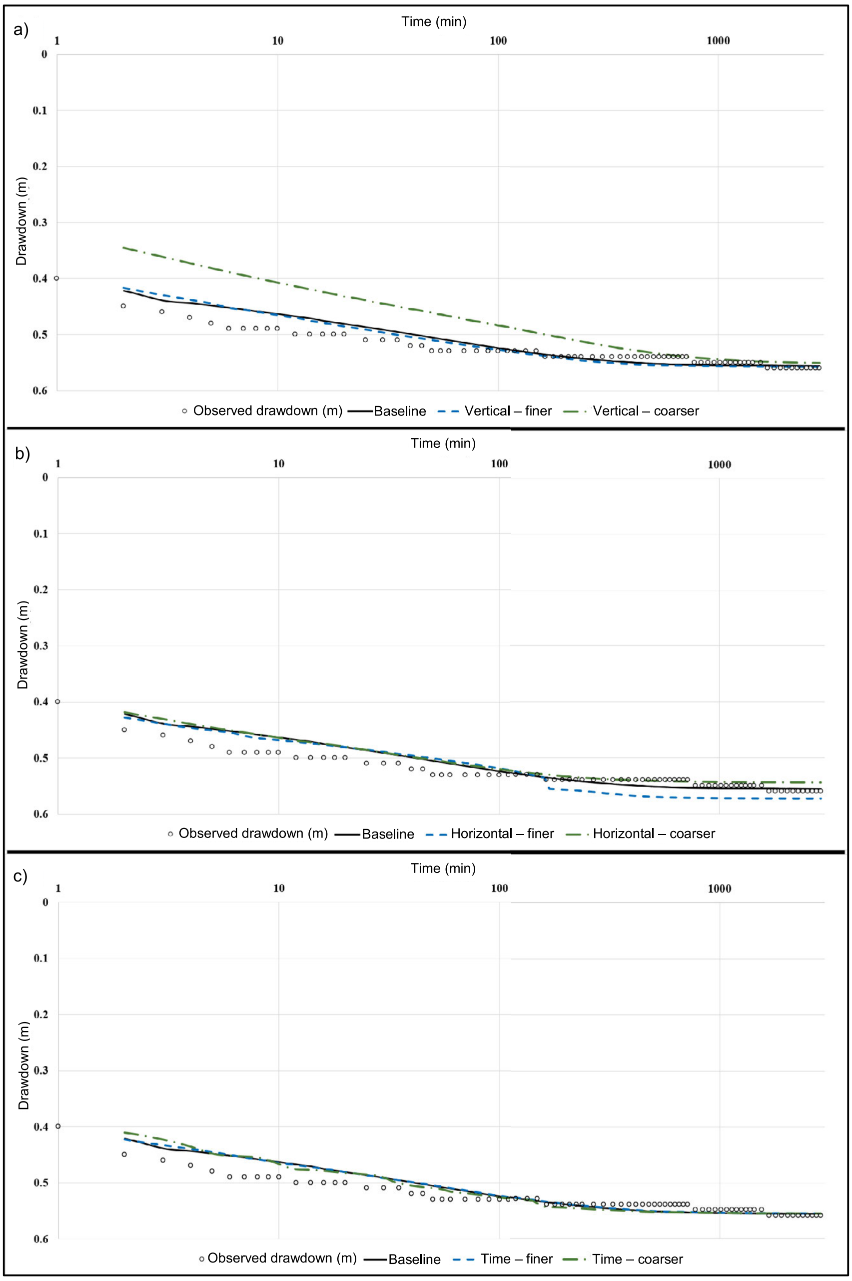

Fig. 5. Observed versus simulated drawdown under varying numerical discretisation settings: a) vertical layering, b) horizontal grid spacing, c) time step resolution; all plots are shown on a semi-logarithmic scale to highlight transient behaviour at early times; source: own study

grid led to a doubling of total residual error and a 9% reduction in the radius of influence, highlighting the necessity of fine vertical layering to resolve steep vertical gradients and early-time drawdown behaviour. In contrast, varying the temporal resolution resulted in minimal changes to root mean square error, while reducing simulation time by up to 85%. This suggests that coarser time steps may be suitable for preliminary assessments, although finer steps remain preferable for capturing transient behaviour in detail.

LIMITATIONS AND FUTURE WORK

This study assumes radial homogeneity within each geological layer, meaning that *Hk* is uniform in the horizontal direction across each stratum. While this assumption is justified by the confined nature of the aquifer and the axisymmetric modelling framework, it may not fully capture lateral variations in hydraulic properties that could exist in more geologically complex settings. Extending the modelling approach to account for lateral heterogeneity – through the use of unstructured grids or fully three-dimensional models – could enhance the applicability of the method to more variable aquifer systems.

Moreover, wellbore storage effects and partially saturated conditions were not included in the current model configuration. Although these assumptions are reasonable for the deep, fully saturated confined aquifer considered here, their incorporation may be necessary when applying the methodology to shallower or unconfined systems where early-time drawdown is more sensitive to storage effects. Future research could also explore the integration of observation well responses and multi-well interference to further validate the robustness of the inverse modelling approach.

CONCLUSIONS

This study developed and applied a Python-based, axisymmetric numerical model to simulate pumping test data in a vertically stratified confined aquifer located in Bahariya, Egypt. The model integrates MODFLOW 6 and FloPy in a reproducible workflow and incorporates both pumping and recovery phases, enabling the estimation of hydraulic parameters in heterogeneous conditions.

The numerical approach successfully captured the aquifer's observed behaviour, with the root mean square error between simulated and observed drawdown reaching 0.01 m. The hydraulic conductivity estimates ranged from $5.5 \cdot 10^{-5}$ to 27.2 $m \cdot d^{-1}$ across stratified sandstone, limestone, and shale layers, which aligns with values reported for comparable hydrogeological units in the Bahariya region. The estimated specific storage of $4 \cdot 10^{-8}$ m^{-1} is also consistent with typical values for confined aquifers.

Sensitivity analysis showed that vertical grid refinement has the most significant influence on simulation accuracy. Coarser vertical layering increased residual error by nearly 90% and reduced the estimated radius of influence by approximately 9%. Conversely, adjusting time discretisation had minimal impact on model accuracy but substantially reduced computation time – up to 85% faster in the coarser case. These findings underscore the need to prioritise spatial refinement, especially near well screens, while allowing more flexibility in temporal resolution during early simulation phases.

Unlike analytical methods that assume homogeneity, the proposed model explicitly considers lithological layering and thus provides more reliable estimates of aquifer parameters for complex geological settings. The model's ability to integrate calibration, validation, and sensitivity analysis in a single script-driven framework supports its application to both research and practical hydrogeological assessments.

Future work should aim to overcome some of the model's assumptions, particularly the lateral uniformity of hydraulic properties within each layer. Incorporating horizontal heterogeneity using unstructured or 3D grids may further enhance accuracy in more complex hydrogeological units. In addition, the inclusion of wellbore storage effects and unsaturated zone processes would expand the model's applicability to shallow or partially confined aquifers. These results reinforce the suitability of numerical modelling for estimating hydraulic parameters in vertically layered confined aquifer systems, particularly where analytical methods fall short due to assumptions of homogeneity.

CONFLICT OF INTERESTS

The authors declare that they have no conflict of interest.

REFERENCES

Bakker, M. *et al.* (2016) "Scripting MODFLOW model development using Python and FloPy," *Groundwater*, 54(5), pp. 733–739. Available at: https://doi.org/10.1111/gwat.12413.

Bedekar, V., Scantlebury, L. and Panday, S. (2019) "Axisymmetric modeling using MODFLOW-USG," *Groundwater*, 57(5), pp. 772–777. Available at: https://doi.org/10.1111/gwat.12861.

Boulton, N.S. (1963) "Analyse of data from non-equilibrium pumping tests allowing for delayed yield from storage," *Proceedings of the Institution of Civil Engineers*, 26(3), 6693, pp. 469–482. Available at: https://doi.org/10.1680/iicep.1963.10409.

Chen, C. and Jiao, J.J. (1999) "Numerical simulation of pumping tests in multilayer wells with non-Darcian flow in the wellbore," *Groundwater*, 37, pp. 465–474. Available at: https://doi.org/10.1111/j.1745-6584.1999.tb01126.x.

Cooper, H.H. and Jacob, C.E. (1946) "A generalized graphical method for evaluating formation constants and summarizing well-field history," *Eos, Transactions American Geophysical Union*, 27(4), pp. 526–534. Available at: https://doi.org/10.1029/TR027i004 p00526.

Driscoll, F.G. (1986) *Groundwater and wells*. 2nd edn. St. Paul: Johnson Screens.

Freeze, R.A. and Cherry, J.A. (1979) *Groundwater*. Englewood Cliffs: Prentice-Hall.

Gunawardhana, L.N. *et al.* (2021) "Analytical and numerical analysis of constant-rate pumping test data considering aquifer boundary effect," *Environmental Earth Sciences*, 80(17), 543. Available at: https://doi.org/10.1007/s12665-021-09833-x.

Hamdan, A.M. and Sawires, R.F. (2013) "Hydrogeological studies on the Nubian sandstone aquifer in El-Bahariya Oasis, Western Desert, Egypt," *Arabian Journal of Geosciences*, 6(5), pp. 1333–1347. Available at: https://doi.org/10.1007/s12517-011-0439-8.

Hantush, M.S. and Jacob, C.E. (1955) "Non-steady radial flow in an infinite leaky aquifer," *Transactions, American Geophysical Union*, 36(1), pp. 95–100. Available at: https://doi.org/10.1029/TR036i001p00095.

Harbaugh, A.W. (2005) *MODFLOW-2005, the U.S. Geological Survey modular ground-water model – The ground-water flow process. Techniques and Methods 6–A16*. Reston: U.S. Geological Survey. Available at: https://doi.org/10.3133/tm6A16.

Harbaugh, A.W. *et al.* (2000) *MODFLOW-2000, the U.S. Geological Survey modular ground-water model: User guide to modularization concepts and the ground-water flow process. Open-File Report 00–92*. Reston: U.S. Geological Survey. Available at: https://doi.org/10.3133/ofr200092.

Hughes, J.D. *et al.* (2024) "FloPy workflows for creating structured and unstructured MODFLOW models," *Groundwater*, 60(1), pp. 124–139. Available at: https://doi.org/10.1111/gwat.13327.

Langevin, C.D. (2008) "Modeling axisymmetric flow and transport," *Groundwater*, 46(4), pp. 579–590. Available at: https://doi.org/10.1111/j.1745-6584.2008.00445.x.

Langevin, C.D. *et al.* (2017) *MODFLOW 6 modular hydrologic model: U.S. Geological Survey Software*. [Computer software]. Available at: https://doi.org/10.5066/F76Q1VQV.

Neuman, S.P. (1972) "Theory of flow in unconfined aquifers considering delayed response of the water table," *Water Resources Research*, 8(4), pp. 1031–1045. Available at: https://doi.org/10.1029/WR008i004p01031.

Neuman, S.P. (1975) "Analysis of pumping test data from anisotropic unconfined aquifers considering delayed gravity response," *Water Resources Research*, 11(2), pp. 329–342. Available at: https://doi.org/10.1029/WR011i002p00329.

Saqr, A.M. *et al.* (2023) "Optimal solution for increasing groundwater pumping by integrating MODFLOW-USG and particle swarm optimization algorithm: A case study of Wadi El-Natrun, Egypt," in X. Chen (ed.) *Proceedings of the 2022 12th International Conference on Environment Science and Engineering (ICESE 2022)*. Beijing, China, 2–5 Sep 2022. Singapore: Springer. Available at: https://doi.org/10.1007/978-981-99-1381-7_6.

Saqr, A.M. *et al.* (2025) "Integrating MODFLOW-USG and Walrus optimizer for estimating sustainable groundwater pumping in arid regions subjected to severe drawdown," in S. Ali and A. Negm (eds.) *Groundwater in Developing Countries. Case Studies from MENA, Asia and West Africa*. *Springer Water*, pp. 127–143. Cham: Springer. Available at: https://doi.org/10.1007/978-3-031-79122-2_5.

Theis, C.V. (1935) "The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground-water storage" *Eos, Transactions American Geophysical Union*, 16(2), pp. 519–524. Available at: https://doi.org/10.1029/TR016i002p00519.

Wu, Y.X. *et al*. (2017) "Semi-analytical solution to pumping test data with barrier, wellbore storage, and partial penetration effects," *Engineering Geology*, 226, pp. 44–51. Available at: https://doi.org/10.1016/j.enggeo.2017.05.011.

Zhu, Q. *et al*. (2020) "Optimization strategies for in situ groundwater remediation by a vertical circulation well based on particle-tracking and node-dependent finite difference methods," *Water Resources Research*, 56(11). Available at: https://doi.org/10.1029/2020WR027396.