Ain Shams Engineering Journal 15 (2024) 103127

Contents lists available at ScienceDirect

Ain Shams Engineering Journal

journal homepage: www.sciencedirect.com

Full Length Article

Integrating of Bayesian model averaging and formal likelihood function to enhance groundwater process modeling in arid environments

Ahmad Jafarzadeh [a,*], Abbas Khashei-Siuki [b], Mohsen Pourreza-Bilondi [b], Kwok-wing Chau [c]

[a] Civil Engineering and Architecture Department, Engineering Faculty, University of Torbat Heydarieh, Torbat Heydarieh, Iran
[b] Water Engineering Department, Faculty of Agriculture, University of Birjand, Birjand, Iran
[c] Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China

ARTICLE INFO

Keywords:
Ensemble Groundwater Modeling
Mesh Less
Residual Error Assumptions
Heteroscedasticity

ABSTRACT

Predictive uncertainty has influenced by traditional assumptions about the residual error. This study attempts to perform an uncertainty analysis of ensemble groundwater modeling through Bayesian Model Averaging- BMA in conditions that these assumptions are violated. This study hired a framework accompanied by BMA to generate an anticipative inference of numerical groundwater contents with non-stationary, dependent, and non-Gaussian errors. Groundwater levels were numerically simulated using three different methods for an arid aquifer in Iran. Subsequently, the BMA approach generated an improved estimate of groundwater levels by incorporating various likelihood contexts (i.e., formal and informal) to address assumptions related to residual errors. Results showed that the formal likelihood function deals with residual assumptions well, primarily for stationary and normality. Additionally, the results of the uncertainty analysis revealed that the formal function-based BMA outperforms the informal function-based BMA. Furthermore, the final predictions generated by the formal function-based BMA are comparable to the outputs of the Mesh free method in terms of RMSE.

1. Introduction

Uncertainty analysis is an essential part of any hydrogeology study where there is a porous media with high complexities [45]. In these mediums, several uncertainty sources affect the reliability of groundwater model prediction, in particular numerical models. These sources may be related to conceptual models, parameters estimation, and existing bias of input data. There is a widespread literature in groundwater uncertainty background with different purposes: parameter and recharge uncertainty [10,34], geological structure and parameter uncertainty (e.g., [14,46,16]), parameter and input data uncertainty (e.g., [49,26]), and conceptual model uncertainties [31,37].

The literature, from above point of view, suggests two key conclusions: First, direct uncertainty assessment towards integrated approaches combining all different uncertainty sources using a unique framework. Second, more related studies have broadly discussed the uncertainty of input data, parameters, and conceptual models, but mathematical uncertainty coming from numerical methods (i.e., solver) has received less attention. This fact that the numerical uncertainty is often associated with a high level of technical complexity has been approved and reported frequently [49]. As a results, although the analysis of numerical model uncertainty can provide valuable insights, making many researchers (e.g., [32,20,6,33,50]) have opted to relax this aspect of uncertainty. How can this issue can be resolved, actually? Further, there are yet some challenge, epistemic uncertainty due to incomplete knowledge [13].

The related studies have reported that the probability and evidence theories are possible response to epistemic uncertainty [38]. Nevertheless, this uncertainty and its coincident assessment with others cannot be completely eliminated, instead, ensemble modeling context such as probabilistic Bayesian Model Averaging (BMA) can quantify diverse uncertainty sources. BMA's ability to incorporate uncertainty has made it a valuable tool for decision-making and policy analysis in many fields especially groundwater modeling field [22,42]. The current version of BMA is equipped with a more efficient a Markov Chain Monte Carlo (MCMC)-based algorithm called DiffeRential Evolution Adaptive Metropolis (DREAM), which enhances exploring the parameter space and improving convergence. BMA's specifications lead most last researchers to develop comprehensive frameworks to address various sources of uncertainty, with a particular emphasis on jointly examining

* Corresponding author.
E-mail addresses: mnt.jafarzadeh@torbath.ac.ir (A. Jafarzadeh), abbaskhashei@birjand.ac.ir (A. Khashei-Siuki), mohsen.pourreza@birjand.ac.ir (M. Pourreza-Bilondi), dr.kwok-wing.chau@connect.polyu.hk (K.-w. Chau).

https://doi.org/10.1016/j.asej.2024.103127
Received 11 May 2024; Received in revised form 3 October 2024; Accepted 16 October 2024
Available online 28 October 2024

the uncertainties of input and influencing parameters. Han and Zheng [18] proposed a holistic uncertainty approach to depict the quality of watershed water, while Mustafa et al. [31] integrated BMA technique with PMWIN (Processing MODFLOW for Windows) to simulate groundwater fluctuation. In conclusion, literature confirmed that BMA can integrate diverse uncertainty sources, even epistemic one, but a practical and reliable response for numerical model uncertainty has not yet been raised. Further, relevance of BMA in quantifying residual error uncertainty (discussed in following) has not examined?

A state-of-the-art discussion about the residual errors' effect on the final prediction has appeared in recent years. Literature cited a challenging debate about applying the concept of likelihood functions (i.e., formal and informal) to determine predictive uncertainty accompanied by residual errors through BMA (e.g., [30,47]). Pioneer researchers of informal function (e.g., [11,12]) acknowledged that the residual errors' probability density function (pdf) have a pre-specified statistical form as default: 1- they have been well distributed with a Gaussian form in which mean is zero and variance is always constant, 2- they are away from time autocorrelation, and 3- they possess an inherent stationary. These characteristics create the required form of likelihood function to derive parameter estimation. For example, Standard Least Square (SLS) function explains the density of independent residual errors in Gaussian pdf (e.g., informal context). While, Kuczera [26] and Sorooshian and Dracup [44] have also violated these assumptions. For example, Kuczera [26] has reported that residual errors are non-Gaussian and non-stationary and indicated that the least-squares do not provide an acceptable explanation of parameter uncertainty. Also, Sorooshian and Dracup [44] demonstrated that as the streamflow discharge increases, residual error variance propagates (i.e., heteroscedasticity emerges). However, for better judgment, a flexible likelihood function is required to address the effectiveness of both likelihood functions simultaneously. Schoups and Vrugt [41] suggested an adjustable structure to identify the residual error assumption along with parameter and total uncertainty in rainfall-runoff modeling in five catchments of the US. This function can relax traditional error assumptions by treating heteroscedasticity and non-normality. Also, formal and informal likelihood functions can be monitored through a few adjustments. Their findings revealed this likelihood function has flexibility and significant ability to deal with some situations where residual errors show temporal correlation, heteroscedasticity (non-constant variance), and non-Gaussian distribution concurrently. Some other works, such as Smith et al. [43] and Nourali et al. [35], Olyaei, and Karamouz [36], Choi et al., [11], and Liu et al. [27] tried to account for the residual error assumptions through some separate likelihood function. For instance, Smith et al. [43] presented a framework to examine four typical widely used likelihood functions and suggested some justifications to address likelihood function assumptions in dry catchments. Also, Mustafa et al. [31] have examined just the heteroscedasticity along with other uncertainty sources and dismissed non-normality and correlation. As a result, residual error uncertainty may extremely influence the final prediction and its specifications may be quantified through a suitable likelihood function along with a robust uncertainty tool, such as BMA. However, coincident dealing with all residual error assumptions in uncertainty analysis of groundwater modelling is generally dismissed. Therefore, this question is raised in mind about how to incorporate the formal likelihood functions with BMA for uncertainty assessment of groundwater modeling.

Finally, after reviewing all comparative studies, it was found that there are limited studies in applying formal likelihood functions involved with BMA for uncertainty assessment and enhancing predictions in groundwater modeling. The Han and Zheng [18] accomplished an experiment (an especial and rare study) to couple a formal likelihood function with BMA to concurrently address the uncertainty of the water quality model along with error model. However, the combination of formal function and BMA technique in uncertainty assessment of groundwater level fluctuations has been neglected. Further, literature indicated that numerical model uncertainty has been generally dismissed. From this perspective, this study is one of the first attempts to address this part uncertainty by employing and joining three different numerical methods just like applied strategy for conceptual model uncertainty. Recent research efforts, Jafarzadeh et al. [20] and Jafarzadeh et al. [21], have shed light on this aspect of uncertainty and underscored the importance of including it in the overall analysis. However, they did not point out to coincident analysis of residual error uncertainty with numerical one and applicability of BMA was ignored. The major contribution of this effort is to investigate the effectiveness of the formal likelihood function proposed by Schoups and Vrugt [41] in the uncertainty assessment of groundwater fluctuations through BMA in relevance of multi numerical models. Specially, this study affair to overcome some mentioned research bugs: 1- how to account for the mathematical uncertainty coming from numerical methods? 2- how to set up the all residual error assumptions in uncertainty analysis of groundwater process? 3- how much the benefits of BMA' application in enhancing total prediction of groundwater fluctuations? This study does not aim to present a new formal function, but it is the first attempt to apply formal functions-based BMA in the numerical modeling of groundwater levels. Also, the outputs of three different numerical models were considered as input models of the BMA framework, which others have yet to discuss. The current study offers an applied framework to generate an enhanced skilful groundwater prediction through BMA, equipped with the formal likelihood function. Also, our study performs an uncertainty assessment of mathematical models in a real-world case study in an arid region.

2. Methods

2.1. Study area

Birjand, the center of Southern Khorasan in the east of Iran (See Fig. 1), with low quantity of annual rainfall is (<100 mm), has an arid climate pattern, and its temperature (24.5 centigrade) causes a considerable amount of potential evaporation (i.e., 2600 mm per year). The industry's inadequate growth has made the main focus of regional development turn to the agriculture section, leading to groundwater overexploitation [21,33]. This aquifer has various operational challenges and limitations that make the planning and water resources management activities especially a challenge. Hence, assessment and knowledge of the groundwater process in this region are critical. It can be expressed that the presented framework is an appropriate approach that can be applied to other arid groundwater modeling problems.

2.2. Groundwater modeling

This part explains the methodology for groundwater flow modeling. The conceptual model (and its characteristics), and numerical models (their formulation and validity), are introduced and described in detail.

2.3. Conceptual model

Since the uncertainty assessment of Birjand's groundwater resources has been previously in some studies, this study builds the conceptual model based on their findings. Hamraz et al. [15] defined a conceptual model and examined its different uncertainty sources well and in proceed, Sadeghi-Tabas et al. [40] modified and boosted their model. This paper appropriately constructed the geology structure with a one-layer in which thickness is ranges from 8 m to 255 m, and it includes some input and output underground pathways at which groundwater levels during simulation process take the constant values (Dirichlet conditions). Also, some others key features including, surface recharge, drain flows, and return flow are imposed on conceptual models. Moreover, a sever annually drawdown took place (i.e., 0.7 m) due to several wells drilled into aquifer. Besides, multiple observation wells (e.g., Piez212 and Piez760) measure the fluctuation of groundwater level (see Fig. 2). Ultimately, 17 unique regions were added to conceptual model in order

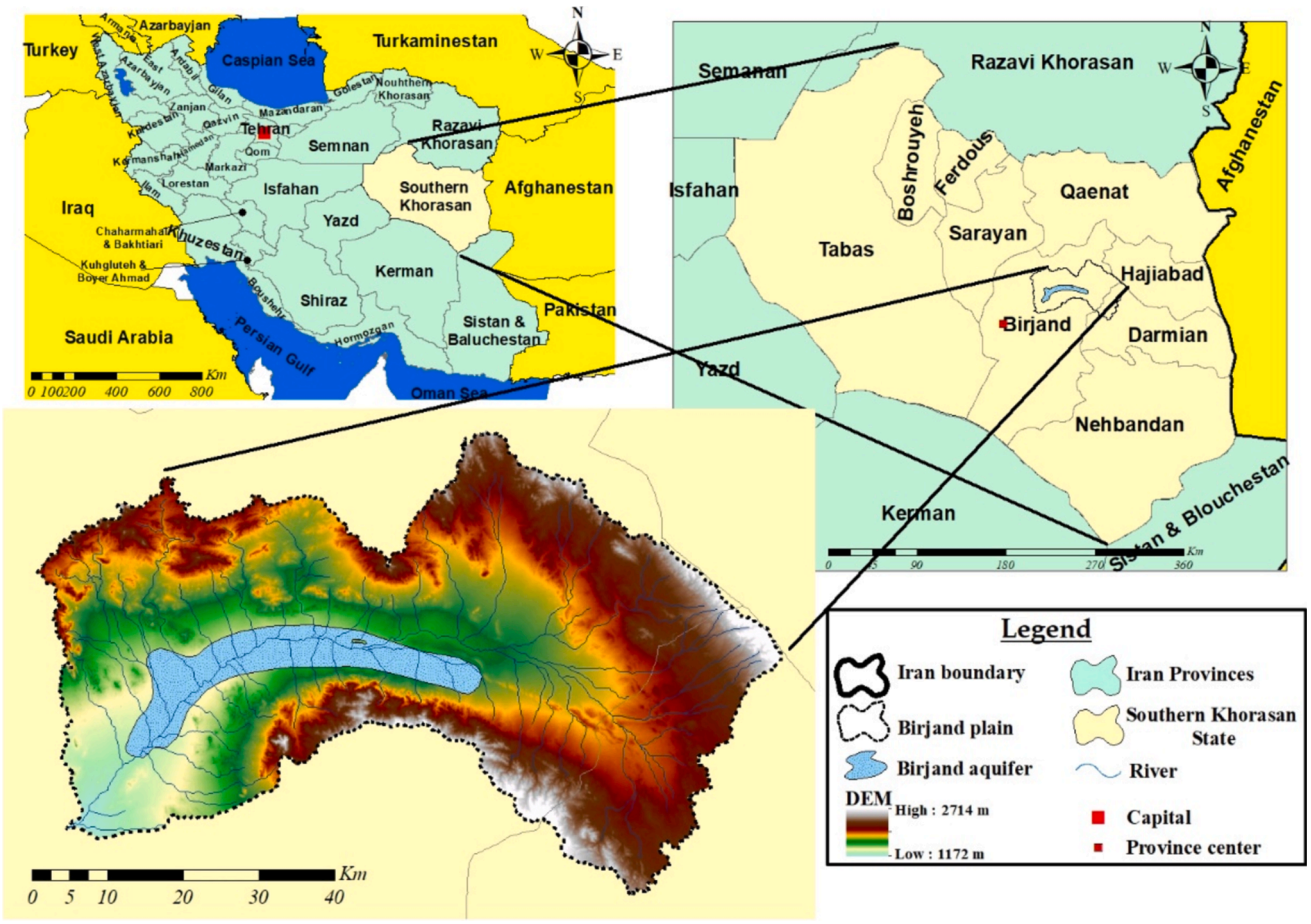


Fig. 1. Location of Birjand Plain and its aquifer: In this figure, Birjand Plain and its aquifer have been located on the left bottom slide; the right slide belongs to Southern Khorasan Province boundary of Iran and its international neighbors.

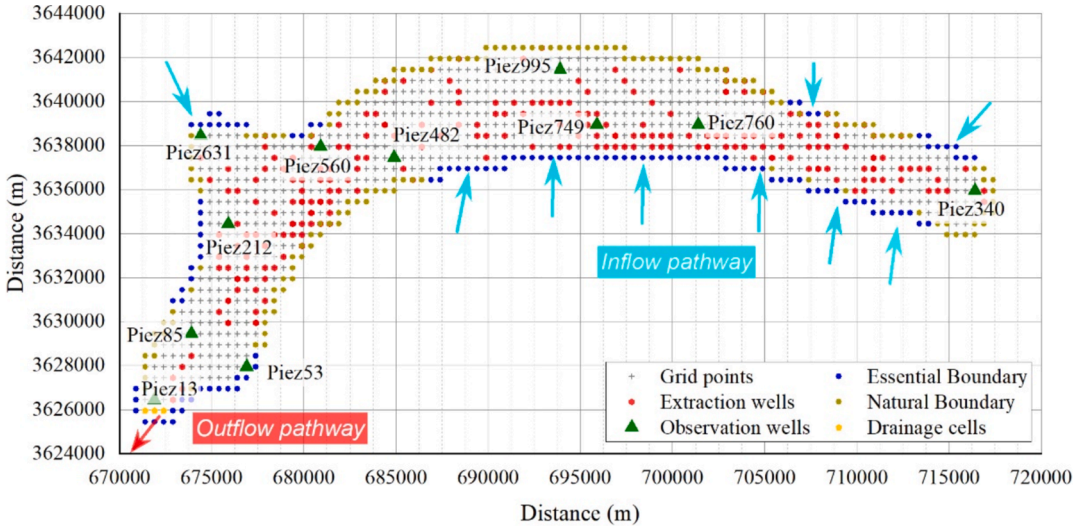


Fig. 2. Schematic representation delineating the components of the grid model used in the study. This includes essential (constant) and natural (no-flow) boundaries, extraction wells, observation wells, inflow and outflow pathways, and drainage locations, providing a comprehensive overview of the model's geometry.

to explain the anisotropy and spatial heterogeneous of hydrodynamic parameters (Fig. 3). All used datasets were implanted on a mesh with regular horizontal and vertical distance (500 m) and 1159 nodes.

2.4. Numerical modeling

Based on numerical methods' strategy for solving the Partial differential equations (PDEs), these methods can be categorized into two groups: strong-form and weak-form. In strong-form methods, such as Finite Difference (FD), an approximate function (i.e., Taylor series) is

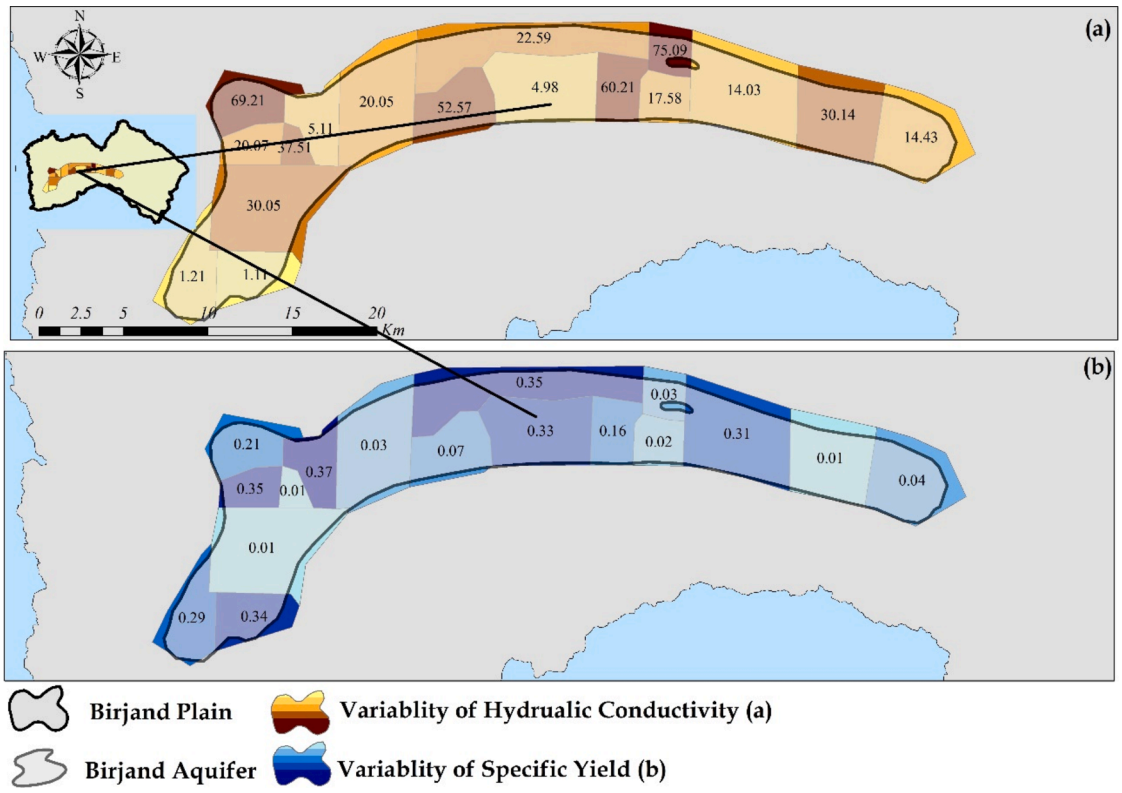


Fig. 3. Spatial conceptual model of various specific yield zones (panel a, using an orange palette) and hydraulic conductivity zones (panel b, using a blue palette) in the Birjand aquifer area. This figure depicts the spatial pattern of the parameters elaborated by Sadeghi-Tabas et al. [40].

required to discrete PDEs directly. While, in weak-form ones, such as Finite Element (FE) and Meshfree (Mfree), using the integral operation, the order of the derivatives is decreased, and PDEs' form changed from strong to weak [1,2,4,8,28,29]. The weak form-based numerical techniques are usually from Weighted Residual Methods (WRMs) in which an arithmetic act is applied to zero the weighted error in an integral context. In this concept, an essential position was defined for weight functions and different approaches exist for each numerical method such that Galerkin and Petrov- Galerkin were generally hired, respectively, for FE and Mfree methods.

2.5. Finite Difference (FD)

The groundwater process can be physically descried at 2D domain:

$$K.h\frac{\partial^2 h}{\partial x^2}+K.h\frac{\partial^2 h}{\partial y^2}+Q(i,j)=S_y\frac{\partial h}{\partial t}.$$
$$where,$$
$$\left(Q(i,j)=q_d+\sum_{i=1}^{n}Qc_i\delta(x_o-x_i,y_o-y_i)\right) \quad (1)$$

where t, h, K, Sy denote time (day), potential head (m), hydraulic conductivity (m/day), and specific yield (dimension less). Additionally, δ, q_d, and Qc represent, respectively, Dirac Delta function, distributed surface force (i.e., recharge or evaporation, m/day), and concentrated force for extraction/injection wells (i.e., source or sink context, m^3/day). Further, the boundary conditions are given by:

$$\begin{aligned}\frac{\partial h}{\partial \Gamma t}&=\frac{q_t}{K}\Rightarrow on\ \ \Gamma=\Gamma t\\ h(x,y,t)&=\overline{h}\Rightarrow on\ \ \Gamma=\Gamma u\\ h(x,y,o)&=h_0\Rightarrow on\ \ \Omega\end{aligned} \quad (2)$$

where, the initial and constant values are denoted by $h_0, \overline{h}$, so the global boundary (Γ) is split to essential and natural boundaries respectively (i.e., Γu and Γt indicating Dirichlet with constant head and Neuman conditions). While qt reflects the known inflow (m^3/day), and, Ω reflects the aquifer domain.

Taking the Taylor series approximation and applying a change of variable ($v = h^2$) the Eq.1 simplify is solved and the approximated solution was presented as following:

$$v_{ij}^{t+1}=\left(\frac{1}{\omega+1}\right)^*\left(\left(\frac{v_{i+1j}^{t+1}+v_{i-1j}^{t+1}+v_{ij+1}^{t+1}+v_{ij-1}^{t+1}}{4}\right)+\left[\omega^* v_{ij}^{t}\right]+\left[\frac{Q_{ij}}{4^*K}\right]\right).$$
$$where,$$
$$\omega=\frac{S_y{}^*a^2}{4^*K^*dt.\sqrt{v_{ij}^{t}}}, a=\Delta x=\Delta y \quad (3)$$

Where, dt indicates time step (day) and Δx , Δy represent mesh size (m). Ultimately, the potential head is obtained through the use of $h = \sqrt{v}$.

2.6. Finite Element (FE)

We constructed a triangular element matrix comprising 1842 elements across the aquifer domain. The linear triangular function was hired similarly to the shape and weight functions. The approximated solution of groundwater level in Eq.1 was achieved by minimizing the integral of the weighted average of residual, a mathematic trick, and

some simplifications (such as integration by part) as following:

$$\left([G]+\frac{[P]}{\Delta t}\right).\{h^{t+1}\}=\left(\frac{[P]}{\Delta t}\right).\{h^{t}\}+\{B_L\}+\{L_L\}. \tag{4}$$

Where [G] and [P] indicate the conductance and mass matrixes, while the boundary flux and load vector are represented through {B} and {L} respectively (other symbols are defined as Eq. (2), 3). The standard form and matrix representation of above equation may be arranged as the following:

$$\begin{gathered} KU=P. \\ where, \\ K=\left([G]+\frac{[P]}{dt}\right),\ U=\{h^{t+1}\},\ F=\frac{[P]}{dt}.\{h^{t}\}+\{B\}+\{L\}. \end{gathered} \tag{5}$$

Where K is stiffness matrix, U is unknown vector and F is force vector. The more and detailed descriptions about FD, FE, and Mfree formulations can be found in Jafarzadeh et al. [20,21]).

2.7. Mesh free (Mfree)

Mfree determines the influencing nodes of the interested point through the support domain (Ω*s*). The support domain can be local or general and can hire different shapes and sizes (see Fig. 4). The local support domain (used in this study) is calculated based on the nodal distancing (*dc*) and support domain size (r_s). Besides that, the weight function is employed to define the importance degree of support domain nodes [18,24,19,25]. This research hired the quartic-spline weight function and shape function was built through Radial Point Interpolation Method (RPIM). The local support domain is divided into some partitions comprising Gauss points for more accuracy. Just like FE methods, Mfree can gain the approximated solution. A more detailed description of Mfree implementation may be found in Jafarzadeh et al. [23].

2.8. Numerical Models' validity

This work accomplished the groundwater modeling through a numerical framework (open-source script implemented in "MATrix LABoratory" MATLAB environment) introduced by Jafarzadeh et al. [23], where the verification of the developed numerical models has been carefully tested in some benchmarks case studies. They declared that the outputs of three numerical models have close deals with analytical solutions, and consequently, we can confirm their validity.

2.9. Bayesian model Averaging (BMA)

BMA, a robust ensemble modeling tools, is a consensus view of different models (numerical models) in which a probabilistic likelihood process is adopted to perform model averaging. Given measured values as $Y=[y_1^{obs},y_2^{obs},\ldots,y_T^{obs}]$ and $\{S_1,S_2,\ldots,S_K\}$ as numerical simulation, a correction stage through linear regression fit is first imposed on the inputs (i.e., S series is converted to f series). The probabilistic density of observation based on probability's law, may be approximated as:

$$p(y|f_1,f_2\ldots,f_k,Y)=\sum_{i=1}^{K}p(f_i|Y).p_i(y|f_i,Y). \tag{6}$$

Where $p(f_i|Y)$ indicates the posterior probability under input corrected member f_i (i.e., relative closeness quantity of model compared to observation, hereafter the weight), so, their sum must be unique. While, the conditional pdf of *ith* model is represented by $p_i(y|f_i,Y)$. Rafteri et al. [39] assumed that Gaussian distribution with zero average and a constant variance (σ_i) can explain this conditional PDF in forecasting context well. Therefore, the reworded version of Eq.6 can be expressed as follow:

$$p(y|f_1,f_2\ldots,f_k,Y)=\sum_{i=1}^{K}p(f_i|Y).p_i(y|f_i,Y)=\sum_{i=1}^{K}w_i.g(y|f_i,\sigma_i^2). \tag{7}$$

The input members' weights and σ_i are considered as parameters set ($\theta=\{w_i,\sigma_i,\ i=1,\ldots,K\}$) and can be inferred from data using Monte Carlo Main Chain (MCMC) based DiffeRential Evolution Adaptive Metropolis (DREAM) algorithm. DREAM algorithm is widely used (perfectly well-known) in uncertainty assessment studies, so the summarized description of this algorithm is presented here (see Fig. 5). First, the prespecified numbers of Markov chains (i.e., the sets of parameter realizations) are produced through Latin Hypercube Sampling (LHS), and their fitness is calculated. Then each chain introduces a new point

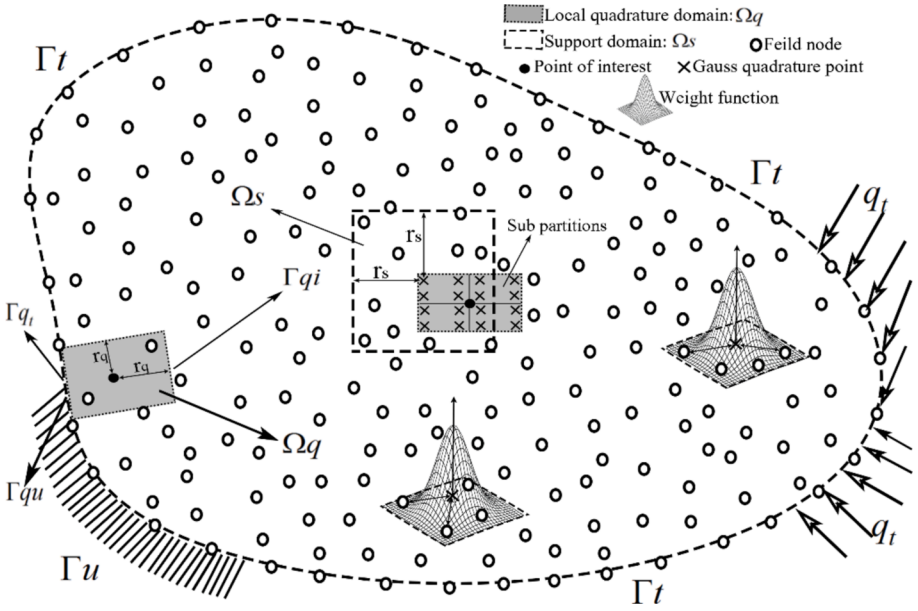


Fig. 4. A plan view of main Mfree features, reprinted from [23]. What can be understood from this figure is a conceptualization of naturally and inherently connected boundaries that define the domain: the weight function influences the interpolation, while the shape and size of the support and the local domains become crucial to the correct spatial feature representations of the model.

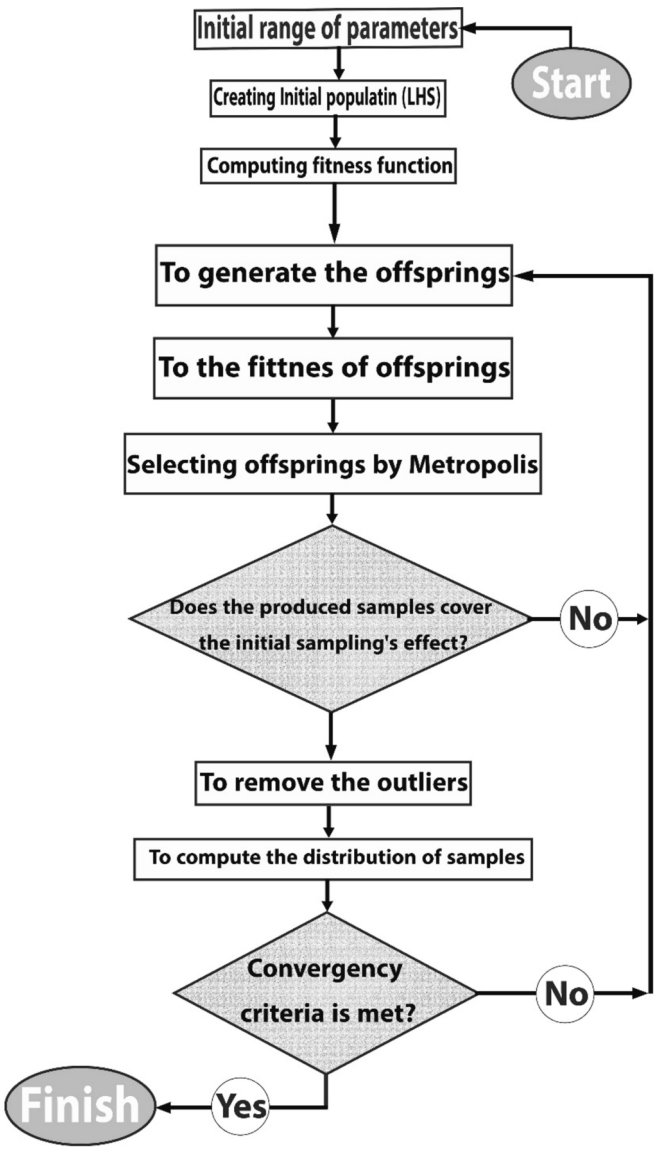


Fig. 5. Schematic flow of the DREAM algorithm. The diagram outlines certain steps followed within the dream process. Key features considered in this view include initialization, which involves setting the starting parameters; the assessment of fitness of models developed; the generation of offspring, involving the generation of new candidate solutions; the Metropolis-Hastings Algorithm that will help in realizing the selection of such candidates in light of their fitness; post-processing to refine the results; and convergence, which describes the point at which the algorithm terminates.

based on the differences between randomly chosen pairs of chains. The Metropolis ratio determines whether this candidate point is accepted or not. All chains apply this repeat to a prespecified count (burn-in period). Now the R index, proposed by Gelman and Rubin [14], is employed to determine the convergence of posterior distribution (the values smaller than 1.2 are desirable). Someone can use the samples produced after convergence to represent the posterior distribution (a comprehensive description is available on [48]). The standard BMA evaluates the parameter set's strength through the log-likelihood function proposed by Raftery et al. [39] as follow:

$$\ell(\theta = w_i, \sigma_i | Y) = \sum_{t=1}^{n} \log \left(\sum_{i=1}^{k} w_i . \frac{\exp(-\frac{1}{2}(\frac{e_i^t}{\sigma})^2}{\sigma\sqrt{2\pi}} \right). \quad (8)$$

Where e_i^t indicates the residual errors arising from ith prediction ($e_i^t = f_i^t - Y^t$).

2.10. Formal likelihood function

To relax the residual error assumptions, we joint the BMA with a formal function to consider the dependence, non-Gaussian, and stationary status of residual errors. The focus of this study is not to provide a new formal function, but it addresses the applicability of formal functions accompanied with BMA in numerical modeling of groundwater levels. As discussed previously, some related studies assessed the effect of the formal likelihood functions on the predictive uncertainty to track incompletely the residual error through separate likelihood functions. While this study, considering the research performed by Schoups and Vrugt, [41], and later verified by other studies [16,36], applied a flexible framework to consider all residual error assumptions in numerical modeling of groundwater context.

The mentioned likelihood function in Eq.8 is an informal function assuming that the residual errors are intrinsically independent (non-correlated) and mimic a Gaussian distribution with constant variance (homoscedastic). The formal likelihood function used in this study can directly parameterize residual error assumptions in the groundwater modelling using the presented methodology (proposed by [41]). Assuming the residual error is descripted as the $e_t = y_t - Y_t$, where y_t and Y_t denote time series of the simulation and observation, the correlated error may be defined as following:

$$E_t = AR(\varphi(p)).e_t. \quad (9)$$

where E_t is the correlated error, $AR(\varphi(p))$ is the pth order autoregressive model to consider correlation φ. A correlated error with any order of autoregressive can be produced using this equation. Literature indicated that residual error dependency could be explained through a polynomial autoregressive model even in complex simulation such as rainfall-runoff context. For example, Bates and Campbell [7] mentioned that polynomial AR works efficiently better than AR Moving Average (ARMA) to prevent the challenging problems in finding the optimum values of orders ARMA. Also, Schoups and Vrugt [41] showed that AR with lower order could express correlated residual error in the practical experimental. Considering more time dependence of stream flow events than groundwater level and findings of related works, the correlated error model defined in Eq.9 was hired in this study to account for the time dependence of residual error.

Also, Heteroscedasticity issue can be explained by a linear relation with simulated value (y_t) as the following:

$$\sigma_t = \sigma_0 + \sigma_1 y_t. \quad (10)$$

where σ_t is error standard deviation, σ_0 and σ_1 are the intercept and slope of above equation. The model residuals will be stable if σ_0 and σ_1 limited to minimum values. Indeed, these two parameters regulate the stationary of errors. The assumption of the linear relation (between residual variance and predicted values) was previously examined for stream flow [41], in which there are more inherent dynamics than groundwater. Therefore, applying this assumption in a study like groundwater context may also be perfectly logical. Also, some other studies, such as Mustafa et al. [31] and Mustafa et al. [32], reported formerly that the heteroscedastic error can be described by a linear model similar to Eq.10.

Now, we able to generate a residual error that its correlation and stationary represented through E_t and σ_t. As followings:

$$a_t = \frac{E_t}{\sigma_t}. \quad (11)$$

Where a_t is an independent error distributed identically. Recently, Han and Zheng [16] used some probabilities distribution sensitive to skewness and kurtosis parameters (Skew Exponential Power- 'SEP') to examine the normality of prediction error in groundwater media.

Therefore, this study according to their findings assumed that a_t follows SEP distribution function to explain the normality situation of the residuals. The SEP density function of a_t can be estimated as follows:

$$p(a_t|\beta,\xi) = \sigma_t^{-1}\frac{2\sigma_\xi}{\xi+\xi^{-1}}\omega_\beta \exp\left(-c_\beta\left|\frac{\mu_\xi+\sigma_\xi a_t}{\xi^{\mu_\xi+\sigma_\xi a_t}}\right|^{2/1+\beta}\right). \quad (12)$$

where β and ξ are the kurtosis and skewness, and σ_ξ, μ_ξ, ω_β and c_β are SEP's parameters calculated from β and ξ. Based on the SEP function, residual errors have normal distribution when β and ξ are set, respectively, to zero and one. Interested readers are referred to Schoups and Vrugt [41] for more discussion. Using mentioned procedure, the parameterization of autocorrelation, heteroscedasticity and normality of residual errors can be defined through parameter set η_e, that is, $\eta_e = \{\varphi, \sigma_0, \sigma_1, \beta, \xi\}$. Since $\ell(\theta|Y) \equiv p(Y|\theta) = p(e|\theta)$ and considering $\eta_b = \{w_i,\ i = 1, 2, \ldots, n\}$ the formal likelihood function of BMA and error model can be formulated as following:

$$\ell(\theta = \eta_e, \eta_b|Y) = \sum_{t=1}^{n}\log\left[\sum_{i=1}^{k} w_i.\left\{\sigma_t^{-1}\frac{2\sigma_\xi}{\xi+\xi^{-1}}\omega_\beta \exp\left(-c_\beta\left|\frac{\mu_\xi+\sigma_\xi a_t}{\xi^{\mu_\xi+\sigma_\xi a_t}}\right|^{2/1+\beta}\right)\right\}\right]. \quad (13)$$

Where η_b and η_e represent the parameters sets of BMA and error model. If the skewness and kurtosis parameters (β and ξ) are set, respectively, to zero and one, and $\sigma_1 = 0$ as well as $\varphi = 0$ this equation can represent the parameter set for a homoscedastic, independent, and Gaussian errors (i.e., an informal likelihood function).

2.11. Model setup

An overview of used procedure in this study is described in this section for a better understanding for readers. Valuable input. Firstly, a conceptual groundwater model was developed for the Birjand aquifer using multiple datasets, which included rainfall and evaporation data, hydrodynamic parameters, observations from extraction wells, topographical information, and relevant boundary and initial conditions. This conceptual model was subsequently input into three numerical models to simulate groundwater levels. To generate a consensus prediction of groundwater modeling, a BMA alongside a formal likelihood function was employed. Additionally, the DREAM algorithm to quantify the uncertainty associated with the weights of the groundwater models and the error models. Table 1 summarizes the parameters used in the BMA and error models, along with their prior uncertainty ranges. As previously discussed, the likelihood function defined in Equation (13) establishes the conditions for comparing the outcomes of the BMA with both formal and informal likelihood approaches. Finally, Fig. 6 illustrates the steps and strategies employed in the current study.

Furthermore, since there are a low dynamical in groundwater level context, the variance-based indexes (e.g., Kling-Gupta or Nash Sutcliff), by which model performance is evaluated, cannot provide insights into the model proficiency. Also, the correlated-based criteria are not recommended because they merely account for the percentage of observation that are explained by the predictive models. Hence, this study utilized Root Mean Squared Error "RMSE" criterion to benchmark results. Also, different uncertainty metrics such as P-factor (percentage of data bracketed by a 95 % prediction uncertainty band), R-factor (the average thickness of 95 % prediction uncertainty band divided by the measurement standard deviation), D-factor (distance or thickness of the prediction uncertainty band), and Total Uncertainty Index- 'TUI' (i.e., the fraction of P-factor per R-factor) were hired to address uncertainty quantification and its performance.

Table 1
Details of the uncertain parameters utilized in the DREAM algorithm. This table includes the prior ranges and units of the parameters employed in the BMA model (Wi) as well as the error model (φ, σ_0, σ_1, β, ξ). It aims to provide a comprehensive overview of the uncertainty analysis, highlighting the variability and assumptions underlying the modeling process.

Model	Parameter	Symbols	Range	Units
BMA model	Weights of contributing numerical models	w_i	[0, 1]	–
Error model	Autoregressive coefficient	φ	[0, 1]	–
	The intercept of Heteroscedasticity model	σ_0	[0, 1]	m
	The slope of Heteroscedasticity model	σ_1	[0, 1]	m
	The kurtosis of normal distribution	β	[-1, 1]	–
	The skewness of normal distribution	ξ	[0.1, 10]	–

3. Results and discussion

The presented discussion in this part will focus on verification of developed numerical models and their uncertainty assessment, tracking the validity of residual error assumptions, and obtaining multi-model predictions using BMA and numerical models.

3.1. Efficiency of numerical models

The obtained results of the numerical simulation in separate different techniques throughout the Birjand's arid aquifer system have been presented here. Table 2 outlines the RMSE quantity calculated for the FD, Mfree, as well as FE methods, separately for different observation wells. Also, the total RMSE of each numerical model is presented in the last row, which reveals that the Mfree method outperforms both FD and FE methods. Also, these results confirm the spatial dependency of models' skills. For example, Mfree prediction in most piezometers is closer to measured values, but in Piez749 and Piez85, the FE simulations are more accurate than Mfree ones. Moreover, the worst prediction of Mfree was for Piez995, where is away from boundaries, and complicated groundwater fluctuation take apart due to abundant extraction wells [5,9]. These results align with other groundwater case studies (e.g., [3] and with general conclusions from uncertainty analysis researches (e.g., [15]; previously carried out in same current regions).

Fig. 7 gives a visual comparison between the simulated and measured groundwater levels at select locations within the Birjand aquifer. For brevity, the comparisons from the remaining observation wells have been omitted. The results illustrate a notable agreement between the simulated fluctuations produced by the Mfree model and the actual measured groundwater levels. In contrast, the simulations derived from the Finite Element (FE) and Finite Difference (FD) methods exhibit less accuracy in capturing the observed trends. This discrepancy highlights the effectiveness of the Mfree approach in modeling groundwater fluctuations in this arid region, suggesting its potential for more reliable predictions in similar hydrogeological contexts. Overall, the visual representation underscores the importance of selecting appropriate numerical modeling techniques to enhance the understanding of groundwater dynamics.

3.2. Uncertainty assessment

Fig. 9 illustrates the evolution of the R-Gelman index [14] for parameters associated with both informal and formal likelihood functions in the Piez13 and Piez212 piezometers (note that results for the remaining piezometers are not displayed here). As depicted, all parameters influencing both the informal and formal likelihood functions converged after 50,000 runs. Notably, the parameters associated with the informal function-based BMA (Fig. 8-a) achieved convergence criteria after approximately 1,200 runs (with only two piezometers shown), whereas the parameters related to the formal function (Fig. 8-b)

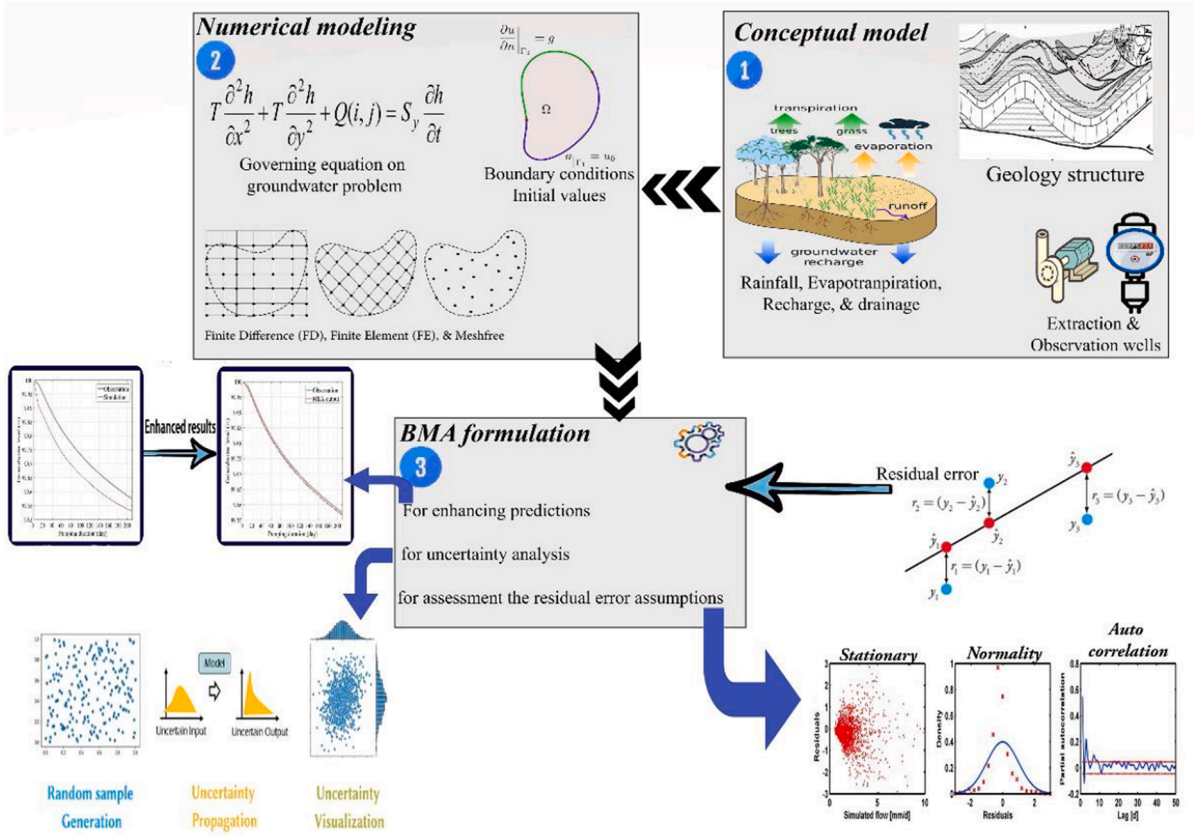


Fig. 6. Below depicts the broad overview of methodology adopted for this research. the main steps are as follows: 1. conceptual model: an idealized representation of the theoretical framework with assumptions from the groundwater system; 2. numerical modeling: application of different techniques with simulations of groundwater behaviors using numerical techniques; and 3. bma formulation: integrating the outcomes from the numerical modeling techniques into a bayesian framework so that high accuracy and reliability can be guaranteed in the prediction.

Table 2
The rmse quantity results in separate of numerical methods.

Observation wells	FD	FE	Mfree
Piez13	0.1547	0.1949	**0.0662**
Piez53	0.5163	0.2716	**0.1983**
Piez85	0.7490	**0.1585**	0.1960
Piez212	0.4104	0.2655	**0.1426**
Piez340	0.3554	0.4098	**0.1223**
Piez482	4.0149	0.4697	**0.0486**
Piez560	1.5278	0.3417	**0.2107**
Piez631	0.1738	0.1204	**0.0643**
Piez749	0.3447	**0.1487**	0.2222
Piez760	0.4961	0.2004	**0.1087**
Piez995	0.4633	0.3434	**0.1022**
Total RMSE (m)	1.3554	0.2872	**0.1477**

required around 15,000 runs to reach similar convergence. This indicates that the parameters in the formal function-based BMA necessitate a longer duration to achieve convergence compared to those in the informal function.

The violin plot of inferred values for the Piez631 piezometers, which represents both formal and informal functions, is presented to investigate the influencing parameters (Fig. 9). This plot illustrates the 95 % variation (indicated by the grey rectangle) and the median (represented by the solid circle) for each violin, providing a comprehensive view of the data distribution.

Notably, the Bayesian Model Averaging (BMA) approach appears to assign greater weights to the output from the Mfree model, suggesting its prominence in the analysis. Furthermore, the plot indicates that the uncertainty band associated with the weights of the numerical methods in the formal case is wider compared to that in the informal function, highlighting a greater variability in the formal approach.

Additionally, the inferred parameters derived from the formal likelihood function reveal that the two stationary parameters, (σ_0 and σ_1, exhibit lower values. Interestingly, the majority of the generated parameter realizations favor higher values for the first-order correlation parameter (φ_1), indicating a stronger correlation in the inferred relationships.

A summary comparison among different likelihood functions is presented in Table 3. It is argued from derived results that the formal likelihood function works better than the informal in almost all piezometers (high P-factor values). The parameters inferred using formal likelihood functions provide a more accurate representation of measurement values compared to those derived from informal likelihood functions. This distinction is effectively illustrated in Fig. 10, which displays the observed data (represented by black squares) alongside the 95 % predictive uncertainty band (depicted as a green shaded area). The figure specifically compares the total uncertainty associated with both the formal and informal likelihood functions for the Piez631 piezometer, highlighting the enhanced predictive capability of the formal approach. This plot confirms that uncertainty assessment using the formal likelihood function has more appealing outcomes, so all uncertainty metrics in formal results have been improved than in informal ones. For example, the values of the P-factor in these two likelihood functions indicate that uncertainty band with the formal likelihood function includes 91.67 % of measurements, while this band for the informal is almost 16.67 %.

Considering the derived results of this part, which can reflect the effect of employing different numerical methods to quantify mathematical uncertainty, it is noteworthy that this aspect is often overlooked

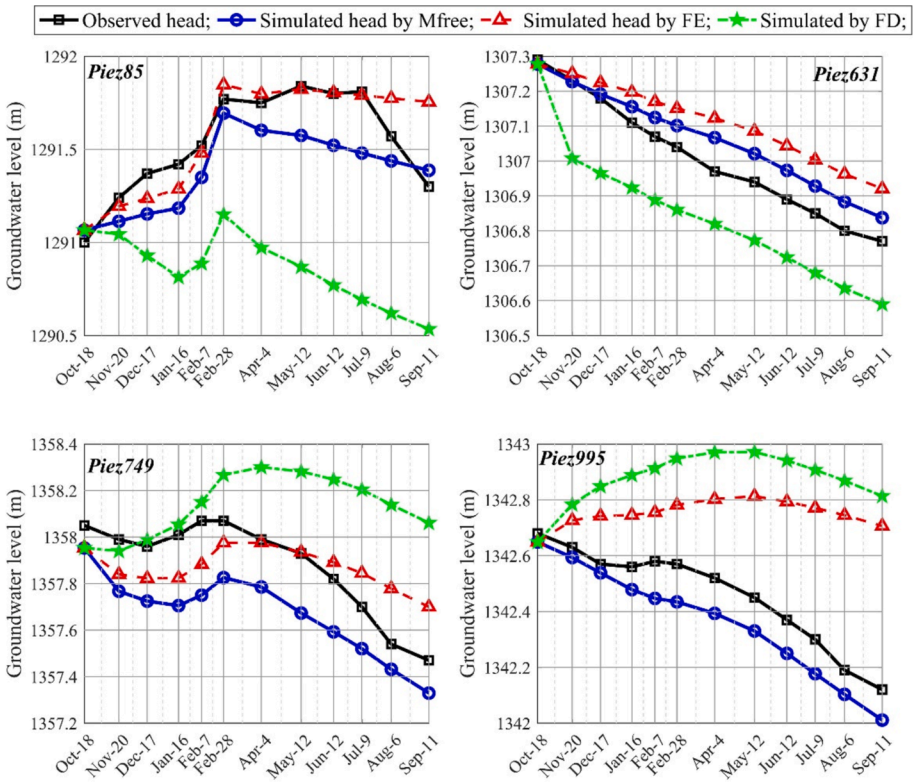


Fig. 7. Visual comparison of the accuracy of various numerical methods for simulating groundwater levels. The black square line represents the measured groundwater levels, while the blue circular line indicates the outcomes from the Mfree model. The red triangular markers correspond to the results of the Finite Element (FE) simulation, and the green star markers denote the predictions from the Finite Difference (FD) method. The simulation period spans from October 2011 to September 2012, providing a comprehensive view of the performance of each method over this timeframe.

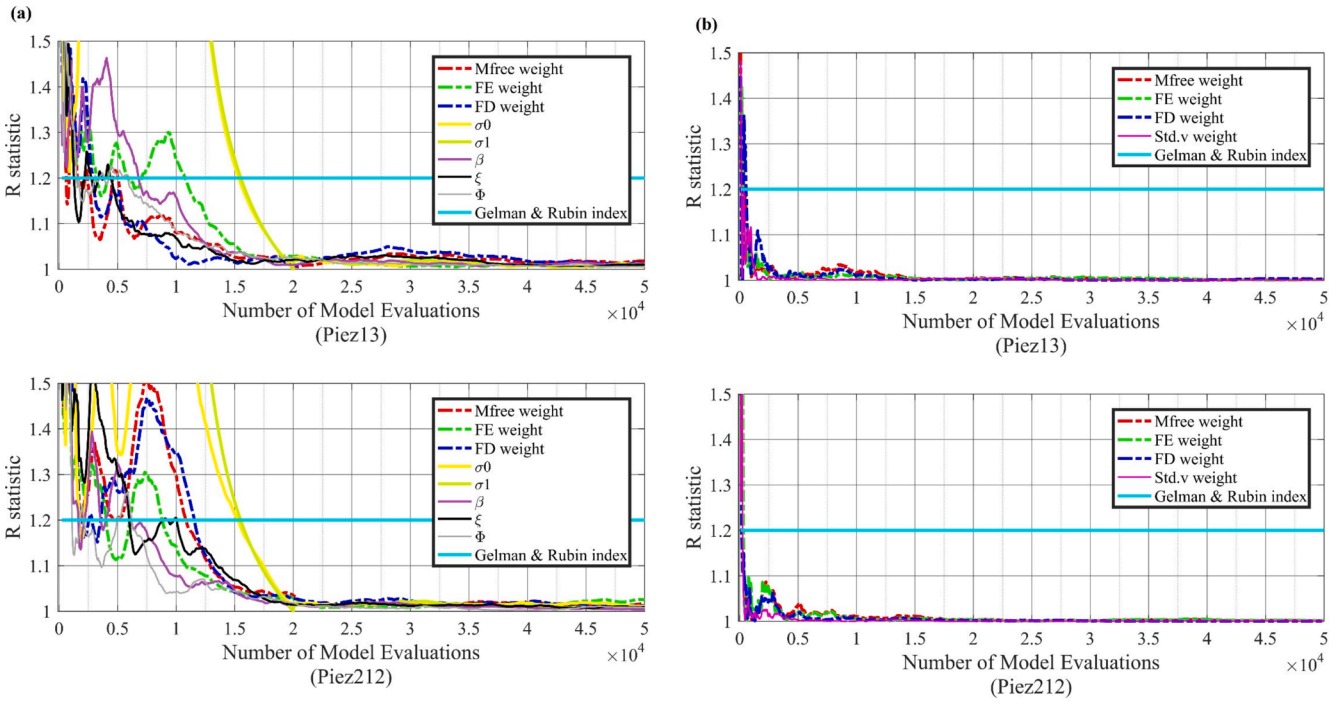


Fig. 8. Evolution of the R statistic for BMA and error model parameters, presented for (a) formal likelihood functions and (b) informal likelihood functions across different piezometers (Piez 13 and Piez 212). This figure illustrates the consistency of the models in capturing groundwater fluctuations over time, demonstrating how both formal and informal likelihood functions perform in relation to the R statistic.

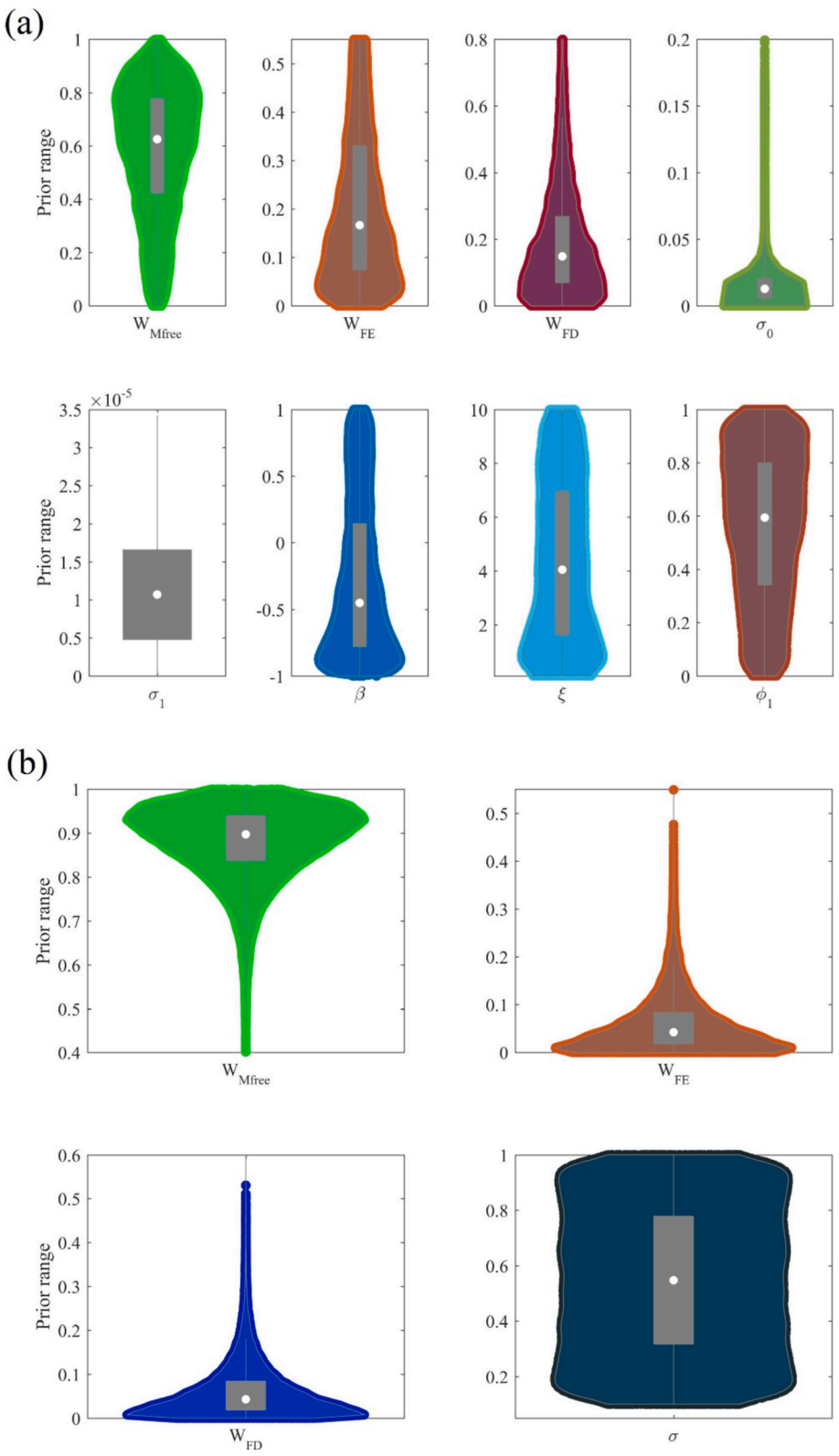


Fig. 9. Violin plots illustrating the posterior realizations generated by the BMA model for the Piez631 piezometer. Panel (a) displays the results obtained using the formal likelihood function, while panel (b) presents the outcomes from the informal likelihood function. These plots provide a visual representation of the distribution and variability of the posterior estimates, highlighting differences in the parameters estimation.

in many studies as mentioned in first paragraph of Introduction (e.g., [32,16,1,33,51]). The challenge lies in identifying and demonstrating this type of uncertainty, because it requires the implementation of a comprehensive process similar to that undertaken in this research. The results provided in section 3.1 reveal that, despite feeding a similar conceptual model into several numerical models, their performance varies significantly, confirming the existing and effects of mathematical uncertainty arising from numerical techniques. Further, the outcomes of

Table 3
Results of uncertainty performance indices, specifically the R-factor and P-factor, comparing the effectiveness of formal and informal likelihood functions. This table summarizes key performance metrics that emphasize the differences in uncertainty quantification between the two approaches, providing insights into their respective efficacy in modeling and analysis.

Piezometer	Informal likelihood function		Formal likelihood function	
	R_factor	P_factor	R_factor	P_factor
Piez13	0.40	16.67	0.93	25.00
Piez53	0.30	16.67	0.94	8.33
Piez85	0.48	8.33	1.06	33.33
Piez212	0.39	25.00	0.66	41.67
Piez340	0.34	25.00	0.82	25.00
Piez482	4.25	50.00	6.56	58.33
Piez560	0.91	41.67	3.09	50.00
Piez631	0.30	16.67	0.87	91.67
Piez749	0.52	16.67	1.32	58.33
Piez760	0.36	16.67	0.73	8.33
Piez995	0.78	83.34	1.51	91.67

different test showed that integrating multi numerical model along with residual error uncertainty could derive more appealing results indicating the positive feedback of proposed strategy for quantifying numerical modelling.

3.3. Likelihood functions assessment

This section investigates how the formal likelihood function defined in equation (13) explains the residual errors in groundwater modelling applications. Figs. 11 and 12 summarize the validity of error assumptions in BMA under formal and informal functions. Close statistical inspection of BMA's residual realizes the model discrimination in expressing residual model behaviour so that: 1: heteroscedasticity issue was controlled well in formal case (Fig. 11-a) compared to informal one (Fig. 11-d) in which increase variance quantity is a subordinate of simulated groundwater levels; 2: actual pdf of residual error in the formal case (Fig. 11-b) has more agreements with experimental one, while this result is not attained for informal function (Fig. 11-e). These results confirmed that some residual error assumptions such as stationary and normality had been violated in the informal case. Some related studies focused on rainfall-runoff (e.g., [13,17,26,27,47,51]) have reported the same violations in their findings. Although it is not possible to give a definite conclusion about correlation through current results (Fig. 11-c, f), it is clear that the temporal independence of the residues cannot be accepted by default. Further, results denoted that the proposed formal likelihood function could not control correlation well. It may be because correlation was not considered directly in the formulation of the formal likelihood function leading to the lower sensitivity of φ. In contrast, heteroscedasticity and Gaussian parameters (i.e.,$\sigma_0, \sigma_1, \beta, \xi$) regulate the stationary and normality of residual errors. From this perspective, the high density area of actual pdf is very close to the experimental one, and the density of actual pdf has been positively skewed due to inferred skewness and kurtosis parameters ($\beta = 0.71$, $\xi = 8.67$).

In proceed, groundwater fluctuations generated by BMA using different likelihood functions are analyzed visually and statistically (See Figs. 13, 14) at Piez631 and Piez995 piezometers. Results reveal that the difference of the produced potential head by BMA between formal and informal likelihood functions in terms of RMSE is very slight. The right way of deciding the fitness of two likelihood functions is likely to analyze their proficiency in maximum likelihood values instead of the RMSE criterion. For the considered simulation period and Piez631 piezometer, the groundwater level generated by BMA with the formal likelihood function received a more significant maximum likelihood (32.77) than the informal function (−11.03), which confirms the superiority of the formal likelihood function.

Further, comparing simulated groundwater fluctuations between BMA equipped to formal likelihood function and numerical models' outputs indicates that BMA produces a more accurate simulation than even a sophisticated numerical model such as the FE and even the Mfree method (see Table 2 and Fig. 15). The superiority of BMA in producing skilful groundwater fluctuation and its flexibility to relax residual error assumptions confirm the effectiveness of the proposed framework in the groundwater modelling field.

Another point that may be inferred from the results is that when the competing models' outputs include the measurement values, one can expect that BMA yields better than all models. For example, the total range of numerical simulations in the Piez631 and Piez995 piezometers cover the measured groundwater levels (i.e., the observed values are always between simulations), resulting in the formal function-based BMA provides a complete consensus simulation that is more reliable and residual error assumptions are met well. Therefore, as much as this matter is violated, the BMA performance is reduced.

The superiority of formal function over informal is not limited to maximum likelihood. Its posterior distribution for influencing parameters is well identified so that the stationary and normality of residual errors have more logical situation in the formal likelihood function than those informal ones. The posterior histograms of the influencing parameters of the formal likelihood function are demonstrated in Fig. 16 (true values are displayed in the red circle). As shown, all eight parameters are well identified. The weight distributions of numerical models (WMfree, WFE, and WFD) indicate that BMA prefers more weight for Mfree outputs to generate a more accurate simulation. Also, distribution of σ_0,σ_1 denotes that DREAM tries to remove the heteroscedasticity. Moreover, kurtosis and skewness parameters (i.e., β and ξ) represent a positive skewness and more picked density than Gaussian pdf. These results have sound agreement with the findings of other related studies [20,41].

4. Conclusion

This paper examined the effectiveness of a formal likelihood function to enhance groundwater numerical modelling through the BMA framework in the arid region of east Iran (Birjand plain). Three numerical models- FD, FE, and Mfree- were developed and employed to simulate groundwater fluctuations. Subsequently, The BMA framework was used to produce a prediction, which represent a weighted average of the outputs from the numerical models. Furthermore, a formal likelihood function was utilized to account for residual the specifications of residual error assumptions, including autocorrelation, heteroscedasticity, and normality. A different scientific examinations were numerically drawn based on a real-world investigation in the Birjand aquifer to test the proposed plan of the current study.

The implemented test exercise highlighted the BMA's flexibility in producing an enhanced weighted prediction. Also, the BMA has a good performance in quantifying and covering the uncertainty of mathematical models, which others generally dismissed. In this concept, this study revealed the effects of mathematical uncertainty and demonstrated that it cannot be overlooked due to its complexity; rather, it can be effectively explored using the appropriate framework, as proposed by this study.

The results confirmed that addressing residual error assumptions is a proper tool for discriminating between different likelihood functions. Indeed, the findings obtained Contradicted previous assumptions regarding the characteristics of residual errors and clearly demonstrated their failure to hold. This study showed that it cannot be assumed a priori that residuals follow a normal distribution, are homoscedastic, or are always uncorrelated. Therefore, it is advisable that future studies make a concerted effort to address the characteristics of residual errors. The proposed tools in this study offer the necessary capability and flexibility to apply both formal and informal approaches. Consequently, after pre-processing, it is up to the user to determine whether the

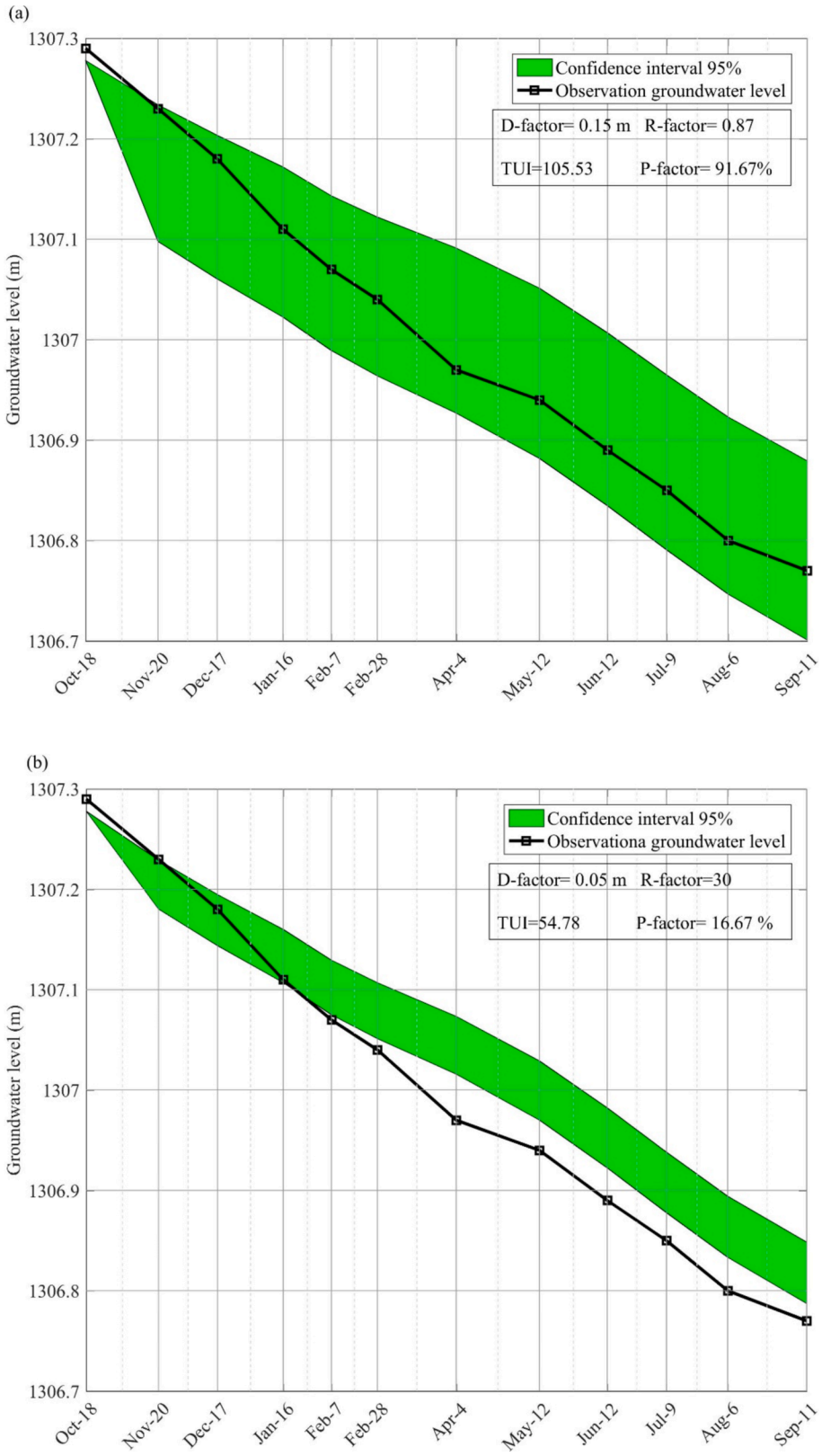


Fig. 10. 95% prediction uncertainty ranges for the Piez631 piezometer, presented in (a) formal and (b) informal likelihood functions. The plots include the observed data represented by black squares, along with the 95% predictive uncertainty band illustrating total uncertainty, shown as a green shaded area. Additionally, uncertainty performance metrics are displayed, including the D-factor, P-factor, R-factor, and TUI, to evaluate the performance of the uncertainty assessments.

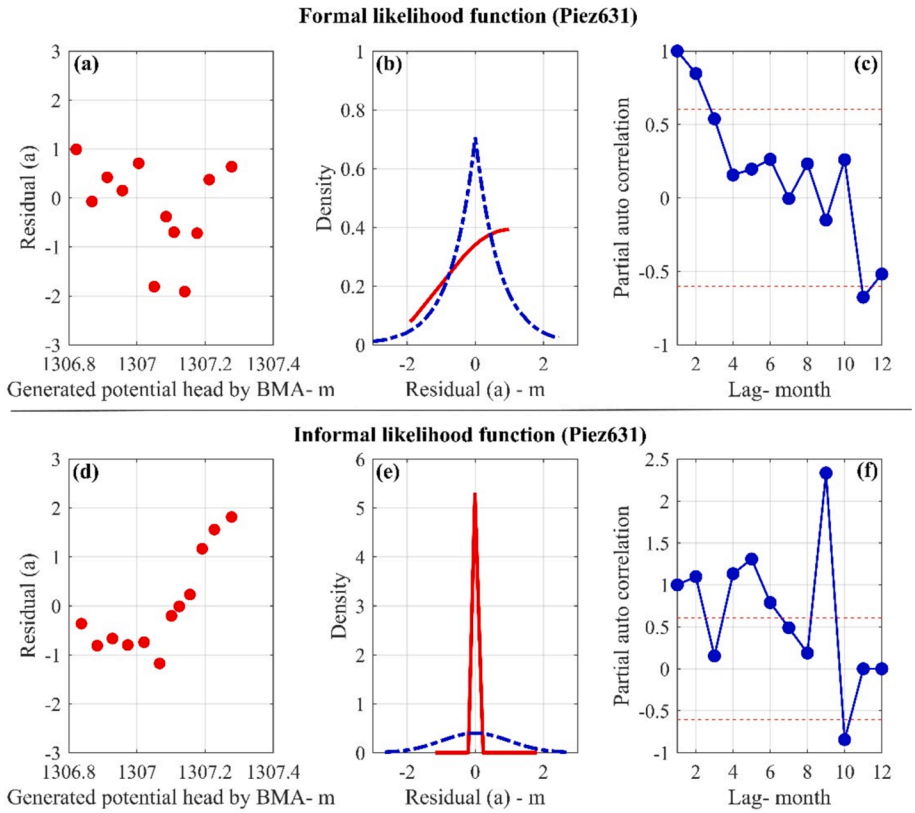


Fig. 11. Illustration of residual errors in formal (first three panels) and informal (second three panels) functions for Piez631 piezometer: (a,d) residual error as a function of simulated head, (b,e) experimental (blue dash) and actual (red line) pdf of residual, and (c,f) residual independency with 95% significant level.

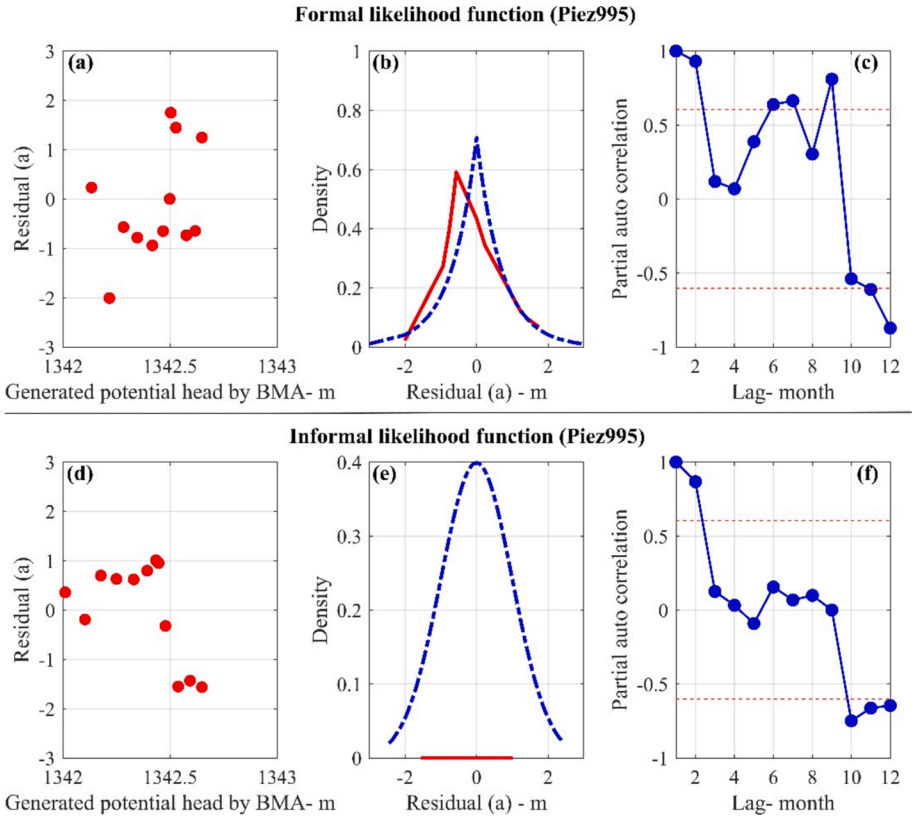


Fig. 12. Illustration of residual errors in formal (first three panels) and informal (second three panels) functions for Piez995 piezometer: (a,d) residual error as a function of simulated head, (b,e) experimental (blue dash) and actual (red line) pdf of residual, and (c,f) residual independency with 95% significant level.

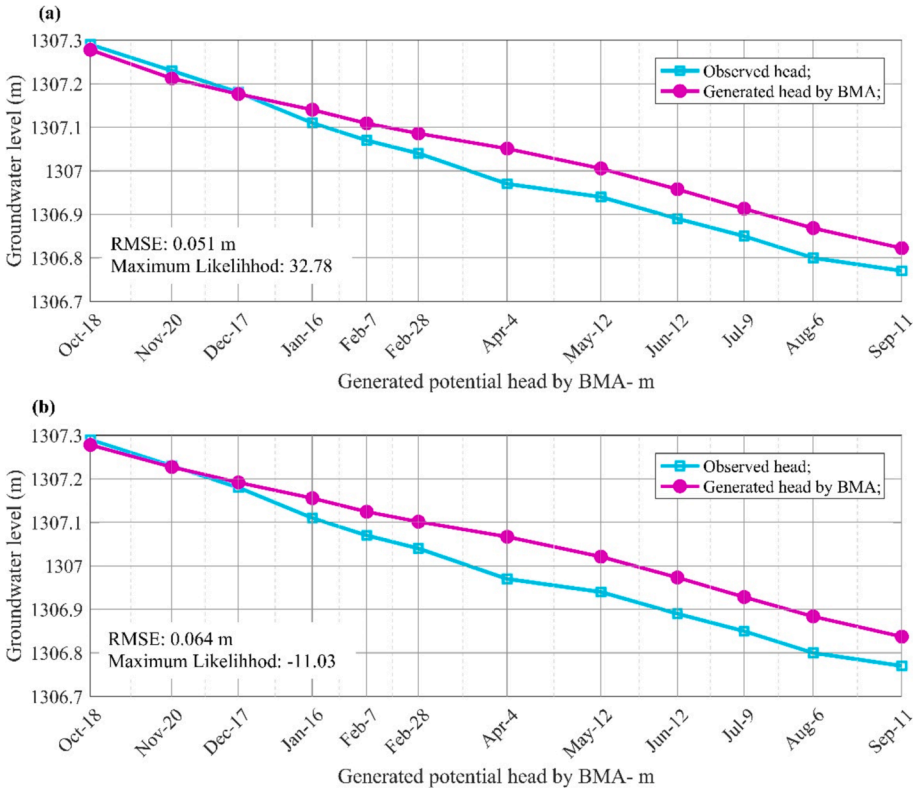


Fig. 13. Comparison of groundwater level simulations at the Piez631piezometer: (a) Top panel displays the results from the formal likelihood function, while (b) the bottom panel presents the informal likelihood function. The figure also includes the corresponding RMSE and Maximum Likelihood values for both methods.

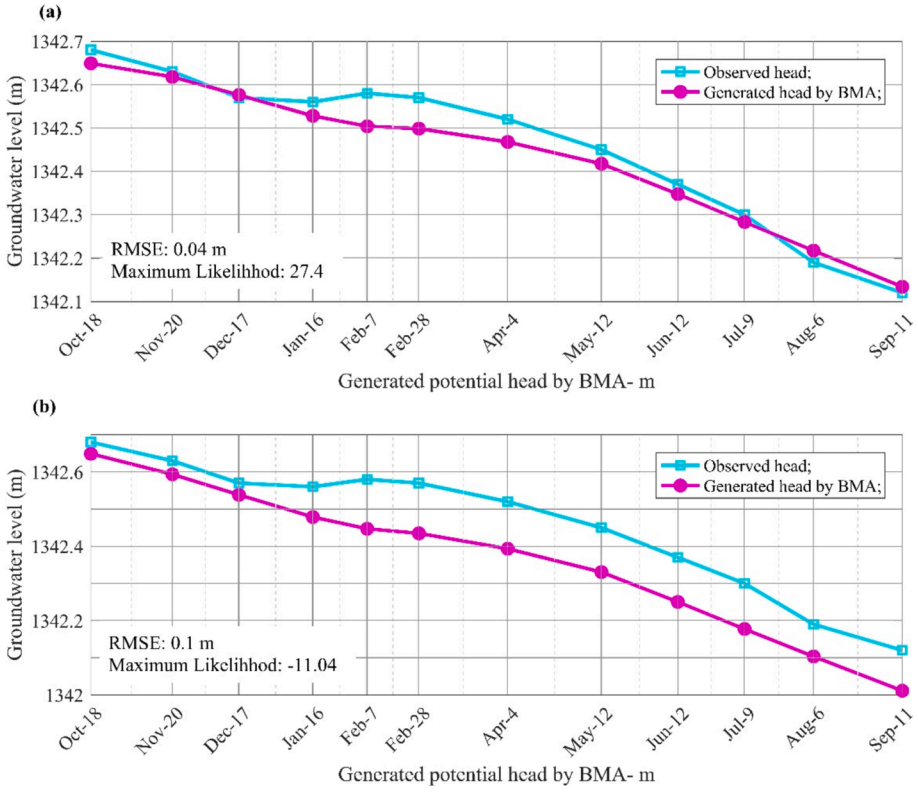


Fig. 14. Comparison of groundwater level simulations at the Piez995 piezometer: (a) Top panel displays the results from the formal likelihood function, while (b) the bottom panel presents the informal likelihood function. The figure also includes the corresponding RMSE and Maximum Likelihood values for both methods.

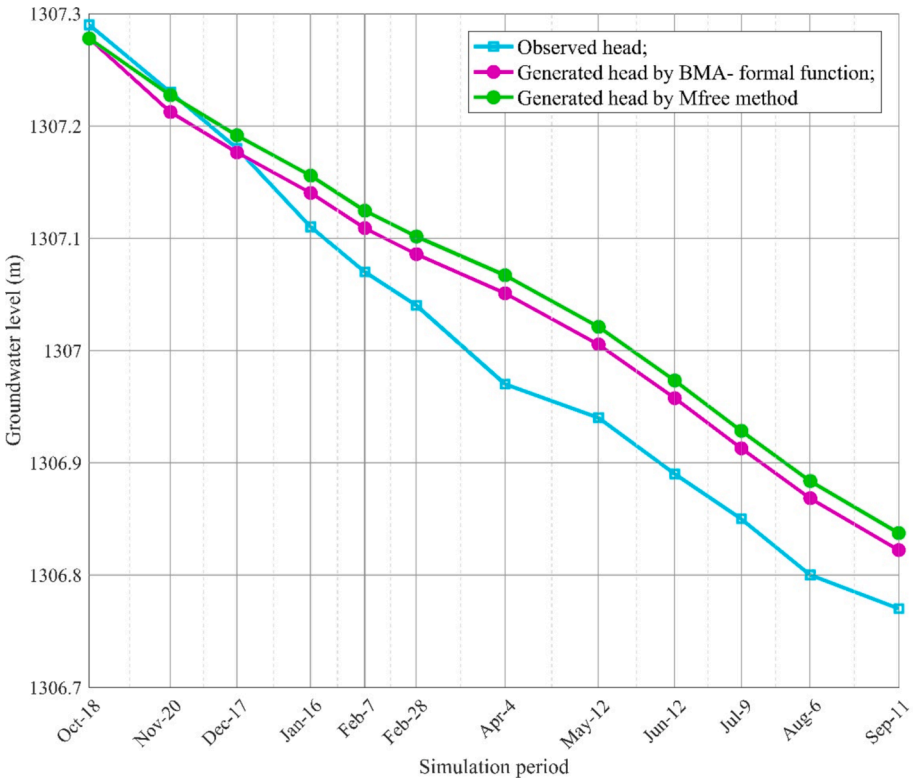


Fig. 15. Comparison of observed groundwater fluctuations (turquoise line) with predictions made using the formal likelihood function (pink line) and the Mfree method (green line) in the Piez631 piezometer.

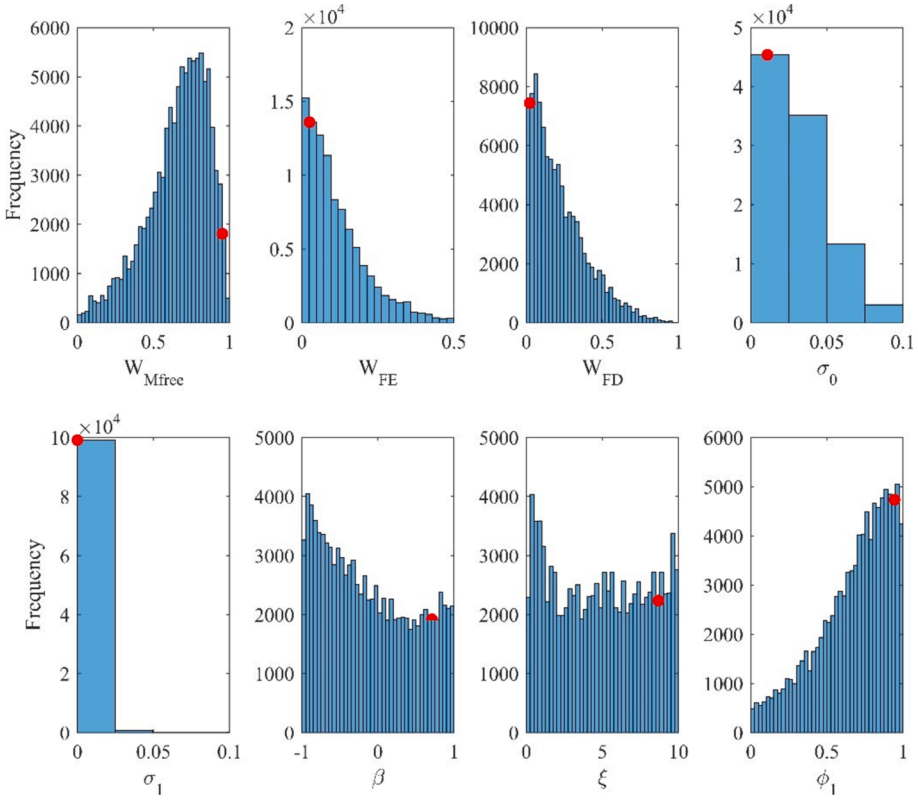


Fig. 16. Posterior histograms of error model parameters using formal-based Bayesian Model Averaging (BMA) in the Piez631 piezometer. The figure includes weight distributions of numerical models (WMfree, WFE, and WFD), heteroscedasticity parameters (σ_1 and σ_0), kurtosis (β) and skewness (ζ) parameters, as well as autocorrelation parameter (φ).

characteristics of residual errors should be incorporated into the modeling process or not.

Derived results reveal that the formal likelihood context applied in the current work can address residual assumptions well, mainly stationary and normality. Also, formal function-based BMA performs better than informal function-based BMA considering different uncertainty performance metrics such as R-factor, P-factor, D-factor, and TUI. Further, the final prediction of formal function-based BMA is comparable to FE and even Mfree's outputs.

Here our findings and recommendations for future scopes researches are listed (it is emphasized that these findings are specific to current case study, and they are obtained through limited data):

The obtained results suggest including more numerical models for tracking and discovering the uncertainty of mathematical models.

The in-depth analysis revealed that the performance of individual competing models (in this case, numerical models) varied over the simulation timeline. However, the existing BMA methodology applies constant weights to each competing model throughout the simulation period. As a result, future research will focus on enhancing the BMA framework by incorporating dynamic weights instead of using fixed weights during the simulation.

It was inferred that the greater the variety of input models, the more reasonable BMA's output will attain. Hence, it is advised to set competing models in a way that their outputs encompass measurements.

Considering different functions like something introduced by Han and Zheng [16] can be conducted at next studies.

CRediT authorship contribution statement

Ahmad Jafarzadeh: Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Abbas Khashei-Siuki:** Writing – review & editing, Validation, Methodology, Conceptualization. **Mohsen Pourreza-Bilondi:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation. **Kwok-wing Chau:** Writing – review & editing, Visualization, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Afan HA, Osman AIA, Essam Y, Ahmed AN, Huang YF, Kisi O, et al. Modeling the fluctuations of groundwater level by employing ensemble deep learning techniques. Engineering Applications of Computational Fluid Mechanics 2021;15 (1):1420–39.

[2] Akram T, Abbas M, Riaz MB, Ismail AI, Ali NM. An efficient numerical technique for solving time fractional Burgers equation. Alex Eng J 2020;59(4):2201–20.

[3] Anshuman A, Eldho TI. Modeling of transport of first-order reaction networks in porous media using meshfree radial point collocation method. Comput Geosci 2019;23(6):1369–85.

[4] Arqub OA. Numerical solutions for the Robin time-fractional partial differential equations of heat and fluid flows based on the reproducing kernel algorithm. Int J Numer Meth Heat Fluid Flow 2018;28(4):828–56.

[5] Badawi H, Arqub OA, Shawagfeh N. Well-posedness and numerical simulations employing Legendre-shifted spectral approach for Caputo-Fabrizio fractional stochastic integrodifferential equations. Int J Mod Phys C 2023;34(06):2350070.

[6] Band SS, Heggy E, Bateni SM, Karami H, Rabiee M, Samadianfard S, et al. Groundwater level prediction in arid areas using wavelet analysis and Gaussian process regression. Engineering Applications of Computational Fluid Mechanics 2021;15(1):1147–58.

[7] Bates BC, Campbell EP. A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. Water Resour Res 2001;37(4):937–47.

[8] Beven KJ, Smith PJ, Freer JE. So just why would a modeller choose to be incoherent? J Hydrol 2008;354(1–4):15–32.

[9] Beven K, Young P. A guide to good practice in modeling semantics for authors and referees. Water Resour Res 2013;49(8):5092–8.

[10] Bjerre E, Kristensen LS, Engesgaard P, Højberg AL. Drivers and barriers for taking account of geological uncertainty in decision making for groundwater protection. Sci Total Environ 2020;746:141045.

[11] Choi SY, Seo IW, Kim YO. Parameter uncertainty estimation of transient storage model using Bayesian inference with formal likelihood based on breakthrough curve segmentation. Environ Model Softw 2020;123:104558.

[12] Dai H, Liu Y, Guadagnini A, Yuan S, Yang J, Ye M. Comparative assessment of two global sensitivity approaches considering model and parameter uncertainty. Water Resour Res 2024;60(2). e2023WR036096.

[13] Feyen L, Caers J. Quantifying geological uncertainty for flow and transport modeling in multi-modal heterogeneous formations. Adv Water Resour 2006;29 (6):912–29.

[14] Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci 1992;7(4):457–72.

[15] Hamraz B, Akbarpour A, Bilondi MP, Tabas SS. On the assessment of ground water parameter uncertainty over an arid aquifer. Arab J Geosci 2015;8(12):10759–73.

[16] Han F, Zheng Y. Joint analysis of input and parametric uncertainties in watershed water quality modeling: A formal Bayesian approach. Adv Water Resour 2018;116: 77–94.

[17] Hoege M, Guthke A, Nowak W. The hydrologist's guide to Bayesian model selection, averaging and combination. J Hydrol 2019;572:96–107.

[18] Iqbal MK, Abbas M, Wasim I. New cubic B-spline approximation for solving third order Emden-Flower type equations. Appl Math Comput 2018;331:319–33.

[19] Iqbal A, Siddiqui MJ, Muhi I, Abbas M, Akram T. Nonlinear waves propagation and stability analysis for planar waves at far field using quintic B-spline collocation method. Alex Eng J 2020;59(4):2695–703.

[20] Jafarzadeh A, Khashei-Siuki A, Pourreza-Bilondi M. Performance Assessment of Model Averaging Techniques to Reduce Structural Uncertainty of Groundwater Modeling. Water Resour Manag 2022:1–25.

[21] Jafarzadeh A, Pourreza-Bilondi M, Akbarpour A, Khashei-Siuki A, Samadi S. Application of multi-model ensemble averaging techniques for groundwater simulation: synthetic and real-world case studies. J Hydroinf 2021;23(6):1271–89.

[22] Jafarzadeh A, Pourreza-Bilondi M, Siuki AK, Moghadam JR. Examination of Various Feature Selection Approaches for Daily Precipitation Downscaling in Different Climates. Water Resour Manag 2021;35(2):407–27.

[23] Jing M, Heße F, Kumar R, Kolditz O, Kalbacher T, Attinger S. Influence of input and parameter uncertainty on the prediction of catchment-scale groundwater travel time distributions. Hydrol Earth Syst Sci 2019;23(1):171–90.

[24] Khalid N, Abbas M, Iqbal MK. Non-polynomial quintic spline for solving fourth-order fractional boundary value problems involving product terms. Appl Math Comput 2019;349:393–407.

[25] Khalid N, Abbas M, Iqbal MK, Singh J, Ismail AIM. A computational approach for solving time fractional differential equation via spline functions. Alex Eng J 2020; 59(5):3061–78.

[26] Kuczera G. Improved parameter inference in catchment models: 1. Evaluating Parameter Uncertainty *Water Resources Research* 1983;19(5):1151–62.

[27] Liu Y, Fernández-Ortega J, Mudarra M, Hartmann A. Pitfalls and a feasible solution for using KGE as an informal likelihood function in MCMC methods: DREAM (ZS) as an example. Hydrol Earth Syst Sci Discuss 2021:1–22.

[28] Maayah B, Arqub OA. Uncertain M-fractional differential problems: existence, uniqueness, and approximations using Hilbert reproducing technique provisioner with the case application: series resistor-inductor circuit. Phys Scr 2024;99(2): 025220.

[29] Mantovan P, Todini E. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. J Hydrol 2006;330:368–81.

[30] Mustafa SMT, Hasan MM, Saha AK, Rannu RP, Van Uytven E, Willems P, et al. Multi-model approach to quantify groundwater level prediction uncertainty using an ensemble of global climate models and multiple abstraction scenarios. Hydrol Earth Syst Sci 2019;23(5):2279–303. https://doi.org/10.5194/hess-23-2279-2019.

[31] Mustafa SMT, Nossent J, Ghysels G, Huysmans M. Integrated Bayesian Multi-model approach to quantify input, parameter and conceptual model structure uncertainty in groundwater modeling. Environ Model Softw 2020;126:104654.

[32] Mustafa SMT, Nossent J, Ghysels G, Huysmans M. Estimation and impact assessment of input and parameter uncertainty in predicting groundwater flow with a fully distributed model. Water Resour Res 2018;54(9):6585–608.

[33] Nazeri Tahroudi M., Ramezani Y., De Michele C., Mirabbasi R. Wind speed monitoring using entropy theory and a copula-based approach. Probabilistic Engineering Mechanics 2024;75:103582.

[34] Nguyen GT, Huynh NTH. Seasonal variations in groundwater quality under different impacts using statistical approaches. Civil Engineering Journal 2023;9 (03).

[35] Nourali M, Ghahraman B, Pourreza-Bilondi M, Davary K. Effect of formal and informal likelihood functions on uncertainty assessment in a single event rainfall-runoff model. J Hydrol 2016;540:549–64.

[36] Olyaei MA, Karamouz M. Bayesian approach for estimating biological treatment parameters under flooding condition. J Environ Eng 2020;146(8):04020083.

[37] Peach D, Taylor A. The development of a hydrogeological conceptual model of groundwater and surface water flows in the Silala River Basin. Wiley Interdiscip Rev Water 2024;11(1):e1676.

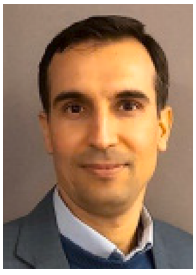
[38] Possantti I, Barbedo R, Kronbauer M, Collischonn W, Marques G. A comprehensive strategy for modeling watershed restoration priority areas under epistemic uncertainty: A case study in the Atlantic Forest. Brazil *Journal of Hydrology* 2023; 617:129003.

[39] Raftery AE, Gneiting T, Balabdaoui F, Polakowski M. Using Bayesian model averaging to calibrate forecast ensembles. Mon Weather Rev 2005;133(5): 1155–74.

[40] Sadeghi-Tabas S, Samadi SZ, Akbarpour A, Pourreza-Bilondi M. Sustainable groundwater modeling using single-and multi-objective optimization algorithms. J Hydroinf 2017;19(1):97–114.
[41] Schoups G, Vrugt JA. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. Water Resour Res 2010;46(10).
[42] Schübl M, Stumpp C, Brunetti G. A Bayesian perspective on the information content of soil water measurements for the hydrological characterization of the vadose zone. J Hydrol 2022;613:128429.
[43] Smith T, Sharma A, Marshall L, Mehrotra R, Sisson S. Development of a formal likelihood function for improved Bayesian inference of ephemeral catchments. Water Resour Res 2010;46(12).
[44] Sorooshian S, Dracup JA. Stochastic parameter estimation procedures for hydrologie rainfall-runoff models: Correlated and heteroscedastic error cases. Water Resour Res 1980;16(2):430–42.
[45] Sukri AS, Saripuddin M, Karama R, Talanipa R, Kadir A, Aswad NH. Utilization management to ensure clean water sources in coastal areas. Journal of Human, Earth, and Future 2023;4(1):23–35.
[46] Taşan M, Taşan S, Demir Y. Estimation and uncertainty analysis of groundwater quality parameters in a coastal aquifer under seawater intrusion: a comparative study of deep learning and classic machine learning methods. Environ Sci Pollut Res 2023;30(2):2866–90.
[47] Thyer M., Renard B., Kavetski D., Kuczera G., Franks S. W., Srikanthan S. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. Water Resources Research 2009;45(12).
[48] Vrugt JA, Ter Braak CJ, Gupta HV, Robinson BA. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stoch Env Res Risk A 2009;23(7):1011–26.
[49] Wu J, Zeng X. Review of the uncertainty analysis of groundwater numerical simulation. Chin Sci Bull 2013;58:3044–52.
[50] Xu T, Valocchi AJ, Ye M, Liang F, Lin YF. Bayesian calibration of groundwater models with input data uncertainty. Water Resour Res 2017;53(4):3224–45.
[51] Yin J, Tsai FTC, Kao SC. Accounting for uncertainty in complex alluvial aquifer modeling by Bayesian multi-model approach. J Hydrol 2021;601:126682.

Ahmad Jafarzadeh earned his Ph.D. (2022) from Univerty of Birjand, Birjand, Iran. His position in currently is Assistant Professor, Dept. of Civil and Architecture Engineering, University of Torbat Heydarieh, Torbat Heydarieh, Iran. His research interests are uncertainty analysis, groundwater mondeling, ensemble modeling, optimization, calibration sensitivity analysis, and machine learning.

Abbas Khashei-Siuki earned his Ph.D. (2011) from Tarbiat Modares Univesity in Iran. He is currently a Professor in the Department of Civil and Architecture Engineering, University of Torbat Heydarieh, Torbat Heydarieh. He has authored more than 150 international refereed journal articles. His research interests include investigation of water resources management, irrigation and drainage and optimization algorithms.



Mohsen Pourreza-Bilondi earned his Ph.D. (2012) from Shahid Chamran University of Ahvaz, Ahvaz, Iran under Hydrology and Water Resources Management. His domain interests are Advanced Hydrology Water Resources Management, Optimization, Risk and uncertainty analysis.

Kwok-wing Chau Professor at The Hong Kong Polytechnic University. His domain interests are Civil Engineering, Computer Engineering, Ocean Engineering, Skills and expertise, Rainfall Runoff Modelling, Water Quality Modeling, and Hydrological Modeling.