Title

Machine-learning-assisted long-term G functions for bidirectional aquifer thermal energy storage system operation

Permalink

https://escholarship.org/uc/item/50v0z603

Authors

Chen, Kecheng

Sun, Xiang

Soga, Kenichi

et al.

Publication Date

2024-08-01

DOI

10.1016/j.energy.2024.131638

Copyright Information

Peer reviewed
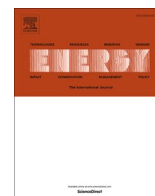
Machine-learning-assisted long-term G functions for bidirectional aquifer thermal energy storage system operation

Kecheng Chen [a,*], Xiang Sun [c], Kenichi Soga [a], Peter S. Nico [b], Patrick F. Dobson [b]

[a] *University of California, Berkeley, CA, USA*
[b] *Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
[c] *Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan, China*

ARTICLE INFO

Handling editor: X Zhao

ABSTRACT

Optimization of aquifer thermal energy storage (ATES) performance in a building system is an important topic for maximizing the seasonal offset between energy demand and supply and minimizing the building's primary energy consumption. To evaluate ATES performance with bidirectional operation, this study develops an analytical solution-based model to simulate the spatiotemporal thermal response in an aquifer. The model consists of three temperature response functions, similar to the G functions in borehole thermal energy storage (BTES), to estimate the transient temperature profile in the aquifer during seasonally varying injection and extraction of hot/cold water. Applying machine learning (ML) based data classification and regression techniques to the results of a series of finite element (FE) benchmark simulations of typical ATES configurations, model input parameters are linked to the subsurface thermal, hydrogeological, and ATES operational properties. Compared to the benchmark simulation results, the errors of the proposed model in estimating the annual energy storage and locating the thermally affected area are about 3 % and 1 %, respectively. The model was applied to a previous short-term case study, and the error in the transient production temperature estimation is about 1 %. The long-term heat recovery ratio estimated from the model also compares well to those calculated from the previous study and the validated numerical model. Because of its fast computation, the proposed model can be coupled with the individual building system simulation and used for preliminary ATES design, and this will allow for greater exploration of ATES operational space and, therefore, better choices of ATES operating conditions. The proposed model can also be coupled with the district heating and cooling network simulation for computationally efficient city-scale long-term ATES potential assessment.

1. Introduction

Heating and cooling occupy half of the world's energy consumption [1]. The seasonal offset between thermal energy demand and supply is a critical challenge in increasing renewables' share in heating and cooling. Underground thermal energy storage (UTES) has captured interest in dealing with this seasonal mismatch because of its high storage efficiency and capacity [2]. Aquifer thermal energy storage (ATES) is an open-loop UTES system that stores heat and cold in the subsurface through groundwater injection and extraction [3]. Compared to other seasonal thermal energy storage methods, ATES offers several distinct advantages [3–6]. Firstly, ATES benefits from the vast storage capacity of aquifers, which typically exceeds that of borehole thermal energy

storage (BTES) or other sensible heat storage techniques. Secondly, using aquifers enables efficient heat transfer due to groundwater's relatively high thermal conductivity, which means that ATES can achieve excellent energy efficiency during storage and retrieval. ATES typically has low environmental impacts as it avoids chemical additives or high-energy consumption processes. Moreover, ATES is known for its long-term reliability and durability. Properly designed and maintained ATES installations can provide decades of stable heating and cooling solutions. These merits make ATES an attractive option for sustainable heating and cooling solutions in various applications.

Although the adoption of ATES has been promoted to reduce emissions and save energy [6–8], it is often not considered in future energy plans due to (i) the lack of awareness of this technology, (ii) the high early-stage investment costs and (iii) inadequate knowledge to arrange

Nomenclature

A/V	surface-to-volume ratio
a, b	fitting parameters
$a(\varepsilon)$	implicit solution of the equation $F(a; \varepsilon) = 0$
a_n	coefficient ($n = 0, 1, 2, \ldots$)
c	aquifer layer heat capacity [J /kgK]
c_f	water heat capacity [J /kgK]
D	aquifer layer diffusion coefficient [m^2 /s]
D_2	impermeable layer diffusion coefficient [m^2 /s]
d_p	L^p distance
E_{store}	annual energy storage after each injection process [kWh]
$erfc$	complementary error function
$F(a; \varepsilon)$	target function
$F_n(a_n, \ldots, a_1, a_0)$	subfunction ($n = 0, 1, 2, \ldots$)
f	heat loss term [K /s]
$f(r)$	dimensionless temperature distribution after the extraction process
$g(r)$	dimensionless temperature distribution after the injection process
H	aquifer layer thickness [m]
h	heat transfer coefficient [W /m^{2}K]
h^*	dimensionless heat transfer coefficient
O	big O notation
Q_i	the injection flow rate during the injection period [m^3 /s]
Q_e	the extraction flow rate during the extraction period [m^3 /s]
Q_{base}	base flow rate [m^3 /s]
R_{th}	thermal radius [m]
r	radial distance to the well [m]
T	aquifer layer temperature [°C]
$\widehat{T}$	internal layer function
T_i	injection water temperature [°C]
T_e	extraction water temperature [°C]
T_0	undisturbed ground temperature [°C]
T_w	injection water temperature [°C]
T^*	dimensionless aquifer layer temperature
T^{*-}	left-hand side term
T^{*+}	right-hand side term
T_n^*	perturbed term ($n = 0, 1, 2, \ldots$)
$T_{inject}^* n$	dimensionless temperature distribution after the n-th year injection process
$T_{extract}^* n$	dimensionless temperature distribution after the n-th year extraction process
t	operation time [year]
t^*	dimensionless time
t_i	injection period [year]
t_e	extraction period [year]
V_{in}	seasonal water injection [m^3]
V_{ex}	seasonal water extraction [m^3]
w_0	heat flux at the interface [W /m^2]
z	spatial coordinate along the depth axis [m]
α	dimensionless flow rate
α_Q	ratio between Q_i and Q_e
Γ	gamma function
ε	small parameter
ε_e	the relative error rate of energy estimation
$\varepsilon_e\vert_{advection, t=i}$	the relative error rate of energy estimation for the advection-dominated solution at the i-th year
$\varepsilon_e\vert_{diffusion, t=i}$	the relative error rate of energy estimation for the diffusion-dominated solution at the i-th year
ε_{20}	20-year relative error rate of energy estimation
ε_r	relative error rate of temperature profile prediction along the radial direction
ε_t	the relative error rate of the transient production temperature estimation
η	heat recovery ratio
θ	Heaviside step function
λ	aquifer layer thermal conductivity [W /mK]
λ_2	impermeable layer thermal conductivity [W /mK]
$\mu_n(\varepsilon)$	gauge function ($n = 0, 1, 2, \ldots$)
ν	aquifer layer advective flow velocity [m /s]
ξ	stretched variable
ρc	aquifer layer volumetric heat capacity [J /m^{3}K]
$(\rho c)_2$	impermeable layer volumetric heat capacity [J /m^{3}K]
$(\rho c)_f$	water volumetric heat capacity [J /m^{3}K]

the subsurface space [9]. Its limitations are also related to the local environment and other issues, such as long-term efficiency decrease [10, 11]. The unfamiliarity with ATES further leads to the absence of related heating and cooling facilities. Therefore, an intuitive and systematic methodology to evaluate ATES potential is needed to promote this technology [12].

ATES can be performed in a unidirectional or bidirectional manner. With unidirectional operation, cool or warm water is injected from one well, and the other well is used for water extraction, as shown in Fig. 1a. It extracts constant-temperature groundwater to cool buildings in summer and heat buildings in winter, where the heat pump helps transfer the heat [13]. With the bidirectional operation, cool and warm wells act as both injection and extraction wells, as shown in Fig. 1b. Under this operational mode, cold groundwater stored during the winter is extracted to cool buildings in summer, and warm groundwater stored in the summer is extracted to heat buildings in winter.

The mathematical methods used to simulate ATES's thermal performance are generally classified into three types, including numerical, analytical, and data-driven methods. Table 1 lists prior studies that utilized numerical models to evaluate the ATES potential. With increased computing power, 3D numerical modeling has become commonplace. Software like MODFLOW [14–18], TOUGH2 [19], which

use the finite difference (FD) method, FEFLOW [20–22], COMSOL [23–26], SUTRA [27], OpenGeoSys [28], MOOSE [29], which use the finite element (FE) method, DuMux [30], and CSMP++ [31], which use the finite volume (FV) method, have been successfully applied to model subsurface flow and heat transport processes. The main objective of these studies is to predict recovery temperature as a function of heterogeneous hydrogeological properties, injection temperature, well placement, pumping rate, regional groundwater flow, and buoyancy flow. Most studies investigate the performance of bidirectionally operated ATES systems with operation time under 20 years. Although numerical methods can solve problems with complex geometries and boundary conditions, the increased computation time caused by mesh discretization in FD, FV, and FE methods limits their practical application, particularly for ATES system optimization, which requires a fast algorithm.

Compared to numerical models, analytical models can be attractive for conducting many scenario-testing simulations with sufficient accuracy [36]. Table 2 summarizes prior studies that utilized analytical models to evaluate the ATES potential. These studies can be divided into two groups. One group is the analytical model derived from the governing equation; the other is the analytical model learned from numerical simulation. For the first group, assumptions were made to
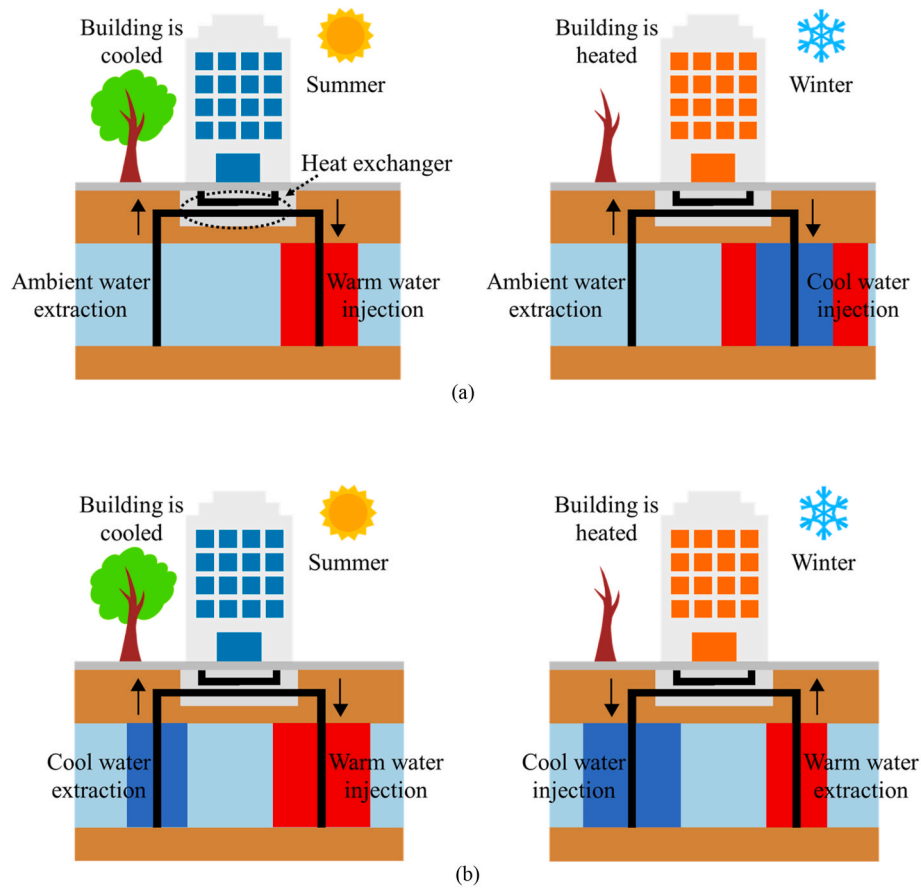
Fig. 1. Operations of (a) the unidirectional and (b) the bidirectional ATES systems in summer and winter. In the aquifer, the blue area represents the cool zone, and the red area represents the warm zone. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Past studies on numerical modeling of ATES.

Dimension		Operation mode		Operation years	Numerical method (FE: Finite element; FD: Finite difference; FV: Finite volume)	Reference
2D	3D	Unidirectional	Bidirectional			
	✓		✓	10	FD	Sommer et al. [14]
	✓	✓		20	FD	Todorov et al. [15]
	✓		✓	10	FD	Bakr et al. [16]
✓	✓		✓	50	FD	Bloemendal et al. [17]
✓			✓	5	FD	Beernink et al. [18]
	✓		✓	1	FD	Guo et al. [19]
	✓		✓	20	FE	Xiao et al. [20]
	✓	✓		10	FE	Major et al. [21]
	✓		✓	0.05	FE	De et al. [22]
	✓		✓	15	FE	Kim et al. [23]
✓			✓	4	FE	Jeon et al. [24]
	✓		✓	5	FE	Gao et al. [25]
	✓		✓	30	FE	Stemmle et al. [26]
	✓		✓	30	FE	Burns et al. [27]
	✓		✓	5	FE	Perez Silva et al. [28]
	✓		✓	10	FE	Dashti et al. [29]
	✓		✓	1	FV	Ganguly et al. [30]
	✓		✓	1	FE + FV	Yapparova et al. [31]
✓			✓	0.5	FD	Buscheck et al. [32]
✓			✓	0.5	FD	Tsang et al. [33]
	✓		✓	3	FD	Xue et al. [34]
	✓		✓	20	FV	Collignon et al. [35]

simplify complex 3D scenarios into 1D or 2D to ensure the existence of the solution. The goal is to estimate the heat recovery ratio or the aquifer temperature distribution. Most describe the unidirectional ATES operation [37–49]. The rest examine the bidirectional process without considering heat loss into the surrounding impermeable layers and seasonally varying water injection and extraction [50–54]. However, it has been proved by simulations [27,55,63] and experiments [64–66] that the effect of the heat loss into the surrounding impermeable layers on the ATES recovery performance is noticeable and even dominates when the injected thermal volume is large. For the realistic ATES

Table 2
Past studies on analytical modeling of ATES.

Dimension			Operation mode		Analytical model description	Reference
1D	2D	3D	Unidirectional	Bidirectional		
✓			✓		Aquifer temperature distribution with the hot water injected	Li et al. [37]
✓			✓		Vertical elevations of the aquifer thermal fronts during the injection	Nordbotten [38]
✓			✓		Aquifer temperature distribution with water extracted	Yeh et al. [39]
✓	✓		✓		Aquifer temperature distribution around injection and production wells	Schulz [40]
✓			✓		Aquifer temperature distribution during the hot water injection	Lauwerier [41]
✓			✓		Extend the concept of Lauwerier [39] by including the thermal bleeding effect	Barends [42]
✓			✓		Aquifer temperature distribution during nonisothermal immiscible two-phase flow injection	LaForce et al. [43]
	✓		✓		Aquifer temperature distribution during the injection	Pophillat et al. [44]
✓			✓		Aquifer solute distribution during the injection	Chen [45]
✓			✓		Aquifer temperature distribution during the injection with the mechanical dispersion and the molecular diffusion	Aichi et al. [46]
	✓		✓		Planar reservoir temperature distribution when fluid circulates in it	Wu et al. [47,48]
✓			✓		Heterogeneous aquifer temperature distribution with the cold water injected into it	Ganguly and Kumar [49]
✓				✓	Stabilized heat recovery ratio	Bloemendal et al. [50]
✓				✓	Aquifer solute distribution during the injection/extraction without solute loss into surrounding impermeable layers	Veling [51]
✓				✓	Solute and heat recovery coefficients without solute/heat loss into surrounding impermeable layers	Tang and Van [52,53]
✓				✓	Heat recovery efficiency and aquifer temperature distribution during the injection/storage/extraction without heat loss into surrounding impermeable layers when time spent in the injection/extraction counts as $\frac{d}{3d-2}$ times time spent in the storage, where d is the radial dimension	Tang and Rijnaarts [54]
	✓			✓	Heat recovery efficiency with a five-year operation as a function of the dimensionless parameters	Doughty et al. [55]
✓		✓	✓		Aquifer temperature distribution during the injection with conduction, dispersion, and a Robin-type boundary condition	Lin et al. [56]
	✓			✓	Heat recovery efficiency in the fourth year, considering aquifer properties and operational variables	Schout et al. [57]
	✓			✓	Heat recovery efficiency with a fifty-year operation considering the density-driven flow	Oerlemans [58]
		✓		✓	Effect of well placement on the system performance with a five-year operation	Duijff et al. [59]
		✓		✓	Heat recovery efficiency with a thirty-year operation considering the thermal interference effect	Sommer et al. [60]
		✓		✓	Heat recovery coefficient with a four-year operation considering the salinity contrast	Van et al. [61]
	✓			✓	Heat loss fraction considering the buoyancy-driven flow	Beernink et al. [62]

operation, seasonal water injection and extraction flow rates and durations may differ due to the interplay between geological characteristics and energy demand fluctuations [20]. For the second group, these studies consider 2D or 3D bidirectional operated ATES systems, and the resulting heat recovery ratio is given in charts, graphs, or analytical descriptions. However, none of the simulation-derived relationships provides the production temperature during a long-term ATES operation, which is essential when performing building-ATES coupled analysis. The current limitations of analytical models and the absence of a computationally efficient method in ATES potential assessment, like the G function in BTES [67], are the primary motivations for carrying out the present work.

Recently, there has been a growing interest in leveraging data-driven approaches for ATES potential assessment because of their ability to capture complex physical dynamics based on large datasets [68]. Table 3 lists prior studies that utilized data-driven models to evaluate the ATES potential. Data for training the data-driven model is from experiments or simulations. Popular data-driven methods are deep neural networks and random forests. Most of these studies investigate the unidirectional ATES operation because the thermal pattern in this operation condition is more stable compared to the bidirectional one and more accessible to be learned by the data-driven method [73–83]. For the rest of the studies that use the data-driven method to learn the bidirectional operation, some predict the critical metrics of ATES performance [68,71], including heat recovery ratio and annual energy storage; some predict the production temperature for a specific case [69, 70], and others indicate the uncertainty on the aquifer thermal pattern [72]. The operation time for these studies is less than 20 years. Because physical principles are not incorporated in the data-driven methods used in these ATES studies, the resulting products are case-specific and hard to generalize as they need tons of data inputs for learning underlying physics, where the data generation and training consume a significant amount of time. The physics-informed neural network is a possible approach to deal with the mentioned limitation, but it needs expertise and a careful structure design [84,85]. Inspired by the data-driven method's ability to discover patterns and relationships within data, the present work uses machine learning (ML) to predict the complicated thermal interaction between the aquifer and the surrounding impermeable layers with analytical solutions as a base, which is a tradeoff between interpretability/generalizability and solvability.

To develop a computationally efficient tool for building-coupled ATES potential assessment and optimization, this paper describes an analytical solution-based model that gives the spatiotemporal aquifer thermal response in a bidirectional ATES system. The initial model consists of three 1D analytical solutions, with unknown heat loss terms, that estimate the aquifer temperature profile during seasonally varying hot/cold water injection and extraction. The assumptions and definitions will be discussed in Section 2. The derivation of the analytical solutions is discussed in Section 3. The K-nearest neighbors (KNN) algorithm is adopted to learn the rules behind the thermal interaction between the aquifer and the impermeable layers from a 2D axisymmetric FE benchmark model to quantify the heat loss terms of the 1D analytical solutions, which is given in Section 4. The results of the fitting exercise between the benchmark model and the analytical solutions and the proposed model's cross-validation is provided in Section 5. In Section 6, the practical application possibilities of the proposed model are explored based on a short-term field test dataset and a long-term simulation dataset. The limitations of the current work are discussed in Section 7. The resulting software for this study has been released.[1]

---

[1] https://github.com/Kecheng-Chen/gfunction-ates.

Table 3
Past studies on data-driven modeling of ATES.

Data source		Operation mode		Operation years	Data-driven method	Prediction	Reference
Experiment	Simulation	Unidirectional	Bidirectional				
	✓		✓	10	Feedforward neural network	Recovery efficiency, charging time, operating time, annual storage	Jin et al. [68]
✓			✓	4	Long short-term memory network	Production temperature	Parya [69]
	✓		✓	16	Gaussian process regression	Production temperature	Rohmer et al. [70]
	✓		✓	5	Convolutional neural network	Recovery efficiency	Sheldon et al. [71]
✓	✓		✓	0.17	Bayesian evidential learning	Aquifer temperature distribution uncertainty	Hermans et al. [72]
	✓	✓		50	Random forest	Production temperature, water table	Wang et al. [73]
	✓	✓		2	Bidirectional gated recurrent unit network	Production temperature	Ullah et al. [74]
	✓	✓		20	Feedforward neural network	Production temperature, production pressure	Duplyakin et al. [75]
	✓	✓		10	Convolution neural network and long short-term memory recurrent network	Production rate, production temperature	Wang et al. [76]
	✓	✓		10	Feedforward neural network	Production temperature, production pressure	Aydin et al. [77]
	✓	✓		10	Feedforward neural network	Injected tracer concentration	Gudmun and Horne [78]
	✓	✓		30	Feedforward neural network	Production temperature	Pandey and Singh [79]
	✓	✓		20	Feedforward neural network, convolution neural network, and long short-term memory recurrent network	Production temperature	Beckers et al. [80]
	✓	✓		30	Feedforward neural network, linear regression, random forest, and support vector machine	Production temperature	Gudala and Govindarajan [81]
	✓	✓		30	Feedforward neural network and long short-term memory recurrent network	Production temperature	Shi et al. [82]
✓		✓		13	Feedforward neural network	Production temperature, production pressure, flow rate	Ariturk [83]

2. Problem formulation

In the context of BTES, the G function represents the dimensionless temperature response of the ground to a unit heat flux over time, considering the conductive heat transfer [67]. It can be spatially and temporally superposed to quantify the ground temperature distribution due to heat injection and extraction of borehole heat exchangers (BHE). It serves as a fundamental tool for engineers to analyze and predict the behavior of ground source heat pump (GSHP) systems, guiding decisions regarding BHE configuration, system layout, and operational strategies. Several assumptions are typically made by the G function, including homogeneous and constant ground properties, steady-state, infinite boundary, negligible groundwater flow, and insignificant latent heat effect.

Following the spirit of the G function in BTES, this study describes some assumptions and scenarios specifically for ATES to derive dimensionless thermal response of the aquifer to a unit well temperature over time considering the conductive and convective heat transfer, the aquifer-impermeable layer interface thermal interaction and the seasonal cyclic operation. The obtained thermal response functions can be spatially and temporally superposed to quantify the aquifer temperature distribution due to hot/cold water injection and extraction of the well.

2.1. Problem description

The geometry of the ATES system considered by the present model is shown in Fig. 2a. The configuration consists of a cool well and a warm well that thoroughly penetrates a horizontal aquifer with thickness H. The aquifer is bounded above and below by impermeable layers. Water is injected through the warm well and extracted from the cool well with the same volumetric flow rate during the summer operation. This process is reversed during winter operation. The vital mechanical

components enabling the ATES system's operation include a ground loop circulation pump, a pump from the heat pump to the buffer tank, and a pump from the buffer tank to the building's air handling unit [86, 87]. The ground loop circulation pump maintains a consistent flow rate to enhance the system reliability and minimize the operational cost. The buffer tank, also known as an expansion or thermal storage tank, is a container used in various heating, cooling, and plumbing systems to manage fluid volumes and pressure fluctuations. In the ATES system, the buffer tank decouples the heat pump from the building's immediate heating or cooling demands. It acts as a reservoir, preventing short cycling and allowing the heat pump to operate optimally for long periods.

Some assumptions are made for the present model. The aquifer is homogeneous for volumetric heat capacity (ρc) and heat conductivity (λ) (or diffusion coefficient $D = \lambda/\rho c$). The impermeable layers are homogeneous for volumetric heat capacity (ρc)$_2$ and heat conductivity λ_2 (or diffusion coefficient $D_2 = \lambda_2/(\rho c)_2$). The thicknesses of the impermeable layers are large enough to ignore the boundary effects. All material properties are temperature-independent. There is radial symmetry for both warm and cool wells. The well spacing is more extensive than three times the thermal radius R_{th} so that the thermal behaviors of the cool and warm wells are not significantly affected by each other [88,89]. The thermal radius R_{th} can be calculated from Eq. (1), where c_f is the water heat capacity and V_{in} is the seasonal water injection. Seasonal aquifer injection and extraction are considered to cover the building's seasonal heating and cooling loads. The length of each ATES cycle is one year. The radial flow rate Q and the injection water temperature T_w remain constant [90]. For the initial condition, the entire system is at a constant temperature T_0. Typically, T_w is close to T_0, so buoyancy flow (or density-driven flow) is ignored in the present study [91].

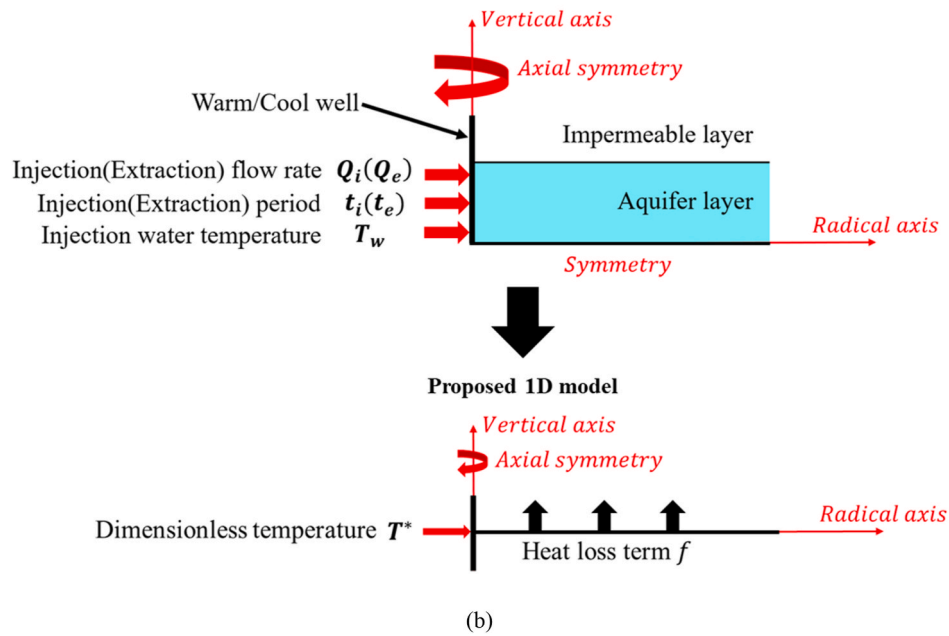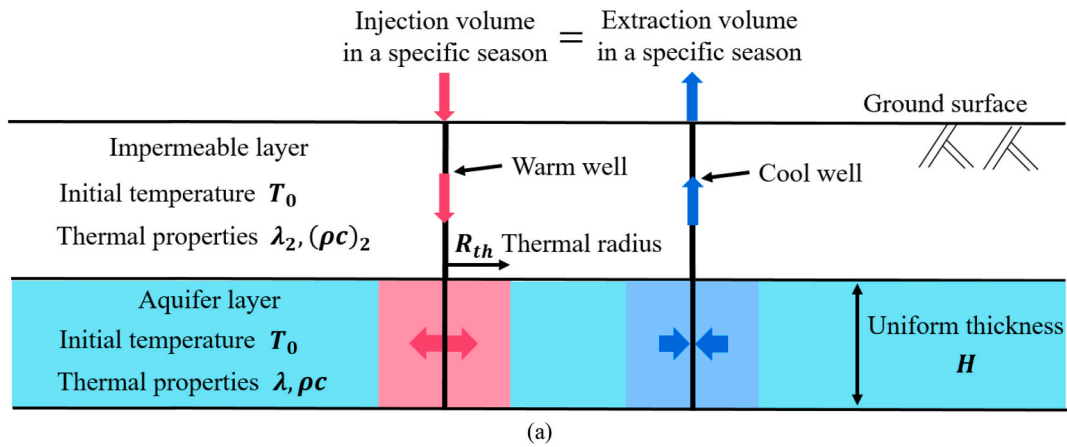$$R_{th} = \sqrt{\frac{c_f V_{in}}{cH\pi}} \tag{1}$$

Fig. 2. (a) Side view of ATES operation in summer. The injection volume equals the extraction volume in the same season. In the aquifer, the blue area represents the cool zone, and the red area represents the warm zone. (b) Explanation of the proposed model development. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
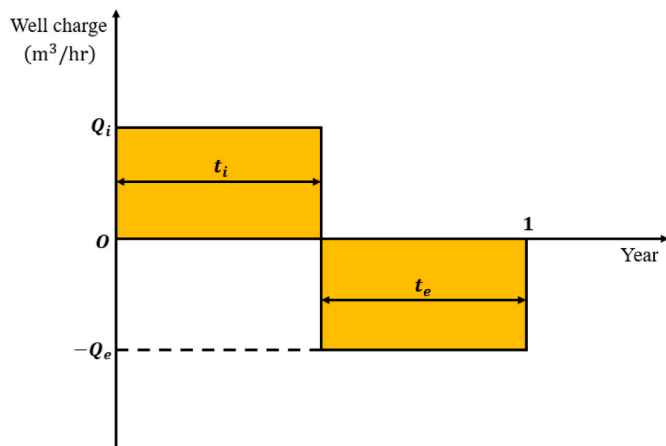

Fig. 3. Parameters related to ATES cyclic operation. The operation graph is a periodic step function graph, and the period is one year.

$$T^* = \frac{T - T_0}{T_W - T_0} \tag{2}$$

Some definitions are made for the present model. As shown in Fig. 3, the duration of the injection period and extraction period are t_i and t_e respectively, and the sum of t_i and t_e equals one year. Because the heating and cooling loads of buildings are variable, t_i is not always equal to t_e [92] and the ratio of t_i/t_e is defined as α_t. For the same reason, the injection flow rate Q_i during the injection period may not be equal to the extraction flow rate Q_e during the extraction period. The ratio of Q_i/Q_e is defined as α_Q. The value of Q_i is defined as the base flow rate Q_{base}. A well is injection-dominated when the injection volume during the injection period is larger than the extraction volume during the extraction period for a given 12 months ($\alpha_t \bullet \alpha_Q > 1$); otherwise, the well is extraction-dominated ($\alpha_t \bullet \alpha_Q < 1$). Since T_w of the cool well is different from that of the warm well, dimensionless temperature T^* is defined as shown in Eq. (2) to keep the results uniform, where T is the temperature of the aquifer.

Based on the definition, to obtain a single analytical expression of aquifer temperature distribution for both warm and cool wells, the proposed model simplifies the 2D axisymmetric cases into 1D axisym-

Table 4
Model parameters.

Parameter	Sampling range
Aquifer thermal conductivity λ [W/mK]	[2.51, 2.71, 3.01]
Impermeable layer thermal conductivity λ_2 [W/ mK]	3.77
Aquifer volumetric heat capacity ρc [MJ /m^{3}K]	[1.64, 2.04, 2.44, 2.74]
Impermeable layer volumetric heat capacity ρc_2 [MJ /m^{3}K]	2.05
Base flow rate Q_{base} [m^3 /hr]	[20, 85, 150]
α_Q	[0.5, 0.75, 1, 1.5, 2]
α_t	[0.67,0.82,1,1.22,1.5]
Aquifer thickness [m]	[10,15,20,25,30,35,40]
Injection water temperature T_w [°C]	[8.8, 25.8]
Undisturbed ground temperature T_0 [°C]	14.8

Table 5
Numerical model parameters.

Parameter	Value
Heating season total groundwater flow rate [m^3/ hr]	26
Cooling season total groundwater flow rate [m^3/ hr]	21
Heating season injection water temperature [°C]	6
Cooling season injection water temperature [°C]	13
Impermeable layer thermal conductivity [W /mK]	3
Aquifer layer thermal conductivity [W /mK]	2.5
Impermeable layer volumetric heat capacity [MJ/ m^{3}K]	2
Aquifer layer volumetric heat capacity [MJ/ m^{3}K]	2.25
Undisturbed temperature [°C]	10

metric dimensionless cases by averaging and nondimensionalizing the vertical temperature in the aquifer, as shown in Fig. 2b. The thermal interaction between the aquifer and impermeable layers in the 2D case is substituted by a heat loss term f in the 1D case. The analytical solutions of the simplified governing equation can then be obtained by considering three different ATES operation conditions. KNN, a non-parametric ML algorithm, is used to learn the unknown heat loss term in analytical solutions. To prepare the training data for learning, a 2D axisymmetric FE model is built. Table 4 lists the input parameters for the FE simulation, which are deduced from typical ATES operational and hydrogeological conditions. The parameter sampling ranges are determined based on the previous field tests [93–95] and the parametric sensitivity analysis [36]. This study considers 6300 scenarios to reveal the thermal behavior in an ATES system, where 3276 are injection-dominated and 3024 are extraction-dominated scenarios. The modeled ATES operation time is 50 years because ATES systems are generally designed to work during the lifetime of the building (30–50 years) [96].

2.2. Governing equation and perturbation

Based on the previous assumptions, the equation of temperature distribution in the aquifer is given in Eq. (3), which describes a 1D convection-diffusion process in the polar coordinate system. f is the heat loss term. This is an inhomogeneous parabolic partial differential equation, which rarely has an analytical solution. However, based on the model parameters listed in Table 4, D is small and varies between $5.44 \bullet 10^{-7}$ and $1.09 \bullet 10^{-6}$ m^2/s. Assuming the exact solution can be expressed as a power series in D, an approximate solution to the governing equation may be obtained through the perturbation method.

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rD\frac{\partial T}{\partial r}\right) - v\frac{\partial T}{\partial r} = \frac{\partial T}{\partial t} + f \tag{3}$$

$$v = \frac{Q}{2\pi rH}\frac{(\rho c)_f}{\rho c} \tag{4}$$

$$D = \frac{\lambda}{\rho c} \tag{5}$$

The method to solve the problem depending on a typically small parameter ε is called the perturbation method [97,98]. If $a(\varepsilon)$ is the implicit solution of an equation shown in Eq. (6) and both $a(\varepsilon)$ and $F(a;\varepsilon)$ have an asymptotic series expansion with the same gauge functions μ_n, $a(\varepsilon)$ may be determined asymptotically by the perturbation method. It involves (i) expanding $a(\varepsilon)$ as shown in Eq. (7), (ii) substituting this expansion in $F(a;\varepsilon)$, and (iii) expanding F to obtain Eq. (8). Because an asymptotic expansion vanishes only if the coefficients vanish [97], the sequence of coefficients (a_n) can be determined by induction, as shown in Eq. (9).

$$F(a;\varepsilon) = 0 \tag{6}$$

$$a(\varepsilon) = a_0\mu_0(\varepsilon) + a_1\mu_1(\varepsilon) + \dots \tag{7}$$

$$F(a;\varepsilon) = F_0(a_0)\mu_0(\varepsilon) + F_1(a_1, a_0)\mu_1(\varepsilon) + F_2(a_2, a_1, a_0)\mu_2(\varepsilon) + \dots = 0 \tag{8}$$

$$F_0(a_0) = 0, F_1(a_1, a_0) = 0, F_2(a_2, a_1, a_0), etc. \tag{9}$$

2.3. ATES operation scenarios

To make suitable assumptions when deriving analytical solutions, this study introduces three cases for the ATES operation: two cases for the injection-dominated scenario (the advection-dominated case and the diffusion-dominated case) and one for the extraction-dominated scenario. Fig. 4 explains the seasonal operation of a single well (take the warm well as an example, but the cool well operates similarly) in these three cases.

In the injection-dominated scenario, the warm well has more hot water injection in summer than extraction in winter. The advection-dominated case occurs when ATES operates with a thin aquifer and/or high flow rate, whereas the diffusion-dominated case occurs when ATES operates with a thick aquifer and/or low flow rate [36]. In the extraction-dominated scenario, the warm well has less hot water injection in summer than extraction in winter. There are no two cases in the extraction-dominated scenario because the thermal interaction between the aquifer and the impermeable layer is more significant than the diffusion effect in the aquifer layer.

3. Analytical solutions

The analytical expressions of the dimensionless temperature distribution of the aquifer during the seasonal injection and extraction processes can be derived based on the definitions and assumptions in Section 2.1. The exact solution for the first-year injection process can be deduced because of the constant initial temperature conditions. However, for the following extraction/injection process, the starting temperature distribution depends on the previous process, leading to a complicated radial function. Therefore, different assumptions are made to modify the governing equation in Section 2.2 for the three ATES operation cases defined in Section 2.3 to attain the approximate solutions.

3.1. The first-year injection

For the injection in the first year, it is assumed that the heat loss from the aquifer into the impermeable layers obeys Newton's heat loss model because the diffusion length is small [36,99]. f shown in Eq. (3) becomes $\frac{w_0}{\rho cH}$ shown in Eq. (10), where $w_0 = h(T - T_0)$ is the heat flux at the aquifer-impermeable layer interface, and h is the convective heat transfer coefficient. The advantage of introducing such a boundary is that the influence of the surrounding impermeable layer on the aquifer can be compacted into a single coefficient h. Eqs. (12)–(14) are the initial and boundary conditions.
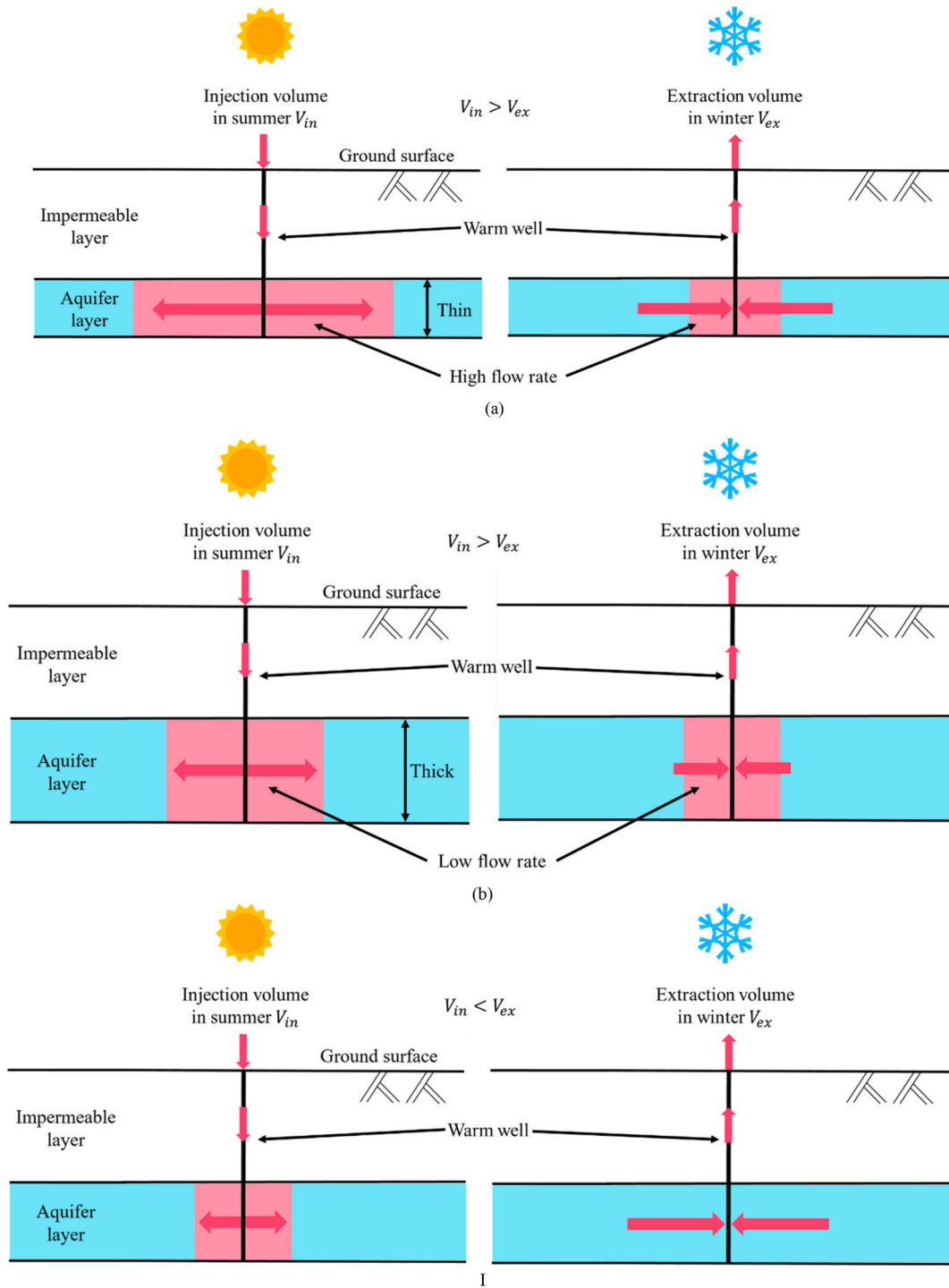
Fig. 4. Cases of analytical solutions (warm well): (a) the advection-dominated and (b) the diffusion-dominated cases in the injection-dominated scenario, and (c) the extraction-dominated scenario. The red arrow indicates the direction of injection and extraction, and its length qualifies the magnitude of the flow rate. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rD\frac{\partial T}{\partial r}\right) - v\frac{\partial T}{\partial r} = \frac{\partial T}{\partial t} + \frac{w_0}{\rho cH} \tag{10}$$

$$w_0 = h(T - T_0) \tag{11}$$

$$r = 0, T = T_w \tag{12}$$

$$r = \infty, T = T_0 \tag{13}$$

$$t = 0, T = T_0 \tag{14}$$

Through the Laplace transform, the solution for the dimensionless temperature T^* is obtained as shown in Eq. (15), where α is the dimensionless injection rate [99] and $(\rho c)_f$ is the volumetric heat capacity of water.

$$T^* = \frac{1}{\Gamma(\alpha)} \int_{\frac{r^2}{4Dt}}^{\infty} e^{-\left(x + \frac{hr^2}{4\rho cDH} * \frac{1}{x}\right)} x^{\alpha - 1} dx \tag{15}$$
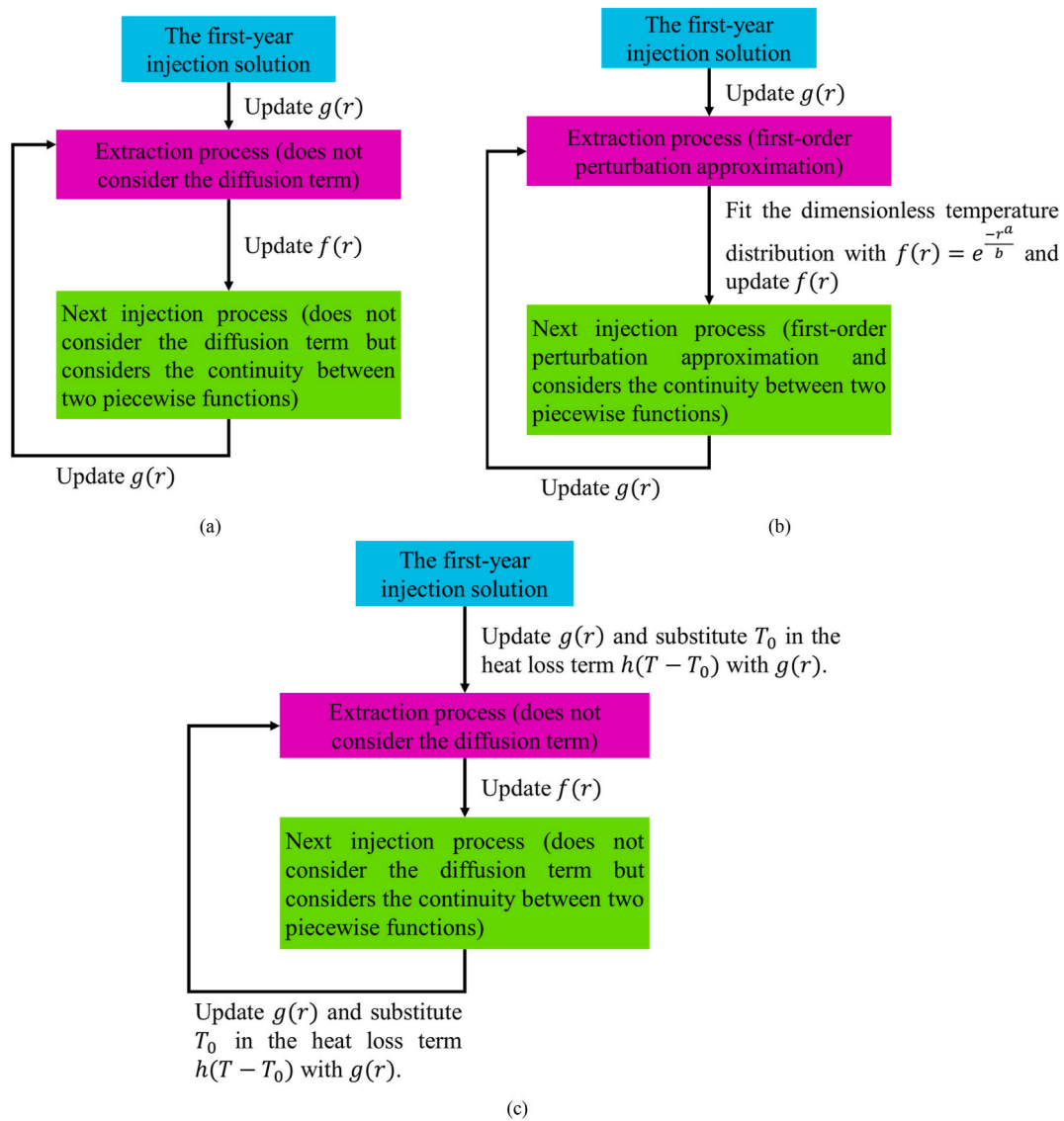
Fig. 5. (a) Advection-dominated case, (b) diffusion-dominated case, and (c) extraction-dominated analytical schemes.

$$\alpha = \frac{(\rho c)_f}{4\pi}\left(\frac{1}{H\lambda}\right)Q \tag{16}$$

The computed spatial temperature field from the first year injection is used for the subsequent annual extraction/injection cycles, which is described in the following sections.

3.2. The injection-dominated scenario ($\alpha_t \cdot \alpha_Q > 1$)

After the injection in the first year changing the subsurface temperature field, this scenario is used when a well injects more water than it extracts for a given 12 months. Two cases are considered: (i) an advection-dominated case for a thin aquifer and/or fast flow rate and (ii) a diffusion-dominated case for a thick aquifer and/or slow flow rate.

3.2.1. The advection-dominated case for thin aquifer and/or fast flow rate

3.2.1.1. Extraction.
The governing equation of the extraction process in the advection-dominated case is given by Eq. (17), which ignores the second-order diffusion term. It is because when v is large, the coefficient of the second-order diffusion term D^2 is much less than the coefficient of the convection term. Eqs. (18)–(20) establish the initial and boundary

conditions, where $g(r)$ is the dimensionless temperature distribution after the previous injection process.

$$\left(\frac{D}{r}+v\right)\frac{\partial T}{\partial r} = \frac{\partial T}{\partial t} + \frac{w_0}{\rho c H} \tag{17}$$

$$r=0, T=T_w \tag{18}$$

$$r=\infty, T=T_0 \tag{19}$$

$$t=0, T=(T_w-T_0)g(r)+T_0 \tag{20}$$

The solution is

$$T^* = e^{-\frac{ht}{H\rho c}}g\left(\sqrt{r^2+2\gamma t}\right) \tag{21}$$

where parameter γ is defined as follows.

$$\gamma = \frac{(\rho c)_f}{2\pi}\left(\frac{1}{H\rho c}\right)Q+D \tag{22}$$

3.2.1.2. Injection.
The governing equation of the injection process is Eq. (23). Again, it ignores the second-order diffusion term. Eqs. (24)–(26)

establish the initial and boundary conditions, where $f(r)$ is the dimensionless temperature distribution after the last extraction process.

$$\left(\frac{D}{r} - v\right)\frac{\partial T}{\partial r} = \frac{\partial T}{\partial t} + \frac{w_0}{\rho cH} \tag{23}$$

$$r = 0, T = T_w \tag{24}$$

$$r = \infty, T = T_0 \tag{25}$$

$$t = 0, T = (T_w - T_0)f(r) + T_0 \tag{26}$$

The solution is a piecewise function, where θ is the Heaviside step function.

$$T^* = \theta\left(-\frac{r^2}{2\beta} + t\right) \bullet T^{*-} + \theta\left(\frac{r^2}{2\beta} - t\right) \bullet T^{*+} \tag{27}$$

Where functions T^{*-}, T^{*+} and parameter β are defined as follows.

$$T^{*-} = e^{-\frac{hr^2}{2\beta H\rho c}} \tag{28}$$

$$T^{*+} = e^{\frac{-ht}{H\rho c}} * f\left(-i \bullet \sqrt{-r^2 + 2\beta t}\right) \tag{29}$$

$$\beta = \frac{(\rho c)_f}{2\pi}\left(\frac{1}{H\rho c}\right)Q - D \tag{30}$$

Shih et al.'s [100] singular perturbation approach to moving sharp fronts is adopted. Eliminating the discontinuity along the curve caused by ignoring the second-order diffusion term, the stretched variable ξ along the characteristic curve $r = \sqrt{2\beta t}$ is defined as follows.

$$\xi = \frac{r - \sqrt{2\beta t}}{\sqrt{D}} \tag{31}$$

With the variable ξ and t, Eq. (23) becomes Eq. (32), where $\widehat{T}$ is the internal layer function.

$$\frac{\partial \widehat{T}}{\partial t} - \frac{\partial^2 \widehat{T}}{\partial \xi^2} + \frac{h\widehat{T}}{\rho cH} + \left(\frac{1}{\sqrt{D}\,\frac{\xi}{\beta} + \sqrt{\frac{2t}{\beta}}} - \frac{1}{\sqrt{\frac{2t}{\beta}}}\right)\frac{\partial \widehat{T}}{\partial \xi} = 0 \tag{32}$$

Denoting $\widehat{T}(\xi, t)$ by $\widehat{T}^+(\xi, t)$ and $\widehat{T}^-(\xi, t)$ for $0 < \xi < \infty$ and $-\infty < \xi < 0$, respectively, the relationship to overcome discontinuity is shown as follows.

$$T^{*-}\left(\sqrt{2\beta t}, t\right) + \widehat{T}^-(0, t) = T^{*+}\left(\sqrt{2\beta t}, t\right) + \widehat{T}^+(0, t) \tag{33}$$

Because $\widehat{T}$ overcomes the discontinuity around $\xi = 0$, the rightmost term in Eq. (32) can be ignored, leading to the following.

$$\frac{\partial \widehat{T}}{\partial t} - \frac{\partial^2 \widehat{T}}{\partial \xi^2} + \frac{h\widehat{T}}{\rho cH} = 0 \tag{34}$$

The final solution of $\widehat{T}$ is shown in Eq. (35).

$$\widehat{T} = \begin{cases} \frac{1 - f(0)}{2}e^{-\frac{ht}{H\rho c}}erfc\left(\frac{r - \sqrt{2\beta t}}{2\sqrt{Dt}}\right), r \geq \sqrt{2\beta t} \\ -\frac{1 - f(0)}{2}e^{-\frac{ht}{H\rho c}}erfc\left(-\frac{r - \sqrt{2\beta t}}{2\sqrt{Dt}}\right), r < \sqrt{2\beta t} \end{cases} \tag{35}$$

The solution for the injection process is revised to be Eq. (36).

$$T^* = \begin{cases} T^{*+} + \frac{1 - f(0)}{2}e^{-\frac{ht}{H\rho c}}erfc\left(\frac{r - \sqrt{2\beta t}}{2\sqrt{Dt}}\right), r \geq \sqrt{2\beta t} \\ T^{*-} - \frac{1 - f(0)}{2}e^{-\frac{ht}{H\rho c}}erfc\left(-\frac{r - \sqrt{2\beta t}}{2\sqrt{Dt}}\right), r < \sqrt{2\beta t} \end{cases} \tag{36}$$

3.2.1.3. Summary. Fig. 5a shows the calculation flow chart of the advection-dominated case. First, Eq. (15) is used to evaluate the dimensionless aquifer temperature distribution after the first-year injection $T^*_{inject}1$. Second, with $g(r) = T^*_{inject}1$, Eq. (21) is used to evaluate the dimensionless aquifer temperature distribution after the first-year extraction $T^*_{extract}1$. Third, with $f(r) = T^*_{extract}1$, Eq. (36) is used to evaluate the dimensionless aquifer temperature distribution after the second-year injection $T^*_{inject}2$. For the subsequent extractions and injections, Eqs. (21) and (36) are used to evaluate the dimensionless aquifer temperature distribution iteratively by updating $g(r)$ and $f(r)$.

3.2.2. The diffusion-dominated case for a thick aquifer and/or slow flow rate

3.2.2.1. Extraction. When a system has a thick aquifer and/or slow flow rate, the second-order diffusion term in Eq. (10) cannot be ignored. Based on regular perturbation theory, the solution is approximated as follows.

$$T^* = T^*_0 + D \bullet T^*_1 + O(D^2) \tag{37}$$

Eq. (38) below gives the modified governing function of the extraction process. Substituting Eq. (37) into Eq. (38) results in two equation systems: (i) Eq. (39) with initial and boundary conditions given by Eqs. (40)–(42), and (ii) Eq. (43) with the initial state given by Eq. (44).

$$D\frac{\partial^2 T^*}{\partial r^2} + \frac{\gamma}{r}\frac{\partial T^*}{\partial r} = \frac{\partial T^*}{\partial t} + \frac{hT^*}{\rho cH} \tag{38}$$

$$\frac{\partial T^*_0}{\partial t} - \frac{\gamma}{r}\frac{\partial T^*_0}{\partial r} + \frac{hT^*_0}{\rho cH} = 0 \tag{39}$$

$$r = 0, T^*_0 = 1 \tag{40}$$

$$r = \infty, T^*_0 = 0 \tag{41}$$

$$t = 0, T^*_0 = g(r) \tag{42}$$

$$\frac{\partial T^*_1}{\partial t} - \frac{\gamma}{r}\frac{\partial T^*_1}{\partial r} + \frac{hT^*_1}{\rho cH} = \frac{\partial^2 T^*_0}{\partial r^2} \tag{43}$$

$$t = 0, T^*_1 = 0 \tag{44}$$

The final solutions of Eq. (39) and Eq. (43) are the following.

$$T^*_0 = e^{-\frac{ht}{H\rho c}}g\left(\sqrt{r^2 + 2\gamma t}\right) \tag{45}$$

$$T^*_1 = e^{\frac{hr^2}{2\gamma H\rho c}}\left(\int_{-\sqrt{r^2 + 2\gamma t}}^{r} -\frac{e^{-\frac{hx^2}{2\gamma H\rho c}} \bullet I\left(x, \frac{r^2 + 2\gamma t - x^2}{2\gamma}\right) \bullet x}{\gamma}dx\right) \tag{46}$$

$$I(r, t) = \frac{\partial^2 T^*_0}{\partial r^2} \tag{47}$$

The final analytical solution of the extraction process is given as

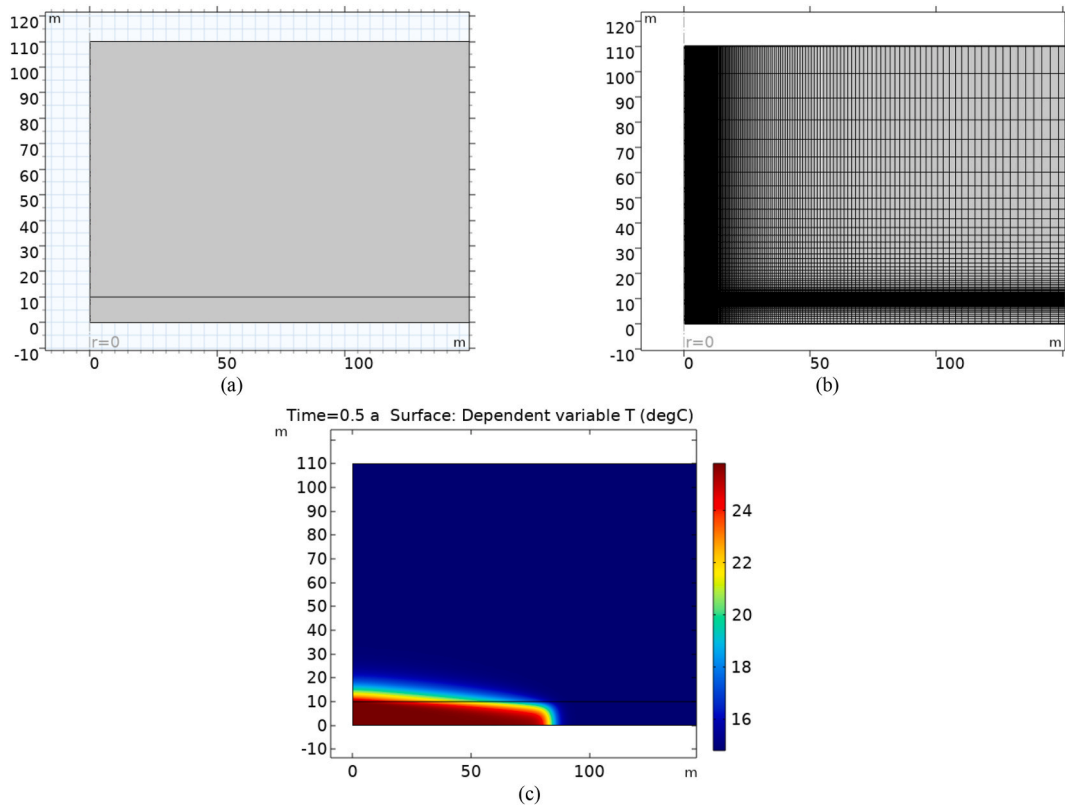$$T^* = T^*_0 + D \bullet T^*_1 \tag{48}$$

Fig. 6. (a) A 2D benchmark simulation model with half-modeled aquifer layer (bottom) coupled with impermeable layer (above) (b) Mesh (c) Example temperature distribution after the first-year injection.

3.2.2.2. Injection. The dimensionless aquifer temperature distribution after the last extraction process $f(r)$ is defined as $f(0)e^{-\frac{r^a}{b}}$, where a, b, and $f(0)$ are fitting parameters. Substitute Eq. (37) into the modified governing equation Eq. (49) results in two equation systems: (i) Eq. (50) with the initial and boundary conditions given by Eqs. (51)–(53), and (ii) Eq. (54) with the initial condition given by Eq. (55).

$$D\frac{\partial^2 T^*}{\partial r^2} - \frac{\beta}{r}\frac{\partial T^*}{\partial r} = \frac{\partial T^*}{\partial t} + \frac{hT^*}{\rho cH} \tag{49}$$

$$\frac{\partial T_0^*}{\partial t} + \frac{\beta}{r}\frac{\partial T_0^*}{\partial r} + \frac{hT_0^*}{\rho cH} = 0 \tag{50}$$

$$T_{11}^* = -\frac{\left(af(0)e^{-\frac{ht}{H\rho c}}t\left(-i\sqrt{-r^2+2\beta t}\right)^a\left((-1+a)br^2 - ab\beta t - a\left(-i\sqrt{-r^2+2\beta t}\right)^a(r^2-\beta t)\right)\right)}{b^2(r^2-2\beta t)^2} \tag{59}$$

$$r=0, T_0^* = 1 \tag{51}$$

$$r=\infty, T_0^* = 0 \tag{52}$$

$$t=0, T_0^* = f(r) \tag{53}$$

$$\frac{\partial T_1^*}{\partial t} + \frac{\beta}{r}\frac{\partial T_1^*}{\partial r} + \frac{hT_1^*}{\rho cH} = \frac{\partial^2 T_0^*}{\partial r^2} \tag{54}$$

$$t=0, T_1^* = 0 \tag{55}$$

The following shows the final solutions of Eq. (50) and Eq. (54). Both T_0^* and T_1^* are stepwise functions. T_{00}^* and T_{01}^* constitute T_0^*. T_{10}^* and T_{11}^* constitute T_1^*.

$$T_{00}^* = e^{-\frac{hr^2}{2\beta H\rho c}} \tag{56}$$

$$T_{01}^* = f(0)e^{-\frac{ht}{H\rho c}} \tag{57}$$

$$T_{10}^* = -\frac{e^{-\frac{hr^2}{2\beta H\rho c}}ht(-hr^2 + \beta H\rho c + \beta ht)}{\beta^2 H^2 \rho c^2} \tag{58}$$

The diffusion-dominated case also considers the discontinuity elimination along the curve. The final analytical solution for the injection process is given below.

$$T^* = \begin{cases} T_{01}^* + D \bullet T_{11}^* + \dfrac{1-f(0)}{2}e^{-\frac{ht}{H\rho c}}erfc\left(\dfrac{r-\sqrt{2\beta t}}{2\sqrt{Dt}}\right), r \geq \sqrt{2\beta t} \\[4mm] T_{00}^* + D \bullet T_{10}^* - \dfrac{1-f(0)}{2}e^{-\frac{ht}{H\rho c}}erfc\left(-\dfrac{r-\sqrt{2\beta t}}{2\sqrt{Dt}}\right), r < \sqrt{2\beta t} \end{cases} \tag{60}$$

Table 6
Properties of each layer used in the numerical model.

No	Layer [0: Aquitard; 1: Aquifer]	Depth [m]	Porosity	Thermal conductivity [W/mK]	Volumetric heat capacity [J/m³K]	Ambient groundwater temperature [°C]
1	0	20	0.3	1.409	1312440	12.5
2	1	30	0.3	1.574	2512080	12.5
3	0	5	0.3	1.409	1312440	12.5
4	1	20	0.3	1.574	2512080	12.5
5	0	10	0.3	1.409	1312440	12.5
6	1	10	0.3	1.574	2512080	12.5
7	0	40	0.3	1.409	1312440	12.5
8	1	20	0.3	1.574	2512080	12.5
9	0	45	0.3	1.409	1312440	12.5

3.2.2.3. Summary. Fig. 5b shows the calculation flow chart of the diffusion-dominated case. First, Eq. (15) is used to evaluate the dimensionless aquifer temperature distribution after the first-year injection $T^*_{inject}1$. Second, with $g(r) = T^*_{inject}1$, Eq. (48) is used to evaluate the dimensionless aquifer temperature distribution after the first-year extraction $T^*_{extract}1$. Third, $f(0)e^{-\frac{r^a}{b}}$ is used to fit $T^*_{extract}1$ and update $f(r)$. Eq. (60) is used to evaluate the dimensionless aquifer temperature distribution after the second-year injection $T^*_{inject}2$. For the subsequent extractions and injections, Eqs. (48) and (60) are used to evaluate the dimensionless aquifer temperature distribution iteratively by updating $g(r)$ and $f(r)$.

3.3. The extraction-dominated scenario ($\alpha_t \bullet \alpha_Q < 1$)

This extraction-dominated scenario is used when a well extracts more water than it injects in 12 months. The injection process utilizes the advection-dominated case in the injection-dominated scenario because the thermal interaction between the aquifer and the impermeable layers is more significant than the diffusion in the aquifer layer during the injection. For the extraction process, because heat is extracted from the impermeable layers, it is assumed that the impermeable layers have a temperature distribution $g(r)$ same as the aquifer layer after the last injection. The governing equation then becomes as follows.

$$\left(\frac{D}{r}+v\right)\frac{\partial T^*}{\partial r} = \frac{\partial T^*}{\partial t} + \frac{h(T^* - g(r))}{\rho cH} \tag{61}$$

$$r = 0, T^* = 1 \tag{62}$$

$$r = \infty, T^* = 0 \tag{63}$$

$$t = 0, T^* = g(r) \tag{64}$$

It is worth noting that the extraction process in the injection-dominated scenario ignores heat extraction from the impermeable layers because, most of the time, the impermeable layers absorb the heat from the aquifer. The solution to the above governing equation is

$$T^* = e^{-\frac{ht}{H\rho c}} \bullet g\left(\sqrt{r^2 + 2\gamma t}\right) + e^{\frac{hr^2}{2\gamma H\rho c}} \int\limits_{\sqrt{r^2+2\gamma t}}^{r} -\frac{e^{\frac{hx^2}{2\gamma H\rho c}}hg(x)x}{\gamma H\rho c}dx \tag{65}$$

3.3.1. Summary

Fig. 5c shows the calculation flow chart of the extraction-dominated scenario. First, Eq. (15) is used to evaluate the dimensionless aquifer temperature distribution after the first-year injection $T^*_{inject}1$. Second, with $g(r) = T^*_{inject}1$, Eq. (65) is used to evaluate the dimensionless aquifer temperature distribution after the first-year extraction $T^*_{extract}1$. Third, with $f(r) = T^*_{extract}1$, Eq. (36) is used to evaluate the dimensionless aquifer temperature distribution after the second-year injection $T^*_{inject}2$. For the subsequent extractions and injections, Eqs. (65) and (36) are used to evaluate the dimensionless aquifer temperature distribution

iteratively by updating $g(r)$ and $f(r)$.

4. ML-assisted heat loss term identification

During seasonal cyclic operation, the thermal interaction between the aquifer and impermeable layer involves complex nonlinear relationships and is compacted into the heat loss coefficient h. This study aims to link h to the subsurface and ATES operational properties. As the intricate relationship is impossible to achieve in this problem, ML algorithms that leverage large datasets are used to identify underlying patterns and nonlinear mappings between input and output variables. In this study, a numerical model of ATES is built and run for different conditions to generate the training datasets for ML.

4.1. FE benchmark model

A full 2D axisymmetric FE benchmark model of the well in ATES is built using COMSOL. Considering symmetry, Fig. 6a shows the geometry of the FE model, where a half-modeled aquifer layer is coupled with an impermeable layer. The thickness of the impermeable layer is set to 100 m, and the right boundary of the model is set to be 1000 m away from the well so that the boundary effects can be ignored. The thickness of the aquifer is set based on the parameters in Table 4. The temperature field in the aquifer layer is governed by the convection-diffusion equation shown in Eq. (66) by assuming an incompressible radial flow. The temperature field in the impermeable layer is governed by the diffusion equation given by Eq. (67). The left boundary of the aquifer layer is set with a constant temperature and a seasonal cyclic injection and extraction flow rate. The heat flux through the left boundary of the impermeable layer is set to zero. Because the aquifer is symmetric to the bottom boundary, the heat flux through the bottom boundary is set to zero. The top boundary of the impermeable layer and the right boundaries of the aquifer and impermeable layers are set with a constant temperature. Fig. 6b shows the mesh of the FE model. The mesh is generated using the 2D mapped meshing with predefined logarithmic distribution nodes on 1D boundary lines. The number of elements is 73,600. Flow rate changes significantly around the well. Sharp spatial gradients in temperature can occur around the aquifer-impermeable layer interface because of the abrupt change in thermal properties. The mesh density is refined near the well and around the aquifer-impermeable layer interface. The resulting mesh size can satisfy the convection computation accuracy and stability by comparing the model with other simulators [101]. Fig. 6c shows the example temperature distribution after the first-year injection. The diffusion effect within the aquifer, coupled with the thermal interaction between the aquifer and the impermeable layer, attenuates the delineation of the injected cylindrical thermal volume, thus facilitating a smooth transition at its periphery.

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rD\frac{\partial T}{\partial r}\right) + D\frac{\partial^2 T}{\partial z^2} - v\frac{\partial T}{\partial r} = \frac{\partial T}{\partial t} \tag{66}$$
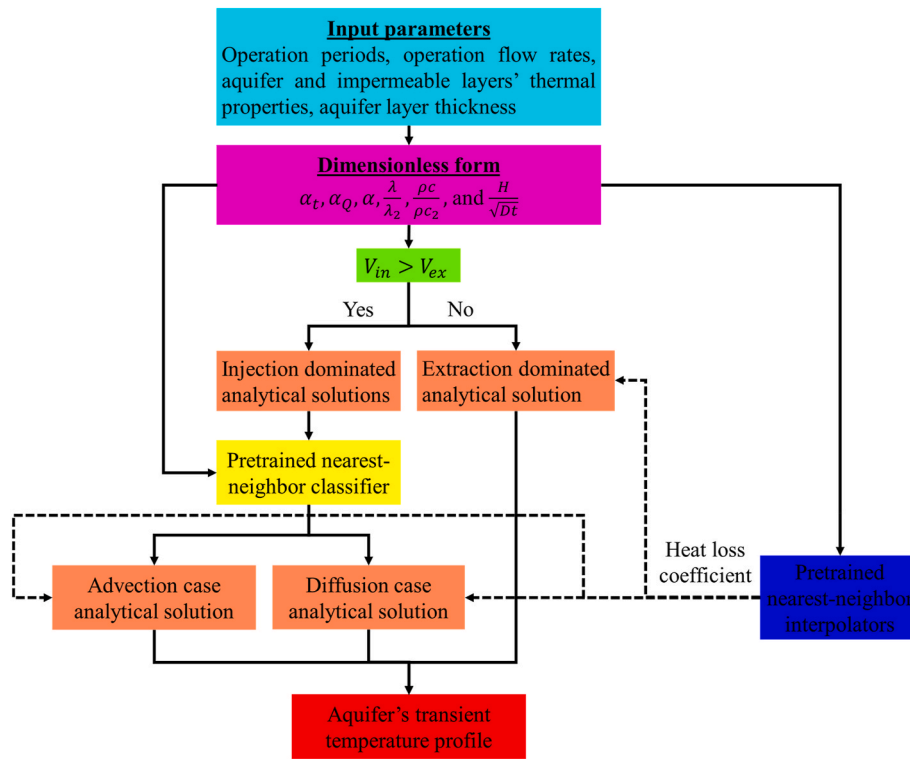
Fig. 7. Flow chart of analytical schemes usage.

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rD_2\frac{\partial T}{\partial r}\right)+D_2\frac{\partial^2 T}{\partial z^2}=\frac{\partial T}{\partial t} \tag{67}$$

The following conditions are assumed for the current numerical modeling: (1) both the aquifer and impermeable layer are homogenous and isotropic; (2) the temperatures of fluid and aquifer are always in the local equilibrium state; (3) the density and thermal properties of fluid are always constant; (4) no regional groundwater flow. Based on the physical properties of the aquifer and impermeable layer and the ATES operational properties listed in Table 4 and Table 6, 300 simulation results of 2D aquifer temperature distribution are generated for training the ML model of heat loss coefficient h, where 3276 simulations are injection-dominated scenarios, and the remaining 3024 simulations are extraction-dominated scenarios. After the ML model is trained, unseen data is needed to provide an unbiased estimate of the model's performance. 1575 simulations, with unseen input properties randomly generated between the minimum and the maximum property limits in Table 4, are run to form the testing dataset. 819 simulations are injection-dominated scenarios, whereas the remaining 756 are extraction-dominated. This way of evaluating the ML model is called hold-out cross-validation. Note that the training and testing dataset sizes conform to the 75%–25 % split ratio, which is the traditional and most common value used in ML. The split ratio can be explained by the bias-variance tradeoff that increasing the training dataset size can help reduce variance. Suppose the training dataset is too large relative to the testing dataset. In that case, there is a risk of overfitting, where the ML model memorizes the training data rather than learning generalizable patterns. In this study, the ML model will have 6 input variables and 1 output variable (see Section 4.2). The present choice of the training dataset size is reasonable compared to the previous study of Ho and Yu [102], which used 13,083 samples to train a KNN interpolator with 19 input variables and 1 output variable, Altay et al. [103] used 161 samples to train several ML classifiers, including a KNN classifier, with 7 input variables and 1 output variable, whereas Xue et al. [104] used 1812 samples to train several ML interpolators, including a KNN interpolator, with 5 input variables and 1 output variable. Cross-validation is

a common way to assess whether the training dataset size is sufficient for training an ML model [68]. This study will cross-validate the model by predicting the aquifer temperature profile and the ATES operation metrics in Section 5.2.

4.2. KNN-based fitting

KNN is a supervised ML algorithm for classification and regression tasks [105]. It belongs to the family of instance-based learning, which means instead of explicitly building a model during the training phase, KNN memorizes the entire training dataset and makes predictions based on the similarity of new instances to known instances. Several advantages of KNN make it a suitable algorithm for the current fitting task: (1) KNN can capture intricate relationships between input and output variables without imposing assumptions about their functional forms, which is crucial in ATES system where various factors can influence the heat transfer process in a nonlinear manner; (2) KNN provides transparent predictions based on the similarity, making it easier to interpret how different properties impact the heat transfer; (3) KNN is a non-parametric method, which means the trained KNN model can be easily extended through incorporating more data without fine-tuning or retraining.

KNN performs classification and regression to predict the class and the value of a new data point by selecting the majority class among and averaging the values of its K closest neighboring data points in the feature space, respectively [106]. The features represent the input variables, while the label is the target variable that KNN aims to predict based on those features. This study uses the KNN algorithm implemented in the scikit-learn library [107] with the Euclidean distance metric and the number of neighbors to be 5. With the training dataset generated from the benchmark model, two KNN-based models are built; (i) one is for regression to predict the heat loss coefficient h, and (ii) another is for classification to choose the better performance analytical model in the injection-dominated scenario. The classifier can help identify when the advection or diffusion effect will dominate the heat
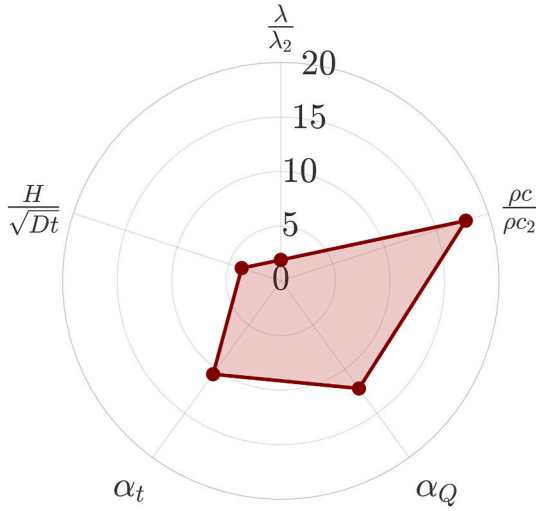
Fig. 8. A linear boundary plane separates the advection- and diffusion-dominated cases with five dimensionless parameters. All axes have the same numeric scale from 0 to 20.

transfer process in the aquifer.

Based on the analytical solutions presented earlier, six dimensionless parameters (α_t, α_Q, α, $\frac{\lambda}{\lambda_2}$, $\frac{\rho c}{\rho c_2}$, and $\frac{H}{\sqrt{Dt}}$) are selected as features for both regression and classification tasks. In the regression model, the dimensionless heat loss coefficient h^* defined by Eq. (68) is selected as the label, where the h value is evaluated by fitting the 1D analytical solution to the 2D FE benchmark results. The 1D benchmark data along the radial direction is calculated from the vertically weighted mean of 2D dimensionless temperature in the aquifer. The L^p (p = 2) distance shown in Eq. (69) measures the difference between the 1D analytical result and the 1D benchmark data. The final optimal h value is found with the trust region reflective algorithm. In the classification model, the name of the better-performance scheme is selected as the label. The benchmark results in the injection-dominated scenario are then compared to those from the advection- and diffusion-dominated analytical solutions with optimal h values. The relative error rate of energy estimation ε_e is given by Eq. (70). Each year's best scheme is the one with a minimum ε_e. The majority voting of each year's best scheme in a specific case during a 50-year operation determines the better performance analytical scheme for this case, as shown by Eq. (71).

$$h^* = \frac{Hh}{2\lambda} \tag{68}$$

$$d_p = \left(\int_0^r \left| T^*_{model}(x) - T^*_{benchmark}(x) \right|^p dx \right)^{\frac{1}{p}} \tag{69}$$

$$\varepsilon_e = \frac{\left| \int_0^r \left(T^*_{model}(x) - T^*_{benchmark}(x) \right) x dx \right|}{\int_0^r T^*_{benchmark}(x) x dx} \tag{70}$$

$$label_{better} = Majority\left(label\left(\min\left(\varepsilon_e|_{advection,t=i}, \varepsilon_e|_{diffusion,t=i} \right) \right)|_{i=1,\dots,50} \right) \tag{71}$$
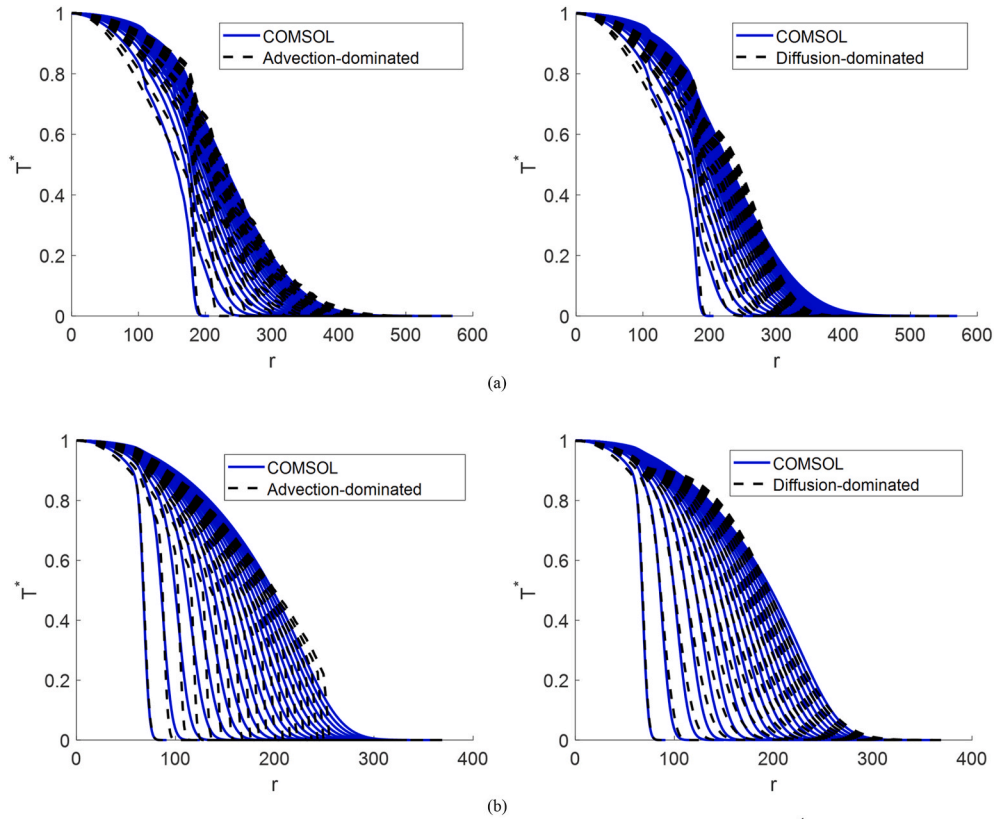

(a)

(b)

Fig. 9. The fitting performance of two analytical solution schemes in (a) an advection-dominated scenario $\left(\alpha_t = 1, \alpha_Q = 1.5, Q_{base} = 0.0417\left(\frac{m^3}{s}\right), \lambda = 2.51\left(\frac{W}{mK}\right), \rho c = 2.74e6\left(\frac{J}{m^3K}\right), H = 10m \right)$ and (b) a diffusion-dominated scenario $\left(\alpha_t = 1.5, \alpha_Q = 2, Q_{base} = 0.0237\left(\frac{m^3}{s}\right), \lambda = 3.01\left(\frac{W}{mK}\right), \rho c = 2.74e6\left(\frac{J}{m^3K}\right), H = 40m \right)$.
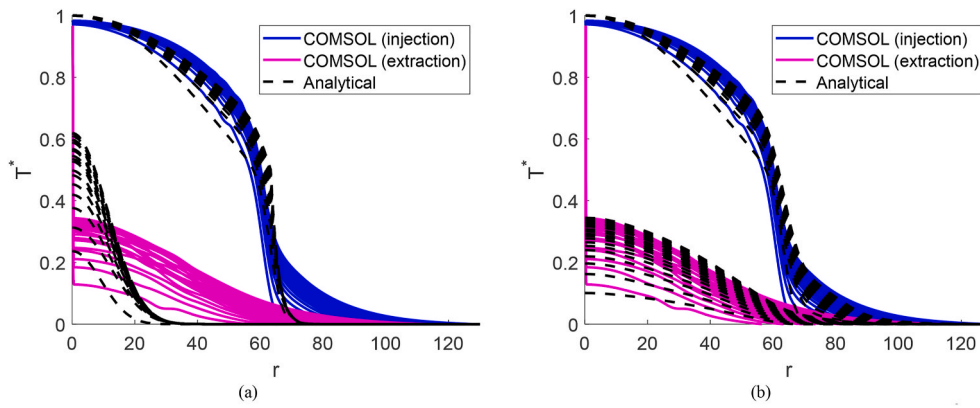
Fig. 10. Fitting performance of analytical schemes in an extraction-dominated scenario $\left(\alpha_t = 0.82, \alpha_Q = 1, Q_{base} = 0.0056 \left(\frac{m^3}{s} \right), \lambda = 2.51 \left(\frac{W}{mK} \right), \rho c = 1.64e6 \left(\frac{J}{m^3 K} \right), H = 10m \right)$ with impermeable layer temperature expressions (a) T_0 and (b) $(T_w - T_0)g(r) + T_0$.

After the KNN-based models are trained and validated, the flow chart to estimate the aquifer temperature profile is shown in Fig. 7. First, when subsurface and ATES operational properties are determined, the values of six dimensionless parameters can be calculated. Second, if the injection volume during the injection period is larger than the extraction volume during the extraction period, the current case is in the injection-dominated scenario; otherwise, it is in the extraction-dominated scenario. Third, if in the injection-dominated scenario, with dimensionless parameters as inputs, the trained KNN classifier selects either the advection-dominated or diffusion-dominated analytical solution as the optimal analytical solution; otherwise, the extraction-dominated analytical solution is chosen as the optimal analytical solution. Fourth, with dimensionless parameters as inputs, the trained KNN interpolator predicts the heat transfer coefficient h. Fifth, with the h value, the optimal analytical solution estimates the transient aquifer temperature distribution during ATES operation.

The trained KNN classifier is used to reveal the heat transfer mechanism in the aquifer. Because the decision boundary learned by the KNN classifier is multi-dimensional and nonlinear, linear discriminant analysis (LDA) is used to distill knowledge to create a linear plane [108], which is shown in Fig. 8. The space outside the boundary represents the diffusion-dominated case. In contrast, the space inside the boundary represents the advection-dominated case. For the long-term operation, the transition between the two cases is related to α_t, α_Q, $\frac{H}{\sqrt{Dt}}$, $\frac{\rho c}{\rho c_2}$, and $\frac{\lambda}{\lambda_2}$. Furthermore, the higher the value of the intersection point location is in a specific variable's axis, the lower the weight value is assigned to this variable. It indicates that the ranking of the variable's relative importance in the classification is $\frac{\lambda}{\lambda_2} > \frac{H}{\sqrt{Dt}} > \alpha_t > \alpha_Q > \frac{\rho c}{\rho c_2}$. This is reasonable as $\frac{\lambda}{\lambda_2}$ describes the magnitude of dispersive effects caused by the flow in the aquifer, $\frac{H}{\sqrt{Dt}}$, α_t, and α_Q describe the geometry of the thermal volume, and $\frac{\rho c}{\rho c_2}$ describes the water content in the aquifer and impermeable layer, which varies within a small range. Doughty et al. [55] commented on similar observations in the significance of variables when analyzing the dependency of the recovery factor on the dimensionless parameters.

5. Results and discussion

The characteristics and fitting performances of the analytical solutions are examined while finding the optimal heat loss coefficients based on the benchmark simulation results. KNN models learn the mappings between the six dimensionless subsurface and ATES operation variables and the dimensionless optimal heat loss coefficient with the training dataset. The prediction accuracy and computational efficiency of the trained KNN-based analytical model are validated on the testing dataset.

5.1. Analytical solution fitting exercise

Fig. 9 shows the 20-year simulation results of the benchmark model and the analytical solutions with optimal heat loss coefficients for the example advection- and diffusion-dominated cases in the injection-dominated scenario for each year. Each line in the figure represents the temperature profile after the injection process. T^* is the dimensionless temperature, and r is the radial distance from the well in meters.

In the example advection-dominated case (Fig. 9a), the 20-year relative error rate of energy estimation $\left(\varepsilon_{20} = \frac{\sum_{i=1}^{20} \varepsilon_e |_{t=i}}{20} \right)$ of the advection-dominated solution is 1.94 %, whereas that of the diffusion-dominated solution is 2.74 %. The profile shape of the advection-dominated solution matches well with that of the benchmark. However, the profile shape of the diffusion-dominated solution exhibits poor fitting; the bulging part due to the perturbation-based approximation causes the mismatch at the profile tail. In the example diffusion-dominated case (Fig. 9b), ε_{20} of the advection-dominated solution is 2.38 %, whereas that of the diffusion-dominated solution is 2.06 %. These results reveal that heat transfer in the aquifer can occur through two main mechanisms: advection and diffusion. Results show that capturing these mechanisms is crucial for accurately predicting the aquifer temperature distribution for heating and cooling applications. Therefore, it is necessary to have two solution schemes to capture aquifer heat transfer mechanisms in the injection-dominated scenario. Although the previous study [36] has tried to qualitatively classify advection- and diffusion-dominated cases based on operation flow rate and aquifer thickness, the finding from this study shows that it is better to use an advanced classifier like KNN because of the nonlinear transition boundary between advection- and diffusion-dominated cases.

Fig. 10 shows the difference in temperature profiles between analytical solutions with two impermeable layer temperature expressions (T_0 and $(T_w - T_0)g(r) + T_0$) and optimal heat loss coefficients in the extraction-dominated scenario for each year, along with 20-year benchmark simulation results. The blue and purple lines in the figure represent the temperature profile after the injection and extraction processes, respectively. In the case of T_0, it fails to capture the benchmark temperature profile's tail caused by the heat extraction from the impermeable layer, resulting in a misfit of the curve. On the other hand, the tail is modeled in the updated temperature case $(T_w - T_0)g(r) + T_0$, and the mismatch from the previous case is fixed. Incorporating the tail of the aquifer temperature distribution profile is crucial for the ATES system design. The tail represents the outer edge of the influenced zone.
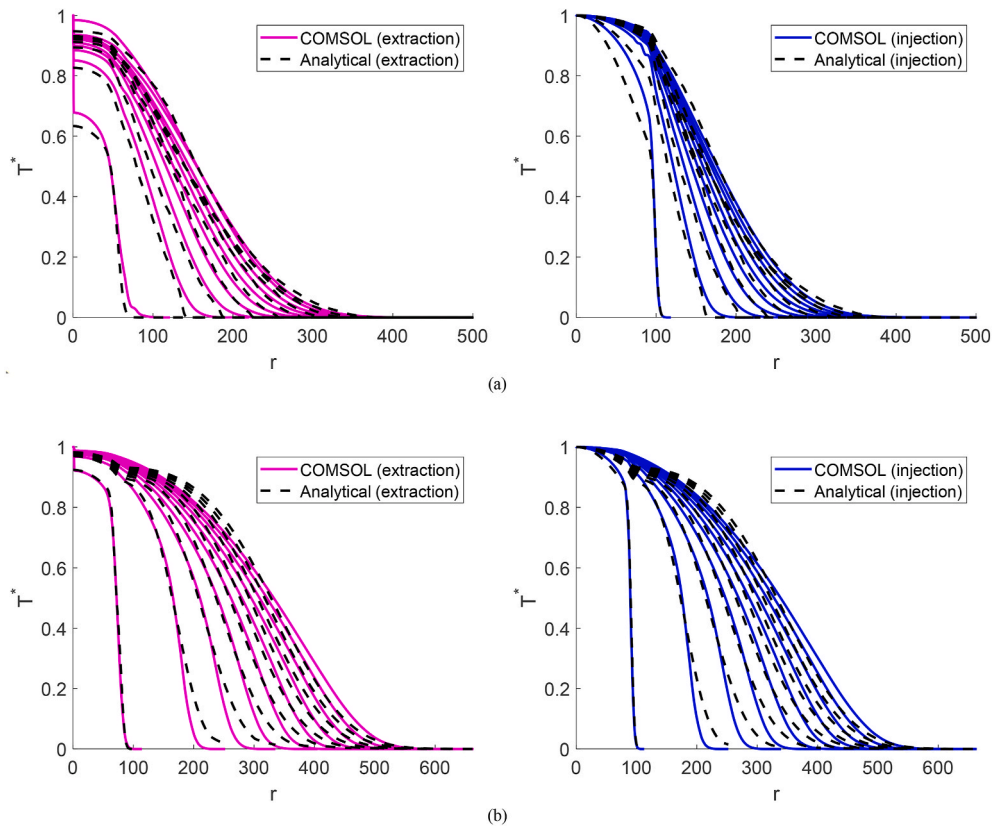
Fig. 11. Temperature profile prediction after extraction (left) and injection (right) processes for (a) an advection-dominated case $\left(\alpha_t = 1.22, \alpha_Q = 1.2, Q_{base} = 0.02\left(\frac{m^3}{s}\right), \lambda = 2.8\left(\frac{W}{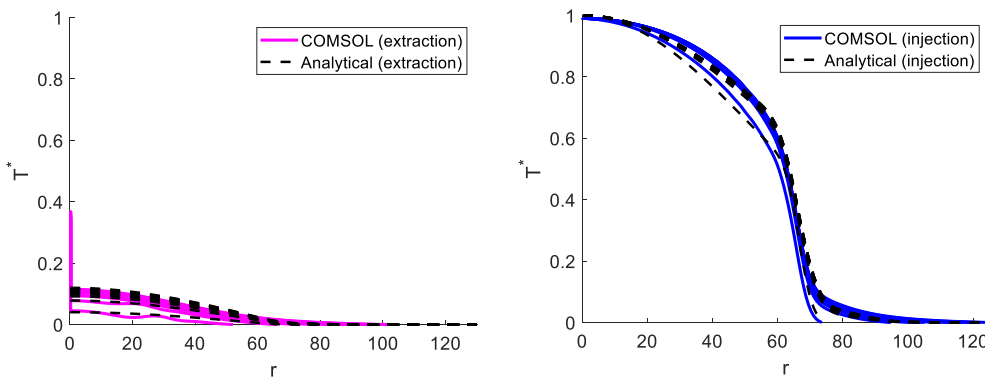mK}\right), \rho c = 2e6\left(\frac{J}{m^3K}\right), H = 16m\right)$ and (b) a diffusion-dominated case $\left(\alpha_t = 1.5, \alpha_Q = 2, Q_{base} = 0.04\left(\frac{m^3}{s}\right), \lambda = 3\left(\frac{W}{mK}\right), \rho c = 2.6e6\left(\frac{J}{m^3K}\right), H = 40m\right)$.

Fig. 12. Temperature profile prediction after extraction (left) and injection (right) processes for an extraction-dominated scenario $\left(\alpha_t = 1, \alpha_Q = 0.5, Q_{base} = 0.006\left(\frac{m^3}{s}\right), \lambda = 2.6\left(\frac{W}{mK}\right), \rho c = 1.7e6\left(\frac{J}{m^3K}\right), H = 10m\right)$.

By modeling the tail, engineers can optimize the placement and spacing of wells to maximize heat recovery efficiency and prevent thermal breakthroughs and depletion of the stored energy.

The reason for the better performance of the updated case in curve fitting can be explained as follows. In the injection-dominated scenario, more water is being injected into the aquifer than extracted through the year, where the impermeable layer acts as a heat sink and absorbs thermal energy from the aquifer. Since the heat continuously diffuses into the impermeable layer and its beyond from the aquifer, it is appropriate to model the impermeable layer temperature as T_0, because the average temperature of the broader area of the subsurface remains close to the initial temperature. However, in the extraction-dominated scenario, more water is being extracted from the aquifer than injected into it throughout the year, where the impermeable layer acts as a heat source and dissipates thermal energy into the aquifer during the water extraction process. Since the temperature of the aquifer is mainly affected by the heat concentration area in the impermeable layer that closes to the aquifer-impermeable layer interface, it is suitable to model the impermeable layer temperature as $(T_w - T_0)g(r) + T_0$, because the temperature distributions of the aquifer and the heat concentration area in the impermeable layer achieve the equilibrium after the injection process. These results indicate the necessity of a modified analytical
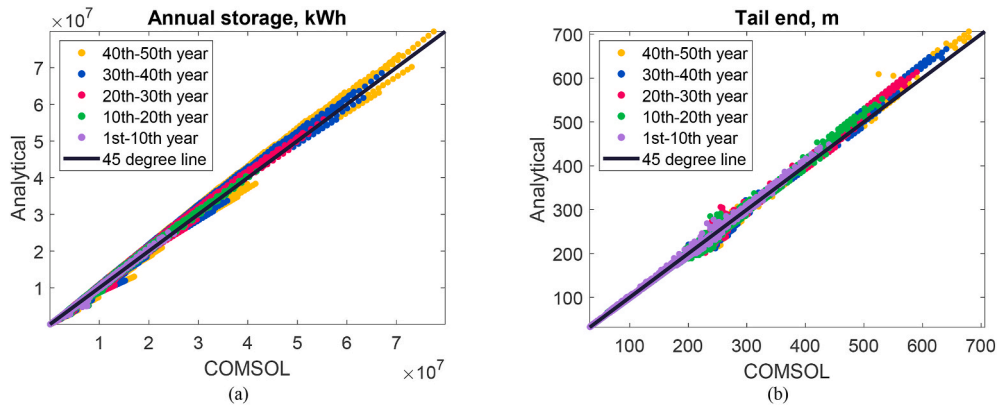
Fig. 13. Cross-validation of the proposed model by comparing the FE results and the analytical scheme predictions in (a) the annual energy storage and (b) the temperature profile tail-end.
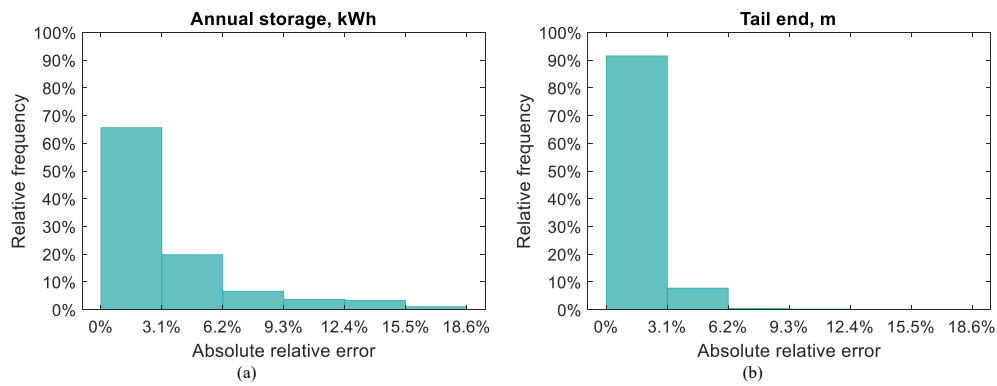

Fig. 14. Histograms of relative absolute error of (a) the annual energy storage and (b) the temperature profile tail-end estimation of the proposed model.
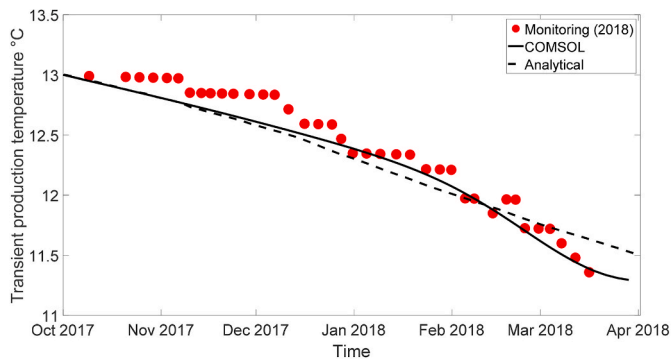

Fig. 15. Comparison of the transient production temperature during the second heating season.

scheme in the extraction-dominated scenario.

5.2. ML-assisted model cross-validation

After the optimal heat loss coefficients are obtained with the fitting exercise, KNN models are implicitly trained by memorizing the entire training dataset. By running on a personal computer with a 6-core Intel® i7-8700K CPU and 16 GB RAM, the average computation time for a 50-year ATES operation of the KNN-based analytical model is 2 min 30 s. In contrast, that of the benchmark model is 3 h 43 min 14 s, which indicates the high computational efficiency of the proposed model.

The trained KNN-based analytical model is tested on the testing dataset. The overall mean of the relative error rate of temperature

profile prediction along the radial direction ($\varepsilon_r = \frac{\int_0^r |T^*_{model}(x) - T^*_{benchmark}(x)| dx}{r}$) of the KNN-based analytical model for the 50-year operation is 2.41 %. More specifically, in the injection-dominated scenario, the mean ε_r for the 50-year operation is 1.91 %. In the extraction-dominated scenario, the mean ε_r for the 50-year operation is 2.96 %. The higher error rate of the extraction-dominated analytical solution in temperature profile estimation can be explained by the analytical solution's underestimation of the tail part of the temperature profile, which is due to the governing equation's lack of accounting for the residual heat accumulation effect in the aquifer and impermeable layer.

Fig. 11 shows example prediction results of the KNN-based analytical model for each 5-year over a 50-year operation in two injection-dominated scenarios. The KNN classifier determines the first parameter set as an advection-dominated case and the second as a diffusion-dominated case. The temperature gradient of the temperature profile can explain the classification results; the former has a sharp temperature profile, whereas the latter has a smooth temperature profile. The profiles in the advection-dominated case match well. The fitting performance in the diffusion-dominated case becomes worse as time increases. This is due to the error caused by the perturbation-based approximation. As time increases, the diffusion effect becomes more significant, and the $O(D^2)$ term in Eq. (37) becomes more unignorable, making it more challenging for the first-order perturbation to capture the pattern. Fig. 12 shows example prediction results of the KNN-based analytical model for each 5-year over a 50-year operation in an extraction-dominated scenario. Because the tail pattern is modeled, the temperature profiles of the proposed model can match well with those of the benchmark model.

This study uses two ATES operation metrics to link the proposed model prediction performance to the ATES system potential evaluation:
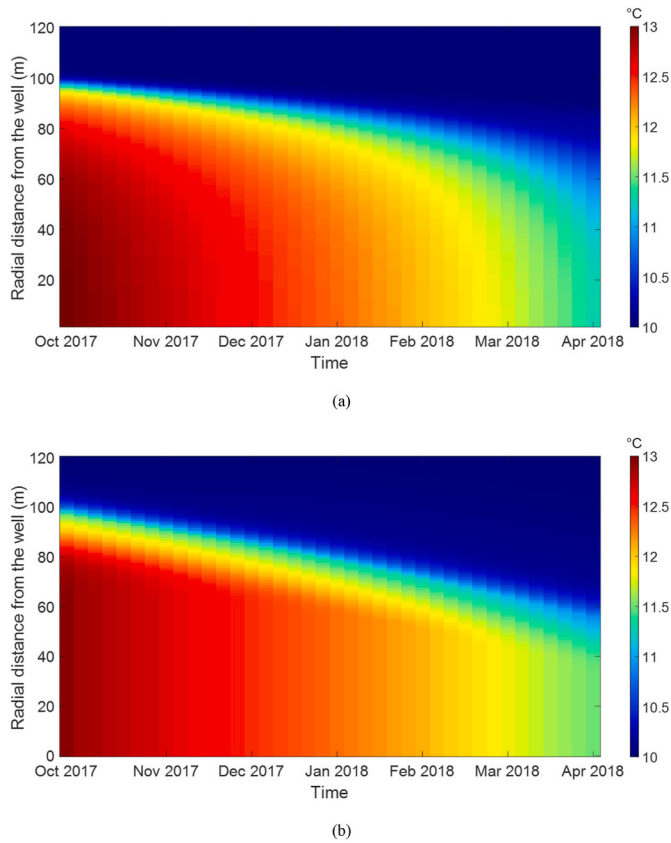
Fig. 16. Spatiotemporal aquifer temperature patterns generated from (a) COMSOL and (b) the analytical solution during the second-year extraction process.
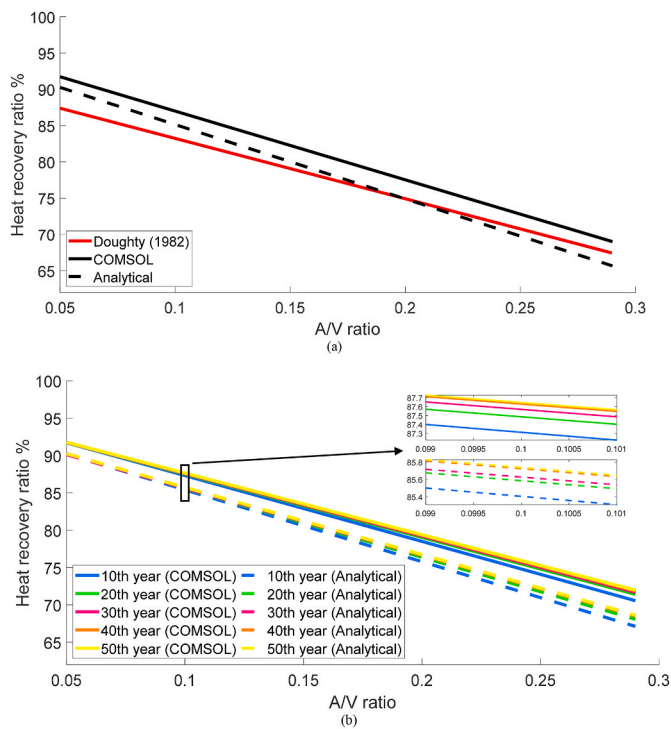

Fig. 17. Comparison of heat recovery ratios from different approaches for (a) the 5th, (b) 10th, 20th, 30th, 40th, and 50th operation year.

(i) annual energy storage and (ii) temperature profile tail end. The annual energy storage, as the yearly total thermal energy injected into the system, is a metric to determine the system's ability to meet the heating/cooling demand [68]. Understanding the annual energy storage helps appropriately size the ATES system components, such as borehole depth and pump capacity. It also allows for estimating the system's contribution to reducing greenhouse gas emissions. On the other hand, the temperature profile tail end is a metric to locate the thermally affected area, from which engineers can plan for maintenance and potential adjustments of the ATES system to prevent thermal breakthroughs and depletion of the stored energy.

The annual energy storage is defined as $E_{store} = \rho c \bullet \int_{0}^{r} |T(x) - T_0| \bullet (2\pi x H) dx$ and the temperature profile tail end is defined as the boundary separates the thermally affected ($T > T_0$) and unaffected area ($T = T_0$). Fig. 13 shows the cross-validation results of the proposed model in two metrics for every ten years with the testing dataset. The horizontal and vertical axes represent the predictions generated by the benchmark and the proposed model. Including a 45-degree line, serving as a reference for perfect alignment between the predictions, facilitates the comparison between the two models. Deviations from this line signify the bias of the proposed model. The comparison results indicate that the proposed model predictions agree with the benchmark simulation results. The proposed model prediction accuracy decays as time increases. This is because the diffusion effects accumulate for the long-term operation, which makes it complicated for the perturbation method to capture.

Fig. 14 shows the histograms of the relative absolute error of the proposed model on estimating the annual energy storage and the temperature profile tail-end. For these two metrics, about 85.4 % and 99.3 % of the proposed model predictions have an error of less than 6.2 %, and the average errors are 3.25 % and 0.83 %, respectively. The more significant error in the annual energy storage estimation can be explained from the definition formula, where the error in energy estimation can be considered as a weighted integral of the temperature difference between the benchmark and the prediction along the radius, and the weight is proportional to the radial distance from the well. Even a slight mismatch in the part of the temperature profile far away from the well can significantly increase the energy estimation error.

This study focuses on matching the temperature profile of the analytical solution to the benchmark so that heat transfer mechanisms in the aquifer can be captured, and the ATES production temperature can be evaluated as precisely as possible, which sacrifices the prediction accuracy of the stored energy. The <5 % error for temperature profile and ATES operation metrics prediction indicates that the chosen training dataset size is sufficient, and the proposed model effectively captures the heat transfer mechanisms in the aquifer. However, to reduce the stored energy estimation error, one can modify the objective function for fitting in Eq. (69) to include the energy difference between the analytical result and the benchmark data with the Lagrange multiplier.

6. Comparison with previous studies

To investigate the practical application possibilities of the proposed model, a short-term field monitoring dataset is used to examine the model's performance in predicting the ATES production temperature, whereas a long-term simulation dataset is used to evaluate the model's ability to estimate the ATES heat recovery efficiency.

6.1. Short-term field data (2018)

This section aims to compare the proposed model predictions with the field monitoring data and use the numerical model to reveal the limitations of the proposed model. Abuasbeh et al. [109] studied a shallow ATES system in the northern part of Stockholm, with a total area
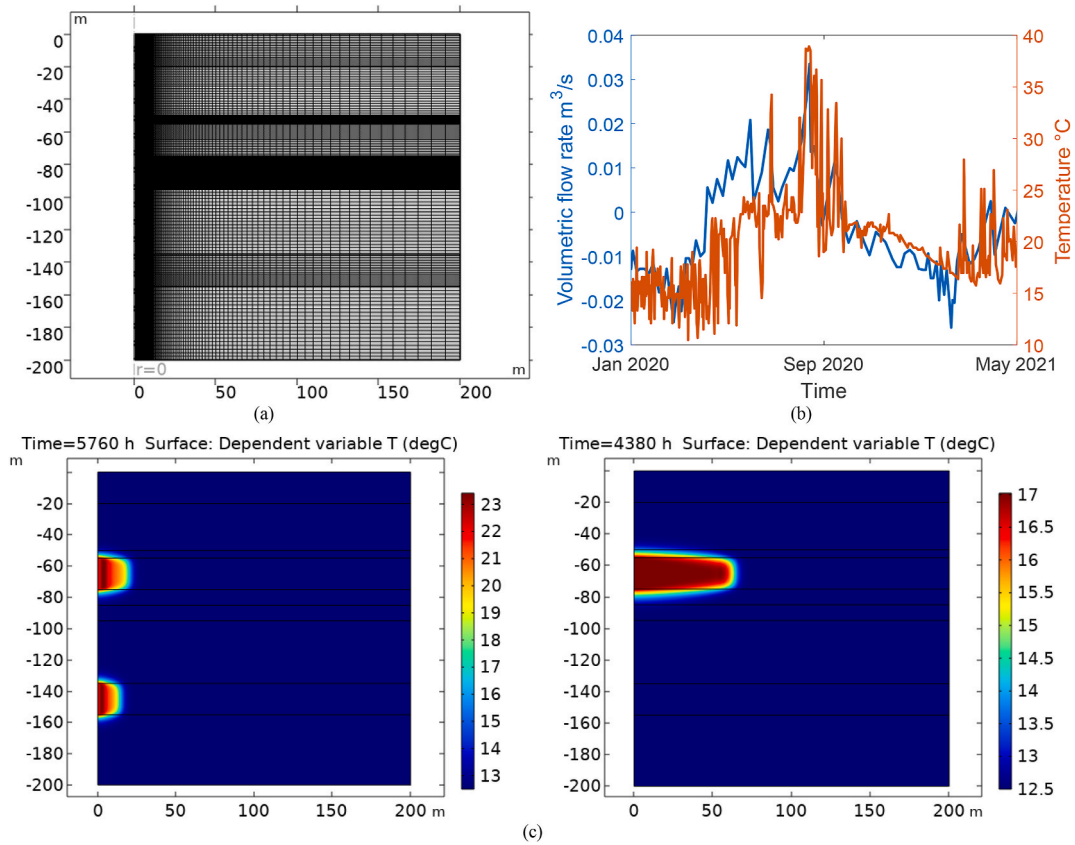
Fig 18. (a) Mesh of the COMSOL model (b) The HT-ATES system's operation water flow rate (blue) and operation water temperature (orange) monitoring data from Jan 2020 to May 2021 (c) Modeled HT-ATES temperature distribution surrounding the warm well in Sep 2020 (d) Modeled hypothetical ATES case temperature distribution surrounding the warm well after the first-year injection
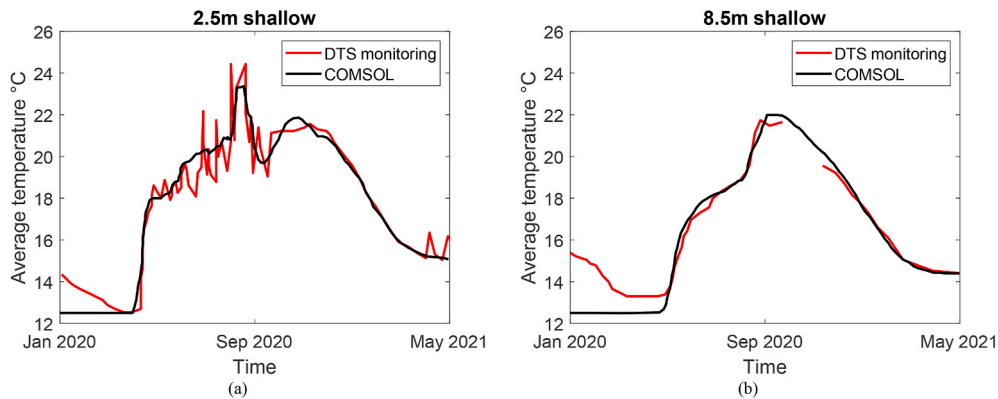

Fig 19. Comparison between modeled (black) and monitored (red) subsurface temperature in the shallow aquifer (Layer 4) at (a) 2.5 m and (b) 8.5 m away from the well

of approximately 30000 m². The ATES system has operated since autumn 2016 and consists of 4 warm and 2 cool wells. The average aquifer thickness is 11.5 m. Drillings have shown that the primary geological material of the studied area comprises sand and gravel. The ATES system was operated seasonally: during the heating season, water from the warm side of the ATES is extracted and reinjected in the cold side, and during the cooling season, the cycle is reversed, and free cooling is utilized. The ATES system has been monitored with temperature sensors and flowmeters. The COMSOL model of one warm well is built in this study to study the aquifer thermal behavior in this ATES system, as in Section 4.1. The model input parameters are listed in Table 5. The duration of the heating season equals that of the cooling season.

Fig. 15 shows the monitored production temperature during the second-year heating season from 2017 to 2018. The trend of the numerical simulation result matches the monitoring result well, indicating that the ATES system's behavior is acceptably represented in the numerical model. The same parameter sets are used as inputs for the proposed model. It shows the average relative error rate of the transient production temperature estimation of the proposed model ($\varepsilon_t = \frac{\int_0^t \frac{|T_{model}(\tau) - T_{benchmark}(\tau)|}{T_{benchmark}(\tau)} d\tau}{t}$) is 0.5 % compared to the numerical result and 1.12 % to the monitoring data.

It is noted that the trend of the proposed model result deviates from

the others. This can be explained by the perturbation-based approximation error, which can be seen from the spatiotemporal temperature patterns of the aquifer shown in Fig. 16. The head part, which is close to the well and has a high temperature, of the aquifer temperature profile of the numerical model shortens faster than that of the proposed model. Similarly, the tail part, which is away from the well and has a low temperature, of the aquifer temperature profile of the numerical model lengthens faster than that of the proposed model. Therefore, the temperature pattern computed from the numerical model shows a more diffusion effect than the proposed model's. The diffusion effect tends to smooth out temperature variation over time, leading to a more uniform temperature distribution. Modeling diffusion effects is crucial for accurately predicting the spread of thermal energy over time and space.

To reduce the difference caused by the diffusion effect, one can add higher-order diffusion-related perturbation terms into the analytical solution. However, as the order increases, it becomes harder to derive the analytical solution. Another option is to use the FE method to numerically solve the 1D governing equation with the trained KNN model, providing the heat loss coefficient, which will cost more computation time but is still faster than the full numerical model. The authors have released software that can be used to achieve this purpose.[2]

Overall, the production temperature can directly impact the heat exchange efficiency of the heat pump and the electricity consumption of the system operation. By knowing the expected production temperature, system operators can adjust various parameters, such as flow rates, to ensure the system operates at its peak efficiency. In this case, the small error in the production temperature prediction of the proposed model shows its potential to be used in the building-coupled modeling to estimate the system's performance.

6.2. Long-term simulation data

The long-term accuracy of the proposed model is examined by evaluating the heat recovery of a hypothetical ATES case and comparing the estimated results to the outputs of the following two studies.

(i) **The Koppert Cress ATES case (2021) -** The configuration of the ATES system of Koppert Cress in the Netherlands is used, which started the operation in 2015 [110]. This ATES system utilizes two aquifers of 20 m thickness and has 4 warm and 4 cool wells. The numerical model of one warm well is developed in this study, where the input parameters are listed in Table 6. The model has been validated by simulating the actual operation and comparing it to the monitoring data. To compare the validated numerical model with the proposed model and the previous study [55], which only considers the single aquifer layer operation, a hypothetical ATES case is studied where only the shallow aquifer (Layer 4) is used for energy production and storage. It assumes the system injects and produces the same warm water yearly. The duration of the heating season equals that of the cooling season. The injection water temperature of the warm well is 17 $^\circ C$. The details of the validation and simulation can be found in Appendix A.

(ii) **Doughty et al. (1982) study -** Doughty et al. [55] conducted numerical simulations of a single-well system with extra storage and rest periods considered. The duration of the system operation is five years. In their steady flow model (SFM), heat transfer by convection is accounted for by translating the aquifer temperature field during the injection period. The diffusion equation describes heat transfer by conduction. During the storage and rest periods, no translation of the temperature field occurs, and heat transfer is purely by conduction. During the production period, the convection is treated as during the injection period. They

assumed the injection, storage, production, and rest periods were of equal duration. The conduction-based heat loss during the storage and rest periods cannot be ignored because these periods last for half a year.

Heat recovery ratio η represents the efficiency of retrieving stored thermal energy from the aquifer. It is defined as $\frac{\int_{t_{start}}^{t_{end}} Q_e(T_e-T_0)dt}{\int_{t_{start}}^{t_{end}} Q_i(T_i-T_0)dt}$, where T_i and T_e represent the injection and extraction water temperature [111]. η directly impacts the system's overall performance and economic viability by determining the amount of stored thermal energy that can be effectively utilized for heating or cooling purposes. Optimizing η not only enhances energy efficiency but reduces operational costs and environmental impact. The previous studies [50,55,61] concluded that the relationship between the heat recovery ratio η and the surface-to-volume ratio A/V is linear when the density-dependent flow is negligible, and no background groundwater flow occurs. A/V is defined as the ratio between the surface and the volume of the heated cylinder region in the aquifer [55].

The linear relationship of the 5th-year η predicted by the proposed model is compared to those estimated from the Koppert Cress ATES case model and Doughty et al.'s model, as shown in Fig. 17a. The figure shows that η decreases as A/V increases. This is because a larger storage volume has a smaller A/V. The heat loss into the impermeable layer in the earlier cycles dampens the loss in the following cycles [55]. Although the absolute heat loss increases with increasing storage volume, the relative heat loss decreases as the large storage volume has a more significant dampening effect.

The average relative error rate of η estimated by the proposed model is 3.06 % compared to the Koppert Cress ATES case result and 1.55 % compared to Doughty et al.'s result. The analytical solution derived in this study is forced to match the temperature profile trend during the fitting exercise and ignore the stored energy estimation. Consequently, a slight mismatch of the temperature profile can result in a noticeable energy estimation error.

The η values of the Koppert Cress ATES case model are larger than those of Doughty et al.'s model because the latter model considers the dissipation during the storage and the rest periods and has differences in parameters and model set-up compared to the former. The difference between the two models decreases as A/V increases. This can be explained by the boundary plane in Fig. 8 that a small ATES system with a larger A/V has a smaller $\frac{H}{\sqrt{Dt}}$, causing its aquifer thermal response closer to the advection-dominated case with less heat diffusion in the storage and rest periods.

The difference between the Koppert Cress ATES case model and the proposed model increases as A/V increases. This can be explained by the pattern shown in Fig. 11, which shows that the proposed model underestimates the extracted energy more for the advection-dominated case than for the diffusion-dominated case. Fig. 17b compares the η values between the proposed model and the Koppert Cress ATES case model for the 10th, 20th, 30th, 40th, and 50th operation years. The average relative error rates of η estimation of the proposed model are 3.12 %, 3.04 %, 2.96 %, 3.04 %, and 3.03 % compared to the numerical model for the 10th, 20th, 30th, 40th, and 50th operation years, respectively. The η values of the Koppert Cress ATES case model are slightly larger than those estimated by the proposed model, as explained earlier. The variation in η with different operation years is slight because the system approaches the steady state. The variation in η becomes smaller as A/V decreases because a large relative storage volume achieves the steady state faster with the dampening effect. The relatively small error in the heat recovery ratio predicted by the proposed model shows its potential to be used in the long-term energy utilization efficiency assessment of the ATES system.

7. Conclusions, limitations, and future research

This study proposes novel 1D analytical solutions with unknown heat loss coefficient from a 1D cylindrical convection-diffusion equation. KNN is used to learn the mappings between the subsurface and ATES operation properties and the optimal heat loss coefficient that is obtained from the fitting exercise between a 2D axisymmetric FE benchmark model and the 1D analytical solutions. The resulting ML-assisted analytical model or G-function allows spatiotemporal prediction of aquifer temperature profiles during seasonal bidirectional ATES operation. As a computationally efficient tool, the ML-assisted G-function can help facilitate the energy system modeling and evaluation. Several conclusions are drawn as follows.

(1) This study uses The KNN algorithm for classification and regression tasks because it provides transparent predictions based on similarity and makes its model decision easy to interpret. The cross-validation of the proposed model shows that for a given set of operational and ground conditions, the KNN classifier could select between the advection- and diffusion-dominated cases, and the KNN interpolators could identify the appropriate heat loss coefficient. This finding can be explained by KNN's ability to handle non-linearities. The resulting KNN model can be easily extended by incorporating more data without fine-tuning or retraining. However, although the KNN model is interpretable, it does not explicitly give the relationship between the input and the output variables. If one wishes to derive an explicit expression or build an ML model of the heat loss coefficient, the training dataset can be retrieved from the trained KNN model with input parameters in Table 4.

(2) The fitting exercise between the benchmark model and the analytical solutions proves the necessity of using three ATES operation conditions to describe the aquifer temperature distribution. The proposed model gives an average relative error rate of 2.41 % on the aquifer temperature profile prediction during cross-validation. To link the proposed model prediction performance to the ATES system potential evaluation, the proposed model is tested on estimating two ATES operation metrics, the annual energy storage and the temperature profile tail-end, and the average relative error rates of the prediction are 3.25 %, and 0.83 %, respectively. The overall <5 % error indicates that the proposed model effectively captures the heat transfer mechanisms in the aquifer. By running on a personal computer, the average computation time for a 50-year ATES operation of the proposed model is 2 min 30 s. In comparison, that of the benchmark model is 3 h 43 min 14 s, which indicates the high computational efficiency of the proposed model.

(3) When the model performance is assessed using the short-term field monitoring dataset, the proposed model gives an average relative error rate of 1.12 % on the transient production temperature estimation. By examining spatiotemporal aquifer temperature patterns from a reference numerical model and the proposed model, it is found that the estimation error is caused by perturbation-based approximation and the failure of the model to capture the diffusion effect comprehensively. In this case, the small error in the production temperature prediction of the proposed model shows its potential to be used in the building-coupled modeling to estimate the system's performance.

(4) When the model performance is assessed using the long-term dataset, the proposed model gives the average relative error rates of 3.06 %, 3.12 %, 3.04 %, 2.96 %, 3.04 %, and 3.03 % on the heat recovery ratio estimation for the 5th, 10th, 20th, 30th, 40th, and 50th operation years. The small error in the proposed model's heat recovery ratio prediction shows its potential to be used in the long-term energy utilization efficiency assessment of the ATES system.

(5) The analytical solutions are derived based on the assumption that the aquifer has a small diffusion coefficient. For an aquifer with a substantial diffusion effect, it is recommended to evaluate the aquifer's thermal behavior numerically. A numerical model is also recommended for an aquifer with a complicated ATES operation, where the temperature and flow rate undergo drastic changes in a short time or with heterogeneous subsurface properties.

There are several opportunities to develop the proposed model further. To consider the effect of the ambient groundwater flow, the proposed model can adopt the moving thermal cylinder model proposed by Ref. [50]. The analytical solution from Ref. [112] can be fused into the proposed model to consider the storage period during the ATES operation.

CRediT authorship contribution statement

Kecheng Chen: Methodology, Software, Validation, Visualization, Writing – original draft. **Xiang Sun:** Software, Writing – review & editing. **Kenichi Soga:** Methodology, Supervision, Writing – review & editing, Funding acquisition. **Peter S. Nico:** Funding acquisition, Writing – review & editing. **Patrick F. Dobson:** Funding acquisition, Writing – review & editing.

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.

Data availability

Data will be made available on request.

Appendix A. Numerical model validation based on the Koppert Cress ATES case (2021)

Koppert Cress is a horticulture company located in the western part of the Netherlands. The ATES system of Koppert Cress started to inject heated groundwater with temperatures above the conventional maximal injection temperature of 25 °C in 2015 [110]. As this is one of the few high-temperature ATES (HT-ATES) locations in the Netherlands, extensive monitoring of the ATES system operation was performed. The ATES system utilizes 2 aquifers of 20 m thickness and has 4 warm and 4 cool wells. The 2D axisymmetric COMSOL model of one warm well is developed. Fig. 18a shows the model mesh. To verify the correctness of the FE model, we simulate the actual HT-ATES system operation and compare the results to the distributed temperature sensing (DTS) data [113]. The simulation settings are listed in Table 7. The warm well penetrates layers 4 and 8. The

simulation only considers the period from 2020 to 2021, for which the system's operation data are complete, as shown in Fig. 18b. The modeled temperature distribution surrounding the warm well in Sep 2020 is shown in Fig. 18c. The modeled temperature of layer 4 at 2.5 m and 8.5 m away from the well matches the monitored temperature, as shown in Fig. 19. The corresponding average relative error rates of the FE model in temperature estimation are 2.69 % and 3.58 %, compared to the DTS monitoring. The residual heat from the previous HT-ATES operation can explain the difference between Jan 2020 and Apr 2020. The <5 % error rate indicates the FE model's ability to reproduce the ATES system's thermal behavior. For the hypothetical ATES case, the modeled temperature distribution surrounding the warm well for the first-year injection, when the annual injection volume is 250000 m^3, is shown in Fig. 18d.

Table 7
HT-ATES case simulation settings

No	Volume distribution [%]	Boundary conditions			
		Top	Bottom	Left	Right
1	–	Constant temperature	–	Zero heat flux	Constant temperature
2	–	–	–	Zero heat flux	Constant temperature
3	–	–	–	Zero heat flux	Constant temperature
4	65	–	–	Varying temperature and flow rate	Constant temperature
5	–	–	–	Zero heat flux	Constant temperature
6	–	–	–	Zero heat flux	Constant temperature
7	–	–	–	Zero heat flux	Constant temperature
8	35	–	–	Varying temperature and flow rate	Constant temperature
9	–	–	Constant temperature	Zero heat flux	Constant temperature

References

[1] Fleuchaus P, Godschalk B, Stober I, Blum P. Worldwide application of aquifer thermal energy storage–A review. Renew Sustain Energy Rev 2018 Oct 1;94:861–76.
[2] Caliskan H, Dincer I, Hepbasli A. Thermodynamic analyses and assessments of various thermal energy storage systems for buildings. Energy Convers Manag 2012 Oct 1;62:109–22.
[3] Mahon H, O'Connor D, Friedrich D, Hughes B. A review of thermal energy storage technologies for seasonal loops. Energy 2022 Jan 15;239:122207.
[4] Office of Environmental Management. Technology readiness assessment guide. Washington, D.C. (USA): United States Department of Energy; 2011.
[5] Council E. Directive on the energy performance of buildings. Brussels (Belgium): European Commission; 2010.
[6] Aneke M, Wang M. Energy storage technologies and real life applications–A state of the art review. Appl Energy 2016 Oct 1;179:350–77.
[7] Yongfu S. The experiment of storing cold water and warm water in aquifer in Shanghai, PR China and its effect. In: Paper presented at: the 6th International conference on thermal energy storage; 1991.
[8] Lu H, Tian P, He L. Evaluating the global potential of aquifer thermal energy storage and determining the potential worldwide hotspots driven by socio-economic, geo-hydrologic and climatic conditions. Renew Sustain Energy Rev 2019 Sep 1;112:788–96.
[9] Hicks RJ, Stewart DL. Environmental assessment of the potential effects of aquifer thermal energy storage systems on microorganisms in groundwater. Richland (USA): Pacific Northwest Lab; 1988.
[10] Bloemendal M, Olsthoorn T, van de Ven F. Combining climatic and geo-hydrological preconditions as a method to determine world potential for aquifer thermal energy storage. Sci Total Environ 2015 Dec 15;538:621–33.
[11] Andersson O, Hellström G, Nordell B. Heating and cooling with UTES in Sweden: current situation and potential market development. 2003 Sep 1-4. In: Proceedings of the 9th International conference on thermal energy storage. Warsaw, Poland. Warszawa: Warsaw University of Technology; 2003. p. 207–15.
[12] Bayer P, Attard G, Blum P, Menberg K. The geothermal potential of cities. Renew Sustain Energy Rev 2019 May 1;106:17–30.
[13] Schüppler S, Fleuchaus P, Blum P. Techno-economic and environmental analysis of an aquifer thermal energy storage (ATES) in Germany. Geoth Energy 2019 Dec;7(1):1–24.
[14] Sommer W, Valstar J, van Gaans P, Grotenhuis T, Rijnaarts H. The impact of aquifer heterogeneity on the performance of aquifer thermal energy storage. Water Resour Res 2013 Dec;49(12):8128–38.
[15] Todorov O, Alanne K, Virtanen M, Kosonen R. Aquifer thermal energy storage (ATES) for district heating and cooling: a novel modeling approach applied in a case study of a Finnish urban district. Energies 2020 May 14;13(10):2478.
[16] Bakr M, van Oostrom N, Sommer W. Efficiency of and interference among multiple aquifer thermal energy storage systems; A Dutch case study. Renew Energy 2013 Dec 1;60:53–62.
[17] Bloemendal M, Vardon PJ, Pijnenborg M, Sudintas G, Medema A, Marif K, Beernink S, van Veldhuizen F, Snelleman S, van Oort T. A techno-economic evaluation of high temperature thermal aquifer storage (HT-ATES) for use with the geothermal well on the TU Delft campus. In: Proceedings of world geothermal Congress 2020+1; 2021 Oct 24-27; Reykjavik, Iceland. Bochum (Germany): International Geothermal Association; 2021, 16013.

[18] Beernink S, Barnhoorn A, Vardon PJ, Bloemendal M, Hartog N. Impact of vertical layering and the uncertainty and anisotropy of hydraulic conductivity on HT-ATES performance. In: Proceedings of European geothermal Congress; 2022 Oct 17-21. Berlin, Germany. Brussels (Belgium): The European Geothermal Energy Council; 2022. p. 442.
[19] Guo C, Zhang K, Pan L, Cai Z, Li C, Li Y. Numerical investigation of a joint approach to thermal energy storage and compressed air energy storage in aquifers. Appl Energy 2017 Oct 1;203:948–58.
[20] Xiao X, Jiang Z, Owen D, Schrank C. Numerical simulation of a high-temperature aquifer thermal energy storage system coupled with heating and cooling of a thermal plant in a cold region, China. Energy 2016 Oct 1;112:443–56.
[21] Major M, Poulsen SE, Balling N. A numerical investigation of combined heat storage and extraction in deep geothermal reservoirs. Geoth Energy 2018 Dec;6(1):1–6.
[22] De Schepper G, Paulus C, Bolly PY, Hermans T, Lesparre N, Robert T. Assessment of short-term aquifer thermal energy storage for demand-side management perspectives: experimental and numerical developments. Appl Energy 2019 May 15;242:534–46.
[23] Kim J, Lee Y, Yoon WS, Jeon JS, Koo MH, Keehm Y. Numerical modeling of aquifer thermal energy storage system. Energy 2010 Dec 1;35(12):4955–65.
[24] Jeon JS, Lee SR, Pasquinelli L, Fabricius IL. Sensitivity analysis of recovery efficiency in high-temperature aquifer thermal energy storage with single well. Energy 2015 Oct 1;90:1349–59.
[25] Gao L, Zhao J, An Q, Liu X, Du Y. Thermal performance of medium-to-high-temperature aquifer thermal energy storage systems. Appl Therm Eng 2019 Jan 5;146:898–909.
[26] Stemmle R, Lee H, Blum P, Menberg K. City-scale heating and cooling with aquifer thermal energy storage (ATES). Geoth Energy 2024 Jan 18;12(1):2.
[27] Burns ER, Bershaw J, Williams CF, Wells R, Uddenberg M, Scanlon D, Cladouhos T, Van Houten B. Using saline or brackish aquifers as reservoirs for thermal energy storage, with example calculations for direct-use heating in the Portland Basin, Oregon, USA. Geothermics 2020 Nov 1;88:101877.
[28] Perez Silva J, McDermott C, Fraser-Harris A. The value of a hole in coal: assessment of seasonal thermal energy storage and recovery in flooded coal mines. Earth Science, Systems and Society 2022:2.
[29] Dashti A, Grimmer JC, Geuzaine C, Bauer F, Kohl T. Developing meshing workflows for geologic uncertainty assessment in high-temperature aquifer thermal energy storage. Geosci Model Dev Discuss (GMDD) 2023 May 31;2023:1–25.
[30] Ganguly S, Kumar MM, Date A, Akbarzadeh A. Numerical investigation of temperature distribution and thermal performance while charging-discharging thermal energy in aquifer. Appl Therm Eng 2017 Mar 25;115:756–73.
[31] Yapparova A, Matthäi S, Driesner T. Realistic simulation of an aquifer thermal energy storage: effects of injection temperature, well placement and groundwater flow. Energy 2014 Nov 1;76:1011–8.
[32] Buscheck TA, Doughty C, Tsang CF. Prediction and analysis of a field experiment on a multilayered aquifer thermal energy storage system with strong buoyancy flow. Water Resour Res 1983 Oct;19(5):1307–15.
[33] Tsang CF, Buscheck T, Doughty C. Aquifer thermal energy storage: a numerical simulation of Auburn University field experiments. Water Resour Res 1981 Jun;17(3):647–58.
[34] Xue Y, Xie C, Li Q. Aquifer thermal energy storage: a numerical simulation of field experiments in China. Water Resour Res 1990 Oct;26(10):2365–75.
[35] Collignon M, Klemetsdal ØS, Møyner O, Alcanié M, Rinaldi AP, Nilsen H, Lupi M. Evaluating thermal losses and storage capacity in high-temperature aquifer

thermal energy storage (HT-ATES) systems with well operating limits: insights from a study-case in the Greater Geneva Basin, Switzerland. Geothermics 2020 May 1;85:101773.

[36] Chen K, Sun X, Soga K, Dobson PF, Nico PS. 1D heat loss models to predict the aquifer temperature profile during hot/cold water injection. 2021. In: Geothermal rising conference Transactions. San Diego, USA. Davis (USA): Geothermal Rising; 2021. p. 405–19.

[37] Li KY, Yang SY, Yeh HD. An analytical solution for describing the transient temperature distribution in an aquifer thermal energy storage system. Hydrol Process 2010 Dec 15;24(25):3676–88.

[38] Nordbotten JM. Analytical solutions for aquifer thermal energy storage. Water Resour Res 2017 Feb;53(2):1354–68.

[39] Yeh HD, Yang SY, Li KY. Heat extraction from aquifer geothermal systems. Int J Numer Anal Methods GeoMech 2012 Jan;36(1):85–99.

[40] Schulz R. Analytical model calculations for heat exchange in a confined aquifer. J Geophys 1987 Apr 23;61(1):12–20.

[41] Lauwerier HA. The transport of heat in an oil layer caused by the injection of hot fluid. Applied Scientific Research, Section A 1955 Mar;5(2):145–50.

[42] Barends FB. Complete solution for transient heat transport in porous media, following Lauwerier's concept. 2020 Sep 20-22. In: Proceedings of the SPE annual technical conference and exhibition. Florence, Italy. Richardson (USA): Society of Petroleum Engineers; 2020. SPE-134670.

[43] LaForce T, Ennis-King J, Paterson L. Semi-analytical solutions for nonisothermal fluid injection including heat loss from the reservoir: Part 1. Saturation and temperature. Adv Water Resour 2014 Nov 1;73:227–41.

[44] Pophillat W, Attard G, Bayer P, Hecht-Méndez J, Blum P. Analytical solutions for predicting thermal plumes of groundwater heat pump systems. Renew Energy 2020 Mar 1;147:2696–707.

[45] Chen CS. Analytical solutions for radial dispersion with Cauchy boundary at injection well. Water Resour Res 1987 Jul;23(7):1217–24.

[46] Aichi M, Akitaya K. Analytical solution for a radial advection-dispersion equation including both mechanical dispersion and molecular diffusion for a steady-state flow field in a horizontal aquifer caused by a constant rate injection from a well. Hydrological Research Letters 2018;12(3):23–7.

[47] Wu B, Lei Z, Zhao G, Jia S, Zhang X, Jeffrey RG. An efficient model for predicting heat extraction from a multiple-well enhanced geothermal system in the presence of areal flow. Appl Therm Eng 2022 Jun 25;210:118399.

[48] Wu B, Zhang G, Zhang X, Jeffrey RG, Kear J, Zhao T. Semi-analytical model for a geothermal system considering the effect of areal flow between dipole wells on heat extraction. Energy 2017 Nov 1;138:290–305.

[49] Ganguly S, Kumar MM. Analytical solutions for movement of cold water thermal front in a heterogeneous geothermal reservoir. Appl Math Model 2014 Jan 15;38 (2):451–63.

[50] Bloemendal M, Hartog N. Analysis of the impact of storage conditions on the thermal recovery efficiency of low-temperature ATES systems. Geothermics 2018 Jan 1;71:306–19.

[51] Veling EJ. Radial transport in a porous medium with Dirichlet, Neumann and Robin-type inhomogeneous boundary values and general initial data: analytical solution and evaluation. J Eng Math 2012 Aug;75(1):173–89.

[52] Tang DW, Van der Zee SE. Dispersion and recovery of solutes and heat under cyclic radial advection. J Hydrol 2021 Nov 1;602:126713.

[53] Tang DW, Van der Zee SE. Macrodispersion and recovery of solutes and heat in heterogeneous aquifers. Water Resour Res 2022 Feb;58(2):e2021WR030920.

[54] Tang DW, Rijnaarts HH. Dimensionless thermal efficiency analysis for aquifer thermal energy storage. Water Resour Res 2023 Nov;59(11):e2023WR035797.

[55] Doughty C, Hellström G, Tsang CF, Claesson J. A dimensionless parameter approach to the thermal behavior of an aquifer thermal energy storage system. Water Resour Res 1982 Jun;18(3):571–87.

[56] Lin YC, Hu TF, Yeh HD. Analytical model for heat transfer accounting for both conduction and dispersion in aquifers with a Robin-type boundary condition at the injection well. Water Resour Res 2019 Aug;55(8):7379–99.

[57] Schout G, Drijver B, Gutierrez-Neri M, Schotting R. Analysis of recovery efficiency in high-temperature aquifer thermal energy storage: a Rayleigh-based method. Hydrogeol J 2014 Feb;22(1):281–91.

[58] Oerlemans PJ. Modelling heat transport in a high temperature ATES system [master's thesis]. Utrecht (NL): Utrecht University; 2018.

[59] Duijff R, Bloemendal M, Bakker M. Interaction effects between aquifer thermal energy storage systems. Groundwater 2023 Mar;61(2):173–82.

[60] Sommer W, Valstar J, Leusbrock I, Grotenhuis T, Rijnaarts H. Optimization and spatial pattern of large-scale aquifer thermal energy storage. Appl Energy 2015 Jan 1;137:322–37.

[61] van Lopik JH, Hartog N, Zaadnoordijk WJ. The use of salinity contrast for density difference compensation to improve the thermal recovery efficiency in high-temperature aquifer thermal energy storage systems. Hydrogeol J 2016 Aug;24 (5):1255–71.

[62] Beernink S, Hartog N, Vardon PJ, Bloemendal M. Heat losses in ATES systems: the impact of processes, storage geometry and temperature. Geothermics 2024 Feb 1; 117:102889.

[63] Ganguly S, Tan L, Date A, Kumar MS. Effect of heat loss in a geothermal reservoir. Energy Proc 2017 Mar 1;110:77–82.

[64] Molz FJ, Melville JG, Güven O, Parr AD. Aquifer thermal energy storage: an attempt to counter free thermal convection. Water Resour Res 1983 Aug;19(4): 922–30.

[65] Sauty JP, Gringarten AC, Fabris H, Thiéry D, Menjoz A, Landel PA. Sensible energy storage in aquifers: 2. Field experiments and comparison with theoretical results. Water Resour Res 1982 Apr;18(2):253–65.

[66] Jenne EA. Aquifer thermal energy (heat and chill) storage. Pacific Northwest Lab., Richland, WA (United States) 1992 Nov 1.

[67] Li M, Lai AC. New temperature response functions (G functions) for pile and borehole ground heat exchangers based on composite-medium line-source theory. Energy 2012 Feb 1;38(1):255–63.

[68] Jin W, Atkinson TA, Doughty C, Neupane G, Spycher N, McLing TL, Dobson PF, Smith R, Podgorney R. Machine-learning-assisted high-temperature reservoir thermal energy storage optimization. Renew Energy 2022 Sep 1;197:384–97.

[69] Parya A. Data-driven modelling of an Aquifer Thermal Energy Storage (ATES) using machine learning [master's thesis]. Turin (ITL): Politecnico di Torino. 2023.

[70] Rohmer J, Armandine Les Landes A, Loschetter A, Maragna C. Fast prediction of aquifer thermal energy storage: a multicyclic metamodelling procedure. Comput Geosci 2023 Apr;27(2):223–43.

[71] Sheldon HA, Wilkins A, Green CP. Recovery efficiency in high-temperature aquifer thermal energy storage systems. Geothermics 2021 Nov 1;96:102173.

[72] Hermans T, Nguyen F, Klepikova M, Dassargues A, Caers J. Uncertainty quantification of medium-term heat storage from short-term geophysical experiments using Bayesian evidential learning. Water Resour Res 2018 Apr;54 (4):2931–48.

[73] Wang J, Zhao Z, Liu G, Xu H. A robust optimization approach of well placement for doublet in heterogeneous geothermal reservoirs using random forest technique and genetic algorithm. Energy 2022 Sep 1;254:124427.

[74] Ullah J, Li H, Soupios P, Ehsan M. Optimizing geothermal reservoir modeling: a unified bayesian PSO and BiGRU approach for precise history matching under uncertainty. Geothermics 2024 May 1;119:102958.

[75] Duplyakin D, Beckers KF, Siler DL, Martin MJ, Johnston HE. Modeling subsurface performance of a geothermal reservoir using machine learning. Energies 2022 Jan 28;15(3):967.

[76] Wang N, Chang H, Kong XZ, Zhang D. Deep learning based closed-loop well control optimization of geothermal reservoir with uncertain permeability. Renew Energy 2023 Jul 1;211:379–94.

[77] Aydin H, Akin S, Senturk E. A proxy model for determining reservoir pressure and temperature for geothermal wells. Geothermics 2020 Nov 1;88:101916.

[78] Gudmundsdottir H, Horne RN. Prediction modeling for geothermal reservoirs using deep learning. In: 45th workshop on geothermal reservoir engineering. Stanford, California: Stanford University; 2020 Feb 10.

[79] Pandey SN, Singh M. Artificial neural network to predict the thermal drawdown of enhanced geothermal system. J Energy Resour Technol 2021 Jan 1;143(1): 010901.

[80] Beckers K, Duplyakin D, Martin M, Johnston H, Siler D. Subsurface characterization and machine learning predictions at brady hot springs. In: 46th workshop on geothermal reservoir engineering. Stanford, California: Stanford University; 2021 February 16-18.

[81] Gudala M, Govindarajan SK. Numerical investigations on a geothermal reservoir using fully coupled thermo-hydro-geomechanics with integrated RSM-machine learning and ARIMA models. Geothermics 2021 Nov;96:102174.

[82] Shi Y, Song X, Song G. Productivity prediction of a multilateral-well geothermal system based on a long short-term memory and multi-layer perceptron combinational neural network. Appl Energy 2021 Jan 15;282:116046.

[83] Ariturk M. Optimizing the production and injection wells flow rates in geothermal field using Artificial Intelligence [master's thesis]. Morgantown (USA): West Virginia University; 2018.

[84] Cai S, Wang Z, Wang S, Perdikaris P, Karniadakis GE. Physics-informed neural networks for heat transfer problems. J Heat Tran 2021 Jun 1;143(6).

[85] Markidis S. The old and the new: can physics-informed deep-learning replace traditional linear solvers? Frontiers in Big Data 2021;92.

[86] Safa AA, Fung AS, Kumar R. Heating and cooling performance characterisation of ground source heat pump system by testing and TRNSYS simulation. Renew Energy 2015 Nov 1;83:565–75.

[87] Sebarchievici C, Sarbu I. Performance of an experimental ground-coupled heat pump system for heating, cooling and domestic hot-water operation. Renew Energy 2015 Apr 1;76:148–59.

[88] Birdsell DT, Adams BM, Saar MO. Minimum transmissivity and optimal well spacing and flow rate for high-temperature aquifer thermal energy storage. Appl Energy 2021 May 1;289:116658.

[89] Drijver B, Willemsen A. Groundwater as a heat source for geothermal heat pumps. In: Proceedings of the International summer School on Direct application of geothermal energy; 2001 Sep. Bad Urach, Germany. Bochum (Germany): International Geothermal Association; 2001. p. 17–20.

[90] Oh J, Sumiyoshi D, Nishioka M, Kim H. Efficient operation method of aquifer thermal energy storage system using demand response. Energies 2021 Jan;14 (11):3129.

[91] Nielsen JE, Sørensen PA. Renewable district heating and cooling technologies with and without seasonal storage. In: Stryi-Hipp G, editor. Renewable heating and cooling. Sawston (UK): Woodhead Publishing; 2016. p. 197–220.

[92] Vanhoudt D, Desmedt J, Van Bael J, Robeyn N, Hoes H. An aquifer thermal storage system in a Belgian hospital: long-term experimental evaluation of energy and cost savings. Energy Build 2011 Dec 1;43(12):3657–65.

[93] Wagner V, Bayer P, Bisch G, Kuebert M, Blum P. Hydraulic characterization of aquifers by thermal response testing: validation by large-scale tank and field experiments. Water Resour Res 2014 Jan;50(1):71–85.

[94] Bakr M, van Oostrom N, Sommer W. Efficiency of and interference among multiple aquifer thermal energy storage systems; A Dutch case study. Renew Energy 2013 Dec 1;60:53–62.

[95] Dang T, Nguyen L, Karim A, Uddameri V. STOAViz: visualizing saturated thickness of ogallala aquifer. In: Proceedings of the workshop on Visualisation in environmental Sciences; 2017 Jun 12. Goslar, Germany. Geneva (Switzerland): The Eurographics Association; 2017. p. 37–41.

[96] Bloemendal M, Olsthoorn T, Boons F. How to achieve optimal and sustainable use of the subsurface for Aquifer Thermal Energy Storage. Energy Pol 2014 Mar 1;66: 104–14.

[97] Mattheij RM, Rienstra SW, Boonkkamp JT. Partial differential equations: modeling, analysis, computation. Philadelphia (USA): Society for Industrial and Applied Mathematics; 2005 Jan 1.

[98] Evangelos C. What is a singular perturbation [lecture notes on Internet]. Albuquerque: University of New Mexico, Faculty of Mathematics and Statistics; 2011 [cited 2022 Oct 5]. Available from: https://math.unm.edu/~vageli/cou rses/Ma570/text.pdf.

[99] Lu X. An engineering analysis method for deep geothermal energy [dissertation]. Cambridge (UK): Cambridge University; 2020.

[100] Shih SD, Tung YK. Internal layers of a transient convection-diffusion problem by perturbation methods. Adv Water Resour 1995 Jan 1;18(2):111–20.

[101] Sun X, Soga K, Cinar A, Su Z, Chen K, Kumar K, Dobson PF, Nico PS. An HPC-based Hydrothermal finite element simulator for modeling Underground response to Community-scale geothermal energy production. In: 46th workshop on geothermal reservoir engineering. Stanford, California: Stanford University; 2021 February 16-18.

[102] Ho WT, Yu FW. Chiller system optimization using k nearest neighbour regression. J Clean Prod 2021 Jun 20;303:127050.

[103] Altay EV, Gurgenc E, Altay O, Dikici A. Hybrid artificial neural network based on a metaheuristic optimization algorithm for the prediction of reservoir temperature using hydrogeochemical data of different geothermal areas in Anatolia (Turkey). Geothermics 2022 Sep 1;104:102476.

[104] Xue Z, Zhang K, Zhang C, Ma H, Chen Z. Comparative data-driven enhanced geothermal systems forecasting models: a case study of Qiabuqia field in China. Energy 2023 Oct 1;280:128255.

[105] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. Scikit-learn: machine learning in Python. J Mach Learn Res 2011 Nov 1;12:2825–30.

[106] Goldberger J, Hinton GE, Roweis S, Salakhutdinov RR. Neighbourhood components analysis. Adv Neural Inf Process Syst 2004;17.

[107] Wahid F, Kim D. A prediction approach for demand analysis of energy consumption using k-nearest neighbor in residential buildings. International Journal of Smart Home 2016 Feb;10(2):97–108.

[108] Izenman AJ. Linear discriminant analysis. Modern multivariate statistical techniques. New York (USA): Springer; 2013.

[109] Abuasbeh M, Acuña J. ATES system monitoring project, first measurement and performance evaluation: case study in Sweden. 2018 Sep 18-19. In: Proceedings of the IGSHPA Research Track 2018. Stockholm, Sweden: Stillwater (USA): International Ground Source Heat Pump Association; 2018.

[110] Bloemendal M, Beernink S, Hartog N, van Meurs B. Transforming ATES to HT-ATES, insights from Dutch pilot project. In: Proceedings of European geothermal Congress; 2019 Jun 11-14. Den Haag, Netherlands. Brussels (Belgium): The European Geothermal Energy Council; 2019. p. 248.

[111] van Lopik JH, Hartog N, Zaadnoordijk WJ, Cirkel DG, Raoof A. Salinization in a stratified aquifer induced by heat transfer from well casings. Adv Water Resour 2015 Dec 1;86:32–45.

[112] Parsons K, Reichanadter T, Vicksman A, Segur H. Explicit solution for cylindrical heat conduction. American Journal of Undergraduate Research 2016 Jun 1;13(2).

[113] Diaz-Maurin F, Saaltink MW. Model validation for subsurface dynamics. Reykjavík (IL): GEOTHERMICA; 2021.