Thermal interactions in large irregular fields of geothermal boreholes: the method of equivalent boreholes

Carlos Prieto [a] and Massimo Cimmino[a]

[a]Department of Mechanical Engineering, Polytechnique Montréal, Case Postale 6079, Succursale "Centre-ville," Montréal, Québec, H3C 3A7, Canada

ARTICLE HISTORY
Compiled April 5, 2022

ABSTRACT
A new method is presented to evaluate thermal interactions between vertical geothermal boreholes. The finite line source (FLS) solution is extended to consider thermal interactions between groups of boreholes. Groups of boreholes that share similar temperatures and heat extraction rates are identified using hierarchical agglomerative clustering, and each group is represented in the model as a single equivalent borehole. Each equivalent borehole is split into segments, and temporal and spatial superposition of the FLS solution are employed to calculate the total temperature change along the length of the equivalent boreholes. The new method is shown to provide an accurate calculation of the g-function, with a mean absolute percentage error below 0.612 % on the g-functions of regular borefields of up to 144 boreholes using only 3 to 5 equivalent boreholes. Calculation times are significantly reduced : the g-function of a borefield of 1024 randomly positioned boreholes is calculated in 3.65 seconds.

KEYWORDS
ground-coupled heat pumps; geothermal boreholes; g-functions; finite line source; thermal interactions; hierarchical agglomerative clustering

1. Introduction

Ground-coupled heat pump (GCHP) systems use the ground as a heat source (or sink) to supply heating (or cooling) to buildings. GCHP systems are composed of one or multiple heat pumps coupled to a set of geothermal boreholes (i.e. a borefield) that allow the exchange of thermal energy with the ground. A geothermal borehole consists in a drilled hole, typically of a diameter ranging from 100 mm to 150 mm and of a length ranging from 15 m to 180 m (ASHRAE, 2019). One or several U-tubes (or alternatively coaxial pipes) are inserted into the borehole and the borehole is subsequently back-filled with grouting material. In some cases, no grouting material is used and the borehole is filled with groundwater. The heat carrier fluid from the heat pump(s) is circulated through the boreholes to exchange heat with the surrounding ground and then fed back to the heat pump(s). Heat extraction and rejection in the borefield cause fluid and ground temperature variations which may compound over the GCHP system's life cycle and impact on its performance and operability.

CONTACT Carlos Prieto. Email: carlos.prieto@polymtl.ca
CONTACT Massimo Cimmino. Email: massimo.cimmino@polymtl.ca

Mathematical models of heat transfer are used in all phases of GCHP system planning, design and operation. These mathematical models allow to predict temperature changes in the fluid and the ground due to fluctuations in the heat extraction and rejection into the borefield. In site characterization applications, physical parameters (e.g. soil thermal conductivity, borehole thermal resistance) can be inferred by fitting mathematical models to measurements from a thermal response test (Spitler & Gehlin, 2015). In the design phase of the system, mathematical models can be used to estimate the required borehole size to satisfy the operation parameters of the equipment, for example lower and upper fluid temperature limits for the operation of the heat pump (Ahmadfard & Bernier, 2019). Heat pump and borefield heat transfer mathematical models can be coupled to predict the energy consumption of the system. Also, these models could allow successfully posing an optimization strategy for the system design and the development of model-based control strategies (Cupeiro Figueroa et al., 2020). Accurate predictions of fluid and ground temperatures require the modeling of both of the long-term and short-term heat transfer effects in geothermal boreholes (Li & Lai, 2015). Long-term heat transfer effects are characterized by the three-dimensional heat transfer in the soil and the thermal interactions between boreholes. Short-term heat transfer effects are characterized by the transit of the fluid through the boreholes and the thermal capacitance of the borehole materials.

Long-term temperature predictions in geothermal borefields can be obtained by temporal superposition of thermal response factors, or g-functions. g-Functions are unit step-response functions of the effective borehole wall temperature change due to a unit heat extraction rate from the borefield. They can be superimposed in time to achieve simulations of geothermal systems (Mitchell & Spitler, 2019). Eskilson (1987) calculated the g-functions of geothermal borefields from the spatial superposition of temperature fields around individual boreholes. The individual temperature fields were calculated using a finite difference method. The inside and outside of the boreholes, delimited by the borehole wall, were uncoupled by considering a uniform borehole wall temperature along the length of the boreholes and equal for all boreholes. In recent work, g-functions are calculated by superimposing suitable analytical heat source solutions to model various physical phenomena, for example the finite line source (FLS) solution to model purely conductive heat transfer in isotropic and homogeneous ground (Eskilson, 1987), the moving finite line source to model the effects of groundwater advection (Molina-Giraldo et al., 2011), and the multilayer finite line source to model layered ground physical properties (Abdelaziz et al., 2014). For pure conduction, the FLS solution was spatially superposed by Zeng et al. (2002) to calculate thermal response factors of geothermal borefields. Lamarche & Beauchamp (2007a) and later Claesson & Javed (2011) obtained simplified expressions for the FLS solution involving a single integral, down from a computationally expensive double integral.

The simple spatial superposition of the FLS solution fails to accurately evaluate g-functions, due to the different boundary conditions at the borehole wall : the FLS solution considers a uniform heat extraction rate and Eskilson's g-functions consider a uniform borehole wall temperature. Cimmino et al. (2013) and Cimmino & Bernier (2014) were able to evaluate the g-functions for a uniform temperature boundary condition by first considering the time variation of the heat extraction rates of individual boreholes, and then by considering the distribution of heat extraction rates along boreholes using a segmented FLS solution. This method was also applied to inclined boreholes by Lazzarotto (2016). Cimmino (2015) coupled the segmented FLS solution to a thermal circuit representation of the inside of geothermal boreholes and showed that the distribution of temperatures and heat extraction rates along boreholes are needed

for accurate predictions of temperature variations. The spatial and temporal superposition of the FLS solution can also be used to assemble so-called network-based models of geothermal systems that account for the hydraulic configuration of the geothermal borefield, where the full distribution of borehole wall temperatures are evaluated every time step of simulations (Cimmino, 2018a; Lamarche, 2017a; Lazzarotto, 2014).

A drawback of analytical models for thermal interactions between boreholes, such as the FLS solution, is the computational effort required to evaluate the solution (due to the integral) and to solve the system of equations. The number of evaluations of the FLS solution scales with the square of the number of line source segments (i.e. the number of segment pairs). Cimmino (2018b) considered similarities between pairs of line sources to reduce the number of evaluations of the FLS solution. Dusseault et al. (2018) approximated the integrand of the FLS solution using Chebyshev polynomials, thereby reducing the time taken for numerical integration. Nguyen & Pasquier (2021) make use of interpolations of a priori calculated FLS values at different distances, thereby reducing the time in the calculation of the borefield thermal response. These last two techniques, however, do not consider the distribution of heat extraction rates over the length of the boreholes. In all cases, the size of the system of equations is unaffected. The size of the system of equations is particularly damaging in network-based methods where it needs to be solved at each time step of the simulation. Lamarche (2017b) proposed a new FLS solution for piecewise linear distributions of heat extraction rates along the boreholes that achieves the same accuracy as the piecewise constant solution using fewer borehole segments and thereby reducing the size of the system of equations.

Short-term effects, caused by the transit of the fluid through the boreholes and the thermal capacitance of the borehole materials, dampen the variations of fluid and ground temperatures during the first few hours of heat extraction and injection. The lower bound of validity of long-term models that neglect short-term effects is estimated to be $t_b = 5r_b^2/\alpha_s$, with r_b the borehole radius and α_s the ground thermal diffusivity (Eskilson, 1987). Neglecting short-term effects leads to an overestimation of the fluid and ground temperature changes, and possibly to an underestimation of the heat pump efficiency in simulations and an overestimation of the required borehole length in design (Ahmadfard & Bernier, 2018; Bernier & Shirazi, 2013). There are several modeling approaches that include short-term effects, some of them are: equivalent pipe models (Brussieux & Bernier, 2019; Gu & O'Neal, 1995; Javed & Claesson, 2011; Lamarche & Beauchamp, 2007b; Xu & Spitler, 2006), thermal resistance and capacitance methods (TRCMs) (Bauer et al., 2011; Minaei & Maerefat, 2017; Pasquier & Marcotte, 2014; Zarella et al., 2011) and other models based on numerical methods (BniLam & Al-Khouri., 2016; Yang & Li, 2014; Yavuzturk, 1999). Recently, transient multipole solutions have been developed in the time domain (Prieto & Cimmino, 2021) and spectral domain (Rivero & Hermanns, 2021) for 2D transient heat conduction. These solutions, that in their mathematical formalism approach exact solutions, have yet to be extended to include end-effects (i.e. axial heat transfer) in the boreholes and include fluid capacity in inside the pipes. There are works in which some of the mentioned methods are used to simulate both short- and long-term effects, such as Claesson & Javed (2011); Laferrière et al. (2020); Li et al. (2014). In these applications, the short- and long-term models are loosely coupled : only one borehole is used in the short-term model to represent the entire borefield, and the distribution of temperatures and heat extraction rates along the length of the borehole from the short-term model is not taken into account for long-term predictions. This approach is justifiable since modeling hundreds boreholes individually, both in the short- and

long-term, would make the computational requirements for simulation impractical. As a result, the impact of short-term effects on the long-term temperature changes is unknown. New computationally-efficient and scalable (with regards to the number of boreholes) approaches to model long-term heat transfer in geothermal borefields are thereby needed to achieve detailed axially-discretized models of geothermal borefields that can be used in multi-annual simulations of GCHP systems.

This paper introduces a new *equivalent borehole* concept to model thermal interactions in geothermal borefields and to reduce the size of the system of equations in thermal models, and enable the efficient modelling of thermal interactions in large geothermal borefields comprised of hundreds of boreholes. A hierarchical agglomerative clustering is applied to identify groups of boreholes that share similar temperatures and heat extraction rates, and the FLS solution is adapted to evaluate thermal interactions between groups of boreholes each represented by a single equivalent borehole. The new approach is validated by calculating the g-function of borefields against a reference method from the literature (Cimmino, 2018b,c).

2. Methodology

2.1. Thermal interaction between boreholes

Figure 1 shows a field of $N_b = 2$ vertical boreholes of equal dimensions. All boreholes have the same length L and radius r_b, and are buried at the same distance D from the ground surface. Each borehole i is located at coordinates (x_i, y_i). The ground has a uniform and isotropic thermal conductivity k_s and thermal diffusivity α_s, and is initially at a uniform temperature $T(x, y, z, t = 0) = T_0$ (i.e. the undisturbed ground temperature).

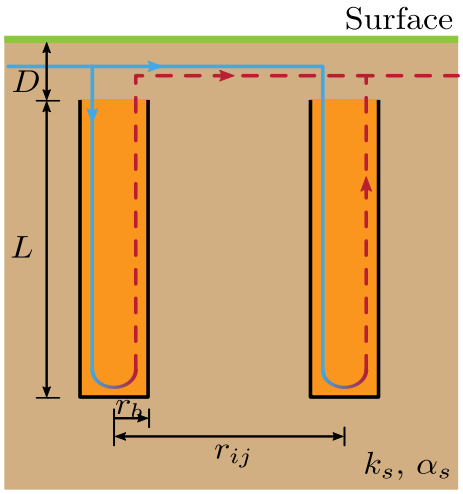


Figure 1.: Field of 2 vertical boreholes of equal dimensions

Following the methodology of Cimmino & Bernier (2014), each borehole is divided into n_q segments, and each borehole segment is modeled as a line segment located along its axis with a uniform heat extraction rate. The temperature drop at the wall of a segment u of a borehole i at time t_k can then be evaluated from the spatial and temporal superposition of the analytical FLS solution :

$$\Delta T_{b,i,u,k} = T_0 - T_{b,i,u,k} = \frac{1}{2\pi k_s} \sum_{j=1}^{N_b} \sum_{v=1}^{n_q} \sum_{p=1}^{k} Q'_{j,v,p} \left(h_{ij,uv}(t_k - t_{p-1}) - h_{ij,uv}(t_k - t_p) \right) \tag{1}$$

where $\Delta T_{b,i,u,k}$ is the temperature drop at the wall of segment u of borehole i at time t_k, T_0 is the initial ground temperature, $T_{b,i,u,k}$ is the temperature at the wall of segment u of borehole i at time t_k, $Q'_{j,v,p}$ is the heat extraction rate per unit length of segment v of borehole j from time t_{p-1} to t_p, and $h_{ij,uv}$ is the segment-to-segment thermal response factor for the borehole wall temperature change over segment u of borehole i caused by heat extraction from segment v of borehole j. The segment-to-segment response factor is given by the FLS solution:

$$h_{ij,uv}(t) = \frac{1}{2L_u} \int_{\frac{1}{\sqrt{4\alpha_s t}}}^{\infty} \frac{1}{s^2} \exp\left(-r_{ij}^2 s^2\right) I_{FLS}(s)\, ds \tag{2a}$$

$$\begin{aligned} I_{FLS}(s) = {} & \operatorname{erfint}\left((D_u - D_v + L_u)s\right) - \operatorname{erfint}\left((D_u - D_v)s\right) \\ & + \operatorname{erfint}\left((D_u - D_v - L_v)s\right) - \operatorname{erfint}\left((D_u - D_v + L_u - L_v)s\right) \\ & + \operatorname{erfint}\left((D_u + D_v + L_u)s\right) - \operatorname{erfint}\left((D_u + D_v)s\right) \\ & + \operatorname{erfint}\left((D_u + D_v + L_v)s\right) - \operatorname{erfint}\left((D_u + D_v + L_u + L_v)s\right) \end{aligned} \tag{2b}$$

$$\operatorname{erfint}(x) = \int_0^x \operatorname{erf}(x')\, dx' = x \operatorname{erf}(x) - \frac{1}{\sqrt{\pi}} \left(1 - \exp(-x^2)\right) \tag{2c}$$

where r_{ij} is the radial distance between boreholes i and j (with $r_{ii} = r_b$), $\operatorname{erf}(x)$ is the error function and $\operatorname{erfint}(x)$ is the integral of the error function.

In non-dimensional form:

$$\theta_{b,i,u,k} = \sum_{j=1}^{N_b} \sum_{v=1}^{n_q} \sum_{p=1}^{k} \phi'_{j,v,p} \left(h_{ij,uv}(\tau_k - \tau_{p-1}) - h_{ij,uv}(\tau_k - \tau_p) \right) \tag{3}$$

where $\theta_{b,i,u,k} = \frac{T_0 - T_{b,i,u,k}}{Q'^* / 2\pi k_s}$ is the dimensionless temperature at the wall of segment u of borehole i at a dimensionless time $\tau_k = \frac{9\alpha_s t_k}{L^2} = \frac{t_k}{t_s}$ with t_s the borefield characteristic time, $\phi'_{j,v,p} = \frac{Q'_{j,v,p}}{Q'^*}$ is the normalized heat extraction rate per unit length of segment v of borehole j from time τ_{p-1} to τ_p, and Q'^* is an arbitrary heat extraction rate per unit length.

Equation (3) can be simplified by introducing matrix notation:

$$\mathbf{\Theta}_{b,k} = \sum_{p=1}^{k} \left(\mathbf{H}(\tau_k - \tau_{p-1}) - \mathbf{H}(\tau_k - \tau_p)\right) \mathbf{\Phi}'_p \tag{4}$$

$$= \sum_{p=1}^{k-1} \left(\mathbf{H}(\tau_k - \tau_{p-1}) - \mathbf{H}(\tau_k - \tau_p)\right) \mathbf{\Phi}'_p + \mathbf{H}(\tau_k - \tau_{k-1})\mathbf{\Phi}'_k \tag{5}$$

$$= \mathbf{\Theta}^0_{b,k} + \mathbf{H}(\tau_k - \tau_{k-1})\mathbf{\Phi}'_k \tag{6}$$

where $\mathbf{\Theta}_{b,k} = \left[\theta_{b,1,1,k}, \theta_{b,1,2,k}, \cdots, \theta_{b,N_b,n_q,k}\right]^T$ is a vector of dimensionless borehole wall temperatures along all segments of all boreholes at time τ_k and $\mathbf{\Phi}'_k = \left[\phi'_{1,1,k}, \phi'_{1,2,k}, \cdots, \phi'_{N_b,n_q,k}\right]^T$ is a vector of dimensionless heat extraction rates per unit length along all segments of all boreholes at time t_k. $\mathbf{H}$ is a $N_b n_q \times N_b n_q$ matrix of segment-to-segment response factors:

$$\mathbf{H}(\tau_k) = \begin{bmatrix} h_{11,11}(\tau_k) & h_{11,12}(\tau_k) & \cdots & h_{1N_b,1n_q}(\tau_k) \\ h_{11,21}(\tau_k) & h_{11,22}(\tau_k) & \cdots & h_{1N_b,2n_q}(\tau_k) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_b1,n_q1}(\tau_k) & h_{N_b1,n_q2}(\tau_k) & \cdots & h_{N_bN_b,n_qn_q}(\tau_k) \end{bmatrix} \tag{7}$$

2.2. Thermal interaction between groups of boreholes

The size of the system of equations presented in Equation (4) increases with the square of the total number of segments in the borefield ($N_b^2 n_q^2$). The calculation time for the simulation of borefields using network-based methods (e.g. Cimmino (2018a); Lamarche (2017a); Lazzarotto (2014)) or for the evaluation of g-functions (e.g. Cimmino (2018b); Dusseault et al. (2018)) thereby increases substantially when the number of boreholes increases. This makes the simulation of geothermal borefields increasingly impractical as the number of boreholes increase. One method for decreasing the size of the system of equations is to consider symmetries in the borefield layout, as done by Cimmino et al. (2013). This method is however only applicable to regular borefield geometries where symmetrical groups of boreholes, which share the same borehole wall temperatures and heat extraction rates, can be easily identified. This section extends the concept of borehole groups to consider groups of non-symmetrically positioned boreholes. Boreholes in the same group are assumed to have similar borehole wall temperature profiles and heat extraction rate profiles, such that all boreholes in a group can be replaced with a single *equivalent* borehole representative of that group.

There exists a fundamental group composed with all the boreholes in the borefield defined as $\mathbb{G} = \{1, 2, 3, \ldots, i, \ldots, j, \ldots, N_b\}$, which is a non-empty finite set where every element in the group corresponds to a borehole index. Since $\mathbb{G}$ is non-empty and finite, it is possible to decompose this group as a set of $\mathcal{G}(\leq N_b)$ non-empty groups such that $\mathbb{G} = \bigcup_{\mathcal{I}=1}^{\mathcal{G}} \mathbb{G}_\mathcal{I}$ (i.e. all boreholes are in a group), $\mathbb{G}_\mathcal{I} \cap \mathbb{G}_\mathcal{J} = \varnothing$ for $\mathcal{I} \neq \mathcal{J}$ (i.e. a borehole is an element of only one group), and consequently $\bigcap_{\mathcal{I}=1}^{\mathcal{G}} \mathbb{G}_\mathcal{I} = \varnothing$. The cardinality of the borefield group is $\#\mathbb{G} = \sum_{\mathcal{I}=1}^{\mathcal{G}} \#\mathbb{G}_\mathcal{I} = \sum_{\mathcal{I}=1}^{\mathcal{G}} N_{b,\mathcal{I}} = N_b$, where $N_{b,\mathcal{I}}(\geq 1)$ corresponds to the number of boreholes belonging to group $\mathbb{G}_\mathcal{I}$. The

formation of borehole groups is described in section 2.3.

The dimensionless borehole wall temperature at the wall of equivalent borehole segments is calculated from the spatial and temporal superposition of the FLS solution:

$$\bar{\theta}_{b,\mathcal{I},u,k} = \sum_{n=1}^{\mathcal{G}} \sum_{v=1}^{n_q} \sum_{p=1}^{k} \bar{\phi}'_{\mathcal{J},v,p} \left(\bar{h}_{\mathcal{I}\mathcal{J},uv}(\tau_k - \tau_{p-1}) - \bar{h}_{\mathcal{I}\mathcal{J},uv}(\tau_k - \tau_p) \right) \tag{8}$$

where $\bar{\theta}_{b,\mathcal{I},u,k}$ is the dimensionless temperature at the wall of segment u of equivalent borehole $\mathcal{I}$ at time τ_k, $\bar{\phi}'_{\mathcal{J},v,p}$ is the normalized heat extraction rate per unit length of segment v of equivalent borehole $\mathcal{J}$ from time τ_{p-1} to τ_p, and $\bar{h}_{\mathcal{I}\mathcal{J},uv}$ is the equivalent segment-to-segment response factor for the borehole wall temperature change over segment u of equivalent borehole $\mathcal{I}$ caused by heat extraction from segment v of equivalent borehole $\mathcal{J}$.

The equivalent segment-to-segment response factor is given by the average segment-to-segment response factors for the borehole wall temperature change along segments u of all boreholes in group $\mathbb{G}_\mathcal{I}$ due to the heat extraction at all segments v of all boreholes in group $\mathbb{G}_\mathcal{J}$:

$$\bar{h}_{\mathcal{I}\mathcal{J},uv}(t) = \frac{1}{2L_u} \frac{1}{N_{b,\mathcal{I}}} \int_{\frac{1}{\sqrt{4\alpha_s t}}}^{\infty} \sum_{i \in \mathbb{G}_\mathcal{I}} \sum_{j \in \mathbb{G}_\mathcal{J}} \frac{1}{s^2} \exp\left(-r_{ij}^2 s^2\right) I_{FLS}(s)\, ds \tag{9}$$

$$\bar{h}_{\mathcal{J}\mathcal{I},vu}(t) = \frac{N_{b,\mathcal{I}}}{N_{b,\mathcal{J}}} \frac{L_u}{L_v} \bar{h}_{\mathcal{I}\mathcal{J},uv}(t) \tag{10}$$

In matrix notation:

$$\bar{\mathbf{\Theta}}_{b,k} = \sum_{p=1}^{k} \left(\bar{\mathbf{H}}(\tau_k - \tau_{p-1}) - \bar{\mathbf{H}}(\tau_k - \tau_p) \right) \bar{\mathbf{\Phi}}'_p \tag{11}$$

$$= \sum_{p=1}^{k-1} \left(\bar{\mathbf{H}}(\tau_k - \tau_{p-1}) - \bar{\mathbf{H}}(\tau_k - \tau_p) \right) \bar{\mathbf{\Phi}}'_p + \bar{\mathbf{H}}(\tau_k - \tau_{k-1}) \bar{\mathbf{\Phi}}'_k \tag{12}$$

$$= \bar{\mathbf{\Theta}}^0_{b,k} + \bar{\mathbf{H}}(\tau_k - \tau_{k-1}) \bar{\mathbf{\Phi}}'_k \tag{13}$$

where $\bar{\mathbf{\Theta}}_{b,k} = \left[\bar{\theta}_{b,1,1,k}, \bar{\theta}_{b,1,2,k}, \cdots, \bar{\theta}_{b,\mathcal{G},n_q,k}\right]^T$ is a vector of dimensionless borehole wall temperatures along all segments of all equivalent boreholes at time τ_k and $\bar{\mathbf{\Phi}}'_k = \left[\bar{\phi}'_{1,1,k}, \bar{\phi}'_{1,2,k}, \cdots, \bar{\phi}'_{\mathcal{G},n_q,k}\right]^T$ is a vector of dimensionless heat extraction rates per unit length along all segments of all boreholes at time τ_k. $\bar{\mathbf{H}}$ is a $\mathcal{G}n_q \times \mathcal{G}n_q$ matrix of equivalent segment-to-segment response factors:

$$\bar{\mathbf{H}}(\tau_k) = \begin{bmatrix} \bar{h}_{11,11}(\tau_k) & \bar{h}_{11,12}(\tau_k) & \cdots & \bar{h}_{1\mathcal{G},1n_q}(\tau_k) \\ \bar{h}_{11,21}(\tau_k) & \bar{h}_{11,22}(\tau_k) & \cdots & \bar{h}_{1\mathcal{G},2n_q}(\tau_k) \\ \vdots & \vdots & \ddots & \vdots \\ \bar{h}_{\mathcal{G}1,n_q1}(\tau_k) & \bar{h}_{\mathcal{G}1,n_q2}(\tau_k) & \cdots & \bar{h}_{\mathcal{G}\mathcal{G},n_qn_q}(\tau_k) \end{bmatrix} \tag{14}$$

Comparing Equations (3) and (8), it can be seen that the size of the system of equations using equivalent boreholes is reduced by a factor $N_b^2/\mathcal{G}^2$.

2.3. Equivalent boreholes

This section introduces a systematic way to compute the finite set of groups for the borefield group $\mathbb{G} = \bigcup_{\mathcal{I}=1}^{\mathcal{G}} \mathbb{G}_{\mathcal{I}}$. A hierarchical clustering approach is introduced to define the groups. Then, an optimal number of groups is defined based on the maximum dissimilarity between the identified groups.

2.3.1. Hierarchical agglomerative clustering

A hierarchical agglomerative clustering method is applied to form groups of boreholes that share similar borehole wall temperatures. Agglomerative clustering methods present advantages compared to other clustering methods (e.g. k-means, k-medoids), as they do not require the number of groups to be known prior to clustering. In hierarchical agglomerative clustering, each element (i.e. borehole) is initialized as their own cluster (i.e. group). At each iteration of the algorithm, the two closest clusters are merged according to a *linkage* criterion evaluated from a *distance metric* based on the distance between elements of the two clusters. The product of the hierarchical agglomerative clustering can be visualized on a dendrogram tree from which the clusters can be identified for a known number of clusters or based on a criterion to identify the optimal number of groups (as will be presented in section 2.3.2).

Recalling from section 2.2 that boreholes in the same group should have similar borehole wall temperatures and heat transfer rate, the absolute temperature difference between boreholes is used as the distance metric. A complete linkage is proposed, meaning the distance between two groups of boreholes is given by the maximum absolute temperature difference between boreholes of the two groups:

$$\mathcal{L}(\mathcal{I}, \mathcal{J}) = \max_{i \in \mathbb{G}_{\mathcal{I}}, j \in \mathbb{G}_{\mathcal{J}}} \mathcal{M}(i, j) \tag{15}$$

$$\mathcal{M}(i, j) = |\theta_i - \theta_j| \tag{16}$$

where $\mathcal{L}(\mathcal{I}, \mathcal{J})$ is the complete linkage criterion between groups $\mathcal{I}$ and $\mathcal{J}$ (i.e. the distance between the groups), $\mathcal{M}(i, j)$ is the distance metric between two boreholes, and θ_i is the dimensionless temperature at the wall of borehole i.

Dimensionless borehole wall temperatures are calculated considering a unit normalized heat extraction rate equal for all boreholes (i.e. $\phi' = 1$). For the purpose of the hierarchical agglomerative clustering, the single segment steady-state finite line source (FLS) is used to evaluate the dimensionless borehole wall temperatures, rather than the FLS solution used in sections 2.1 and 2.2. It was found that the single segment steady-state FLS solution provides sufficient accuracy for clustering with the advantage of being much less computationally intensive to evaluate than Equation (2) since there exists a formulation free of integrals. The average dimensionless temperature change at the location of a borehole i due to a unit normalized heat extraction rate per unit length at a borehole j is given by:

$$\begin{aligned} w_l = \{&D_i - D_j + L_i, D_i - D_j, D_i - D_j - L_j, D_i - D_j + L_i - L_j, \\ &D_i + D_j + L_i, D_i + D_j, D_i + D_j + L_j, D_i + D_j + L_i + L_j\} \end{aligned} \tag{17a}$$

$$\theta_{ij} = \frac{1}{2L_i} \sum_{l=1}^{8} (-1)^{l+1} \left\{ w_l \ln \left[\left(w_l + \sqrt{w_l^2 + r_{ij}^2} \right) / r_{ij} \right] - \sqrt{w_l^2 + r_{ij}^2} \right\} \tag{17b}$$

where θ_{ij} is the steady-state dimensionless temperature change at the location of a borehole i, located at a distance r_{ij} from a borehole j (with $r_{ii} = r_b$), caused by a unit normalized heat extraction rate per unit length.

The dimensionless borehole wall temperature change is given by the spatial superposition of the temperature changes caused by all boreholes:

$$\theta_i = \sum_{j=1}^{N_b} \theta_{ij} \tag{18}$$

2.3.2. Optimal number of groups

The result of the hierarchical agglomerative clustering can be visualized on a dendrogram. Figure 2 shows an example dendrogram generated from the clustering of a field of 16 boreholes. The dendrogram shows the complete clustering process, starting with all boreholes in separate groups until all groups are merged into a single group. The leaf nodes at the bottom of the dendrogram represent the initial groups. At each iteration, the two closest groups (according to Equations (15) and (16)) are merged. The created group is represented as an internal node, with branches connecting the node to the two merged groups, and its height represents the distance between the two groups. The root node, at the top of the dendrogram, contains all boreholes.

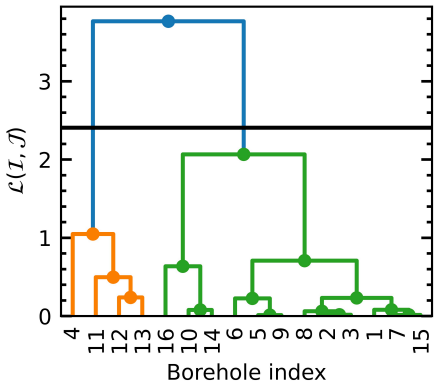


Figure 2.: Dendrogram Example

A minimum number of groups, expressed as $\mathcal{G}_{\min}$, can be identified from a cut-off threshold on the dendrogram. A simple technique to identify the minimum number of groups is to cut the dendrogram at the half-height with an horizontal line at the longest distance between two consecutive nodes on the tree, as shown on Figure 2. The number of branches crossed by the cut-off line corresponds to the minimum number of groups. In the example, $\mathcal{G}_{\min} = 2$. The optimal number of groups is greater than the minimum. However, it is not possible to identify the optimal number of groups on rigorous formalism (not based on nonlinear programming) but particular techniques have

been studied (Jung et al., 2003; Patil & Baidari, 2019; Zhou et al., 2017). Therefore, the number of groups can be chosen by increasing the minimum number of groups:

$$\mathcal{G} = \mathcal{G}_{\min} + \mathcal{K} \tag{19}$$

where $\mathcal{K}$ is a precision increment from which the number of groups can be increased from the minimum. It will be shown later that the proposed method achieves acceptable accuracy for $\mathcal{K} = 1$.

2.4. g-function calculation

The proposed method is validated by calculating the g-function. The g-function represents the effective borehole wall temperature variation in a field of thermally interacting boreholes for a constant total heat extraction rate. Temporal superposition can later be applied to the g-function to simulate borefields with varying heat extraction rates. As shown by Cimmino & Bernier (2014), the accurate evaluation of the g-function requires detailed modeling of the borefield where the variation of heat extraction rates along the boreholes is considered.

In the proposed method, each group of boreholes is modeled using a single equivalent borehole and the variation of the heat extraction rate along its length is applied to all boreholes of the same group, thereby limiting the degrees of freedom of the borefield. If the equivalent boreholes are indeed representative of their respective groups and the number of groups is sufficient, the proposed method should produce the same g-function as if all boreholes are modeled individually. The proposed method will thus be validated by evaluating the g-function and comparing the variation of heat extraction rates in the borefield.

For a field of boreholes connected in parallel, the g-function is typically evaluated by imposing a uniform borehole wall temperature equal for all boreholes (Eskilson, 1987), based on the assumptions that the fluid temperature variation inside boreholes is small and that fluid and borehole wall temperatures are sufficiently close. At each time step, a system of equations is built and solved to identify the required borehole wall temperature to achieve the desired constant total heat extraction rate:

$$\begin{bmatrix} \bar{\mathbf{H}}(\tau_k - \tau_{k-1}) & \mathbf{1} \\ \bar{\mathbf{L}} & 0 \end{bmatrix} \begin{bmatrix} \bar{\mathbf{\Phi}}'_k \\ \bar{\theta}_{b,k} \end{bmatrix} = \begin{bmatrix} -\bar{\mathbf{\Theta}}^0_{b,k} \\ 1 \end{bmatrix} \tag{20a}$$

$$\bar{\mathbf{L}} = \begin{bmatrix} \frac{N_{b,1}}{N_b n_q} & \frac{N_{b,1}}{N_b n_q} & \cdots & \frac{N_{b,2}}{N_b n_q} & \frac{N_{b,2}}{N_b n_q} & \cdots & \frac{N_{b,\mathcal{G}}}{N_b n_q} & \frac{N_{b,\mathcal{G}}}{N_b n_q} \end{bmatrix} \tag{20b}$$

where $\bar{\theta}_{b,k}$ is the uniform dimensionless borehole wall temperature at time τ_k, corresponding to the g-function, and $\bar{\mathbf{L}}$ is a vector of length ratios that imposes a constant total unit normalized heat extraction rate per unit length. A similar system of equations can be built from the relations in section 2.1 to evaluate the g-function considering each borehole individually. The method is not limited to the calculation of g-functions using a uniform temperature boundary condition. The same hierarchical agglomerative clustering approach can be considered to evaluate g-functions using an equal inlet fluid temperature boundary condition by extending Equation (20) to include the interaction between the fluid and the borehole walls (Cimmino, 2015).

Boreholes of unequal lengths can also be considered by applying the clustering approach to each subset of boreholes of equal lengths in the borefield.

3. Results

The proposed method is implemented in Python, using the `SciPy` library (Virtanen et al., 2020) for the implementation of the hierarchical agglomerative clustering, the evaluation of integrals and the evaluation of special functions (e.g. the error function), and the `NumPy` library (Van Der Walt et al., 2011) to handle matrix and vector manipulations. The reference g-functions, evaluated with all boreholes modeled individually as presented in section 2.1, are calculated using the `pygfunction` (version 2.0.0) library (Cimmino, 2018c) which implements the similarities method of Cimmino (2018b) to accelerate calculations. The `pygfunction` library was extended to accommodate the proposed method.

The proposed method is first applied to the calculation of the g-function of a field of 20 boreholes in a rectangular configuration and compared to the reference method with regards to the g-function values and the predicted heat extraction rates. Then, the accuracy and computational time of the proposed method is assessed on various regular borefield configurations with up to 144 boreholes. Finally, the robustness and scalability of the method, in terms of accuracy and calculation time, is tested on borefields of randomly positioned boreholes with up to 1024 boreholes. Simulation parameters are presented in Table 1, where the borehole spacing B only applies to regular borefield configurations.

Table 1.: Borehole and ground parameters

Parameter	**Value**
Borehole buried depth	$D = 4\ m$
Borehole length	$L = 150\ m$
Borehole spacing	$B = 7.5\ m$
Borehole radius	$r_b = 0.075\ m$
Ground thermal diffusivity	$\alpha_s = 1.0 \times 10^{-6}\ m^2/s$

In all cases, the number of equivalent boreholes is chosen according to the criterion of section 2.3.2 with $\mathcal{K} = 1$, the number of segments per borehole is $n_q = 12$ as per the recommendation of Cimmino & Bernier (2014), and the g-functions are calculated at $N_t = 25$ time values in geometrically expanding time-steps in the range $-10 \leq \ln(t/t_s) \leq 5$. All calculations were done on a personal computer with 16 GB of RAM and a 6-core processor (12 threads) running at normal speed of 2.60 GHz and maximum speed of 4.80 GHz.

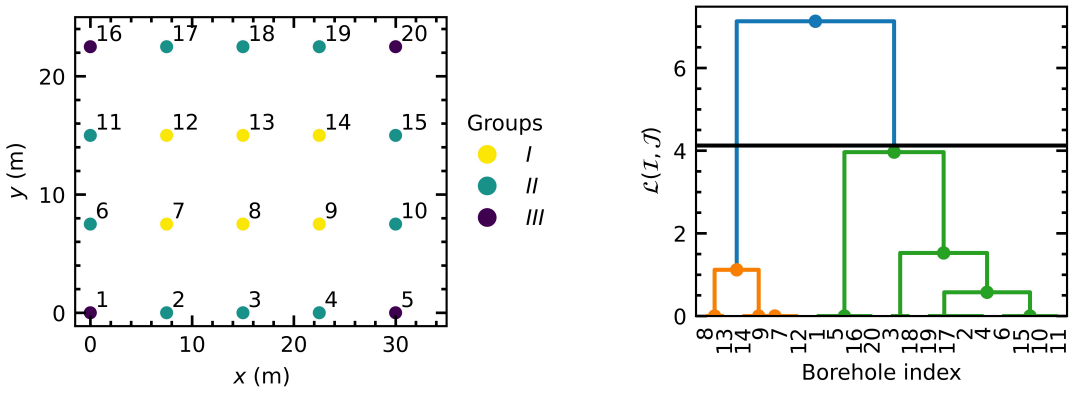
The accuracy of the proposed method is quantified using the mean absolute percentage error:

$$MAPE = \frac{100}{N_t} \sum_{k=1}^{N_t} \left| \frac{g_{\text{reference}}(\tau_k) - g(\tau_k)}{g_{\text{reference}}(\tau_k)} \right| \tag{21}$$

where $MAPE$ is the mean absolute percentage error and $g_{\text{reference}}$ is the g-function evaluated using the reference method.

3.1. Rectangular borefield of 20 boreholes

Figure 3 shows the results of the hierarchical agglomerative clustering method applied to a borefield of 20 boreholes in a $N_x \times N_y = 5 \times 4$ rectangular configuration. Figure 3a shows the borefield layout, with markers representing the positions of the boreholes and different marker colors representing the identified borehole groups. Figure 3b shows the dendrogram resulting from the clustering method. It can be seen that the minimum number of groups is $\mathcal{G}_{\text{min}} = 2$ since the cut-off line, at half-height of maximum $\mathcal{L}(\mathcal{I}, \mathcal{J})$ between two consecutive nodes, crosses two branches of the dendrogram, and therefore the number of groups is $\mathcal{G} = 3$ for a precision increment $\mathcal{K} = 1$. For simplicity, groups will be referred to using roman numerals (e.g. group III refers to $\mathbb{G}_3$). In this example, group I contains 6 boreholes, group II contains 10 boreholes, and group III contains the remaining 4 boreholes. It should be noted that, due to symmetries in the borefield layout, the borefield could be modeled with no error (compared to the reference) using only 6 groups. Figure 3a shows that the clustering approach successfully places symmetrically positioned boreholes in the same group.



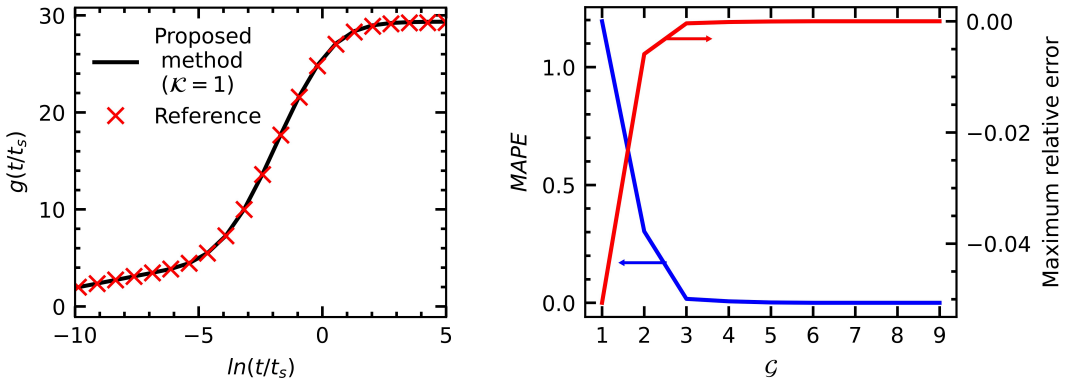
(a) Hierarchical clustering

(b) Dendrogram

Figure 3.: Hierarchical agglomerative clustering method for a 5×4 borefield

Figure 4a compares the g-function of the 5×4 borefield evaluated using the proposed method to the reference method of Cimmino (2018b) for a $\mathcal{K} = 1$. Figure 4b shows the $MAPE$ and the maximum relative error as a function of the number of groups $\mathcal{G}$ used for the calculation of the g-function (with $\mathcal{G}_{\text{min}} + 1 = 3$). It is shown that the errors are maximum when only one group is considered, with a $MAPE$ (left y-axis) of 1.2% and a maximum relative error (right y-axis) of -0.05. A negative sign on the maximum relative error means that the g-function calculated by the proposed method overestimates the value given by the reference method. As expected, the errors reach 0 when $\mathcal{G} = 6$. According to these results, a precision increment $\mathcal{K} = 1$ is sufficient to estimate the g-function. For a precision increment $\mathcal{K} = 1$ ($\mathcal{G} = 3$), the $MAPE$ is equal to 0.017 % and the maximum relative error is equal to -3.75×10^{-4} and is found at the last time step $\ln(t/t_s) = 5$. It is thus found that thermal interactions in

a 5×4 borefield totaling 20 boreholes can be accurately represented by 3 equivalent boreholes.



(a) g-function of a 5×4 borefield

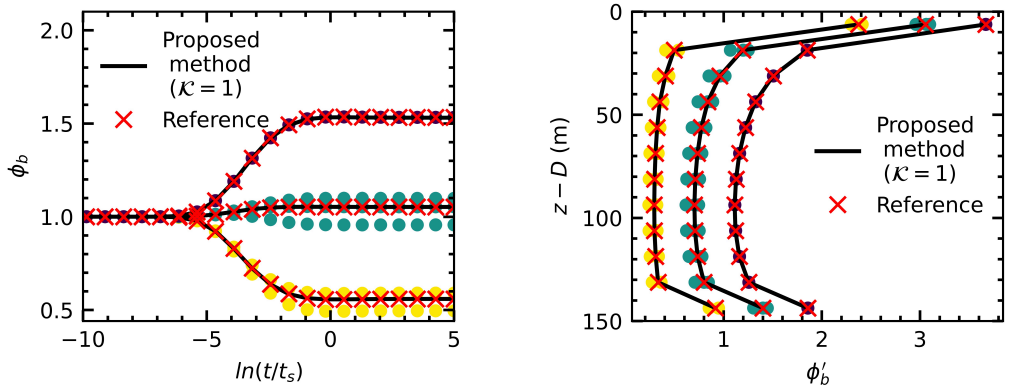
(b) Errors on the g-function

Figure 4.: g-function comparison for a 5×4 borefield

The heat extraction rates of the equivalent boreholes are compared to the heat transfer rates of the individual boreholes calculated using the reference method to verify that equivalent boreholes are representative of their respective group. Figure 5a presents the time-variation of the average over-the-length heat extraction rate profile and Figure 5b presents the vertical steady-state heat extraction rate profile along for the same 5×4 borefield evaluated at $\ln(t/t_s) = 5$. Colored markers correspond to the heat transfer rates calculated by the reference method and their colors correspond to the groups identified in Figure 3a. Solid black lines correspond to the heat transfer rates of the equivalent boreholes calculated using the proposed method while the red crosses correspond to the average of the heat transfer rates calculated by the reference method for each group. It is shown that core boreholes in group I experiment lower heat transfer rates than perimeter boreholes in group III. The good agreement between the heat transfer rates of equivalent boreholes and the group averages shows that the proposed hierarchical agglomerative clustering method is able to identify equivalent boreholes which are representative of their group. The maximum relative error on the average over-the-length heat extraction rates of equivalent boreholes is less than -7×10^{-4} for all groups when compared to the group average evaluated using the reference method. The maximum relative error on the steady-state heat extraction rate profiles is -5×10^{-4} for group III.

3.2. Regular borefield configurations

The accuracy and calculation time of the method is assessed for various regular borefield configurations with up to 144 boreholes. Figure 6 presents schematic representations of rectangular, L-, Box- and U-shaped borefields. For each of the configurations, the method is applied to borefields in all possible combinations of 2×1 up to 12×12 boreholes on the horizontal and vertical directions, respectively. For these simulations, the number of groups varies between 2 and 4 for $\mathcal{K} = 1$ and between 2 and 5 for $\mathcal{K} = 2$. A number of groups $\mathcal{G} = 2$ is only found in the case of the 2×1 configurations. Figure 7 compares the $MAPE$ for $\mathcal{K} = 1$ and $\mathcal{K} = 2$ for regular borefield configurations



(a) Average borehole heat transfer rate

(b) Steady-state borehole heat transfer rate profiles

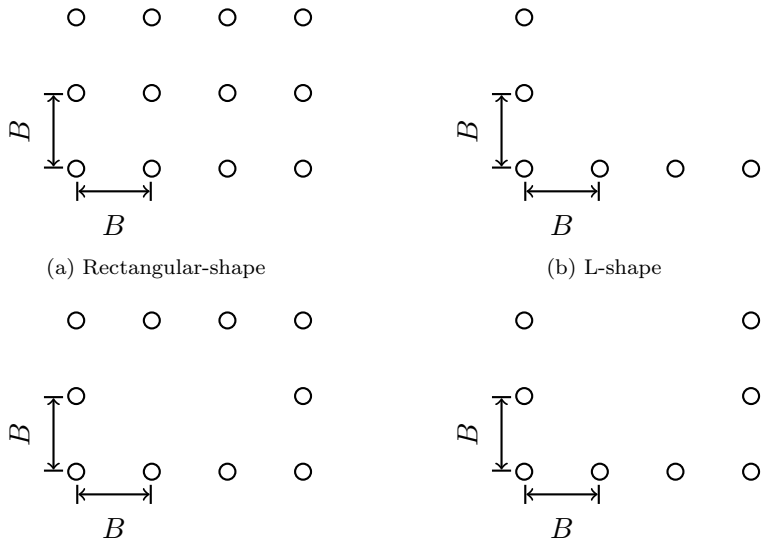
Figure 5.: Heat extraction rates of a 5×4 borefield

from 2×1 up to 12×12 boreholes. *MAPE* values lower than 10^{-4} are shown on a linear scale. Markers are colored according to the number of equivalent boreholes used for the calculation of the g-function. The largest *MAPE* is equal to 0.612 % and found for the 12×9 rectangular borefield totaling 108 boreholes for $\mathcal{K} = 1$ ($\mathcal{G} = 4$), and is equal to 0.490% for a 12×7 rectangular borefield totaling 84 boreholes with $\mathcal{K} = 2$ ($\mathcal{G} = 5$). For the largest borefield, totaling 144 boreholes in a 12×12 rectangular configuration, the *MAPE* is equal to 0.493 and 0.254 % with 3 ($\mathcal{K} = 1$) and 4 groups ($\mathcal{K} = 2$), respectively. It is thus observed that the precision increment $\mathcal{K}$ has a positive but not significant impact on the *MAPE* when increasing its value by one. A value of $\mathcal{K} = 1$ is found to provide sufficient accuracy across all regular configurations. The borefield configuration was not found to have any significant impact on the *MAPE*. It should be noted that the *MAPE* reaches zero when the number of groups is equal to the number of boreholes in the borefield, or earlier if there are any symmetries in the borefield layout.

The method presents substantial improvements in calculation time due to the decrease in the number of modeled boreholes for large borefields. Figure 8 presents the calculation times of the reference method (using `pygfunction`) and the proposed method. As can be seen, the calculation time starts increasing rapidly in the reference method once the number of boreholes reaches approximately 100 boreholes. Comparatively, the calculation time in the proposed method increases very slowly with the number of boreholes. As an example, for the 12×12 rectangular borefield with 144 boreholes, the calculation time is 7.4 sec and 0.9 sec for the reference method and the proposed method, respectively.

3.3. Irregular borefield configurations

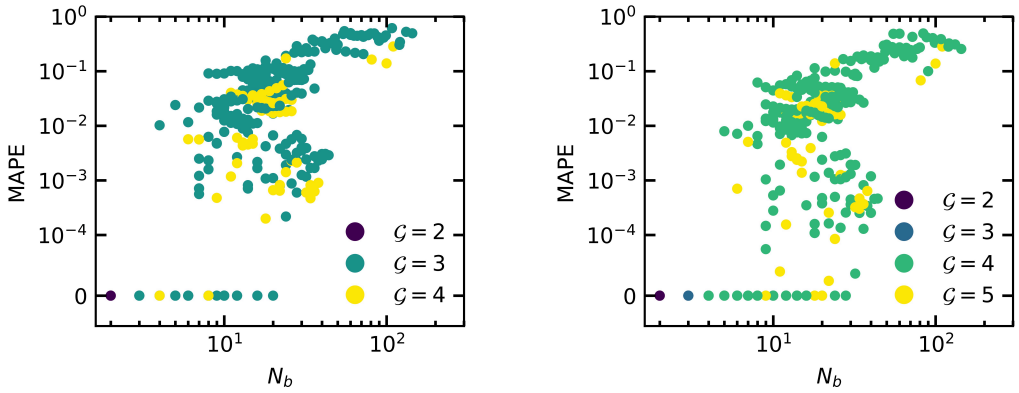
The proposed methodology is not limited to regular configurations such as those presented in the previous section but also to non-conventional configurations. The applicability of the method to irregular configurations is tested in this section by evaluating the g-functions of 100 randomly positioned boreholes in different borefield shapes. Two series of cases are presented: (i) four configurations with boreholes randomly positioned



(a) Rectangular-shape (b) L-shape

(c) Box-shape (d) U-shape

Figure 6.: Regular borefield configurations



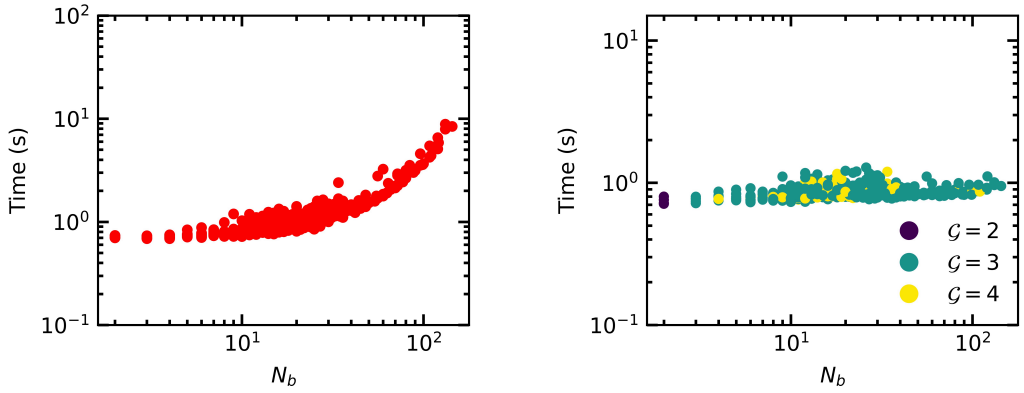
(a) $\mathcal{K} = 1$ (b) $\mathcal{K} = 2$

Figure 7.: Mean absolute percentage error on the g-functions

in a rectangular domain with different form factors (ii) four configurations with the same overall rectangular domain but with a geometric constraint in the center of it (e.g. a building). The scalability of the method to very large borefields is then tested in section 3.3.2.

3.3.1. Mean absolute percentage error

The accuracy of the proposed method is now evaluated by calculating g-functions of borefields of randomly positioned boreholes. 100 boreholes are randomly positioned in rectangular and hollow-rectangular domains as shown on figure 9. In all cases, the overall domain area is 40,000 m^2 and the minimum spacing between boreholes is 7.5 m. Configurations A, B, C and D have varying shape factors L_y/L_x of 1, 2, 4 and

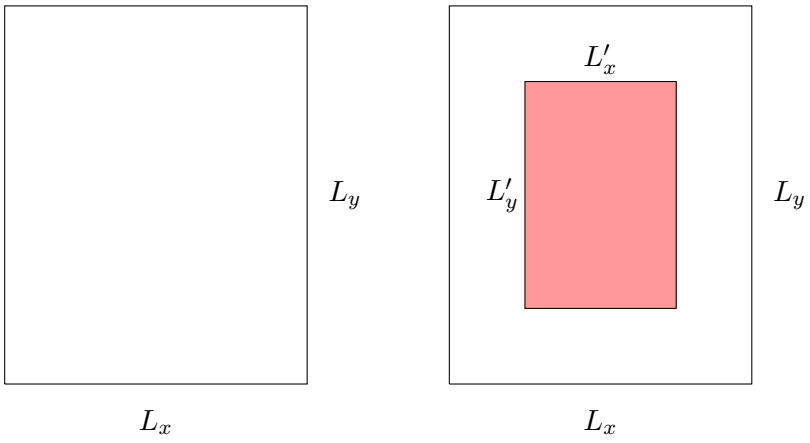


(a) Reference method

(b) Proposed method

Figure 8.: Calculation time for the evaluation of g-functions of regular borefields

8, respectively. Configurations E, F, G and H have the same overall dimensions as configuration A but boreholes cannot be placed in a rectangular area at the center of the field (which could represent a building in practical cases). The physical dimensions of the borefield domains are presented in table 2.



(a) Rectangular domains: configurations A to D

(b) Hollow-rectangular domains: configurations E to H

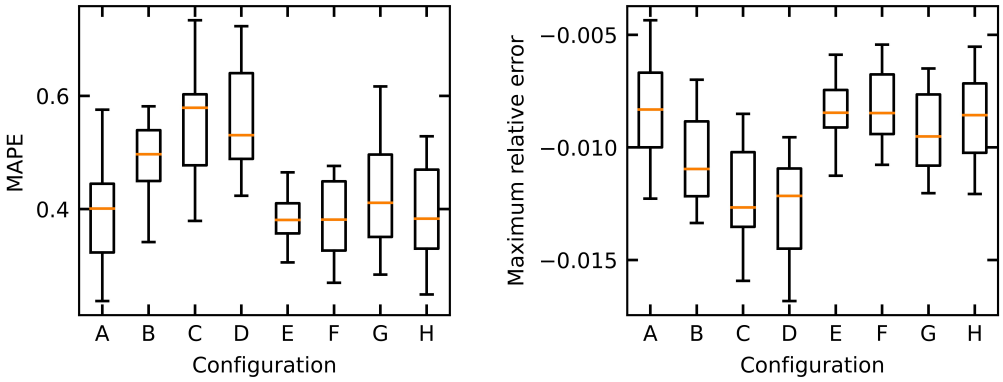
Figure 9.: Shapes of the irregular borefield domains

Figure 10 shows the *MAPE* and maximum relative error for 20 random borefields in each of the 8 configurations. The number of groups identified from the clustering method varies from $\mathcal{G} = 3$ and $\mathcal{G} = 5$ depending on the positions of the boreholes in the randomly generated borefields. The proposed method maintains a *MAPE* below 0.72 % in all cases, and below 0.62% when considering the *MAPE* under the third quartile. The maximum relative error (in magnitude) is -0.0168 and found in configuration D at $\ln(t/t_s) = 5$. Again, the negative sign means the proposed method overpredicts the reference value. An overestimation of the g-function is consistent with results obtained when decreasing the axial discretization of boreholes, i.e. using fewer segments per borehole to evaluate the g-function (Cimmino & Bernier, 2014). As the number of

Table 2.: Dimensions of the irregular borefield domains

Parameter	Configuration							
	A	B	C	D	E	F	G	H
L_x $[m]$	200	141	100	71	200	200	200	200
L_y $[m]$	200	238	400	566	200	200	200	200
L'_x $[m]$	-	-	-	-	160	128	102	82
L'_y $[m]$	-	-	-	-	160	128	102	82

segments decreases, the temperature at the borehole walls moves away from an equal and uniform temperature and the g-function is overpredicted. In the same way, when using fewer equivalent boreholes, the borehole wall temperatures of the boreholes (as opposed to the *equivalent* boreholes) move away from an equal and uniform value and the g-function should likewise be overpredicted.



(a) g-function $MAPE$ on random field configurations

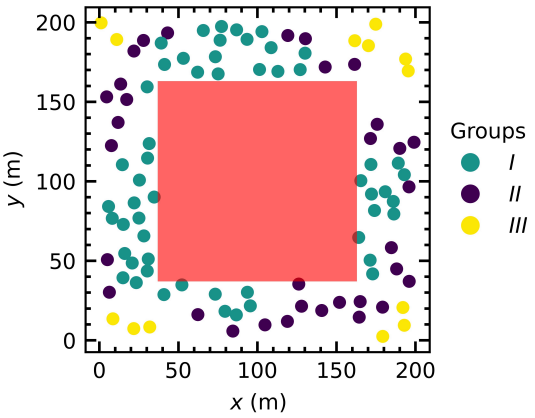
(b) g-function maximum relative error on random field configurations

Figure 10.: $MAPE$ on the g-function and number of groups for the irregular borefields

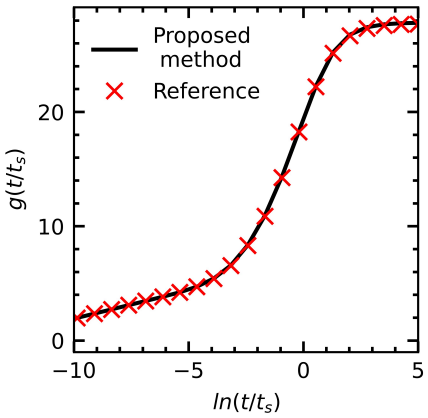
Sample results for a borefield of 100 boreholes on a hollow-rectangular domain in configuration F are shown on Figure 11. The g-function, calculated using only 3 equivalent boreholes and shown on Figure 11b, is evaluated with a $MAPE = 0.334$ % and with a maximum relative error of -0.006 (at $\ln(t/t_s) = 5$). The numbers of boreholes per group are 54, 33 and 13 (I, II and III), respectively. Despite the large number of boreholes and the irregular configuration, Figures 11c and 11d show that the heat extractions rates of equivalent boreholes are in good agreement with the average heat extraction rates calculated with the reference method. The maximum relative error on the heat extraction rate profiles at $\ln(t/t_s) = 5$ is -0.002 for group I.

3.3.2. Scalability of the proposed method

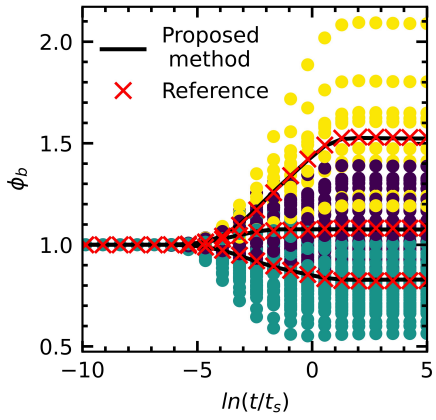
Figure 12 shows the g-function computation based on the proposed and reference methods for different random borefields with numbers of boreholes in increments of



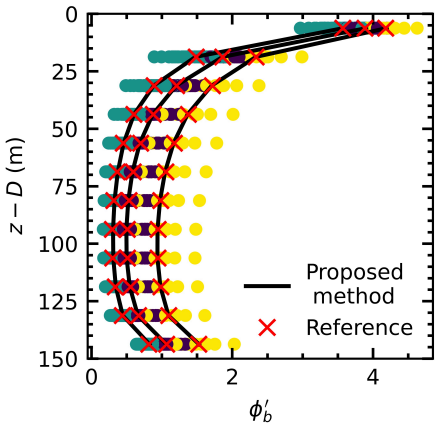
(a) Hierarchical clustering



(b) g-function



(c) Average borehole heat extraction rates



(d) Steady-state borehole heat extraction rate profiles

Figure 11.: Sample results for configuration F

powers of 2, starting from 8 up to 1024 in a limited region with dimensions $L_x \times L_y = 500 \times 500\ m^2$. Table 3 presents the $MAPE$, the maximum relative error and the calculation time for each of the cases. The results show good agreement with the reference method. In general, the maximum relative error is found at $\ln(t/t_s) = 5$. The calculation time was significantly reduced for large borefields. For instance, the calculation time of the reference method for 1024 boreholes is 5125 sec. For the calculation of the g-function of the field of 1024 boreholes, `pygfunction` had to be configured to use single precision (float32) instead of double precision (float64) so that the available memory was sufficient to do the computation.

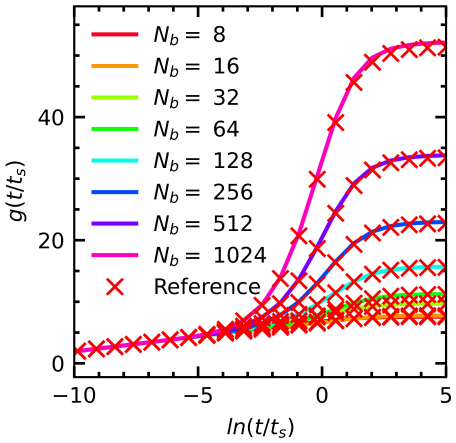


Figure 12.: g-function for different random borefield configurations

Table 3.: Scalability of the proposed method

N_b	8	16	32	64	128	256	512	1024
$MAPE(\%)$	0.002	0.007	0.092	0.115	0.224	0.405	0.540	0.690
Maximum relative error	6.95×10^{-5}	2.69×10^{-4}	-0.003	-0.003	-0.005	-0.009	-0.011	-0.015
Calculation time $[s]$	0.79	0.84	0.86	0.87	0.93	1.24	1.86	3.65

4. Conclusions

A new method is proposed to evaluate thermal interactions between geothermal boreholes. The finite line source (FLS) solution is extended to consider thermal interactions between groups of boreholes rather than between individual boreholes. It is shown that a small number of representative groups of boreholes, each modeled using a single equivalent borehole, can accurately represent the heat transfer process in a bore field. Hierarchical agglomerative clustering allows the identification of representative groups using a simplified heat transfer model, based on the steady-state finite line source (FLS) solution. A cut-off line is drawn on the borefield dendrogram to estimate the minimum number of groups and incremented by a precision increment ($\mathcal{K} = 1$).

The accuracy of the proposed method is assessed by calculating g-functions of geothermal borefields and comparing the results to the reference method (Cimmino, 2018b) implemented in `pygfunction` (Cimmino, 2018c). A comparison of regular borefield configurations with up to 144 boreholes shows that the proposed method achieves a mean absolute percentage error ($MAPE$) of less than 0.612 % using 4 or fewer equivalent boreholes. A marginal increase of the precision increment (i.e. from $\mathcal{K} = 1$ to $\mathcal{K} = 2$) is shown not to have a significant impact on the clustering method and the accuracy of the heat transfer model, however the error approaches 0 as the number of equivalent boreholes increases. The proposed method also presents important savings in calculation time : for a rectangular borefield of 12×12 boreholes, the calculation time is decreased from 7.4 sec to 0.9 sec.

The proposed method maintains accuracy even when considering irregular borefield configurations. Comparisons of g-functions of borefields of 100 randomly positioned boreholes shows a third-quartile $MAPE$ below 0.62% and a third-quartile maximum relative error below -0.015 in the worst configuration studied (configuration D). For a borefield of 1024 randomly positioned boreholes, the new method presents a $MAPE$ of 0.69 %, a maximum relative error of -0.015 and a calculation time of 3.65 sec. In comparison, the reference method has a calculation time of 5125 sec. The method thereby demonstrates efficient scaling properties and makes it possible to model thermal interactions in very large borefields comprising hundreds of boreholes, which was previously impractical due to the large computational requirements.

Using suitable analytical solutions, the new method could also be extended to consider more complex borefield geometries (e.g. inclined boreholes (Lazzarotto, 2016)), groundwater flow (Diao et al., 2004; Molina-Giraldo et al., 2011), or layered ground physical properties (Abdelaziz et al., 2014). Fields of mixed series and parallel connections between boreholes (Cimmino, 2019) could be considered but require a new distance metric to substitute Equation (16) and extend the equivalent *borehole* concept to an equivalent *loop* (or *branch*) concept. The presented method is only applicable to parrallel-connected boreholes but is not only limited to the calculation of g-functions using a uniform borehole wall temperature condition. The same clustering approach can be applied to the calculation of g-functions using an equal inlet fluid temperature boundary condition (Cimmino, 2015). Most importantly, it has been shown that thermal interactions between geothermal boreholes can be represented using a small number of equivalent boreholes. This approach can thereby be applied to network-based methods (Cimmino, 2018a; Lamarche, 2017a; Lazzarotto, 2014) for parallel-connected boreholes and coupled to accurate short-term borehole models (Prieto & Cimmino, 2021; Rivero & Hermanns, 2021) to obtain detailed axially discretized borefield models that are valid at all time scales.

References

Abdelaziz, S.L., T.Y. Ozudogru, C.G. Olgun and J.R. Martin. 2014. Multilayer Finite Line Source Model for Vertical Heat Exchangers. *Geothermics,* 51:406–416.

Ahmadfard, M., and M. Bernier. 2018. Modifications to ASHRAE's Sizing Method for Vertical Ground Heat Exchangers. *Science and Technology for the Built Environment,* 24(7):803–817.

Ahmadfard, M., and M. Bernier. 2019. A Review of Vertical Ground Heat Exchanger Sizing Tools Including an Inter-Model Comparison. *Renewable and Sustainable Energy Reviews,* 110(August):247-–265.

ASHRAE. 2019. *Chapter 35 : Geothermal Energy.* In HVAC Applications, 35.1–35.51. Atlanta GA, USA.

Bauer, D., W. Heidemann and H.-J. Diersch. 2011. Transient 3D analysis of borehole heat exchanger modeling. *Geothermics,* 40 (4):250–260.

Bernier, M., and A. Salim Shirazi 2013. Thermal capacity effects in borehole ground heat exchangers. *Energy and Buildings*, 67:352–364.

BniLam, N. and R. Al-khouri. 2016. A spectral model for heat transfer with friction heat gain in geothermal borehole heat exchangers. *Applied Mathematical Modeling,* 40 (15-16):7410–7421.

Brussieux, Y., and M. Bernier 2019. Universal short-time g*-functions: Generation and application. *Science and Technology for the Built Environment,* 25 (8):993–1006.

Cimmino, M. 2015. The effects of borehole thermal resistances and fluid flow rate on the g-functions of geothermal bore fields. *International Journal of Heat and Mass Transfer,* 91:1119–1127.

Cimmino, M. 2018a. A finite line source simulation model for geothermal systems with series- and parallel-connected boreholes and independent fluid loops. *Journal of Building Performance Simulation,* 11:4, 414–432.

Cimmino, M. 2018b. Fast calculation of the g-functions of geothermal borehole fields using similarities in the evaluation of the finite line source solution. *Journal of Building Performance Simulation,* 11:6, 655–668.

Cimmino, M. 2018c. *pygfunction* : an open-source toolbox for the evaluation of thermal response factors for geothermal borehole fields. *Proceedings of eSim 2018,* Montreal QC, Canada, 492–501.

Cimmino, M. 2019. Semi-Analytical Method for g-Function Calculation of Bore Fields with Series- and Parallel-Connected Boreholes. *Science and Technology for the Built Environment,* 25(8):1007–1022.

Cimmino, M., M. Bernier and F. Adams. 2013. A contribution towards the determination of g-functions using the finite line source. *Applied Thermal Engineering,* 51, 401–412.

Cimmino, M. and M. Bernier. 2014. A semi-analytical method to generate g-functions for geothermal borefields. *International Journal of Heat and Mass Transfer,* 70, 641–650.

Claesson, J. and S. Javed. 2011. An analytical method to calculate borehole fluid temperatures for time-scales from minutes to decades. *ASHRAE Transactions,* 117:279–288.

Cupeiro Figueroa, I., M. Cimmino and L. Helsen. 2020. A Methodology for Long-Term Model Predictive Control of Hybrid Geothermal Systems: The Shadow-Cost Formulation. *Energies,* 13(23):6203.

Diao, N., Q. Li and Z. Fang. 2004. Heat transfer in ground heat exchangers with groundwater advection. *International Journal of Thermal Sciences,* 43 (12):1203–1211.

Dussault, B., Pasquier, P., and D. Marcotte. 2018. A block matrix formulation for efficient g-function construction. *Renewable Energy,* 121, 249–260.

Eskilson, P. 1987. Thermal analysis of heat extraction boreholes. University of Lund, Sweden.

Gu, Y. and D. L. O'Neal. 1995. Analytical solution to transient heat conduction in a composite region with a cylindrical heat source. *Journal of Solar Energy Engineering, Transactions of the ASME* 117 (3):242–248.

Ingersoll, L., O. Zobel and A. Ingersoll. 1954. Heat conduction : with engineering, geological and other applications. University of Wisconsin, Madison.

Javed, S., and J. Claesson. 2011. New analytical and numerical solutions for the short-term analysis of vertical ground heat exchangers. *ASHRAE Transactions* 117 (1):3–12.

Jung, Y., H. Park, D. Ding-Zhu and L. Barry. 2003. A Decision Criterion for the Optimal Number of Clusters in Hierarchical Clustering. *Journal of Global Optimization,* 25, 91–111.

Laferrière, A., M. Cimmino, D. Picard and L. Helsen. 2020. Development and validation of a full-time-scale semi-analytical model for the short- and long-term simulation of vertical geothermal bore fields. *Geothermics,* 86:101788.

Lamarche, L. 2017. Mixed Arrangement of Multiple Input-Output Borehole Systems. *Applied Thermal Engineering,* 124, 466–476.

Lamarche, L. 2017. *g*-function generation using a piecewise-linear profile applied to ground heat exchangers. *International Journal of Heat and Mass Transfer,* 115(December):354–360.

Lamarche, L. and B. Beauchamp. 2007a. A new contribution to the finite line-source model for geothermal boreholes *Energy and Buildings* 39 (2):188–198.

Lamarche, L. and B. Beauchamp 2007b. New solutions for the short-time analysis of geothermal vertical boreholes. *International Journal of Heat and Mass Transfer* 50 (7-8):1408–1419.

Lazzarotto, A. 2014. A Network-Based Methodology for the Simulation of Borehole Heat Storage Systems. *Renewable Energy,* 62, 265–275.

Lazzarotto, A. 2016. "A Methodology for the Calculation of Response Functions for Geothermal Fields with Arbitrarily Oriented Boreholes – Part 1." Renewable Energy 86: 1380–93. https://doi.org/10.1016/j.renene.2015.09.056.

Li, M., P. Li, V. Chang and A. C. K. Lai. 2014. Full-scale temperature response function (G-function) for heat transfer by borehole ground heat exchangers (GHEs) from sub-hour to decades. *Applied Energy,* 136:197–205.

Li, M. and A.C.K. Lai. 2015. Review of Analytical Models for Heat Transfer by Vertical Ground Heat Exchangers (GHEs): A Perspective of Time and Space Scales. *Applied Energy,* 151(August):178–91.

Minaei, A., and M. Maerefat. 2017. A new analytical model for short-term borehole heat exchanger based on thermal resistance capacity model. *Energy and Buildings* 146:233–242.

Mitchell, M.S. and J.D. Spitler. 2019. Characterization, Testing, and Optimization of Load Aggregation Methods for Ground Heat Exchanger Response-Factor Models. *Science and Technology for the Built Environment,* 25(8):1036–51.

Molina-Giraldo, N., P. Blum, K. Zhu, P. Bayer and Z. Fang. 2011. A moving finite line source model to simulate borehole heat exchangers with groundwater advection. *International Journal of Thermal Sciences,* 50 (12):2506–2513.

Nguyen, A. and P. Pasquier. 2021. A successive flux estimation method for rapid g-function construction of small to large-scale ground heat exchanger. *Renewable Energy,* 165:359–368.

Pasquier, P., and D. Marcotte 2014. Joint use of quasi-3D response model and spectral method to simulate borehole heat exchanger. *Renewable Energy,* 51:281–299.

Patil, C. and I. Baidari. 2019. Estimating the Optimal Number of Clusters k in a Dataset Using Data Depth. *Data Science and Engineering,* 4 (2):132–140.

Prieto, C. and M. Cimmino. 2021. Transient multipole expansion for heat transfer in ground heat exchangers. *Science and Technology for the Built Environment,* 27 (3):253–270.

Rivero, J. M. and M. Hermanns. 2021. Enhanced multipole method for the transient thermal response of slender geothermal boreholes. *International Journal of Thermal Sciences,* 164:106531.

Spitler, J.D. and S.E.A. Gehlin. 2015 Thermal response testing for ground source heat pump systems-An historical review. *Renewable and Sustainable Energy Reviews,* 50:1125–1137.

Van Der Walt, S., S. C. Colbert and G. Varoquaux. 2011. The NumPy Array: A Structure for Efficient Numerical Computation. *Journal of Computing in Science & Engineering,* 13 (2):22–30.

Virtanen, P., R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, et al. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods,* 17 (3):261–272.

Xu, X., and J. Spliter. 2006. Modeling of vertical ground loop heat exchangers with variable convective resistance and thermal mass of the fluid. *Proceedings of Ecostock America*: Pomona.

Yang, Y., and M. Li. 2014. Short-time performance of composite-medium line-source model for predicting responses of ground heat exchangers with single U-shaped tube. *International Journal of Thermal Sciences* 82:130–137.

Yavuzturk, C. 1999. *Modelling of vertical ground loop heat exchangers for ground source heat pump systems*. PhD Thesis. Oklahoma State University. Stillwater OK, United States of America.

Zarrella, A., M. Scarpa and M. De Carli. 2011. Short time step analysis of vertical ground-coupled heat exchangers: The approach of CaRM. *Renewable Energy* 36 (9):2357–2367.

Zeng, H. Y., N.R. Diao, and Z. H. Fang. 2002 A finite line-source model for boreholes in

geothermal heat exchangers. *Heat Transfer—Asian Research*, 31 (7):558–567.

Zhou, S., Z. Xu and F. Liu. 2017. Method for Determining the Optimal Number of Clusters Based on Agglomerative Hierarchical Clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 28 (12):3007–3017.