15TH IEA HEAT PUMP CONFERENCE
MAY 26–29, 2026 | HOFBURG VIENNA 2026
DECARBONISATION THROUGH INNOVATION

An efficient grouping strategy for modeling thermal interactions in large borefields

Alain Nguyen[a,*], Philippe Pasquier[b]

[a]Natural Resources Canada, 1615 Lionel-Boulet, P.O. Box 4800, Varennes, J3X 1S6, Canada
[b]Polytechnique Montréal, P.O. Box 6079 Station Centre-Ville, Montréal, H3C 3A7, Canada

Abstract

Accurately modeling thermal interactions between ground heat exchangers (GHEs) is essential for the design and simulation of geothermal heat pump (GHP) systems, particularly in large and irregular borefields. However, the computational cost of evaluating these interactions increases rapidly with the number of boreholes, often becoming a limiting factor. This work introduces a novel, efficient, method to alleviate this burden through a similarity-based grouping approach for g-function construction. The proposed method identifies groups of boreholes exhibiting similar thermal behavior—characterized by comparable heat extraction rates—by solving a single linear system derived from the superposition principle. Each group is then represented in the model by a single representative borehole, significantly reducing the problem complexity while maintaining high accuracy. In contrast to previously proposed aggregation or clustering methods that depend on spatial proximity or geometric symmetry, this approach directly leverages thermal behavior and avoids iterative refinement or convergence steps, resulting in a computationally robust and scalable solution. In a case study with 1,958 randomly positioned GHEs, the grouped model produces a g-function with negligible deviation from the full-resolution solution, while reducing computation time to just 0.8 seconds on a standard workstation laptop.

Selection and/or peer-review under the responsibility of the organizers of the 15th IEA Heat Pump Conference 2026.

Keywords: Geothermal heat pump; Ground heat exchanger; Thermal interaction; Grouping

1. Introduction

Accurately modeling thermal interactions between ground heat exchangers (GHE) is essential for the design and analysis of geothermal heat pump (GHP) systems, particularly in large and irregular borefields. The standard analytical tool for this purpose is the g-function, first introduced several decades ago by [1], which characterizes the temporal thermal response of a borefield to a prescribed heat extraction rate. Since then, a variety of methods have been proposed to extend and improve the computation of g-functions for different borefield configurations and operating conditions. Nevertheless, evaluating g-functions for large-scale systems remains computationally demanding, as the calculation of pairwise interactions between boreholes has quadratic complexity with respect to the number of boreholes (i.e., $O(n^2)$). This computational burden can become a significant limitation in iterative design, optimization, and real-time monitoring and control applications.

Existing approaches to reduce computational cost include spatial aggregation, symmetry-based clustering, and simplified equivalent representations. For example, [2] accelerated g-function calculations by grouping segment pairs by distance, then applying computed factors to all pairs within each group. Later, [3] employed hierarchical agglomerative clustering to replace groups of boreholes with single equivalent representations. While effective, these approaches rely on geometric approximations, heuristic grouping, or iterative refinement, which can limit both accuracy and efficiency in irregular or heterogeneous borefields, particularly when

* Corresponding author. *E-mail address:* tuananhalain.nguyen@nrcan-rncan.gc.ca

additional physical processes—such as groundwater advection, stratified ground properties, or short-term thermal effects—are present.

To address these limitations, this work introduces a simple response-based grouping strategy for efficient g-function computation in large irregular borefields. Unlike previous methods, it identifies borehole groups with similar thermal behavior—based solely on comparable heat extraction rates—by efficiently solving a linear system derived from the superposition principle. The specific objectives are to (i) formulate the grouping procedure within a tensor-based framework, (ii) quantify the trade-off between accuracy and computational efficiency, and (iii) demonstrate the scalability of the method to large, irregular borefield configurations.

2. Methodology

2.1. Source Modelling

In this work, the finite line source (FLS) model [4] is employed to evaluate the unitary transfer function of a single borehole. For a borehole of index i, with length H, buried at depth D in homogeneous ground, its impulse response function g, evaluated at radial distance r_j and time t_k, is given by:

$$g_{k,i\to j} = \frac{1}{4\pi\lambda} \int_{1/\sqrt{(4\lambda t_k)/C_s}}^{\infty} \exp\left(-r_j^2 s^2\right) \frac{Y(Hs, Ds)}{Hs^2} ds \tag{1}$$

with

$$Y(h,d) = 2\,\text{ierf}(h) + 2\,\text{ierf}(h+2d) - 2\,\text{ierf}(2h+2d) - \text{ierf}(2d) \tag{2}$$

and

$$\text{ierf}(X) = X\,\text{erf}(X) - \frac{1}{\sqrt{\pi}}\left(1 - \exp\left(-X^2\right)\right) \tag{3}$$

In Equation (1), λ is the ground thermal conductivity and C_s is the ground specific heat. It is worth emphasizing that the construction of a g-function is not limited to the FLS model; it can be applied to any type of GHE—such as vertical boreholes, standing column wells, open-loop systems, or energy piles—modeled using analytical, numerical, or experimental approaches, provided that a physically consistent impulse response function capable of representing thermal interactions between the various sources and receptors is available.

2.2. Proposed Grouping Strategy

Let $\mathbf{G} \in \mathbb{R}^{n\times n\times l}$ denote the interaction tensor of a borefield containing n boreholes over l time steps:

$$\mathbf{G_k} = \begin{bmatrix} g_{k,1\to1} & g_{k,1\to2} & \cdots & g_{k,1\to n} \\ g_{k,2\to1} & g_{k,2\to2} & \cdots & g_{k,2\to n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k,n\to1} & g_{k,n\to2} & \cdots & g_{k,n\to n} \end{bmatrix} \tag{4}$$

and let $\mathbf{A_k}$ be an augmented matrix:

$$\mathbf{A_k} = \begin{bmatrix} g_{k,1\to1} & g_{k,1\to2} & \cdots & g_{k,1\to n} & 1 \\ g_{k,2\to1} & g_{k,2\to2} & \cdots & g_{k,2\to n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{k,n\to1} & g_{k,n\to2} & \cdots & g_{k,n\to n} & 1 \\ 1 & 1 & \ldots & 1 & 0 \end{bmatrix} \tag{5}$$

For a given time step k, a solution vector $\mathbf{x_k}$ containing the contributions of all sources and the resulting response $\hat{g}_k$ can be assembled as follows:

$$\mathbf{x_k} = \begin{bmatrix} x_{k,1} & x_{k,2} & \ldots & x_{k,n} & -\hat{g}_k \end{bmatrix}^T \tag{6}$$

where $\mathbf{x_k}$ satisfies the constraint vector $\mathbf{c}$:

$$\mathbf{c} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 \end{bmatrix}^T \tag{7}$$

Here, $\mathbf{c}$ enforces two conditions: (1) the thermal response of all sources is identical and equal to $\hat{g}_k$, and (2) the total contribution satisfies $\sum x_k = 1$. In compact matrix form, the linear system can be expressed as:

$$\mathbf{A_k x_k} = \mathbf{c}, \quad \forall k = 1, \ldots, l \tag{8}$$

Solving for the vector of contributions $\mathbf{x_k}$ provides the basis for identifying components that exhibit thermal similarity. Indeed, two entries of the solution vector, $x_{k,i}$ and $x_{k,j}$, are considered similar if their relative difference satisfies:

$$\frac{|x_{k,i} - x_{k,j}|}{\max(|x_{k,i}|, |x_{k,j}|)} < \delta, \quad i, j = 1, \dots, n, \tag{9}$$

with δ a prescribed tolerance. This criterion ensures that values differing only within numerical precision are treated as equivalent.

Now, let vector $\mathbf{u_k}$ denote the resulting set of similar values in $\mathbf{x_k}$, and let $\mathbf{p_k}$ be a vector containing, for each unique value $u_{k,z}$ ($z = 1, \dots, m$), the number of entries in $\mathbf{x_k}$ that are within δ of $u_{k,z}$:

$$\mathbf{u_k} = \begin{bmatrix} u_{k,1} & u_{k,2} & \dots & u_{k,m} \end{bmatrix}^T \tag{10}$$

$$\mathbf{p_k} = \begin{bmatrix} p_{k,1} & p_{k,2} & \dots & p_{k,m} \end{bmatrix}^T \tag{11}$$

A keen eye might note that the identification of similarities is inherently time dependent, hence the index k in Equations (9)-(11). Indeed, at early times, before thermal interactions between boreholes develop, more similarities are observed, whereas at later times these similarities diminish. Consequently, solving the system exclusively at the final step (i.e., $k = l$) ensures that the grouping captures components that remain thermally similar over the entire prescribed simulation horizon.

Accordingly, by considering only the final time step (i.e. $k = l$), each entry of $\mathbf{x_l}$ can be associated with one of these unique values through a weight mapping matrix $\mathbf{M} \in \mathbb{R}^{n \times m}$, defined as:

$$\mathbf{M}_{j,z} = \begin{cases} 1/p_{l,z}, & \text{if } x_{l,j} \text{ is within } \delta \text{ of } u_{l,z}, \\ 0, & \text{otherwise,} \end{cases} \tag{12}$$

Here, $\mathbf{M}$ and $\mathbf{M^T}$ represent the row and column normalized contributions, respectively. Finally, the aggregation is performed as follows:

$$\overline{\mathbf{G}}_k = \mathbf{M}^\mathsf{T} \mathbf{G}_k \mathbf{M}, \quad \forall k = 1, \dots, l \tag{13}$$

The resulting tensor $\overline{\mathbf{G}} \in \mathbb{R}^{m \times m \times l}$ represents the interactions between the m groups, effectively reducing the size of the problem while preserving the weighted contributions from the original entries. Equation (13) can be evaluated efficiently by exploiting the sparsity of $\mathbf{M}$, thereby minimizing the number of nonzero operations. Additional efficiency is achieved through broadcasting, which allows all $k = 1, \dots, l$ to be computed simultaneously.

In addition, notice that this formulation is not tied to the explicit geometry of the borefield, nor to the specific characteristics of individual boreholes, but rather to the emergence of aggregated thermal interactions. As a result, the approach remains broadly applicable—scaling efficiently to large configurations and adapting naturally to irregular layouts or site-specific conditions—provided that an appropriate impulse response function is available.

Finally, it should be emphasized that Equations. (9)–(11) can be implemented with minimal computational cost using standard numerical computing tools (e.g., the `uniquetol` function in MATLAB). A numerical example is provided in Appendix A to facilitate reader comprehension.

2.3. g-function Construction

To construct the g-function of a borefield composed of m groups over a time span of l time steps, it is necessary to determine the $m \times l$ contributions of each group that ensure a consistent thermal response for all groups. However, the groups thermally interact with each other, as each impacts the borefield differently. Consequently, to maintain the same output across all groups, the contribution from each individual group must vary over time.

In this work, the g-function is constructed efficiently using the successive flux estimation algorithm [5, 6], an iterative procedure for simultaneously estimating the heat flux signals of multiple interacting sources to reproduce prescribed temperatures. For a detailed description of the algorithm, the reader is referred to the original publications.

It should be noted that other approaches—either sequential [7, 8, 9] or direct [10]—can also be used to construct the g-function, provided they offer a consistent representation of the interactions between sources within a tensor-based framework. Consequently, the method presented here is not limited to the successive flux estimation algorithm.

3. Test Cases

To evaluate the performance and accuracy of the proposed grouping strategy, four playful borefield configurations are considered:

1. **Irregular Square Layout:** Boreholes are distributed randomly within a square area, resulting in uneven spacing and varying distances between neighboring boreholes ($n = 625$).
2. **Word-Shaped Layout:** Boreholes are arranged to draw the word "HELLO". The positions are generated using Delaunay triangulation to ensure well-distributed spacing along the letters ($n = 1,045$).
3. **Star Layout:** Boreholes are positioned along radial lines emanating from a central point, forming a star-like pattern. Delaunay triangulation is applied to control spacing along each radial line ($n = 1,339$).
4. **Circle Layout:** Boreholes are placed along the circumference of a circle. Delaunay triangulation is used to maintain uniform spacing and connectivity among boreholes ($n = 1,958$).

These test cases capture a wide range of illustrative borefield geometries, from irregular to highly structured configurations, allowing a thorough assessment of the proposed methodology under diverse spatial arrangements.

For all borefield configurations, g-functions are constructed both without grouping and with grouping applied using tolerances of $\delta = 10^{-3}, 10^{-2}$, and 10^{-1}. This approach enables a systematic assessment of how the grouping tolerance influences the resulting g-function across different geometrical layouts. Finally, in all these cases, the parameters used for generating the impulse response functions with the FLS are D=2 m, H=50 m, r_b=0.075 m, λ=4 W $^\circ$C^{-1} m^{-1} and C_s= 2 MJ $^\circ$C^{-1} m^{-3}.

4. Results and discussion

Tables 1–2 summarize the performance of the proposed grouping strategy across the four borefield configurations. Three metrics are reported: the number of groups after grouping, the mean absolute percentage error (MAPE) and the corresponding computation time with respect to the reference case (no grouping).

Table 1 shows that the grouping strategy substantially reduces the number of groups as the tolerance δ increases. For example, in Case 4, the number of groups decreases from 1,958 in the reference case to only 30 when $\delta = 10^{-1}$. This trend is consistent across all configurations, with reductions by one to two orders of magnitude. Such results demonstrate the efficiency of the proposed method in condensing large-scale borefields into more compact equivalent representations, thereby lowering the dimensionality of the problem.

Table 1: Number of groups (-)

δ	**Case 1**	**Case 2**	**Case 3**	**Case 4**
–	625	1,045	1,339	1,958
10^{-3}	365	616	466	658
10^{-2}	147	222	153	185
10^{-1}	35	42	28	30

What is more important, however, is whether accuracy is preserved as the number of groups decreases. Table 2 reports the mean absolute percentage error (MAPE) between the grouped and reference g-functions. The results show that even for aggressive grouping (e.g., $\delta = 10^{-1}$), the error remains below 0.3% across all cases. For smaller tolerances such as $\delta = 10^{-3}$ and $\delta = 10^{-2}$, the errors are even lower, typically well below 0.2%. This high accuracy stems from the weight mapping matrix $\mathbf{M}$, which preserves the weighted contributions of the original boreholes within each group. This is illustrated in Figure 1, where the g-functions for all borefield configurations appear visually indistinguishable across cases.

While these reductions in error are measurable, they may not translate into significant practical gains, as the deviations are already negligible for most engineering applications. What naturally follows is an assessment of the computation time for each case, as summarized in Table 3. All computations were performed using MATLAB R2024b on a standard workstation laptop.

The results show that in all cases, grouping reduces computational cost. Moreover, increasing the tolerance δ further amplifies these savings, as illustrated in Case 4, where the computation time decreases from 6.82 s to 0.80 s. These times encompass all steps, including the generation of the initial tensor $\mathbf{G}$, the grouping procedure, and the construction of the g-function.

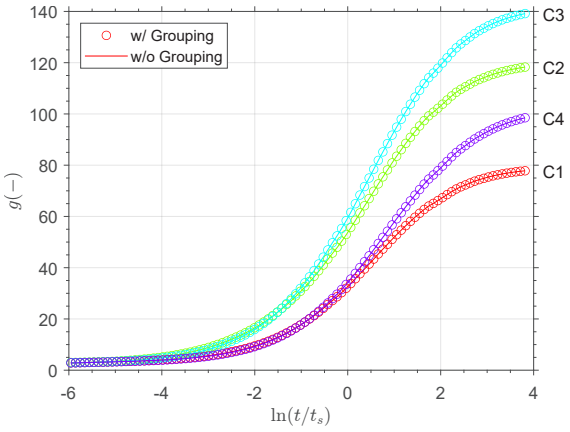

Figure 1: g-Functions for all borefield configurations, showing the thermal response over the simulation horizon (with $\delta = 10^{-1}$).

Table 2: Mean absolute percentage error (%)

δ	**Case 1**	**Case 2**	**Case 3**	**Case 4**
10^{-3}	0.0167	0.0434	0.0039	0.0003
10^{-2}	0.0273	0.1525	0.0097	0.0006
10^{-1}	0.0431	0.2449	0.0186	0.0010

As shown, even with moderate tolerances such as $\delta = 10^{-2}$, the computation time is reduced by nearly a factor of four-to-six compared to the ungrouped scenario, while maintaining errors below 0.2%. These results highlight the practical benefit of the proposed grouping strategy: by reducing the dimensionality of the problem, it significantly accelerates g-Function evaluation without compromising accuracy. To the best of the authors' knowledge, these computation times for such large borefields are unprecedented, demonstrating that the proposed grouping strategy enables rapid g-function evaluation even for configurations with thousands of boreholes.

Table 3: Computation time (s)

δ	**Case 1**	**Case 2**	**Case 3**	**Case 4**
–	0.55338	2.0281	3.5813	6.8189
10^{-3}	0.50061	1.2713	1.2901	2.8082
10^{-2}	0.15234	0.5164	0.67736	1.2119
10^{-1}	0.11589	0.25961	0.37819	0.7983

At last, to illustrate the effects of the proposed grouping strategy, Figures 2 and 3 show the resulting groups for the four borefield configurations at a tolerance of $\delta = 10^{-1}$ and $\delta = 10^{-3}$, respectively. Each colored dot represents a borehole assigned to a group based on the similarity of its thermal contribution, providing a visual complement to the quantitative results reported in Tables 1–3.

5. Conclusion

The objective of this work was to develop an efficient framework for reducing the computational complexity of g-function evaluations in large-scale borefields. To this end, a grouping strategy was introduced, whereby boreholes are aggregated into representative groups to construct a reduced-order tensor formulation. This method preserves the weighted thermal interactions of the original system while significantly lowering the dimensionality of the problem. The results demonstrated that the proposed approach achieves substantial gains in computational efficiency without compromising accuracy. In particular, computation times were reduced from 6.82 s to 0.80 s for the largest case (i.e. 1,958 boreholes) when applying a grouping tolerance of $\delta = 10^{-1}$, while maintaining relative errors below 0.001%. To our knowledge, such performance improvements in g-function evaluation are unprecedented, establishing a new benchmark for large-scale borefield simulations.

Ultimately, the proposed formulation offers a key advantage: rather than being tied to the detailed geometry of the borefield or the properties of individual boreholes, it captures the aggregated thermal interactions that govern system behavior. This makes the approach naturally suited to highly irregular layouts and scalable to very large borefields. In addition, because the formulation is agnostic to local subsurface processes, it

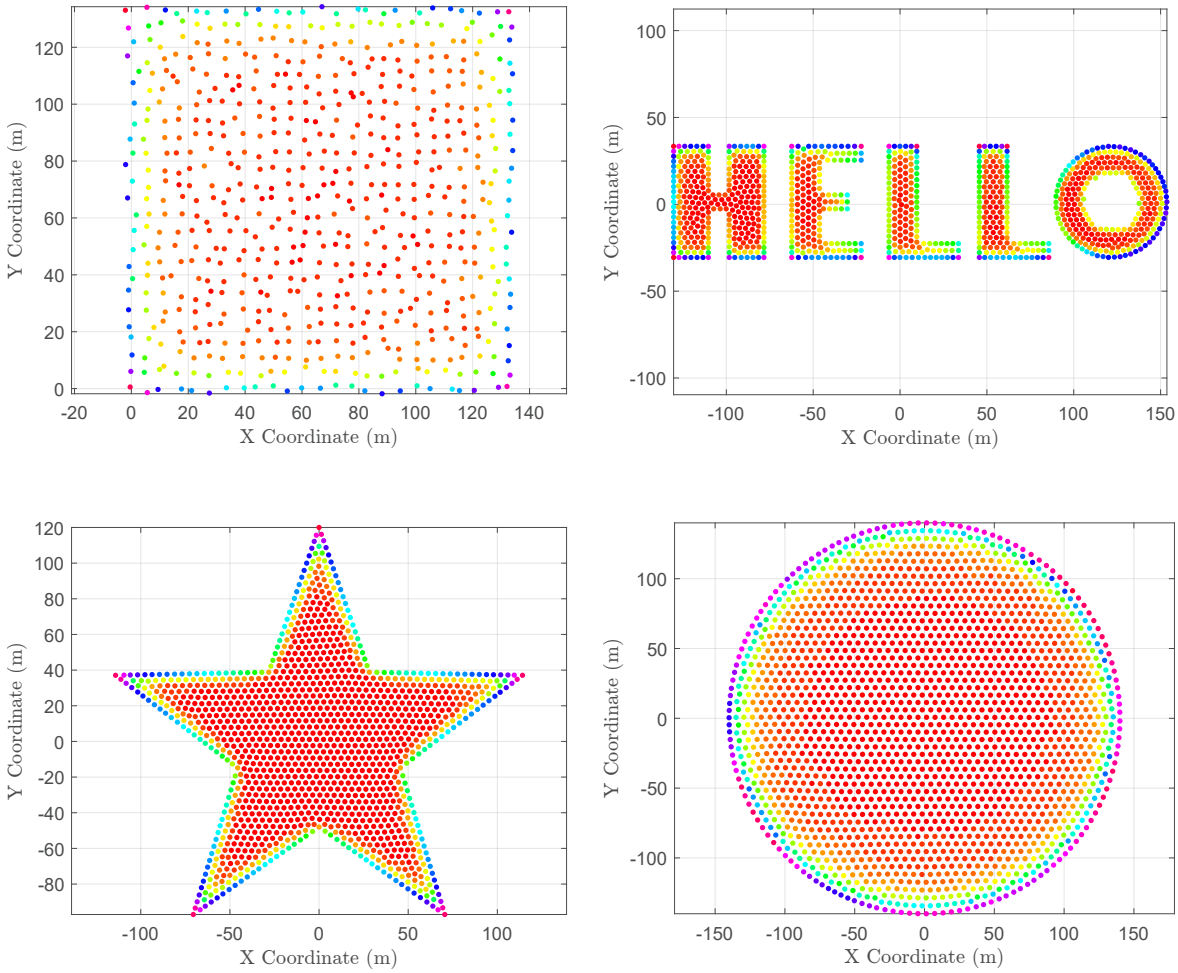

Figure 2: Illustration of borehole grouping at the final step for the four test cases (with $\delta = 10^{-1}$). Each colored dot represents a borehole assigned to a group based on thermal similarity. Case 1: Irregular rectangle layout. Case 2: HELLO layout. Case 3: Star layout. Case 4: Circle layout.

can readily incorporate the influence of groundwater flow, stratified thermal properties, or heterogeneous soil conditions. Future work will focus on integrating these processes more explicitly, further validating the framework under diverse field conditions.

Acknowledgments

The authors wish to thank the anonymous referees who reviewed this work and provided helpful and constructive comments and suggestions. The authors would like to acknowledge the funding received by the Office of Energy Research and Development (OERD) of Canada through the Energy Innovation Program (EIP) to undertake this research.

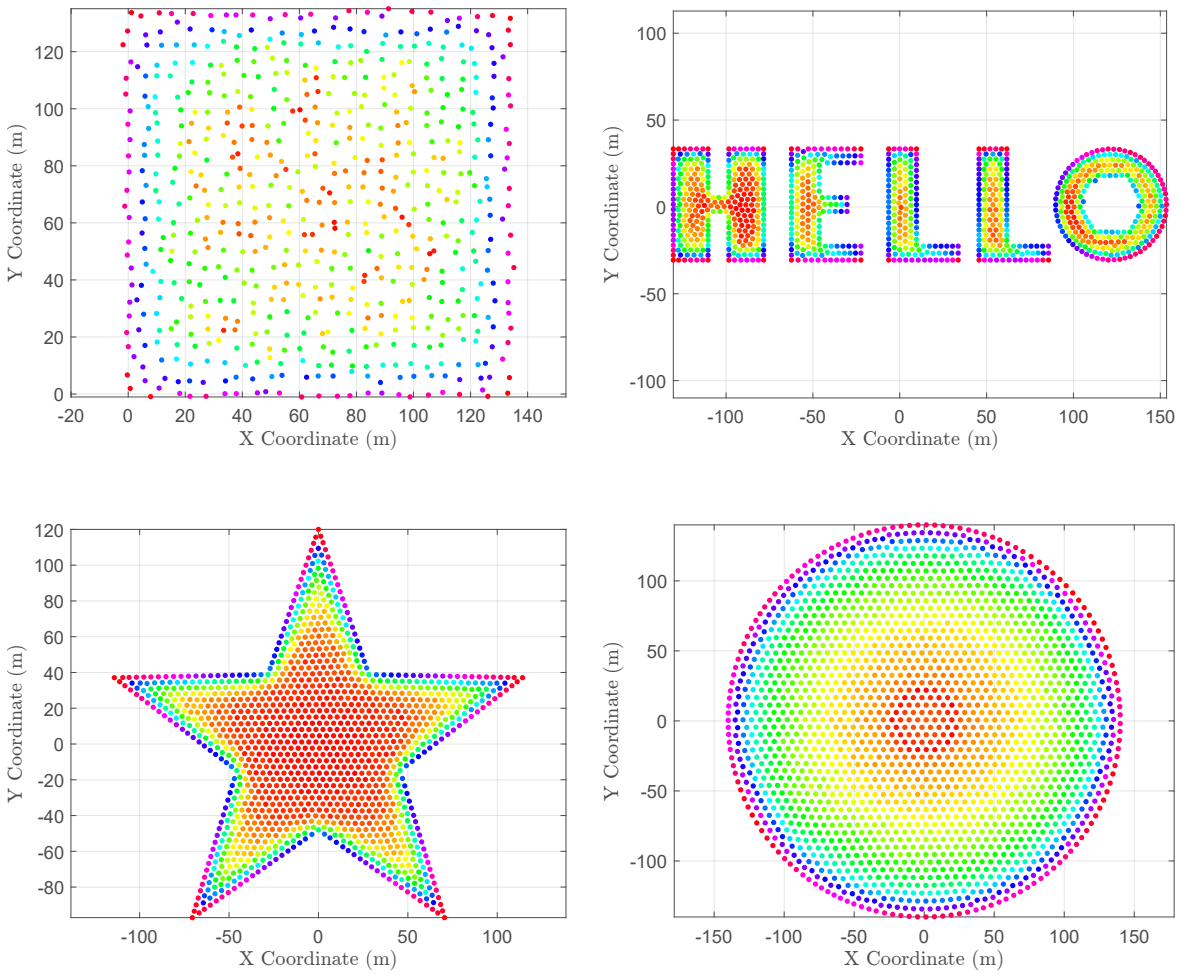

Figure 3: Illustration of borehole grouping at the final step for the four test cases (with $\delta = 10^{-3}$). Each colored dot represents a borehole assigned to a group based on thermal similarity. Case 1: Irregular rectangle layout. Case 2: HELLO layout. Case 3: Star layout. Case 4: Circle layout.

Nomenclature

g	=	Response function (mK/W)
λ	=	Thermal conductivity (W/(mK))
r	=	Radial distance from the line source (m)
t	=	Time (s)
C_s	=	Volumetric heat capacity (J/(m^3K))
D	=	Buried depth (m)
H	=	Borehole length (m)
n	=	number of sources (-)
m	=	number of groups (-)
δ	=	Similarity tolerance (-)
$\mathbf{G}$	=	Interaction tensor (mK/W)
$\overline{\mathbf{G}}_k$	=	Aggregated interaction tensor (mK/W)
$\mathbf{A}$	=	Augmented matrix (mK/W)
$\mathbf{x}$	=	Solution vector (-)
$\mathbf{u}$	=	Similarity vector (-)
$\mathbf{p}$	=	Similarity count vector (-)
$\mathbf{c}$	=	Constraint vector (mK/W)
$\mathbf{M}$	=	Mapping matrix (-)

Appendix A. Numerical example

The following section presents a numerical example illustrating the application of the proposed grouping approach to a borefield consisting of five identical, equally spaced boreholes arranged in a 5×1 configuration. Let the interaction tensor G_k be defined as:

$$\mathbf{G_k} = \begin{bmatrix} 0.0045 & 0.0017 & 0.0012 & 0.0009 & 0.0008 \\ 0.0017 & 0.0045 & 0.0017 & 0.0012 & 0.0009 \\ 0.0012 & 0.0017 & 0.0045 & 0.0017 & 0.0012 \\ 0.0009 & 0.0012 & 0.0017 & 0.0045 & 0.0017 \\ 0.0008 & 0.0009 & 0.0012 & 0.0017 & 0.0045 \end{bmatrix} \tag{A.1}$$

The solution to Equation 8 is then given by:

$$\mathbf{x_k} = [0.2409 \quad 0.1769 \quad 0.1645 \quad 0.1769 \quad 0.2409 \quad -0.0019]^T \tag{A.2}$$

For $\delta = 10^{-3}$, it follows that $m = 3$ and the corresponding grouping is described by:

$$\mathbf{u_k} = [0.1645 \quad 0.1769 \quad 0.2409]^T \tag{A.3}$$

$$\mathbf{p_k} = [1 \quad 2 \quad 2]^T \tag{A.4}$$

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & 0.5 \\ 0 & 0.5 & 0 \\ 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \tag{A.5}$$

The resulting aggregated interaction tensor $\overline{\mathbf{G}}_k$ is then given by:

$$\overline{\mathbf{G}}_k = \begin{bmatrix} 0.0045 & 0.0017 & 0.0012 \\ 0.0017 & 0.0028 & 0.0013 \\ 0.0012 & 0.0013 & 0.0026 \end{bmatrix} \tag{A.6}$$

References

[1] Eskilson, P. “Thermal analysis of heat extraction boreholes”. In: *PhD Thesis – Department of Mathematical Physics, University of Lund* (1987).

[2] Cimmino, M. “Fast calculation of the g-functions of geothermal borehole fields using similarities in the evaluation of the finite line source solution”. In: *Journal of Building Performance Simulation* 11.6 (2018), pp. 655–668.

[3] Prieto, C. and Cimmino, M. “Thermal interactions in large irregular fields of geothermal boreholes: the method of equivalent boreholes”. In: *Journal of Building Performance Simulation* 14.4 (2021), pp. 446–460.

[4] Claesson, J. and Javed, S. “An Analytical Method to Calculate Borehole Fluid Temperatures for Time-scales from Minutes to Decades.” In: *ASHRAE Transactions* 117.2 (2011).

[5] Nguyen, A. and Pasquier, P. “A successive flux estimation method for rapid g-function construction of small to large-scale ground heat exchanger”. In: *Renewable Energy* 165 (2021), pp. 359–368.

[6] Pasquier, P. and Marcotte, D. “Efficient computation of heat flux signals to ensure the reproduction of prescribed temperatures at several interacting heat sources”. In: *Applied Thermal Engineering* 59.1 (2013), pp. 515–526.

[7] Cimmino, M., Bernier, M., and Adams, F. “A contribution towards the determination of g-functions using the finite line source”. In: *Applied Thermal Engineering* 51.1-2 (2013), pp. 401–412.

[8] Lazzarotto, A. “A network-based methodology for the simulation of borehole heat storage systems”. In: *Renewable energy* 62 (2014), pp. 265–275.

[9] Marcotte, D. and Pasquier, P. “Unit-response function for ground heat exchanger with parallel, series or mixed borehole arrangement”. In: *Renewable Energy* 68 (2014), pp. 14–24.

[10] Dusseault, B., Pasquier, P., and Marcotte, D. “A block matrix formulation for efficient g-function construction”. In: *Renewable Energy* 121 (2018), pp. 249–260.